MIDTERM EXAM QUESTIONS POOL

1) Which of the following statements is true when using one-way ANOVA?

a) The populations from which the samples are selected have different distributions.
b) The sample sizes are large.
c) The test is to determine if the different groups have the same means. (answer)
d) There is a correlation between the factors of the experiment.
e) All of them.

2) The manager of a sales department decides to measure employee satisfaction by selecting five departments at random, and conducting interviews with all the employees in those departments. What type of survey design is this?

a) cluster (answer)
b) stratified
c) simple random
d) systematic
e) none of them

3) In a left-skewed distribution, which is greater?

a) the mean
b) the median
c) the mode (answer)
d) variance
e) standard deviation

4) You get data from the U.S. FDA Bureau on the median household income for your city, and decide to display it graphically. Which is the better choice for this data?

a) Histogram (answer)
b) Pie chart
c) Bar chart
d) Line graph
e) Boxplot

5) A psychologist is interested in whether the size of tableware (bowls, plates, etc.) influences how much college students eat. He randomly assigns 100 college students to one of two groups: the first is served a meal using normal-sized tableware, while the second is served the same meal, but using tableware that it 20 percent smaller than normal. He records how much food is consumed by each group. What size of the tableware represents?

a) Population
b) Sample
c) Treatment
d) Response variable
e) Explanatory variable (answer)

6) A store is interested in how much money, on average, their customers spend each visit in the produce department. Using their store records, they draw a sample of 1,000 visits and calculate each customer's average spending on produce. "Number of visits per week" is what kind of data?

a) qualitative
b) quantitative-continuous
c) quantitative-discrete (answer)
d) quantitative
e) continuous

7) Which of the following statements is false about the confidence interval?

a) A confidence interval gives a range estimate of values.
b) Takes into consideration variation in sample statistics from sample to sample.
c) Based on all the observations from more than one sample. (answer)
d) Gives information about closeness to unknown population parameters.
e) Can never be 100% confident.

8) Which of the following sampling types, the population first sliced into homogenous groups?

a) simple random
b) systematic
c) distrubution
d) cluster
e) stratified (answer)

9) The correlation coefficient (r) gives us a numerical measurement of the strength of the linear relationship between which variables?

a) explanatory continuos
b) response discrete
c) explanatory response (answer)
d) continuos discrete
e) quantitive quantitive

10) Correlation measures the strength of the linear association between which variables?

a) explanatory continuos
b) response discrete
c) explanatory response
d) continuos discrete
e) quantitive quantitive (answer)

11) A histogram that doesn't appear to have any mode and in which all the bars are approximately the same height is called ……..

a) uniform (answer)
b) binomial

c) geometric
d) exponential
e) poisson

12) A negative z-score tells us that the data value is …… the mean, while a positive z-score tells us that the data value is …….. the mean.
a) above, below (answer)
b) below, above
c) equal to, below
d) equal to, above
e) above, equal to

13) …… is used, when there is a large number of trials, but a small probability of success.
a) Uniform
b) Geometric
c) Exponential
d) Poisson (answer)
e) Binomial

14) If the ……. is known, we can apply z test?
a) Mean
b) Median
c) Mode
f) Standard deviation (answer)
d) Range

15) If a statistically significant difference in blood pressure change at the end of a year for the two activities was found, then:
a. It cannot be concluded that the difference in activity caused a difference in the change in blood pressure because in the course of a year there are lots of possible confounding variables.
b. Whether or not the difference was caused by the difference in activity depends on what else the participants did during the year.
c. It cannot be concluded that the difference in activity caused a difference in the change in blood pressure because it might be the opposite, that people with high blood pressure were more likely to read a book than to walk.
d. It can be concluded that the difference in activity caused a difference in the change in blood pressure because of the way the study was done.(right answer)

16) What is one of the distinctions between a population parameter and a sample statistic?
a. A population parameter is only based on conceptual measurements, but a sample statistic is based on a combination of real and conceptual measurements.
b. A sample statistic changes each time you try to measure it, but a population parameter remains fixed. (right answer)

c. A population parameter changes each time you try to measure it, but a sample statistic remains fixed across samples.
d. The true value of a sample statistic can never be known but the true value of a population parameter can be known.

