Birth weights of babies born to full-term pregnancies follow roughly a Normal distribution. At Meadowbrook Hospital, the mean weight of babies born to full-term pregnancies is 7 lbs with a standard deviation of 14 oz (1 lb = 16 oz).

1. What is the probability that the average weight of the four babies will be more than 7.5 lbs?
   A) 0.0065
   B) 0.1265
   C) 0.2839
   D) 0.4858

2. Dr. Watts (who works at Meadowbrook Hospital) has four deliveries (all for full-term pregnancies) coming up during the night. Assume that the birth weights of these four babies can be viewed as a simple random sample. What is the probability that all four babies will weigh more than 7.5 lbs?
   A) 0.0065
   B) 0.1265
   C) 0.2839
   D) 0.4858

Use the following to answer questions 3-6:

A well-known maker of jams and jellies packages its jams in jars labeled “250 mL”. The process used to fill the jars is known to dispense an amount of jam that is a Normally distributed variable with $\mu = 252$ mL and $\sigma = 0.9$ mL.

3. What proportion of the jars filled by the process will contain less than 250 mL?
   A) 0.5
   B) 0.9868
   C) 0.0068
   D) 0.0131
   E) 0
4. If changing $\sigma$ while keeping $\mu$ the same were possible for this process, what should $\sigma$ be set at so that the percentage of jars filled with less than 250 mL will be at most 0.2%?  
   A) $\sigma = 2.38$  
   B) $\sigma = 1$  
   C) $\sigma = 0.69$  
   D) $\sigma = 0.82$  
   E) Not within ± 0.2 of any of the above.

5. What percentage of jars will be filled with between 251 mL and 254 mL?  
   A) 85.3%  
   B) 1.3%  
   C) 14.7%  
   D) 13.4%  
   E) 8.5%

6. What proportion of jars will be filled with what the label claims is 250 mL?  
   A) 0.9868  
   B) 0  
   C) 0.0068  
   D) 0.0132  
   E) 0.0027

Use the following to answer question 7:

The asking prices (in thousands of dollars) for a sample of 13 houses currently on the market in Neighborville are listed below. For convenience, the data have been ordered. 

175 199 205 234 259 275 299 304 317 345 355 384 549

7. Use the $1.5 \times IQR$ rule to determine if there are any outliers present. What is/are the value(s) of the outlier(s)?  
   A) No outliers present  
   B) One outlier: 175  
   C) One outlier: 549  
   D) Two outliers: 175 and 549
8. Apply the $1.5 \times IQR$ rule to the data in the previous question to check for outlier values. In this case

A) there are no outliers
B) the value 0.75 is the only outlier
C) the values 0.75 and 2.48 are both outliers
D) the value 2.48 is the only outlier
E) the values 1.52 and 2.48 are both outliers

Use the following to answer questions 9-10:

A business has two types of employees: managers and workers. Managers earn either $100,000 or $200,000 per year. Workers earn either $10,000 or $20,000 per year. The number of male and female managers at each salary level and the number of male and female workers at each salary level are given in the table below:

<table>
<thead>
<tr>
<th></th>
<th>Managers</th>
<th>Workers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Income</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>$100,000</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>$200,000</td>
<td>20</td>
</tr>
</tbody>
</table>

9. What is the proportion of female managers who make $200,000 per year?
A) 0.1
B) 0.2
C) 0.4
D) 0.6

10. What proportion of the managers is female?
A) 0.2
B) 0.333
C) 0.5
D) 0.667
11. As Swiss cheese matures, a variety of chemical processes take place. The taste of matured cheese is related to the concentration of several chemicals in the final product. In a study of cheese in a certain region of Switzerland, samples of cheese were analyzed for lactic acid concentration and were subjected to taste tests. The numerical taste scores were obtained by combining the scores from several tasters. A scatterplot of the observed data is shown below:

![Scatterplot](scatterplot.png)

What is a plausible value for the correlation between lactic acid concentration and taste rating?
A) 0.999
B) 0.7
C) 0.07
D) -0.7

Use the following to answer question 12:

Suppose that the random variable $X$ is continuous and takes its values uniformly over the interval from 0 to 2.

12. What is the value of the probability $P\{X \leq 0.4 \text{ or } X > 1.2\}$?
A) 0.40
B) 0.20
C) 0.60
D) 0.80
E) 0.50
13. In a statistics course, a linear regression equation was computed to predict the final exam score from the score on the midterm exam. The equation of the least-squares regression line was

\[ \hat{y} = 10 + 0.9x, \]

where \( y \) represents the final exam score and \( x \) is the midterm exam score. Suppose Joe scores a 90 on the midterm exam. What would be the predicted value of his score on the final exam?
A) 81
B) 89
C) 91
D) Cannot be determined from the information given. We also need to know the correlation.

14. A population of graduate students at a large university is being studied through the use of a sample survey. The researcher is trying to estimate the mean debt carried by the students through student loans, and the proportion of all students that carry a student loan. A stratified random sample of 350 students is selected from the various schools within the university. The sample results showed that 30% held a student loan and the mean debt was $18450. Which of the following statements is TRUE?
A) The result, 30%, is a statistic and the mean debt, $18450, is the value of the parameter.
B) The mean debt held by all students is a parameter, and the 30% who hold student loans is the value of a parameter.
C) The mean debt of all students, the proportion holding a loan, the values 30% and $18450 are all examples of parameters.
D) The mean debt of all students and the proportion holding student loans are both parameters, but the values 30% and $18450 are statistics.
E) The mean debt of all students is a parameter, and its value is $18450; and the proportion holding a student loan is also a parameter whose value is 30%.

Use the following to answer question 15:

The proportion of supermarket customers who do not buy store-brand products is to be estimated.
15. Suppose 500 customers are selected from the roughly 20,000 customers who shop at the stores citywide. The sample proportion of supermarket customers who do not buy store-brand products equals 33.5%. Which value(s) can be labeled as statistic(s)?
   A) 20,000
   B) 500 and 20,000
   C) 33.5%
   D) 20,000 and 33.5%

Use the following to answer question 16:

A researcher is studying the relationship between sugar consumption and weight gain. Twelve volunteers were randomly assigned to one of two groups. The first group of five participants was put on a diet low in sugar and the second group of the remaining seven participants received 10% of their calories from sugar. After 8 weeks, weight gain was recorded from each participant.

16. Each of the 12 volunteers receives a label: 01 through 12. Use the list of random digits below to determine the labels of the first five participants selected who were put on a low-sugar diet. Start at the beginning of the list and use double-digit labels.

   81507 27102 56027 55892 33063 41842 81868 71035 09001 23367 49497

   A) 8, 1, 5, 0, 7
   B) 8, 1, 5, 7, 2
   C) 2, 6, 10, 9, 12
   D) 2, 2, 6, 10, 9

17. In a probability model, all possible outcomes together must have a probability of 1?
   A) True
   B) False

18. Which of the following cannot give informed consent to participate in a study?
   A) Children
   B) Inmates
   C) Mentally ill patients
   D) All of the above.
Use the following to answer questions 19-20:

A system has two components that operate in parallel, as shown in the diagram below. Because the components operate in parallel, at least one of the components must function properly if the system is to work properly. The probabilities of failures for the components 1 and 2 during a particular period of operation are 0.20 and 0.03, respectively. Let F denote the event that component 1 fails during this period of operation and G denote the event that component 2 fails during this period of operation. The component failures are assumed to be independent.

19. Which event corresponds to the event that the above system fails during this particular period of operation?
   A) F and G
   B) F or G
   C) Fc and Gc
   D) Fc or Gc

20. What is the probability that the system fails during this particular period of operation?
   A) 0.006
   B) 0.060
   C) 0.224
   D) 0.230

21. Using the standard Normal distribution tables, what is the area under the standard Normal curve corresponding to \(-0.5 < Z < 1.2\)?
   A) 0.3085
   B) 0.8849
   C) 0.5764
   D) 0.2815

22. Confidence intervals and two-sided significance tests are linked in the sense that a two-sided test at a significance level \(\alpha\) can be carried in the form of a confidence interval with confidence level \(1 - \alpha\).
   A) True
   B) False
Use the following to answer questions 23-24:

John's parents recorded his height at various ages between 36 and 66 months. Below is a record of the results:

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>36</th>
<th>48</th>
<th>54</th>
<th>60</th>
<th>66</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height (inches)</td>
<td>34</td>
<td>38</td>
<td>41</td>
<td>43</td>
<td>45</td>
</tr>
</tbody>
</table>

23. Which of the following is the equation of the least-squares regression line of John's height on age? *(Note: You do not need to directly calculate the least-squares regression line to answer this question.)*
   A) Height = 12 \times (Age)  
   B) Height = Age/12  
   C) Height = 60 - 0.22 \times (Age)  
   D) Height = 22.3 + 0.34 \times (Age)

24. John's parents decide to use the least-squares regression line of John's height on age to predict his height at age 21 years (252 months). What conclusion can we draw?
   A) John's height, in inches, should be about half his age, in months.  
   B) The parents will get a fairly accurate estimate of his height at age 21 years, because the data are clearly correlated.  
   C) Such a prediction could be misleading, because it involves extrapolation.  
   D) All of the above.

25. The probability of an event is defined as the proportion of times an event occurs occasionally.
   A) True  
   B) False
26. The following data are the magnitudes of earthquakes around the world recorded on January 13, 2008:

   4.1  4.8  3.1  5.3  5.1  4.7  3.0  2.9  U  4.6  3.1  3.0  U  2.5.

   Because of equipment problems, two earthquakes were unrecorded although it was known that both had a value less than 2.7. The median magnitude of earthquakes on this date is
   A) 2.95
   B) 3.9
   C) 4.0
   D) 3.1
   E) Cannot determine without knowing the exact value of the missing magnitudes.

Use the following to answer questions 27-28:

The breaking strength of yarn used in the manufacture of woven carpet material is Normally distributed with $\sigma = 2.4$ psi. A random sample of 16 specimens of yarn from a production run were measured for breaking strength, and based on the mean of the sample, $\overline{x}$, a confidence interval was found to be (128.7, 131.3).

27. What is the confidence level, $C$, of this interval?
   A) 0.95
   B) 0.99
   C) 0.90
   D) 0.97
   E) Unable to determine with the information provided.

28. If we wanted to have a greater degree of confidence in the calculated interval, what would have to happen to its length?
   A) The interval would be shorter.
   B) The margin of error would decrease.
   C) The confidence level would have a smaller value.
   D) The sample size would go up.
   E) The interval would increase in length.
Use the following to answer questions 29-31:

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 50 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.

29. Suppose volunteers were first divided by gender, and then half of the men were randomly assigned to the new drug and half of the women were assigned to the new drug. The remaining volunteers received the other drug. What is this an example of?
   A) Replication.
   B) Confounding. The effects of gender will be mixed up with the effects of the drugs.
   C) A block design.
   D) A matched-pairs design.

