Fakültə: Beynəlxalq İqtisadiyyat Məktəbi

Fənn: Statistika

Müəllim Həsən Quliyev

1. 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)

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

a. Mean

b. Interquartile range

c. Standard deviation

d. Range

1. A list of 5 pulse rates is: 70, 64, 80, 74, 92. What is the median for this list?

a. *74*b. 76 c. 77 d. 80

1. Which of the following would indicate that a dataset is not bell-shaped (Normal distribution curve)?

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.

**Questions 5 to 7:** The following histogram shows the distribution of the difference between

the actual and “ideal” weights for 119 female students. Notice that percent is given on the

vertical axis. Ideal weights are responses to the question “What is your ideal weight”? The

difference = actual −ideal. (*Source*: **idealwtwomen** dataset on CD.)



1. What is the approximate shape of the distribution?

a. Nearly symmetric.

b. Skewed to the left.

c. Skewed to the right.

d. Bimodal (has more than one peak).

1. The median of the distribution is approximately

a. −10 pounds.

b. 10 pounds.

c. 30 pounds.

d. 50 pounds

*Scenario for Questions 5 to 7, continued*

1. Most of the women in this sample felt that their actual weight was

a. about the same as their ideal weight.

b. less than their ideal weight.

c. greater than their ideal weight.

d. no more than 2 pounds different from their ideal weight.

1. In a truly normal frequency distribution
        a. the mean always is the same as the standard deviation
         b. the mean is never the same as the mode
        c. the mode is never the same as the median
        d. the mean always is the same as the median
2. How might the standard deviation (S) of a normal distribution be greater  than the mean?
        a. S is given by a square root, and the square root is larger than the fraction.
        b. In a normal distribution, the variance must equal the mean.
        c. If some scores are negative, the mean could be very small despite a large S.
        d. The median would have to be less than the skew.
3. In a class of 100, the mean on a certain exam was 50, the standard  deviation, 0. This means
        a. half the class had scores less than 50
        b. there was a high correlation between ability and grade
        c. everyone had a score of exactly 50
        d. half the class had 0's and half had 50's
4. Suppose the mean on the final exam is 24 (of 40), with a standard deviation of 1.5.  If you get a 21, how well do you do (relative to the rest of the class)?
        a. very poorly--perhaps the lowest score
        b. not well, but somewhere in the C's
        c. OK--about average
        d. nicely--better than the median
5. You can claim that there is a significant difference between scores from two  groups if
       a. the difference between the means is large compared to the standard error
        b. the means are large compared to the standard error
        c. the means are small compared to the standard error
        d. the difference between the standard deviations is large compared to the means
6. The grades of a group of 1000 students in an exam are normally distributed with a mean of 70 and a standard deviation of 10. A student from this group is selected randomly.
a) Find the probability that his/her grade is greater than 80.
b) Find the probability that his/her grade is less than 50.
c) Find the probability that his/her grade is between 50 and 80.
d) Approximately, how many students have grades greater than 80?
7. Consider 3 events A and B with: P(A) = 0.2, P(B)=0.7, and P(A|B) = 0.3. Find the probabilities that

(a) both events occur

(b) at least one of the events occurs

15-18. The following table shows the number of credit hours registered for by 20 randomly selected

students in a class.



1. The sample mean is

 (A) 4.68 (B) 8.9 (C) 9.8 (D) 17.8 (E) None of these

1. The class that contains the median is

 (A) 4 -6 (B) 7- 9 (C) 10-12 (D) 13 - 15

1. The modal class is

 (A) 4 -6 (B) 7 -9 (C) 10 – 12 (D) 13 - 15

1. Which of the following statements is correct?

 (A) The range is a measure of central tendency.

 (B) The median is a measure of dispersion.

 (C) For a symmetric distribution, the mean is equal to the median.

 (D) For a skewed distribution, the variance is a negative number.

 (E) The variance is a measure of central tendency.

**19-20**



1. The probability of at least 3 phone calls is

 (A) .5 (B) .2 (C) .3 (D) .7 (E) None of these



1. The mean number of phone calls is

 (A) 3.2 (B) 3.5 (C) 11.6 (D) 1.36 (E) 1

1. Suppose that 80% of all voters in a city support candidate A. Assume that 40 voters in the city are randomly selected, what is the expected number of voters that will support candidate A in such a group?

 (A) 8 (B) 20 (C) 32 (D) 30 (E) None of these

**22-24**. In order to estimate the mean diameter of a variety of orange, a sample of 25 oranges were selected and the sample mean was found to be 7.5 cm with a sample standard deviation of 1.5 cm.