17) A magazine printed a survey in its monthly issue and asked readers to fill it out and send it in. Over 1000 readers did so. This type of sample is called
a. a cluster sample.
b. a self-selected sample. (right answer)
c. a stratified sample.
d. a simple random sample.

18) Which of the following would be most likely to produce selection bias in a survey?
a. Using questions with biased wording.
b. Only receiving responses from half of the people in the sample.
c. Conducting interviews by telephone instead of in person.
d. Using a random sample of students at a university to estimate the proportion of people who think the legal drinking age should be lowered. (right answer)

19) Which of the following would indicate that a dataset is not bell-shaped?
a. The range is equal to 5 standard deviations.
b. The range is larger than the interquartile range.
c. The mean is much smaller than the median. (right answer)
d. There are no outliers.

20) A scatter plot of number of teachers and number of people with college degrees for cities in
California reveals a positive association. The most likely explanation for this positive association is:
a. Teachers encourage people to get college degrees, so an increase in the number of teachers is causing an increase in the number of people with college degrees.
b. Larger cities tend to have both more teachers and more people with college degrees, so the association is explained by a third variable, the size of the city. (right answer)
c. Teaching is a common profession for people with college degrees, so an increase in the number of people with college degrees causes an increase in the number of teachers.
d. Cities with higher incomes tend to have more teachers and more people going to college, so income is a confounding variable, making causation between number of teachers and number of people with college degrees difficult to prove.

21) What is the effect of an outlier on the value of a correlation coefficient?
a. An outlier will always decrease a correlation coefficient.
b. An outlier will always increase a correlation coefficient.
c. An outlier might either decrease or increase a correlation coefficient, depending on

where it is in relation to the other points. (right answer)
d. An outlier will have no effect on a correlation coefficient.

22) One use of a regression line is
a. to determine if any x-values are outliers.
b. to determine if any y-values are outliers.
c. to determine if a change in x causes a change in y.
d. to estimate the change in y for a one-unit change in x. (right answer)

23) Pick the choice that best completes the following sentence. If a relationship between two variables is called statistically significant, it means the investigators think the variables are
a. related in the population represented by the sample. (right answer)
b. not related in the population represented by the sample.
c. related in the sample due to chance alone.
d. very important.

24) Failing to reject the null hypothesis when it is false is:
a. alpha
b. Type I error
c. beta
d. Type II error(right answer)

25) When asked questions concerning personal hygiene, people commonly lie. This is an example of:
a. sampling bias
b. confounding
c. non-response bias
d. response bias(right answer)

26) Select the order of sampling schemes from best to worst.
a. simple random, stratified, convenience(right answer)
b. simple random, convenience, stratified
c. stratified, simple random, convenience
d. stratified, convenience, simple random

27) A random sample of 5 mosquitos is sampled. The number of mosquitos carrying the West Nile Virus in the sample is an example of which random variable?
a. normal
b. student's t
c. binomial(right answer)
d. uniform
e. none of the above

28) A political scientist is studying voters in California. It is appropriate for him to use a mean to describe
a. the age of a typical voter. (right answer)
b. the party affiliation of a typical voter.
c. the sex of a typical voter.
d. the county of residence of a typical voter.
e. none of the above

29) The long-run average of a random variable is
a. the expected value(right answer)
b. the coefficient of determination
c. the standard deviation
d. the mode
e. none of the above

30) If population A has a larger standard deviation than population B, which of the following is NOT true?
a. Population B has a smaller variance than population A.
b. The mean of a sample of 20 from population A has a larger standard deviation than the mean of a sample of 20 from population B.
c. A typical observation from population A will be farther from the mean of population A than a typical observation from B will be from the mean of population B.
d. The mean of a sample from population A will on average be larger than the mean of a sample from population B. (right answer)
e. none of the above

31) An inspector needs to learn if customers are getting fewer ounces of a soft drink than the 28 ounces stated on
the label. After she collects data from a sample of bottles, she is going to conduct a test of a hypothesis. She
should use
a. a two tailed test.
b. a one tailed test with an alternative to the right.
c. a one tailed test with an alternative to the left. (right answer)
d. either a one or a two tailed test because they are equivalent.
e. none of the above