30. In which situation would this study be double-blind?
   A) Neither drug had any identifying marks on it.
   B) All volunteers were not allowed to see the psychiatrist nor was the psychiatrist allowed to see the volunteers during the session when the psychiatrist evaluated the severity of the depression.
   C) Neither the volunteers nor the psychiatrist knew which treatment any person had received.
   D) All of the above.

31. What is the explanatory variable or factor in this study?
   A) Which drug the volunteers receive.
   B) The use of randomization and the fact that this was a comparative study.
   C) The extent to which the depression was reduced.
   D) The use of a psychiatrist to evaluate the severity of depression.
Use the following to answer question 32:

In the National Hockey League a good predictor of the percentage of games won by a team is the number of goals the team allows during the season. Data were gathered for all 30 teams in the NHL and the scatterplot of their Winning Percentage against the number of Goals Allowed in the 2006/2007 season with a fitted least-squares regression line is provided:

The least-squares regression line and $r^2$ were calculated to be
Winning Percent (%) = 116.95 - 0.26 x Goals Allowed
$r^2 = 0.69$

32. Which of the following provides the best interpretation of the slope of the regression line?
   A) If the Winning Percent increases by 1%, then the number of Goals Allowed decreases by 0.26.
   B) If a team were to allow 100 goals during the season, their Winning % would be 90.95%.
   C) If Goals Allowed increases by one goal, the Winning % increases by 0.26%.
   D) If the Winning % increases by 1%, then the number of Goals Allowed increases by 0.26.
   E) If Goals Allowed increases by one goal, the Winning % decreases by 0.26%.

Use the following to answer question 33:

The distribution of GPA scores is known to be left-skewed. At a large university, an English professor is interested in learning about the average GPA score of the English majors and minors. A simple random sample of 75 junior and senior English majors and minors results in an average GPA score of 2.97. Assume that the distribution of GPA scores for all English majors and minors at this university is also left-skewed with a standard deviation of 0.62.
33. Calculate a 95% confidence interval for the mean GPA of the junior and senior English majors and minors.
   A) (2.786, 3.154)
   B) (2.830, 3.110)
   C) (2.852, 3.088)
   D) (2.954, 2.986)

Use the following to answer questions 34-37:

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

34. Define event $B = \{\text{at least one of the next two babies is a boy}\}$. What is the probability of the complement of event $B$?
   A) 0.125
   B) 0.250
   C) 0.375
   D) 0.500

35. Define events $A = \{\text{the next two babies are boys}\}$ and $B = \{\text{at least one of the next two babies is a boy}\}$. What do we know about events $A$ and $B$?
   A) They are disjoint.
   B) They are complements.
   C) They are independent.
   D) None of the above.

36. What is the probability that at least one of the next three babies is a boy?
   A) 0.125
   B) 0.333
   C) 0.750
   D) 0.875

37. What is the probability that the next three babies are of the same sex?
   A) 0.125
   B) 0.250
   C) 0.375
   D) 0.500
Use the following to answer questions 38-40:

The weight of medium-size tomatoes selected at random from a bin at the local supermarket is a random variable with mean \( \mu = 10 \) oz and standard deviation \( \sigma = 1 \) oz.

38. Suppose we pick four tomatoes from the bin at random and put them in a bag. Define the random variable \( Y \) = the weight of the bag containing the four tomatoes. What is the mean of the random variable \( Y \)?
   A) \( \mu_Y = 2.5 \) oz
   B) \( \mu_Y = 4 \) oz
   C) \( \mu_Y = 10 \) oz
   D) \( \mu_Y = 40 \) oz

39. Suppose we pick two tomatoes at random from the bin. Let the random variable \( V \) = the difference in the weights of the two tomatoes selected (the weight of the first tomato minus the weight of the second tomato). What is the standard deviation of the random variable \( V \)?
   A) \( \sigma_V = 0.00 \) oz
   B) \( \sigma_V = 1.00 \) oz
   C) \( \sigma_V = 1.41 \) oz
   D) \( \sigma_V = 2.00 \) oz

40. Suppose we pick four tomatoes from the bin at random and put them in a bag. Define the random variable \( Y \) = the weight of the bag containing the four tomatoes. What is the standard deviation of the random variable \( Y \)?
   A) \( \sigma_Y = 0.50 \) oz
   B) \( \sigma_Y = 1.0 \) oz
   C) \( \sigma_Y = 2.0 \) oz
   D) \( \sigma_Y = 4.0 \) oz
A study was conducted in a large population of adults concerning eyeglasses for correcting reading vision. Based on an examination by a qualified professional, the individuals were judged as to whether or not they needed to wear glasses for reading. In addition it was determined whether or not they were currently using glasses for reading. The following table provides the proportions found in the study:

<table>
<thead>
<tr>
<th>Judged to need glasses</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.42</td>
<td>0.18</td>
</tr>
<tr>
<td>No</td>
<td>0.04</td>
<td>0.36</td>
</tr>
</tbody>
</table>

41. What is the probability that the selected adult is judged to need eyeglasses but does not use them for reading?
   A) 0.42  
   B) 0.18  
   C) 0.54  
   D) 0.60  
   E) 0.36

42. If a single adult is selected at random from this large population, what is the probability that the adult is judged to need eyeglasses for reading?
   A) 0.46  
   B) 0.42  
   C) 0.78  
   D) 0.60  
   E) 0.40

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

43. What is $P(X > $400)?
   A) 0.2119  
   B) 0.2881  
   C) 0.7881  
   D) 0.8450  
   E) 0.40
44. The probability is approximately 0.6 that on a randomly selected day the store will make less than how much?
   A) $0.30
   B) $347.40
   C) $361.30
   D) $372.60

45. An engineer has designed an improved light bulb. The previous design had an average lifetime of 1200 hours. Based on a sample of 2000 of these new bulbs, the average lifetime was found to be 1201 hours. Although the difference is quite small, the effect was statistically significant. What is the best explanation?
   A) New designs typically have more variability than standard designs.
   B) The sample size is very large, so that even a small difference can be detected.
   C) The mean of 1200 is large.
   D) The relative improvement in average lifetime is 0.000083, which is much smaller than 0.05.

Use the following to answer questions 46-47:

A coin is about to be tossed multiple times. Assume the coin is fair, i.e., the probability of heads and the probability of tails are both 0.5.

46. If the coin is tossed six times, what is the probability that less than ½ of the tosses are heads?
   A) 0.0049
   B) 0.094
   C) 0.109
   D) 0.344

47. If the coin is tossed 60 times, what is the probability that less than ½ of the tosses are heads?
   Hint: Use the Normal Approximation with no continuity assumption.
   A) 0.0049
   B) 0.094
   C) 0.109
   D) 0.344
48. Two variables are confounded when their effects on a response variable can be distinguished from each other.
A) True
B) False

Use the following to answer questions 49-51:

It is known that driving can be difficult in regions where winter conditions involve snow-covered roads. For cars equipped with all-season tires traveling at 90 km/hr, the mean stopping time in fresh snow is known to be 215 meters with a standard deviation of σ = 2.5 meters. It is often advocated that automobiles in such areas should be equipped with special tires to compensate for such conditions, especially with respect to stopping distance. A manufacturer of tires made for driving in fresh snow claims that vehicles equipped with their tires have a decreased stopping distance. A study was done using a random sample of 9 snow tires from the manufacturer on a snow-covered test track. The tests resulted in a mean stopping distance of \( \bar{x} = 212.9 \) meters.

49. What is the P-value?
A) 0.012
B) 0.050
C) 0.025
D) 0.006

50. Using the sample results and assuming that stopping distance is a Normally distributed random variable, what is the value of the test statistic?
A) 0.05
B) -2.52
C) -9.36
D) -1.04
E) -1.96

51. What are the appropriate null and alternative hypotheses to test the manufacturer's claim?
A) \( H_0: \mu = 215 \) against \( H_a: \mu > 215 \)
B) \( H_0: \mu = 215 \) against \( H_a: \mu \neq 215 \)
C) \( H_0: \mu = 215 \) against \( H_a: \mu < 215 \)
D) \( H_0: \bar{x} = 215 \) against \( H_a: \bar{x} < 215 \)
E) Snow tires decrease the stopping distance in loose snow.
A call-in poll conducted by USA Today concluded that Americans love Donald Trump. This conclusion was based on data collected from 7800 calls made by USA Today readers.

52. USA Today reported that of the 7800 calls, 6435 calls were supportive of Donald Trump. This results in a percentage of 82.5%. Of the 6435 supportive calls, about half came from female callers. Which value(s) can be labeled as statistics?
   A) 7800 and 6435
   B) 6435 and 82.5%
   C) 6435, 82.5%, and 50%
   D) 82.5% and 50%

53. What sampling technique is being used?
   A) Simple random sampling
   B) Stratified random sampling
   C) Volunteer sampling
   D) Convenience sampling

54. USA Today later reported that 5640 of the 7800 calls for the poll came from the offices owned by one man, Cincinnati financier Carl Lindner, who is a friend of Donald Trump. What can we conclude about the results of this poll?
   A) They are surprising, but reliable, because it was conducted by a nationally recognized organization.
   B) They are biased, but only slightly because the sample size was quite large.
   C) They are biased understating the popularity of Donald Trump.
   D) They are biased overstating the popularity of Donald Trump.

Use the following to answer questions 55-56:

A population has a distribution with mean $\mu = 50$ and variance $\sigma^2 = 225$. From this population a simple random sample of $n$ observations is to be selected and the mean $\bar{x}$ of the sample values calculated.
55. How big must the sample size $n$ be so that the standard deviation of the sample mean, $\bar{x}$, is equal to 1.4, i.e., $\sigma_{\bar{x}} = 1.4$?
   A) $n = 11$
   B) $n = 161$
   C) $n = 115$
   D) $n = 36$
   E) $n = 21$

56. If the population variable is known to be Normally distributed and the sample size used is to be $n = 16$, what is the probability that the sample mean will be between 48.35 and 55.74, i.e., $P(48.35 \leq \bar{x} \leq 55.74)$?
   A) 0.393
   B) 0.607
   C) 0.937
   D) 0.330
   E) Not within ± 0.010 of any of the above.

57. Influential outliers are easy to detect since the residuals will always be very large compared to the residuals of the other observations.
   A) True
   B) False

58. Confidentiality and anonymity is the same thing when referring to experimental design?
   A) True
   B) False

59. A valid probability is any number between −1 and 1.
   A) True
   B) False

60. Typically, the null hypothesis in a test of significance is a statement of “no difference” in the true means.
   A) True
   B) False
The distribution of GPA scores is known to be left-skewed. At a large university, an English professor is interested in learning about the average GPA score of the English majors and minors. A simple random sample of 75 junior and senior English majors and minors results in an average GPA score of 2.97. Assume that the distribution of GPA scores for all English majors and minors at this university is also left-skewed with a standard deviation of 0.62.

1. Calculate a 95% confidence interval for the mean GPA of the junior and senior English majors and minors.
   A) (2.786, 3.154)
   B) (2.830, 3.110)
   C) (2.852, 3.088)
   D) (2.954, 2.986)

The proportion of supermarket customers who do not buy store-brand products is to be estimated.

2. Suppose 500 customers are selected from the roughly 20,000 customers who shop at the stores citywide. The sample proportion of supermarket customers who do not buy store-brand products equals 33.5%. Which value(s) can be labeled as statistic(s)?
   A) 20,000
   B) 500 and 20,000
   C) 33.5%
   D) 20,000 and 33.5%

A call-in poll conducted by USA Today concluded that Americans love Donald Trump. This conclusion was based on data collected from 7800 calls made by USA Today readers.
3. *USA Today* later reported that 5640 of the 7800 calls for the poll came from the offices owned by one man, Cincinnati financier Carl Lindner, who is a friend of Donald Trump. What can we conclude about the results of this poll?
   A) They are surprising, but reliable, because it was conducted by a nationally recognized organization.
   B) They are biased, but only slightly because the sample size was quite large.
   C) They are biased understating the popularity of Donald Trump.
   D) They are biased overstating the popularity of Donald Trump.

4. *USA Today* reported that of the 7800 calls, 6435 calls were supportive of Donald Trump. This results in a percentage of 82.5%. Of the 6435 supportive calls, about half came from female callers. Which value(s) can be labeled as statistics?
   A) 7800 and 6435
   B) 6435 and 82.5%
   C) 6435, 82.5%, and 50%
   D) 82.5% and 50%

5. What sampling technique is being used?
   A) Simple random sampling
   B) Stratified random sampling
   C) Volunteer sampling
   D) Convenience sampling

6. John's parents decide to use the least-squares regression line of John's height on age to predict his height at age 21 years (252 months). What conclusion can we draw?
   A) John's height, in inches, should be about half his age, in months.
   B) The parents will get a fairly accurate estimate of his height at age 21 years, because the data are clearly correlated.
   C) Such a prediction could be misleading, because it involves extrapolation.
   D) All of the above.
7. Which of the following is the equation of the least-squares regression line of John's height on age? *(Note: You do not need to directly calculate the least-squares regression line to answer this question.)*  
A) Height = 12 x (Age)  
B) Height = Age/12  
C) Height = 60 – 0.22 x (Age)  
D) Height = 22.3 + 0.34 x (Age)

8. Influential outliers are easy to detect since the residuals will always be very large compared to the residuals of the other observations.  
A) True  
B) False

9. The following data are the magnitudes of earthquakes around the world recorded on January 13, 2008:  
4.1 4.8 3.1 5.3 5.1 4.7 3.0 U 4.6 3.1 3.0 U 2.5.  
Because of equipment problems, two earthquakes were unrecorded although it was known that both had a value less than 2.7. The median magnitude of earthquakes on this date is  
A) 2.95  
B) 3.9  
C) 4.0  
D) 3.1  
E) Cannot determine without knowing the exact value of the missing magnitudes.

10. Confidentiality and anonymity is the same thing when referring to experimental design?  
A) True  
B) False

11. Two variables are confounded when their effects on a response variable can be distinguished from each other.  
A) True  
B) False

12. In a probability model, all possible outcomes together must have a probability of 1?  
A) True  
B) False
13. In a statistics course, a linear regression equation was computed to predict the final exam score from the score on the midterm exam. The equation of the least-squares regression line was

\[ \hat{y} = 10 + 0.9x, \]

where \( y \) represents the final exam score and \( x \) is the midterm exam score. Suppose Joe scores a 90 on the midterm exam. What would be the predicted value of his score on the final exam?
A) 81  
B) 89  
C) 91  
D) Cannot be determined from the information given. We also need to know the correlation.

14. A valid probability is any number between -1 and 1.
A) True 
B) False

Use the following to answer questions 15-17:

It is known that driving can be difficult in regions where winter conditions involve snow-covered roads. For cars equipped with all-season tires traveling at 90 km/hr, the mean stopping time in fresh snow is known to be 215 meters with a standard deviation of \( \sigma = 2.5 \) meters. It is often advocated that automobiles in such areas should be equipped with special tires to compensate for such conditions, especially with respect to stopping distance. A manufacturer of tires made for driving in fresh snow claims that vehicles equipped with their tires have a decreased stopping distance. A study was done using a random sample of 9 snow tires from the manufacturer on a snow-covered test track. The tests resulted in a mean stopping distance of \( \bar{x} = 212.9 \) meters.

15. Using the sample results and assuming that stopping distance is a Normally distributed random variable, what is the value of the test statistic?
A) 0.05  
B) -2.52  
C) -9.36  
D) -1.04  
E) -1.96
16. What is the P-value?
   A) 0.012
   B) 0.050
   C) 0.025
   D) 0.006

17. What are the appropriate null and alternative hypotheses to test the manufacturer's claim?
   A) Ho: µ = 215 against Ha: µ > 215
   B) Ho: µ = 215 against Ha: µ ≠ 215
   C) Ho: µ = 215 against Ha: µ < 215
   D) Ho: X̄ = 215 against Ha: X̄ < 215
   E) Snow tires decrease the stopping distance in loose snow.

Use the following to answer questions 18-19:

Birth weights of babies born to full-term pregnancies follow roughly a Normal distribution. At Meadowbrook Hospital, the mean weight of babies born to full-term pregnancies is 7 lbs with a standard deviation of 14 oz (1 lb = 16 oz).

18. Dr. Watts (who works at Meadowbrook Hospital) has four deliveries (all for full-term pregnancies) coming up during the night. Assume that the birth weights of these four babies can be viewed as a simple random sample. What is the probability that all four babies will weigh more than 7.5 lbs?
   A) 0.0065
   B) 0.1265
   C) 0.2839
   D) 0.4858

19. What is the probability that the average weight of the four babies will be more than 7.5 lbs?
   A) 0.0065
   B) 0.1265
   C) 0.2839
   D) 0.4858
20. Confidence intervals and two-sided significance tests are linked in the sense that a two-sided test at a significance level \( \alpha \) can be carried in the form of a confidence interval with confidence level \( 1 - \alpha \).
A) True
B) False

Use the following to answer questions 21-22:

A study was conducted in a large population of adults concerning eyeglasses for correcting reading vision. Based on an examination by a qualified professional, the individuals were judged as to whether or not they needed to wear glasses for reading. In addition it was determined whether or not they were currently using glasses for reading. The following table provides the proportions found in the study:

<table>
<thead>
<tr>
<th>Judged to need glasses</th>
<th>Used glasses for reading</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.42</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0.04</td>
<td>0.36</td>
<td></td>
</tr>
</tbody>
</table>

21. What is the probability that the selected adult is judged to need eyeglasses but does not use them for reading?
A) 0.42
B) 0.18
C) 0.54
D) 0.60
E) 0.36

22. If a single adult is selected at random from this large population, what is the probability that the adult is judged to need eyeglasses for reading?
A) 0.46
B) 0.42
C) 0.78
D) 0.60
E) 0.40

Use the following to answer questions 23-24:

Let the random variable \( X \) represent the profit made on a randomly selected day by a small clothing store on Main Street. Assume \( X \) is Normal with a mean of $360 and a standard deviation of $50.
23. What is \( P(X > 400) \)?
   A) 0.2119
   B) 0.2881
   C) 0.7881
   D) 0.8450

24. The probability is approximately 0.6 that on a randomly selected day the store will make less than how much?
   A) $0.30
   B) $347.40
   C) $361.30
   D) $372.60

25. The probability of an event is defined as the proportion of times an event occurs occasionally.
   A) True
   B) False

26. Using the standard Normal distribution tables, what is the area under the standard Normal curve corresponding to \(-0.5 < Z < 1.2\)?
   A) 0.3085
   B) 0.8849
   C) 0.5764
   D) 0.2815

27. Which of the following cannot give informed consent to participate in a study?
   A) Children
   B) Inmates
   C) Mentally ill patients
   D) All of the above.

Use the following to answer question 28:

Suppose that the random variable \( X \) is continuous and takes its values uniformly over the interval from 0 to 2.
28. What is the value of the probability \( P(X \leq 0.4 \text{ or } X > 1.2) \)?

A) 0.40  
B) 0.20  
C) 0.60  
D) 0.80  
E) 0.50

29. As Swiss cheese matures, a variety of chemical processes take place. The taste of matured cheese is related to the concentration of several chemicals in the final product. In a study of cheese in a certain region of Switzerland, samples of cheese were analyzed for lactic acid concentration and were subjected to taste tests. The numerical taste scores were obtained by combining the scores from several tasters. A scatterplot of the observed data is shown below:

What is a plausible value for the correlation between lactic acid concentration and taste rating?

A) 0.999  
B) 0.7  
C) 0.07  
D) -0.7
30. Apply the $1.5 \times IQR$ rule to the data in the previous question to check for outlier values. In this case ______________.
   A) there are no outliers
   B) the value 0.75 is the only outlier
   C) the values 0.75 and 2.48 are both outliers
   D) the value 2.48 is the only outlier
   E) the values 1.52 and 2.48 are both outliers

Use the following to answer question 31:

The asking prices (in thousands of dollars) for a sample of 13 houses currently on the market in Neighborville are listed below. For convenience, the data have been ordered.

175 199 205 234 259 275 299 304 317 345 355 384 549

31. Use the $1.5 \times IQR$ rule to determine if there are any outliers present. What is/are the value(s) of the outlier(s)?
   A) No outliers present
   B) One outlier: 175
   C) One outlier: 549
   D) Two outliers: 175 and 549
In the National Hockey League a good predictor of the percentage of games won by a team is the number of goals the team allows during the season. Data were gathered for all 30 teams in the NHL and the scatterplot of their Winning Percentage against the number of Goals Allowed in the 2006/2007 season with a fitted least-squares regression line is provided:

The least-squares regression line and \( r^2 \) were calculated to be

\[
\text{Winning Percent (\%)} = 116.95 - 0.26 \times \text{Goals Allowed}
\]

\( r^2 = 0.69 \)

32. Which of the following provides the best interpretation of the slope of the regression line?
   A) If the Winning Percent increases by 1%, then the number of Goals Allowed decreases by 0.26.
   B) If a team were to allow 100 goals during the season, their Winning % would be 90.95%.
   C) If Goals Allowed increases by one goal, the Winning % increases by 0.26%.
   D) If the Winning % increases by 1%, then the number of Goals Allowed increases by 0.26.
   E) If Goals Allowed increases by one goal, the Winning % decreases by 0.26%.