32) You are conducting a one-sided test of the null hypothesis that the population mean is 532 versus the
alternative that the population mean is less than 532. If the sample mean is 529 and the p-value is 0.01, which
of the following statements is true?
a. There is a 0.01 probability that the population mean is smaller than 529.
b. The probability of observing a sample mean smaller than 529 when the population mean is

532 is 0.01. (right answer)
c. There is a 0.01 probability that the population mean is smaller than 532.
d. If the significance level is 0.05, you will accept the null hypothesis.
e. none of the above

33) Half of the observations in a data set are greater than the
a. mean.
b. median. (right answer)
c. mode.
d. standard deviation.
e. none of the above

34) Which of the following is not a discrete random variable:
a) The lifetime of a light bulb. (right answer)
b) The number of checkout lines operating at a large grocery store.
c) The number of defective tires on a car.
d) The number of pages in a book.
e) All of them are discrete.

35) A statistic is:
a) a sample characteristic (correct answer)
b) a population characteristic
c) unknown
d) normally distributed
e) null distribution

36) Which of the following occurs if a researcher incorrectly rejects a true null hypothesis?
a) alpha
b) Type II error
c) beta
d) Type I error (correct answer)
e) variance

37) Which one of these statistics is unaffected by outliers?
a) Mean
b) Interquartile range (correct answer)
c) Standard deviation
d) Range

38) Which one of the following variables is not categorical?
a) Age of a person (correct answer)
b) Gender of a person: male or female.
c) Choice on a test item: true or false.
d) Marital status of a person (single, married, divorced, other)

e) Choice of a question: yes or no

39. The formula E(X) = np applies in which of these situations?

A. For all discrete random variables.
B. For all continuous random variables.
C. For all normal random variables.
D. For all binomial random variables

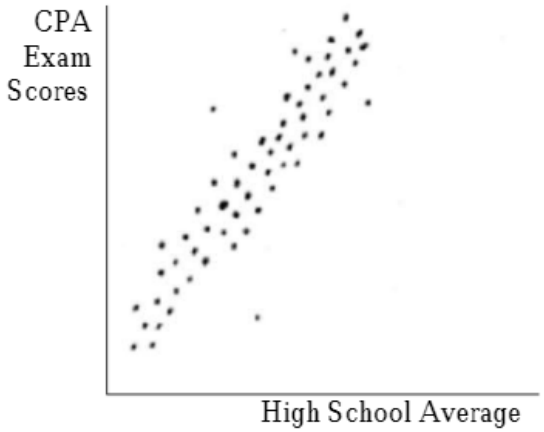
40. Which of the following would be true if people made decisions solely based on maximizing their "expected monetary return" (i.e., the expected value of the money they would make or save)?
A. People wouldn't buy insurance or lottery tickets.
B. People would buy lots of insurance.
C. People would buy lots of lottery tickets.
D. People would always buy an extended warranty if it was offered.

41. One purpose of statistical inference is:
A. To describe sample data with summary statistics from the sample.
B. To describe population data with summary statistics from the population.
C. To make conclusions about populations based on information from a sample.
D. To make conclusions about samples based on information from the population.

42. Which of the following situations results in paired samples as opposed to independent samples?
A. A random sample of Democrats and a separate random sample of Republicans are asked their opinions on legalization of marijuana.
B. Participants in a medical study are randomly assigned to receive either a drug or a placebo.
C. Adults in a random sample are categorized as male/female and asked their opinion on an issue.
D. None of the above situations result in paired samples, they all result in independent samples.

43.

A researcher wants to determine whether there is a relationship between high school average and scores on the CPA exam.



a) The chart above is known as a _____ . b) The data suggest what kind of relationship, if any: _____

What should come into the above spaces in order?