33. Typically, the null hypothesis in a test of significance is a statement of “no difference” in the true means.
   A) True
   B) False
Use the following to answer questions 34-35:

A system has two components that operate in parallel, as shown in the diagram below. Because the components operate in parallel, at least one of the components must function properly if the system is to work properly. The probabilities of failures for the components 1 and 2 during a particular period of operation are 0.20 and 0.03, respectively. Let $F$ denote the event that component 1 fails during this period of operation and $G$ denote the event that component 2 fails during this period of operation. The component failures are assumed to be independent.

34. Which event corresponds to the event that the above system fails during this particular period of operation?
   A) $F$ and $G$
   B) $F$ or $G$
   C) $F^c$ and $G^c$
   D) $F^c$ or $G^c$

35. What is the probability that the system fails during this particular period of operation?
   A) 0.006
   B) 0.060
   C) 0.224
   D) 0.230

Use the following to answer questions 36-37:

A coin is about to be tossed multiple times. Assume the coin is fair, i.e., the probability of heads and the probability of tails are both 0.5.

36. If the coin is tossed 60 times, what is the probability that less than $\frac{1}{2}$ of the tosses are heads?
   Hint: Use the Normal Approximation with no continuity assumption.
   A) 0.0049
   B) 0.094
   C) 0.109
   D) 0.344
37. If the coin is tossed six times, what is the probability that less than \( \frac{1}{2} \) of the tosses are heads?
   A) 0.0049
   B) 0.094
   C) 0.109
   D) 0.344

Use the following to answer questions 38-39:

The breaking strength of yarn used in the manufacture of woven carpet material is Normally distributed with \( \sigma = 2.4 \) psi. A random sample of 16 specimens of yarn from a production run were measured for breaking strength, and based on the mean of the sample, \( \bar{x} \), a confidence interval was found to be \((128.7, 131.3)\).

38. What is the confidence level, \( C \), of this interval?
   A) 0.95
   B) 0.99
   C) 0.90
   D) 0.97
   E) Unable to determine with the information provided.

39. If we wanted to have a greater degree of confidence in the calculated interval, what would have to happen to its length?
   A) The interval would be shorter.
   B) The margin of error would decrease.
   C) The confidence level would have a smaller value.
   D) The sample size would go up.
   E) The interval would increase in length.

Use the following to answer questions 40-43:

A well-known maker of jams and jellies packages its jams in jars labeled “250 mL”. The process used to fill the jars is known to dispense an amount of jam that is a Normally distributed variable with \( \mu = 252 \) mL and \( \sigma = 0.9 \) mL.
40. What proportion of the jars filled by the process will contain less than 250 mL?
   A) 0.5
   B) 0.9868
   C) 0.0068
   D) 0.0131
   E) 0

41. If changing $\sigma$ while keeping $\mu$ the same were possible for this process, what should $\sigma$ be set at so that the percentage of jars filled with less than 250 mL will be at most 0.2%?
   A) $\sigma = 2.38$
   B) $\sigma = 1$
   C) $\sigma = 0.69$
   D) $\sigma = 0.82$
   E) Not within $\pm 0.2$ of any of the above.

42. What percentage of jars will be filled with between 251 mL and 254 mL?
   A) 85.3%
   B) 1.3%
   C) 14.7%
   D) 13.4%
   E) 8.5%

43. What proportion of jars will be filled with what the label claims is 250 mL?
   A) 0.9868
   B) 0
   C) 0.0068
   D) 0.0132
   E) 0.0027

Use the following to answer questions 44-45:

A business has two types of employees: managers and workers. Managers earn either $100,000 or $200,000 per year. Workers earn either $10,000 or $20,000 per year. The number of male and female managers at each salary level and the number of male and female workers at each salary level are given in the table below:

<table>
<thead>
<tr>
<th>Income</th>
<th>Gender</th>
<th>Managers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>$100,000</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>$200,000</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Income</th>
<th>Gender</th>
<th>Workers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>$10,000</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>$20,000</td>
<td>20</td>
<td>80</td>
</tr>
</tbody>
</table>
44. What is the proportion of female managers who make $200,000 per year?
   A) 0.1
   B) 0.2
   C) 0.4
   D) 0.6

45. What proportion of the managers is female?
   A) 0.2
   B) 0.333
   C) 0.5
   D) 0.667

Use the following to answer questions 46-47:

A population has a distribution with mean $\mu = 50$ and variance $\sigma^2 = 225$. From this population a simple random sample of $n$ observations is to be selected and the mean $\bar{x}$ of the sample values calculated.

46. How big must the sample size $n$ be so that the standard deviation of the sample mean, $\bar{x}$, is equal to 1.4, i.e., $\sigma_{\bar{x}} = 1.4$?
   A) $n = 11$
   B) $n = 161$
   C) $n = 115$
   D) $n = 36$
   E) $n = 21$

47. If the population variable is known to be Normally distributed and the sample size used is to be $n = 16$, what is the probability that the sample mean will be between 48.35 and 55.74, i.e., $P(48.35 \leq \bar{x} \leq 55.74)$?
   A) 0.393
   B) 0.607
   C) 0.937
   D) 0.330
   E) Not within $\pm 0.010$ of any of the above.
Use the following to answer question 48:

A researcher is studying the relationship between sugar consumption and weight gain. Twelve volunteers were randomly assigned to one of two groups. The first group of five participants was put on a diet low in sugar and the second group of the remaining seven participants received 10% of their calories from sugar. After 8 weeks, weight gain was recorded from each participant.

48. Each of the 12 volunteers receives a label: 01 through 12. Use the list of random digits below to determine the labels of the first five participants selected who were put on a low-sugar diet. Start at the beginning of the list and use double-digit labels.

81507 27102 56027 55892 33063 41842 81868 71035 09001 23367 49497

A) 8, 1, 5, 0, 7  
B) 8, 1, 5, 7, 2  
C) 2, 6, 10, 9, 12  
D) 2, 2, 6, 10, 9

Use the following to answer questions 49-51:

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 50 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.

49. In which situation would this study be double-blind?
   A) Neither drug had any identifying marks on it.  
   B) All volunteers were not allowed to see the psychiatrist nor was the psychiatrist allowed to see the volunteers during the session when the psychiatrist evaluated the severity of the depression.  
   C) Neither the volunteers nor the psychiatrist knew which treatment any person had received.  
   D) All of the above.

50. What is the explanatory variable or factor in this study?
   A) Which drug the volunteers receive.  
   B) The use of randomization and the fact that this was a comparative study.  
   C) The extent to which the depression was reduced.  
   D) The use of a psychiatrist to evaluate the severity of depression.
51. Suppose volunteers were first divided by gender, and then half of the men were randomly assigned to the new drug and half of the women were assigned to the new drug. The remaining volunteers received the other drug. What is this an example of?
A) Replication.
B) Confounding. The effects of gender will be mixed up with the effects of the drugs.
C) A block design.
D) A matched-pairs design.

Use the following to answer questions 52-54:

The weight of medium-size tomatoes selected at random from a bin at the local supermarket is a random variable with mean $\mu = 10$ oz and standard deviation $\sigma = 1$ oz.

52. Suppose we pick four tomatoes from the bin at random and put them in a bag. Define the random variable $Y =$ the weight of the bag containing the four tomatoes. What is the standard deviation of the random variable $Y$?
A) $\sigma_Y = 0.50$ oz
B) $\sigma_Y = 1.0$ oz
C) $\sigma_Y = 2.0$ oz
D) $\sigma_Y = 4.0$ oz

53. Suppose we pick four tomatoes from the bin at random and put them in a bag. Define the random variable $Y =$ the weight of the bag containing the four tomatoes. What is the mean of the random variable $Y$?
A) $\mu_Y = 2.5$ oz
B) $\mu_Y = 4$ oz
C) $\mu_Y = 10$ oz
D) $\mu_Y = 40$ oz

54. Suppose we pick two tomatoes at random from the bin. Let the random variable $V =$ the difference in the weights of the two tomatoes selected (the weight of the first tomato minus the weight of the second tomato). What is the standard deviation of the random variable $V$?
A) $\sigma_V = 0.00$ oz
B) $\sigma_V = 1.00$ oz
C) $\sigma_V = 1.41$ oz
D) $\sigma_V = 2.00$ oz
55. An engineer has designed an improved light bulb. The previous design had an average lifetime of 1200 hours. Based on a sample of 2000 of these new bulbs, the average lifetime was found to be 1201 hours. Although the difference is quite small, the effect was statistically significant. What is the best explanation?
A) New designs typically have more variability than standard designs.
B) The sample size is very large, so that even a small difference can be detected.
C) The mean of 1200 is large.
D) The relative improvement in average lifetime is 0.000083, which is much smaller than 0.05.

Use the following to answer questions 56-59:

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

56. Define event $B = \{\text{at least one of the next two babies is a boy}\}$. What is the probability of the complement of event $B$?
A) 0.125
B) 0.250
C) 0.375
D) 0.500

57. Define events $A = \{\text{the next two babies are boys}\}$ and $B = \{\text{at least one of the next two babies is a boy}\}$. What do we know about events $A$ and $B$?
A) They are disjoint.
B) They are complements.
C) They are independent.
D) None of the above.

58. What is the probability that the next three babies are of the same sex?
A) 0.125
B) 0.250
C) 0.375
D) 0.500

59. What is the probability that at least one of the next three babies is a boy?
A) 0.125
B) 0.333
C) 0.750
D) 0.875
60. A population of graduate students at a large university is being studied through the use of a sample survey. The researcher is trying to estimate the mean debt carried by the students through student loans, and the proportion of all students that carry a student loan. A stratified random sample of 350 students is selected from the various schools within the university. The sample results showed that 30% held a student loan and the mean debt was $18450. Which of the following statements is TRUE?

A) The result, 30%, is a statistic and the mean debt, $18450, is the value of the parameter.

B) The mean debt held by all students is a parameter, and the 30% who hold student loans is the value of a parameter.

C) The mean debt of all students, the proportion holding a loan, the values 30% and $18450 are all examples of parameters.

D) The mean debt of all students and the proportion holding student loans are both parameters, but the values 30% and $18450 are statistics.

E) The mean debt of all students is a parameter, and its value is $18450; and the proportion holding a student loan is also a parameter whose value is 30%.
1. The probability of an event is defined as the proportion of times an event occurs occasionally.
   A) True
   B) False

Use the following to answer questions 2-3:

A system has two components that operate in parallel, as shown in the diagram below. Because the components operate in parallel, at least one of the components must function properly if the system is to work properly. The probabilities of failures for the components 1 and 2 during a particular period of operation are 0.20 and 0.03, respectively. Let F denote the event that component 1 fails during this period of operation and G denote the event that component 2 fails during this period of operation. The component failures are assumed to be independent.

2. Which event corresponds to the event that the above system fails during this particular period of operation?
   A) F and G
   B) F or G
   C) $F^c$ and $G^c$
   D) $F^c$ or $G^c$

3. What is the probability that the system fails during this particular period of operation?
   A) 0.006
   B) 0.060
   C) 0.224
   D) 0.230
Use the following to answer questions 4-6:

A call-in poll conducted by USA Today concluded that Americans love Donald Trump. This conclusion was based on data collected from 7800 calls made by USA Today readers.

4. USA Today later reported that 5640 of the 7800 calls for the poll came from the offices owned by one man, Cincinnati financier Carl Lindner, who is a friend of Donald Trump. What can we conclude about the results of this poll?
   A) They are surprising, but reliable, because it was conducted by a nationally recognized organization.
   B) They are biased, but only slightly because the sample size was quite large.
   C) They are biased understating the popularity of Donald Trump.
   D) They are biased overstating the popularity of Donald Trump.

5. USA Today reported that of the 7800 calls, 6435 calls were supportive of Donald Trump. This results in a percentage of 82.5%. Of the 6435 supportive calls, about half came from female callers. Which value(s) can be labeled as statistics?
   A) 7800 and 6435
   B) 6435 and 82.5%
   C) 6435, 82.5%, and 50%
   D) 82.5% and 50%

6. What sampling technique is being used?
   A) Simple random sampling
   B) Stratified random sampling
   C) Volunteer sampling
   D) Convenience sampling

Use the following to answer questions 7-10:

A well-known maker of jams and jellies packages its jams in jars labeled “250 mL”. The process used to fill the jars is known to dispense an amount of jam that is a Normally distributed variable with $\mu = 252$ mL and $\sigma = 0.9$ mL.

7. What proportion of jars will be filled with what the label claims is 250 mL?
   A) 0.9868
   B) 0
   C) 0.0068
   D) 0.0132
   E) 0.0027
8. What percentage of jars will be filled with between 251 ml and 254 mL?
   A) 85.3%
   B) 1.3%
   C) 14.7%
   D) 13.4%
   E) 8.5%

9. If changing σ while keeping μ the same were possible for this process, what should σ be set at so that the percentage of jars filled with less than 250 mL will be at most 0.2%?
   A) σ = 2.38
   B) σ = 1
   C) σ = 0.69
   D) σ = 0.82
   E) Not within ± 0.2 of any of the above.

10. What proportion of the jars filled by the process will contain less than 250 mL?
    A) 0.5
    B) 0.9868
    C) 0.0068
    D) 0.0131
    E) 0

11. Two variables are confounded when their effects on a response variable can be distinguished from each other.
    A) True
    B) False

Use the following to answer question 12:

The distribution of GPA scores is known to be left-skewed. At a large university, an English professor is interested in learning about the average GPA score of the English majors and minors. A simple random sample of 75 junior and senior English majors and minors results in an average GPA score of 2.97. Assume that the distribution of GPA scores for all English majors and minors at this university is also left-skewed with a standard deviation of 0.62.
12. Calculate a 95% confidence interval for the mean GPA of the junior and senior English majors and minors.
   A) (2.786, 3.154)
   B) (2.830, 3.110)
   C) (2.852, 3.088)
   D) (2.954, 2.986)

13. Using the standard Normal distribution tables, what is the area under the standard Normal curve corresponding to \(-0.5 < Z < 1.2\)?
   A) 0.3085
   B) 0.8849
   C) 0.5764
   D) 0.2815

14. In a probability model, all possible outcomes together must have a probability of 1?
   A) True
   B) False

15. Influential outliers are easy to detect since the residuals will always be very large compared to the residuals of the other observations.
   A) True
   B) False

16. Confidence intervals and two-sided significance tests are linked in the sense that a two-sided test at a significance level \(\alpha\) can be carried in the form of a confidence interval with confidence level \(1 - \alpha\).
   A) True
   B) False

Use the following to answer questions 17-19:

The weight of medium-size tomatoes selected at random from a bin at the local supermarket is a random variable with mean \(\mu = 10\) oz and standard deviation \(\sigma = 1\) oz
17. Suppose we pick four tomatoes from the bin at random and put them in a bag. Define the random variable \( Y \) = the weight of the bag containing the four tomatoes. What is the mean of the random variable \( Y \)?
   A) \( \mu_Y = 2.5 \) oz
   B) \( \mu_Y = 4 \) oz
   C) \( \mu_Y = 10 \) oz
   D) \( \mu_Y = 40 \) oz

18. Suppose we pick two tomatoes at random from the bin. Let the random variable \( V \) = the difference in the weights of the two tomatoes selected (the weight of the first tomato minus the weight of the second tomato). What is the standard deviation of the random variable \( V \)?
   A) \( \sigma_V = 0.00 \) oz
   B) \( \sigma_V = 1.00 \) oz
   C) \( \sigma_V = 1.41 \) oz
   D) \( \sigma_V = 2.00 \) oz

19. Suppose we pick four tomatoes from the bin at random and put them in a bag. Define the random variable \( Y \) = the weight of the bag containing the four tomatoes. What is the standard deviation of the random variable \( Y \)?
   A) \( \sigma_Y = 0.50 \) oz
   B) \( \sigma_Y = 1.0 \) oz
   C) \( \sigma_Y = 2.0 \) oz
   D) \( \sigma_Y = 4.0 \) oz

Use the following to answer questions 20-21:

A coin is about to be tossed multiple times. Assume the coin is fair, i.e., the probability of heads and the probability of tails are both 0.5.

20. If the coin is tossed six times, what is the probability that less than \( \frac{1}{3} \) of the tosses are heads?
   A) 0.0049
   B) 0.094
   C) 0.109
   D) 0.344
21. If the coin is tossed 60 times, what is the probability that less than \( \frac{1}{2} \) of the tosses are heads?
   Hint: Use the Normal Approximation with no continuity assumption.
   A) 0.0049
   B) 0.094
   C) 0.109
   D) 0.344

Use the following to answer questions 22-25:

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

22. What is the probability that the next three babies are of the same sex?
   A) 0.125
   B) 0.250
   C) 0.375
   D) 0.500

23. Define events \( A = \{ \text{the next two babies are boys} \} \) and \( B = \{ \text{at least one of the next two babies is a boy} \} \). What do we know about events \( A \) and \( B \)?
   A) They are disjoint.
   B) They are complements.
   C) They are independent.
   D) None of the above.

24. What is the probability that at least one of the next three babies is a boy?
   A) 0.125
   B) 0.333
   C) 0.750
   D) 0.875

25. Define event \( B = \{ \text{at least one of the next two babies is a boy} \} \). What is the probability of the complement of event \( B \)?
   A) 0.125
   B) 0.250
   C) 0.375
   D) 0.500
Use the following to answer questions 26-27:

Birth weights of babies born to full-term pregnancies follow roughly a Normal distribution. At Meadowbrook Hospital, the mean weight of babies born to full-term pregnancies is 7 lbs with a standard deviation of 14 oz (1 lb = 16 oz).

26. What is the probability that the average weight of the four babies will be more than 7.5 lbs?
   A) 0.0065
   B) 0.1265
   C) 0.2839
   D) 0.4858

27. Dr. Watts (who works at Meadowbrook Hospital) has four deliveries (all for full-term pregnancies) coming up during the night. Assume that the birth weights of these four babies can be viewed as a simple random sample. What is the probability that all four babies will weigh more than 7.5 lbs?
   A) 0.0065
   B) 0.1265
   C) 0.2839
   D) 0.4858

Use the following to answer question 28:

The proportion of supermarket customers who do not buy store-brand products is to be estimated.

28. Suppose 500 customers are selected from the roughly 20,000 customers who shop at the stores citywide. The sample proportion of supermarket customers who do not buy store-brand products equals 33.5%. Which value(s) can be labeled as statistic(s)?
   A) 20,000
   B) 500 and 20,000
   C) 33.5%
   D) 20,000 and 33.5%
Use the following to answer questions 29-30:

A business has two types of employees: managers and workers. Managers earn either $100,000 or $200,000 per year. Workers earn either $10,000 or $20,000 per year. The number of male and female managers at each salary level and the number of male and female workers at each salary level are given in the table below:

<table>
<thead>
<tr>
<th>Income</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>$100,000</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>$200,000</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Income</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10,000</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>$20,000</td>
<td>20</td>
<td>80</td>
</tr>
</tbody>
</table>

29. What proportion of the managers is female?
A) 0.2
B) 0.333
C) 0.5
D) 0.667

30. What is the proportion of female managers who make $200,000 per year?
A) 0.1
B) 0.2
C) 0.4
D) 0.6

31. Apply the $1.5 \times IQR$ rule to the data in the previous question to check for outlier values. In this case
A) there are no outliers
B) the value 0.75 is the only outlier
C) the values 0.75 and 2.48 are both outliers
D) the value 2.48 is the only outlier
E) the values 1.52 and 2.48 are both outliers

32. Which of the following cannot give informed consent to participate in a study?
A) Children
B) Inmates
C) Mentally ill patients
D) All of the above.
33. A valid probability is any number between $-1$ and $1$.
   A) True
   B) False

Use the following to answer question 34:

In the National Hockey League a good predictor of the percentage of games won by a team is the number of goals the team allows during the season. Data were gathered for all 30 teams in the NHL and the scatterplot of their Winning Percentage against the number of Goals Allowed in the 2006/2007 season with a fitted least-squares regression line is provided:

![Scatterplot of Winning Percentage vs. Goals Allowed]

The least-squares regression line and $r^2$ were calculated to be
\[
\text{Winning Percent (\%) = 116.95 - 0.26 \times \text{Goals Allowed}}
\]
\[
r^2 = 0.69
\]

34. Which of the following provides the best interpretation of the slope of the regression line?
   A) If the Winning Percent increases by 1%, then the number of Goals Allowed decreases by 0.26.
   B) If a team were to allow 100 goals during the season, their Winning % would be 90.95%.
   C) If Goals Allowed increases by one goal, the Winning % increases by 0.26%.
   D) If the Winning % increases by 1%, then the number of Goals Allowed increases by 0.26.
   E) If Goals Allowed increases by one goal, the Winning % decreases by 0.26%.
Use the following to answer question 35:

Suppose that the random variable $X$ is continuous and takes its values uniformly over the interval from 0 to 2.

35. What is the value of the probability $P\{X \leq 0.4 \text{ or } X > 1.2\}$?
   
   A) 0.40  
   B) 0.20  
   C) 0.60  
   D) 0.80  
   E) 0.50

Use the following to answer questions 36-38:

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 50 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.