A. scatter plot / linear and positive
B. box plot / linear and positive
C. scatter plot / linear and negative
D. scatter plot / non-linear and positive

44. ***Use the following information to answer the next five exercises:*** **In a survey of 100 stocks on NASDAQ, the average percent increase for the past year was 9% for NASDAQ stocks.**

1. The "average increase" for all NASDAQ stocks is the:

A. population
B. statistic
C. parameter
D. sample
E. variable

2. All of the NASDAQ stocks are the:

A. population
B. statistics
C. parameter
D. sample
E. variable

3. Nine percent is the:

A. **population**
B. **statistics**
C. **parameter**
D. **sample**
E. **variable**

4. The 100 NASDAQ stocks in the survey are the:

A. **population**
B. **statistic**
C. **parameter**
D. **sample**
E. **variable**

5. The percent increase for one stock in the survey is the:

A. **population**
B. **statistic**
C. **parameter**
D. **sample**
E. **variable**

45. Correct answer: C -> The lowest data value is the median.

The statement that describe the illustration is:

a. the mean is equal to the median.

b. There is no first quartile.

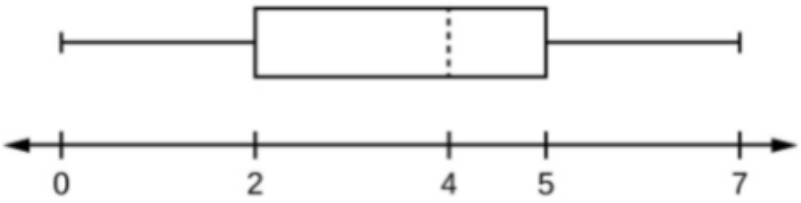
c. The lowest data value is the median.

d. The median equals $\frac{Q_1+Q_3}{2}$.

46. For which distribution is the median not equal to the mean?

A. Uniform
B. Exponential
C. Normal
D. Student *t*

47. Correct answer: B



Which of the following is true for the box plot in (Figure)?

a. Twenty-five percent of the data are at most five.

b. There is about the same amount of data from 4–5 as there is from 5–7.

c. There are no data values of three.

d. Fifty percent of the data are four.

48. Different types of writing can sometimes be distinguished by the number of letters in the words used. A student interested in this fact wants to study the number of letters of words used by Tom Clancy in his novels. She opens a Clancy novel at random and records the number of letters of the first 250 words on the page.
What kind of data was collected?

A. qualitative
B. quantitative continuous
C. quantitative discrete

49. What does it mean when a data set has a standard deviation equal to zero?

A. All values of the data appear with the same frequency.
B. The mean of the data is also zero.
C. All of the data have the same value.
D. There are no data to begin with.

50. Using a sample of 15 Stanford-Binet IQ scores, we wish to conduct a hypothesis test. Our claim is that the mean IQ score on the Stanford-Binet IQ test is more than 100. It is known that the standard deviation of all Stanford-Binet IQ scores is 15 points. The correct distribution to use for the hypothesis test is:

A. Binomial
B. Student's *t*
C. Normal
D. Uniform

51. For ease of data entry into a university database, 1 indicates that the student is an undergraduate and 2 indicates that the student is a graduate student. In this case data are

a) categorical.
b) quantitative.
c) either categorical or quantitative.
d) neither categorical nor quantitative.
e) scale.

52. If a data set has an even number of observations, the median

a) can not be determined.

b) is the average value of the two middle items.
c) must be equal to the mean.
d) is the average value of the two middle items when all items are arranged in ascending order.
e) is the 75th percentage.

53. Which of the following is NOT an assumption of the Binomial distribution?

(a) All trials must be identical.
(b) All trials must be independent.
(c) Each trial must be classified as a success or a failure.
(d) The number of successes in the trials is counted.
(e) The probability of success is equal to 0.5 in all trials.

54. Which one of these statistics is unaffected by outliers?

a) Mean
b) Interquartile range
c) Standard deviation
d) Range

55. Which one of the following variables is not categorical?

a) Age of a person.
b) Gender of a person: male or female.
c) Choice on a test item: true or false.
d) Marital status of a person (single, married, divorced, other)

56. What is one of the distinctions between a population parameter and a sample statistic?
a. A population parameter is only based on conceptual measurements, but a sample statistic is based on a combination of real and conceptual measurements.
b. A sample statistic changes each time you try to measure it, but a population parameter remains fixed.
c. A population parameter changes each time you try to measure it, but a sample statistic remains fixed across samples.
d. The true value of a sample statistic can never be known but the true value of a population parameter can be known.

57. A magazine printed a survey in its monthly issue and asked readers to fill it out and send it in. Over 1000 readers did so. This type of sample is called
a. a cluster sample.
b. a self-selected sample.
c. a stratified sample.

d. a simple random sample.

58. Which of the following would be most likely to produce selection bias in a survey?
a. Using questions with biased wording.
b. Only receiving responses from half of the people in the sample.
c. Conducting interviews by telephone instead of in person.
d. Using a random sample of students at a university to estimate the proportion of people who think the legal drinking age should be lowered.

59. Which of the following would indicate that a dataset is not bell-shaped?
a) The range is equal to 5 standard deviations.
b) The range is larger than the interquartile range.
c) The mean is much smaller than the median.
d) There are no outliers.

60. Statistics is the science of conducting studies to
A) collect, organize, summarize, analyze, and draw conclusions from data.
B) hypothesize, experiment, and form conclusions.
C) solve a system of equations.
D) monitor, study, and report on a subject

61. A _____ is a characteristic or attribute of a subject that can assume different values?
A) datum B) variable C) sample D) exponent

62. Variables with values that are determined by chance are called _____.
A) erratic variables.
B) specialized.
C) random variables.
D) inconsistent variables.

63. Each value in a data set may be referred to as either a data value or a(n) _____.
A) datum B) atom C) point D) subdata

64. Inferential statistics is based on probability.
A) True B) False

65. Which of the following correctly describes the relationship between a sample and a population?
A) A sample is a group of subjects selected from a population to be studied.
B) A population is a group of samples that may or may not be included in a study.
C) A sample is a group of populations that are subject to observation.

D) A population and a sample are not related.

66. In which branch of statistics would a researcher acquire twenty-five 2000 Toyota Celicas, drive them until they had a major mechanical failure, record the final mileage, and then write a report for Car and Driver?
A) descriptive statistics
B) inferential statistics
C) predictive statistics
D) differential statistics

67. Determine which branch of statistics was used to make the following statement. In 2025, the world population is predicted to be 8 billion people.
A) descriptive statistics
B) time series statistics
C) differential statistics
D) inferential statistics

68. What level of measurement classifies data into mutually exclusive categories in which no order or ranking can be imposed on the data?
A) ordinal B) interval C) nominal D) ratio

69. What level of measurement allows for the ranking of data, a precise difference between units of measure, and also includes a true zero?
A) interval B) ratio C) nominal D) ordinal

70. Classifying the fruit in a basket as apple, orange, or banana, is an example of the _____ level of measurement?
A) ordinal B) ratio C) interval D) nominal

71. Rating a restaurant by a number of stars is an example of an ordinal level of measurement.
A) False B) True

72. A person's hair color would be an example of a quantitative variable.
A) False B) True

73. The variable of height is an example of a quantitative variable.
A) True B) False

74. The number of birds in a tree is an example of a continuous variable.
A) True B) False

75. Which one of the following data are continuous?
A) the rankings of the trees, from most numerous to least numerous
B) the number of representatives of each species in the park
C) the number of species of trees in a park
D) the average height of a sample of trees

76. If you were told that four students from a class of twenty were questioned for a poll about study habits, this would be an example of _____.
A) stratified sampling
B) random sampling
C) systematic sampling
D) cluster sampling

77. What type of sampling is being employed if the country is divided into economic classes and a sample is chosen from each class to be surveyed?
A) cluster sampling
B) stratified sampling
C) systematic sampling
D) random sampling

78. Questioning every 14th customer leaving a theatre about the movie they had seen, would be an example of systematic sampling.
A) True B) False

79. An independent variable can also be called a(n)
A) suggestive variable.
B) explanatory variable.
C) outcome variable.
D) free variable.