36. In which situation would this study be double-blind?
   
   A) Neither drug had any identifying marks on it.  
   B) All volunteers were not allowed to see the psychiatrist nor was the psychiatrist allowed to see the volunteers during the session when the psychiatrist evaluated the severity of the depression.  
   C) Neither the volunteers nor the psychiatrist knew which treatment any person had received.  
   D) All of the above.

37. Suppose volunteers were first divided by gender, and then half of the men were randomly assigned to the new drug and half of the women were assigned to the new drug. The remaining volunteers received the other drug. What is this an example of?
   
   A) Replication.  
   B) Confounding. The effects of gender will be mixed up with the effects of the drugs.  
   C) A block design.  
   D) A matched-pairs design.
38. What is the explanatory variable or factor in this study?
   A) Which drug the volunteers receive.
   B) The use of randomization and the fact that this was a comparative study.
   C) The extent to which the depression was reduced.
   D) The use of a psychiatrist to evaluate the severity of depression.

39. As Swiss cheese matures, a variety of chemical processes take place. The taste of matured cheese is related to the concentration of several chemicals in the final product. In a study of cheese in a certain region of Switzerland, samples of cheese were analyzed for lactic acid concentration and were subjected to taste tests. The numerical taste scores were obtained by combining the scores from several tasters. A scatterplot of the observed data is shown below:

![Scatterplot of lactic acid concentration vs. taste rating]

What is a plausible value for the correlation between lactic acid concentration and taste rating?
   A) 0.999
   B) 0.7
   C) 0.07
   D) -0.7

Use the following to answer question 40:

The asking prices (in thousands of dollars) for a sample of 13 houses currently on the market in Neighborville are listed below. For convenience, the data have been ordered.

175 199 205 234 259 275 299 304 317 345 355 384 549
40. Use the $1.5 \times IQR$ rule to determine if there are any outliers present. What is/are the value(s) of the outlier(s)?

A) No outliers present
B) One outlier: 175
C) One outlier: 549
D) Two outliers: 175 and 549

41. In a statistics course, a linear regression equation was computed to predict the final exam score from the score on the midterm exam. The equation of the least-squares regression line was

$$\hat{y} = 10 + 0.9x,$$

where $y$ represents the final exam score and $x$ is the midterm exam score. Suppose Joe scores a 90 on the midterm exam. What would be the predicted value of his score on the final exam?

A) 81
B) 89
C) 91
D) Cannot be determined from the information given. We also need to know the correlation.

Use the following to answer questions 42-43:

A population has a distribution with mean $\mu = 50$ and variance $\sigma^2 = 225$. From this population a simple random sample of $n$ observations is to be selected and the mean $\bar{x}$ of the sample values calculated.

42. If the population variable is known to be Normally distributed and the sample size used is to be $n = 16$, what is the probability that the sample mean will be between 48.35 and 55.74, i.e., $P(48.35 \leq \bar{x} \leq 55.74)$?

A) 0.393
B) 0.607
C) 0.937
D) 0.330
E) Not within ± 0.010 of any of the above.
43. How big must the sample size \( n \) be so that the standard deviation of the sample mean, \( \bar{x} \), is equal to 1.4, i.e., \( \sigma_{\bar{x}} = 1.4 \)?

A) \( n = 11 \)
B) \( n = 161 \)
C) \( n = 115 \)
D) \( n = 36 \)
E) \( n = 21 \)

Use the following to answer questions 44-46:

It is known that driving can be difficult in regions where winter conditions involve snow-covered roads. For cars equipped with all-season tires traveling at 90 km/hr, the mean stopping time in fresh snow is known to be 215 meters with a standard deviation of \( \sigma = 2.5 \) meters. It is often advocated that automobiles in such areas should be equipped with special tires to compensate for such conditions, especially with respect to stopping distance. A manufacturer of tires made for driving in fresh snow claims that vehicles equipped with their tires have a decreased stopping distance. A study was done using a random sample of 9 snow tires from the manufacturer on a snow-covered test track. The tests resulted in a mean stopping distance of \( \bar{x} = 212.9 \) meters.

44. Using the sample results and assuming that stopping distance is a Normally distributed random variable, what is the value of the test statistic?

A) 0.05
B) \(-2.52\)
C) \(-9.36\)
D) \(-1.04\)
E) \(-1.96\)

45. What are the appropriate null and alternative hypotheses to test the manufacturer's claim?

A) \( H_0: \mu = 215 \) against \( H_a: \mu > 215 \)
B) \( H_0: \mu = 215 \) against \( H_a: \mu \neq 215 \)
C) \( H_0: \mu = 215 \) against \( H_a: \mu < 215 \)
D) \( H_0: \bar{x} = 215 \) against \( H_a: \bar{x} < 215 \)
E) Snow tires decrease the stopping distance in loose snow.

46. What is the \( P \)-value?

A) 0.012
B) 0.050
C) 0.025
D) 0.006
Use the following to answer questions 47-48:

A study was conducted in a large population of adults concerning eyeglasses for correcting reading vision. Based on an examination by a qualified professional, the individuals were judged as to whether or not they needed to wear glasses for reading. In addition it was determined whether or not they were currently using glasses for reading. The following table provides the proportions found in the study:

<table>
<thead>
<tr>
<th>Judged to need glasses</th>
<th>Used glasses for reading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Yes</td>
<td>0.42</td>
</tr>
<tr>
<td>No</td>
<td>0.04</td>
</tr>
</tbody>
</table>

47. If a single adult is selected at random from this large population, what is the probability that the adult is judged to need eyeglasses for reading?
   - A) 0.46
   - B) 0.42
   - C) 0.78
   - D) 0.60
   - E) 0.40

48. What is the probability that the selected adult is judged to need eyeglasses but does not use them for reading?
   - A) 0.42
   - B) 0.18
   - C) 0.54
   - D) 0.60
   - E) 0.36

Use the following to answer question 49:

A researcher is studying the relationship between sugar consumption and weight gain. Twelve volunteers were randomly assigned to one of two groups. The first group of five participants was put on a diet low in sugar and the second group of the remaining seven participants received 10% of their calories from sugar. After 8 weeks, weight gain was recorded from each participant.
49. Each of the 12 volunteers receives a label: 01 through 12. Use the list of random digits below to determine the labels of the first five participants selected who were put on a low-sugar diet. Start at the beginning of the list and use double-digit labels.

81507 27102 56027 55892 33063 41842 81868 71035 09001 23367
49497
A) 8, 1, 5, 0, 7
B) 8, 1, 5, 7, 2
C) 2, 6, 10, 9, 12
D) 2, 2, 6, 10, 9

50. The following data are the magnitudes of earthquakes around the world recorded on January 13, 2008:

4.1 4.8 3.1 5.3 5.1 4.7 3.0 2.9 U 4.6 3.1 3.0 U 2.5.

Because of equipment problems, two earthquakes were unrecorded although it was known that both had a value less than 2.7. The median magnitude of earthquakes on this date is
A) 2.95
B) 3.9
C) 4.0
D) 3.1
E) Cannot determine without knowing the exact value of the missing magnitudes.

51. Typically, the null hypothesis in a test of significance is a statement of “no difference” in the true means.
A) True
B) False

Use the following to answer questions 52-53:

The breaking strength of yarn used in the manufacture of woven carpet material is Normally distributed with \( \sigma = 2.4 \) psi. A random sample of 16 specimens of yarn from a production run were measured for breaking strength, and based on the mean of the sample, \( \bar{x} \), a confidence interval was found to be (128.7, 131.3).
52. What is the confidence level, $C$, of this interval?
   A) 0.95
   B) 0.99
   C) 0.90
   D) 0.97
   E) Unable to determine with the information provided.

53. If we wanted to have a greater degree of confidence in the calculated interval, what would have to happen to its length?
   A) The interval would be shorter.
   B) The margin of error would decrease.
   C) The confidence level would have a smaller value.
   D) The sample size would go up.
   E) The interval would increase in length.

54. An engineer has designed an improved light bulb. The previous design had an average lifetime of 1200 hours. Based on a sample of 2000 of these new bulbs, the average lifetime was found to be 1201 hours. Although the difference is quite small, the effect was statistically significant. What is the best explanation?
   A) New designs typically have more variability than standard designs.
   B) The sample size is very large, so that even a small difference can be detected.
   C) The mean of 1200 is large.
   D) The relative improvement in average lifetime is 0.000083, which is much smaller than 0.05.

55. A population of graduate students at a large university is being studied through the use of a sample survey. The researcher is trying to estimate the mean debt carried by the students through student loans, and the proportion of all students that carry a student loan. A stratified random sample of 350 students is selected from the various schools within the university. The sample results showed that 30% held a student loan and the mean debt was $18450. Which of the following statements is TRUE?
   A) The result, 30%, is a statistic and the mean debt, $18450, is the value of the parameter.
   B) The mean debt held by all students is a parameter, and the 30% who hold student loans is the value of a parameter.
   C) The mean debt of all students, the proportion holding a loan, the values 30% and $18450 are all examples of parameters.
   D) The mean debt of all students and the proportion holding student loans are both parameters, but the values 30% and $18450 are statistics.
   E) The mean debt of all students is a parameter, and its value is $18450; and the proportion holding a student loan is also a parameter whose value is 30%.
56. Confidentiality and anonymity is the same thing when referring to experimental design?
A) True
B) False

Use the following to answer questions 57-58:

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

57. The probability is approximately 0.6 that on a randomly selected day the store will make less than how much?
A) $0.30
B) $347.40
C) $361.30
D) $372.60

58. What is $P(X > $400$)?
A) 0.2119
B) 0.2881
C) 0.7881
D) 0.8450

Use the following to answer questions 59-60:

John's parents recorded his height at various ages between 36 and 66 months. Below is a record of the results:

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>36</th>
<th>48</th>
<th>54</th>
<th>60</th>
<th>66</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height (inches)</td>
<td>34</td>
<td>38</td>
<td>41</td>
<td>43</td>
<td>45</td>
</tr>
</tbody>
</table>

59. Which of the following is the equation of the least-squares regression line of John's height on age? (Note: You do not need to directly calculate the least-squares regression line to answer this question.)
A) Height = 12 $\times$ (Age)
B) Height = Age/12
C) Height = 60 - 0.22 $\times$ (Age)
D) Height = 22.3 + 0.34 $\times$ (Age)
60. John's parents decide to use the least-squares regression line of John's height on age to predict his height at age 21 years (252 months). What conclusion can we draw?
A) John's height, in inches, should be about half his age, in months.
B) The parents will get a fairly accurate estimate of his height at age 21 years, because the data are clearly correlated.
C) Such a prediction could be misleading, because it involves extrapolation.
D) All of the above.
1. Use the $1.5 \times IQR$ rule to determine if there are any outliers present. What is/are the value(s) of the outlier(s)?
   A) No outliers present
   B) One outlier: 175
   C) One outlier: 549
   D) Two outliers: 175 and 549