80.Which of the following best defines the relationship between confounding, dependent, and independent variables?
A) The confounding variable influences the dependent variable, but is not separated from the independent variable.
B) The confounding variable influences the independent variable, but has no effect on the dependent variable.
C) The confounding variable may cause the dependent variable to act independently
D) The influence of the confounding variable cannot be separated from the influence of the dependent variable.

81. (Answer C)

The following frequency distribution presents the weights in pounds (lb) of a sample of visitors to a health clinic.

Weight (lb)	Frequency
90-99	1
100-109	4
110-119	4
120-129	3
130-139	7
140-149	6
150-159	4
160-169	2

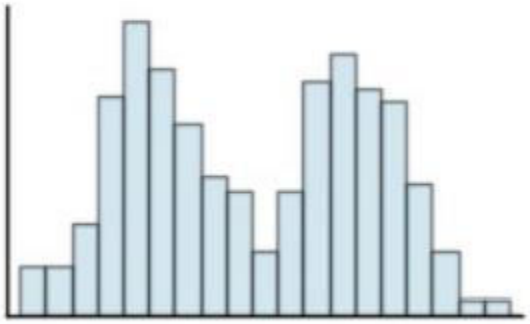
What is the class width?

A) 9 B) 11 C) 10 D) 80

82. Which of the following does not need to be done when constructing a frequency distribution?
A) select the number of classes desired
B) use classes that are mutually exclusive
C) find the range
D) make the class width an even number

83. ANSWER : B

) Classify the histogram as unimodal or bimodal.



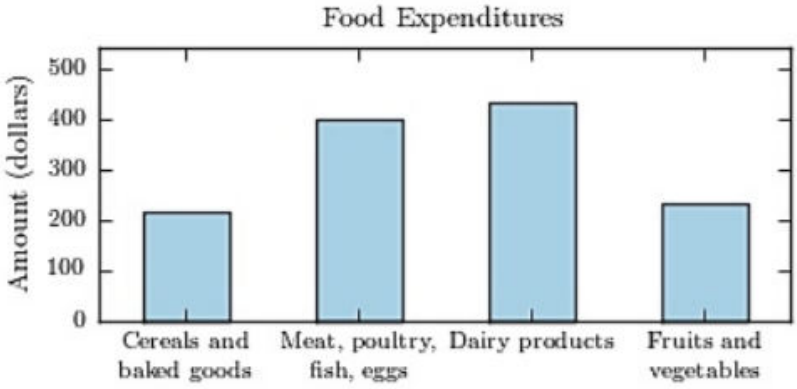
A) unimodal

B) bimodal

84. Answer C

The following bar graph presents the average amount a certain family spent, in dollars, on various food categories in a recent year.

On which food category was the most money spent?



A) Fruits and vegetables

B) Meat poultry, fish, eggs

C) Dairy products

D) Cereals and baked goods

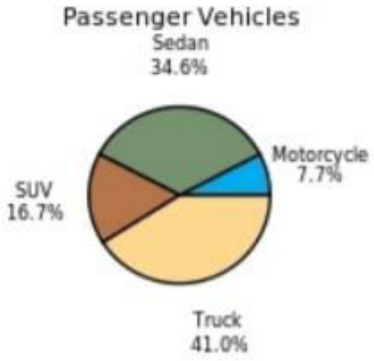
85. ANSWER -> D

The following frequency distribution presents the frequency of passenger vehicles th pass through a certain intersection from 8:00 AM to 9:00 AM on a particular day.

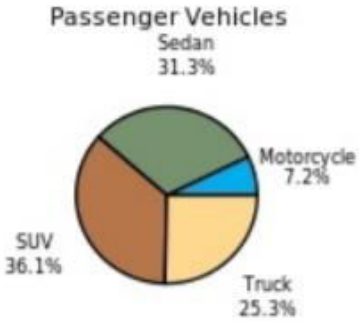
Vehicle Type	Frequency
Motorcycle	9
Sedan	20
SUV	25
Truck	39

Construct a pie chart for the data.