2. Suppose we pick four tomatoes from the bin at random and put them in a bag. Define the random variable $Y$ as the weight of the bag containing the four tomatoes. What is the mean of the random variable $Y$?
   A) $\mu_Y = 2.5$ oz
   B) $\mu_Y = 4$ oz
   C) $\mu_Y = 10$ oz
   D) $\mu_Y = 40$ oz
3. Suppose we pick two tomatoes at random from the bin. Let the random variable \( V \) =
the difference in the weights of the two tomatoes selected (the weight of the first tomato
minus the weight of the second tomato). What is the standard deviation of the random
variable \( V \)?
A) \( \sigma_V = 0.00 \) oz
B) \( \sigma_V = 1.00 \) oz
C) \( \sigma_V = 1.41 \) oz
D) \( \sigma_V = 2.00 \) oz

4. Suppose we pick four tomatoes from the bin at random and put them in a bag. Define
the random variable \( Y \) = the weight of the bag containing the four tomatoes. What is the
standard deviation of the random variable \( Y \)?
A) \( \sigma_Y = 0.50 \) oz
B) \( \sigma_Y = 1.0 \) oz
C) \( \sigma_Y = 2.0 \) oz
D) \( \sigma_Y = 4.0 \) oz

Use the following to answer questions 5-6:

A business has two types of employees: managers and workers. Managers earn either $100,000
or $200,000 per year. Workers earn either $10,000 or $20,000 per year. The number of male
and female managers at each salary level and the number of male and female workers at each
salary level are given in the table below:

<table>
<thead>
<tr>
<th>Income</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>$100,000</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>$200,000</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>Income</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>$10,000</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>$20,000</td>
<td>20</td>
<td>80</td>
</tr>
</tbody>
</table>

5. What proportion of the managers is female?
A) 0.2
B) 0.333
C) 0.5
D) 0.667

6. What is the proportion of female managers who make $200,000 per year?
A) 0.1
B) 0.2
C) 0.4
D) 0.6
Use the following to answer question 7:

A researcher is studying the relationship between sugar consumption and weight gain. Twelve volunteers were randomly assigned to one of two groups. The first group of five participants was put on a diet low in sugar and the second group of the remaining seven participants received 10% of their calories from sugar. After 8 weeks, weight gain was recorded from each participant.

7. Each of the 12 volunteers receives a label: 01 through 12. Use the list of random digits below to determine the labels of the first five participants selected who were put on a low-sugar diet. Start at the beginning of the list and use double-digit labels.

\[ 81507 \quad 27102 \quad 56027 \quad 55892 \quad 33063 \quad 41842 \quad 81868 \quad 71035 \quad 09001 \quad 23367 \quad 49497 \]

A) 8, 1, 5, 0, 7  
B) 8, 1, 5, 7, 2  
C) 2, 6, 10, 9, 12  
D) 2, 2, 6, 10, 9

Use the following to answer question 8:

The proportion of supermarket customers who do not buy store-brand products is to be estimated.

8. Suppose 500 customers are selected from the roughly 20,000 customers who shop at the stores citywide. The sample proportion of supermarket customers who do not buy store-brand products equals 33.5%. Which value(s) can be labeled as statistic(s)?

A) 20,000  
B) 500 and 20,000  
C) 33.5%  
D) 20,000 and 33.5%

9. Typically, the null hypothesis in a test of significance is a statement of "no difference" in the true means.

A) True  
B) False

10. Confidentiality and anonymity is the same thing when referring to experimental design?

A) True  
B) False
Use the following to answer questions 11-12:

A system has two components that operate in parallel, as shown in the diagram below. Because the components operate in parallel, at least one of the components must function properly if the system is to work properly. The probabilities of failures for the components 1 and 2 during a particular period of operation are 0.20 and 0.03, respectively. Let F denote the event that component 1 fails during this period of operation and G denote the event that component 2 fails during this period of operation. The component failures are assumed to be independent.

![Diagram of two components in parallel]

11. What is the probability that the system fails during this particular period of operation?
   A) 0.006
   B) 0.060
   C) 0.224
   D) 0.230

12. Which event corresponds to the event that the above system fails during this particular period of operation?
   A) F and G
   B) F or G
   C) F^c and G^c
   D) F^c or G^c

Use the following to answer questions 13-15:

It is known that driving can be difficult in regions where winter conditions involve snow-covered roads. For cars equipped with all-season tires traveling at 90 km/hr, the mean stopping time in fresh snow is known to be 215 meters with a standard deviation of σ = 2.5 meters. It is often advocated that automobiles in such areas should be equipped with special tires to compensate for such conditions, especially with respect to stopping distance. A manufacturer of tires made for driving in fresh snow claims that vehicles equipped with their tires have a decreased stopping distance. A study was done using a random sample of 9 snow tires from the manufacturer on a snow-covered test track. The tests resulted in a mean stopping distance of \( \bar{x} = 212.9 \) meters.
13. What is the P-value?
A) 0.012
B) 0.050
C) 0.025
D) 0.006

14. What are the appropriate null and alternative hypotheses to test the manufacturer's claim?
A) $H_0: \mu = 215$ against $H_a: \mu > 215$
B) $H_0: \mu = 215$ against $H_a: \mu \neq 215$
C) $H_0: \mu = 215$ against $H_a: \mu < 215$
D) $H_0: \bar{X} = 215$ against $H_a: \bar{X} < 215$
E) Snow tires decrease the stopping distance in loose snow.

15. Using the sample results and assuming that stopping distance is a Normally distributed random variable, what is the value of the test statistic?
A) 0.05
B) -2.52
C) -9.36
D) -1.04
E) -1.96
Use the following to answer question 16:

In the National Hockey League a good predictor of the percentage of games won by a team is the number of goals the team allows during the season. Data were gathered for all 30 teams in the NHL and the scatterplot of their Winning Percentage against the number of Goals Allowed in the 2006/2007 season with a fitted least-squares regression line is provided:

The least-squares regression line and \( r^2 \) were calculated to be

Winning Percent (\%) = 116.95 - 0.26 \times \text{Goals Allowed}

\[ r^2 = 0.69 \]

16. Which of the following provides the best interpretation of the slope of the regression line?

A) If the Winning Percent increases by 1%, then the number of Goals Allowed decreases by 0.26.
B) If a team were to allow 100 goals during the season, their Winning % would be 90.95%.
C) If Goals Allowed increases by one goal, the Winning % increases by 0.26%.
D) If the Winning % increases by 1%, then the number of Goals Allowed increases by 0.26.
E) If Goals Allowed increases by one goal, the Winning % decreases by 0.26%.

Use the following to answer questions 17-18:

Let the random variable \( X \) represent the profit made on a randomly selected day by a small clothing store on Main Street. Assume \( X \) is Normal with a mean of $360 and a standard deviation of $50.
17. What is \( P(X > $400) \)?
   A) 0.2119  
   B) 0.2881  
   C) 0.7881  
   D) 0.8450

18. The probability is approximately 0.6 that on a randomly selected day the store will make less than how much?
   A) $0.30  
   B) $347.40  
   C) $361.30  
   D) $372.60

Use the following to answer questions 19-20:

A study was conducted in a large population of adults concerning eyeglasses for correcting reading vision. Based on an examination by a qualified professional, the individuals were judged as to whether or not they needed to wear glasses for reading. In addition it was determined whether or not they were currently using glasses for reading. The following table provides the proportions found in the study:

<table>
<thead>
<tr>
<th>Judged to need glasses</th>
<th>Used glasses for reading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Yes</td>
<td>0.42</td>
</tr>
<tr>
<td>No</td>
<td>0.04</td>
</tr>
</tbody>
</table>

19. What is the probability that the selected adult is judged to need eyeglasses but does not use them for reading?
   A) 0.42  
   B) 0.18  
   C) 0.54  
   D) 0.60  
   E) 0.36

20. If a single adult is selected at random from this large population, what is the probability that the adult is judged to need eyeglasses for reading?
   A) 0.46  
   B) 0.42  
   C) 0.78  
   D) 0.60  
   E) 0.40
21. The following data are the magnitudes of earthquakes around the world recorded on January 13, 2008:

\[ 4.1 \quad 4.8 \quad 3.1 \quad 5.3 \quad 5.1 \quad 4.7 \quad 3.0 \quad U \quad 4.6 \quad 3.1 \quad 3.0 \quad U \quad 2.5. \]

Because of equipment problems, two earthquakes were unrecorded although it was known that both had a value less than 2.7. The median magnitude of earthquakes on this date is

A) 2.95  
B) 3.9  
C) 4.0  
D) 3.1  
E) Cannot determine without knowing the exact value of the missing magnitudes.

22. In a statistics course, a linear regression equation was computed to predict the final exam score from the score on the midterm exam. The equation of the least-squares regression line was

\[ \hat{y} = 10 + 0.9x, \]

where \( y \) represents the final exam score and \( x \) is the midterm exam score. Suppose Joe scores a 90 on the midterm exam. What would be the predicted value of his score on the final exam?

A) 81  
B) 89  
C) 91  
D) Cannot be determined from the information given. We also need to know the correlation.

23. Using the standard Normal distribution tables, what is the area under the standard Normal curve corresponding to \(-0.5 < Z < 1.2\)?

A) 0.3085  
B) 0.8849  
C) 0.5764  
D) 0.2815
24. Apply the $1.5 \times IQR$ rule to the data in the previous question to check for outlier values. In this case 
A) there are no outliers 
B) the value 0.75 is the only outlier 
C) the values 0.75 and 2.48 are both outliers 
D) the value 2.48 is the only outlier 
E) the values 1.52 and 2.48 are both outliers 

25. Which of the following cannot give informed consent to participate in a study? 
A) Children 
B) Inmates 
C) Mentally ill patients 
D) All of the above. 

26. A valid probability is any number between $-1$ and 1. 
A) True 
B) False 

Use the following to answer questions 27-28:

John's parents recorded his height at various ages between 36 and 66 months. Below is a record of the results:

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>36</th>
<th>48</th>
<th>54</th>
<th>60</th>
<th>66</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height (inches)</td>
<td>34</td>
<td>38</td>
<td>41</td>
<td>43</td>
<td>45</td>
</tr>
</tbody>
</table>

27. Which of the following is the equation of the least-squares regression line of John's height on age? (Note: You do not need to directly calculate the least-squares regression line to answer this question.)
A) Height = 12 x (Age) 
B) Height = Age/12 
C) Height = 60 - 0.22 x (Age) 
D) Height = 22.3 + 0.34 x (Age)
28. John's parents decide to use the least-squares regression line of John's height on age to predict his height at age 21 years (252 months). What conclusion can we draw?
A) John's height, in inches, should be about half his age, in months.
B) The parents will get a fairly accurate estimate of his height at age 21 years, because the data are clearly correlated.
C) Such a prediction could be misleading, because it involves extrapolation.
D) All of the above.

29. A population of graduate students at a large university is being studied through the use of a sample survey. The researcher is trying to estimate the mean debt carried by the students through student loans, and the proportion of all students that carry a student loan. A stratified random sample of 350 students is selected from the various schools within the university. The sample results showed that 30% held a student loan and the mean debt was $18450. Which of the following statements is TRUE?
A) The result, 30%, is a statistic and the mean debt, $18450, is the value of the parameter.
B) The mean debt held by all students is a parameter, and the 30% who hold student loans is the value of a parameter.
C) The mean debt of all students, the proportion holding a loan, the values 30% and $18450 are all examples of parameters.
D) The mean debt of all students and the proportion holding student loans are both parameters, but the values 30% and $18450 are statistics.
E) The mean debt of all students is a parameter, and its value is $18450; and the proportion holding a student loan is also a parameter whose value is 30%.

30. Two variables are confounded when their effects on a response variable can be distinguished from each other.
A) True
B) False

Use the following to answer questions 31-32:

Birth weights of babies born to full-term pregnancies follow roughly a Normal distribution. At Meadowbrook Hospital, the mean weight of babies born to full-term pregnancies is 7 lbs with a standard deviation of 14 oz (1 lb = 16 oz).
31. Dr. Watts (who works at Meadowbrook Hospital) has four deliveries (all for full-term pregnancies) coming up during the night. Assume that the birth weights of these four babies can be viewed as a simple random sample. What is the probability that all four babies will weigh more than 7.5 lbs?
A) 0.0065
B) 0.1265
C) 0.2839
D) 0.4858

32. What is the probability that the average weight of the four babies will be more than 7.5 lbs?
A) 0.0065
B) 0.1265
C) 0.2839
D) 0.4858

33. The probability of an event is defined as the proportion of times an event occurs occasionally.
A) True
B) False

Use the following to answer questions 34-37:

A well-known maker of jams and jellies packages its jams in jars labeled “250 mL”. The process used to fill the jars is known to dispense an amount of jam that is a Normally distributed variable with \( \mu = 252 \text{ mL} \) and \( \sigma = 0.9 \text{ mL} \).

34. If changing \( \sigma \) while keeping \( \mu \) the same were possible for this process, what should \( \sigma \) be set at so that the percentage of jars filled with less than 250 mL will be at most 0.2%?
A) \( \sigma = 2.38 \)
B) \( \sigma = 1 \)
C) \( \sigma = 0.69 \)
D) \( \sigma = 0.82 \)
E) Not within \( \pm 0.2 \) of any of the above.
35. What percentage of jars will be filled with between 251 ml and 254 mL?
   A) 85.3%
   B) 1.3%
   C) 14.7%
   D) 13.4%
   E) 8.5%

36. What proportion of the jars filled by the process will contain less than 250 mL?
   A) 0.5
   B) 0.9868
   C) 0.0068
   D) 0.0131
   E) 0

37. What proportion of jars will be filled with what the label claims is 250 mL?
   A) 0.9868
   B) 0
   C) 0.0068
   D) 0.0132
   E) 0.0027

Use the following to answer questions 38-39:

A coin is about to be tossed multiple times. Assume the coin is fair, i.e., the probability of heads and the probability of tails are both 0.5.

38. If the coin is tossed six times, what is the probability that less than \( \frac{1}{2} \) of the tosses are heads?
   A) 0.0049
   B) 0.094
   C) 0.109
   D) 0.344

39. If the coin is tossed 60 times, what is the probability that less than \( \frac{1}{2} \) of the tosses are heads?
   Hint: Use the Normal Approximation with no continuity assumption.
   A) 0.0049
   B) 0.094
   C) 0.109
   D) 0.344
Use the following to answer questions 40-42:

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 50 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.

40. What is the explanatory variable or factor in this study?
   A) Which drug the volunteers receive.
   B) The use of randomization and the fact that this was a comparative study.
   C) The extent to which the depression was reduced.
   D) The use of a psychiatrist to evaluate the severity of depression.

41. Suppose volunteers were first divided by gender, and then half of the men were randomly assigned to the new drug and half of the women were assigned to the new drug. The remaining volunteers received the other drug. What is this an example of?
   A) Replication.
   B) Confounding. The effects of gender will be mixed up with the effects of the drugs.
   C) A block design.
   D) A matched-pairs design.

42. In which situation would this study be double-blind?
   A) Neither drug had any identifying marks on it.
   B) All volunteers were not allowed to see the psychiatrist nor was the psychiatrist allowed to see the volunteers during the session when the psychiatrist evaluated the severity of the depression.
   C) Neither the volunteers nor the psychiatrist knew which treatment any person had received.
   D) All of the above.

Use the following to answer questions 43-46:

Ignoring twins and other multiple births, assume babies born at a hospital are independent events with the probability that a baby is a boy and the probability that a baby is a girl both equal to 0.5.
43. Define event \( B = \{ \text{at least one of the next two babies is a boy} \} \). What is the probability of the complement of event \( B \)?
   A) 0.125
   B) 0.250
   C) 0.375
   D) 0.500

44. What is the probability that the next three babies are of the same sex?
   A) 0.125
   B) 0.250
   C) 0.375
   D) 0.500

45. What is the probability that at least one of the next three babies is a boy?
   A) 0.125
   B) 0.333
   C) 0.750
   D) 0.875

46. Define events \( A = \{ \text{the next two babies are boys} \} \) and \( B = \{ \text{at least one of the next two babies is a boy} \} \). What do we know about events \( A \) and \( B \)?
   A) They are disjoint.
   B) They are complements.
   C) They are independent.
   D) None of the above.

47. Influential outliers are easy to detect since the residuals will always be very large compared to the residuals of the other observations.
   A) True
   B) False
48. As Swiss cheese matures, a variety of chemical processes take place. The taste of matured cheese is related to the concentration of several chemicals in the final product. In a study of cheese in a certain region of Switzerland, samples of cheese were analyzed for lactic acid concentration and were subjected to taste tests. The numerical taste scores were obtained by combining the scores from several tasters. A scatterplot of the observed data is shown below:

What is a plausible value for the correlation between lactic acid concentration and taste rating?
A) 0.999
B) 0.7
C) 0.07
D) -0.7

49. An engineer has designed an improved light bulb. The previous design had an average lifetime of 1200 hours. Based on a sample of 2000 of these new bulbs, the average lifetime was found to be 1201 hours. Although the difference is quite small, the effect was statistically significant. What is the best explanation?
A) New designs typically have more variability than standard designs.
B) The sample size is very large, so that even a small difference can be detected.
C) The mean of 1200 is large.
D) The relative improvement in average lifetime is 0.000083, which is much smaller than 0.05.
Use the following to answer question 50:

The distribution of GPA scores is known to be left-skewed. At a large university, an English professor is interested in learning about the average GPA score of the English majors and minors. A simple random sample of 75 junior and senior English majors and minors results in an average GPA score of 2.97. Assume that the distribution of GPA scores for all English majors and minors at this university is also left-skewed with a standard deviation of 0.62.

50. Calculate a 95% confidence interval for the mean GPA of the junior and senior English majors and minors.
A) (2.786, 3.154)
B) (2.830, 3.110)
C) (2.852, 3.088)
D) (2.954, 2.986)

Use the following to answer questions 51-52:

The breaking strength of yarn used in the manufacture of woven carpet material is Normally distributed with $\mu = 2.4$ psi. A random sample of 16 specimens of yarn from a production run were measured for breaking strength, and based on the mean of the sample, $\bar{x}$, a confidence interval was found to be (128.7, 131.3).

51. What is the confidence level, $C$, of this interval?
A) 0.95
B) 0.99
C) 0.90
D) 0.97
E) Unable to determine with the information provided.

52. If we wanted to have a greater degree of confidence in the calculated interval, what would have to happen to its length?
A) The interval would be shorter.
B) The margin of error would decrease.
C) The confidence level would have a smaller value.
D) The sample size would go up.
E) The interval would increase in length.
Use the following to answer questions 53-54:

A population has a distribution with mean $\mu = 50$ and variance $\sigma^2 = 225$. From this population a simple random sample of $n$ observations is to be selected and the mean $\bar{x}$ of the sample values calculated.

53. If the population variable is known to be Normally distributed and the sample size used is to be $n = 16$, what is the probability that the sample mean will be between 48.35 and 55.74, i.e., $P(48.35 \leq \bar{x} \leq 55.74)$?
   A) 0.393
   B) 0.607
   C) 0.937
   D) 0.330
   E) Not within $\pm 0.010$ of any of the above.

54. How big must the sample size $n$ be so that the standard deviation of the sample mean, $\bar{x}$, is equal to 1.4, i.e., $\sigma_{\bar{x}} = 1.4$?
   A) $n = 11$
   B) $n = 161$
   C) $n = 115$
   D) $n = 36$
   E) $n = 21$

55. Confidence intervals and two-sided significance tests are linked in the sense that a two-sided test at a significance level $\alpha$ can be carried in the form of a confidence interval with confidence level $1 - \alpha$.
   A) True
   B) False

Use the following to answer question 56:

Suppose that the random variable $X$ is continuous and takes its values uniformly over the interval from 0 to 2.

56. What is the value of the probability $P\{X \leq 0.4 \text{ or } X > 1.2\}$?
   A) 0.40
   B) 0.20
   C) 0.60
   D) 0.80
   E) 0.50
Use the following to answer questions 57-59:

A call-in poll conducted by *USA Today* concluded that Americans love Donald Trump. This conclusion was based on data collected from 7800 calls made by *USA Today* readers.

57. What sampling technique is being used?
   A) Simple random sampling
   B) Stratified random sampling
   C) Volunteer sampling
   D) Convenience sampling

58. *USA Today* later reported that 5640 of the 7800 calls for the poll came from the offices owned by one man, Cincinnati financier Carl Lindner, who is a friend of Donald Trump. What can we conclude about the results of this poll?
   A) They are surprising, but reliable, because it was conducted by a nationally recognized organization.
   B) They are biased, but only slightly because the sample size was quite large.
   C) They are biased understating the popularity of Donald Trump.
   D) They are biased overstating the popularity of Donald Trump.

59. *USA Today* reported that of the 7800 calls, 6435 calls were supportive of Donald Trump. This results in a percentage of 82.5%. Of the 6435 supportive calls, about half came from female callers. Which value(s) can be labeled as statistics?
   A) 7800 and 6435
   B) 6435 and 82.5%
   C) 6435, 82.5%, and 50%
   D) 82.5% and 50%

60. In a probability model, all possible outcomes together must have a probability of 1?
   A) True
   B) False