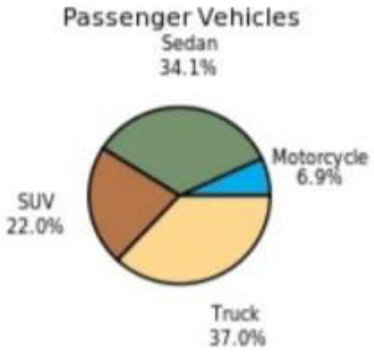
A)



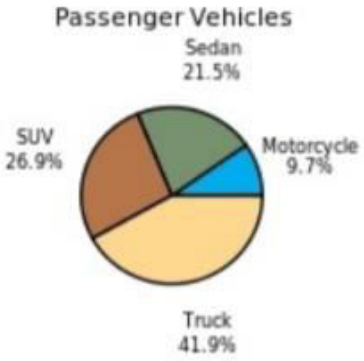
B)



C)

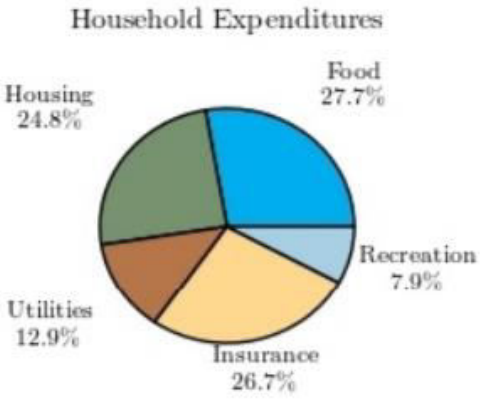


D)



86. Answer B

Following is a pie chart that presents the percentages spent by a certain household on its five largest annual expenditures. What percentage of the money spent was spent on food, housing, and utilities?



A) 60.4% B) 65.4% C) 52.5% D) 47%

87. Suppose that the mean of the sampling distribution for the difference in two sample proportions is 0. This tells us that:

A. The two population proportions are both 0.
B. The two population proportions are equal to each other.
C. The two sample proportions are both 0.
D. The two sample proportions are equal to each other.

88. The expected value of a random variable is

A. always computed as np.
B. the value that has the highest probability of occurring.
C. always one of the possible values for the random variable.
D. the mean value over an infinite number of observations of the variable

89. When a random sample is to be taken from a population and a statistic is to be computed, the statistic can also be thought of as

A. A point estimate
B. A random variable
C. Both of the above
D. None of the above

90. _____ is a decision-making process for evaluating claims about a population, based on information obtained from samples. (Hypothesis testing)

91. If p-value < , reject H0; If p-value , do not reject H0.

92. Type 1 Error() → Reject H0 When H0 is True
Type 2 Error() → Fail to Reject When H0 is False

93. A population is the collection of all items of interest.
A sample is an observed subset of the population.
A parameter is a specific characteristic of a population.
A statistic is a specific characteristic of a sample .

94. ……. protects us from the influences of all the features of our population, even ones that we may not have thought about.(randomizing)

95. Whenever we estimate the standard deviation of a sampling distribution, we call it a …..(standart error)

96. If the population standard deviation "σ" is unknown, we can substitute the sample standard deviation(true or false)

97. When we use t distribution, the ' t ' value depends on degrees of freedom (true)

98. The Null Hypothesis is always about a population parameter(true)

99. When we evaluate hypothesis testing, H0 never accepted.(true)

100. If the standart deviation is known, we can apply z test(true)

[CEVAPLARI YAZARSANIZ iyi olur dedi hocamız]

101. We are 95% confident that the true mean weight of people living in a town is between 65 and 80 kg. According to this information, which assumption is not true?
A)the true mean may be in this interval
B)the true mean may not be in this interval
C)95% of intervals formed in this manner will contain the true mean
D)If we evaluate confidence interval in another sample, the probability of not containing true mean is 1%

102. A group of people have used a drug to lose weight. We have data about weight of this group before using this drug and after using this drug. we want to evaluate whether this drug is effective to lose weight or not. In this situation what kind of test can be applied?
A) One sample t test
B) two sample t test
C) paired t test
D) anova

103. If you print the 5-num summary which of the following you can not reach?

a. Median
b. Standard Deviation
c. Q1
d. Maximum value