1. The point estimate of the population mean is

 (A) 1.2 (B) 25 (C) 1.44 (D) 75 (E) None of these

1. A 95% confidence interval for the population mean is

 (A) (6.91, 8.09) (B) (5.44, 9.56) (C) (6.88, 8.12) (D) (5.54, 9.46)

1. If a 90% confidence interval is constructed, it will be \_\_\_\_\_ the 95% confidence interval.

 (A) wider than (B) narrower than (C) the same as

**25-28.** The average annual medical expense per family in a small city was $750 in 1998. A random sample of 49 families was selected and their expenses for 1999 had a mean of $800 with a standard deviation of $140. Based on this information, can we conclude at α= 5% that the average annual medical expenses had increased from the 1998 average?

1. Let μ represent the population mean expenditure for 1999. Which of the following is the appropriate alternative hypothesis?

 (A) μ≠ 750 (B) μ= 750 (C) μ> 750 (D) μ< 750 (E) μ> 800

1. The average annual medical expense per family in a small city was $750 in 1998. A random sample of 49 families was selected and their expenses for 1999 had a mean of $800 with a standard deviation of $140. Based on this information, can we conclude at α= 5% that the average annual medical expenses had increased from the 1998 average?

Let μ represent the population mean expenditure for 1999. Which of the following is the appropriate alternative hypothesis?

1. The average annual medical expense per family in a small city was $750 in 1998. A random sample of 49 families was selected and their expenses for 1999 had a mean of $800 with a standard deviation of $140. Based on this information, can we conclude at α= 5% that the average annual medical expenses had increased from the 1998 average?

Which of the following is the appropriate rejection region?

1. The average annual medical expense per family in a small city was $750 in 1998. A random sample of 49 families was selected and their expenses for 1999 had a mean of $800 with a standard deviation of $140. Based on this information, can we conclude at α= 5% that the average annual medical expenses had increased from the 1998 average? Which of the following is the value of the test statistic?

1. The average annual medical expense per family in a small city was $750 in 1998. A random sample of 49 families was selected and their expenses for 1999 had a mean of $800 with a standard deviation of $140. Based on this information, can we conclude at α= 5% that the average annual medical expenses had increased from the 1998 average?

Which of the following is the appropriate conclusion?

1. The general partner of a limited partnership firm has told a potential investor that the mean monthly rent for a 3-bedroom home in the area is $500. The investor wants to check out this claim on her own. She obtains the monthly rental charges for a random sample of 9 three-bedroom homes in order to test H0: μ= 500 against Ha: μ≠500, at α= 10%. The sample mean is $520 with a sample standard deviation of $48. Which of the following is the appropriate rejection region?

1. The general partner of a limited partnership firm has told a potential investor that the mean monthly rent for a 3-bedroom home in the area is $500. The investor wants to check out this claim on her own. She obtains the monthly rental charges for a random sample of 9 three-bedroom homes in order to test H0: μ= 500 against Ha: μ≠500, at α= 10%. The sample mean is $520 with a sample standard deviation of $48.

What is the value of the test statistic?

1. The general partner of a limited partnership firm has told a potential investor that the mean monthly rent for a 3-bedroom home in the area is $500. The investor wants to check out this claim on her own. She obtains the monthly rental charges for a random sample of 9 three-bedroom homes in order to test H0: μ= 500 against Ha: μ≠500, at α= 10%. The sample mean is $520 with a sample standard deviation of $48.

Which of the following is the correct conclusion?

1. The general partner of a limited partnership firm has told a potential investor that the mean monthly rent for a 3-bedroom home in the area is $500. The investor wants to check out this claim on her own. She obtains the monthly rental charges for a random sample of 9 three-bedroom homes in order to test H0: μ= 500 against Ha: μ≠500, at α= 10%. The sample mean is $520 with a sample standard deviation of $48.

In order for the above procedure to be valid, what assumption will be necessary?

1. A consumer advocate claims that more than 10% of the bolts from supplier A are defective. To test this claim, the correct alternative hypothesis is

1. For problem, consider this information: On any given day, the probability it will rain is 0.32; the probability the wind will blow is 0.2; and the probability that it will rain and the wind will blow is 0.1.

For a randomly selected day, what is the probability that it will rain or the wind will blow?

1. For problem, consider this information: On any given day, the probability it will rain is 0.32; the probability the wind will blow is 0.2; and the probability that it will rain and the wind will blow is 0.1.

For a randomly selected day, what is the probability that it will NOT rain and the wind will NOT blow?

1. Use the following information for problem. The students in a small high school were surveyed. Each student was asked whether he or she used a safety belt whenever driving. This information and the gender of the student was recorded as follows:



What is the probability that a randomly selected student is a female given that the person is a seat belt user?

1. Use the following information for problem. The students in a small high school were surveyed. Each student was asked whether he or she used a safety belt whenever driving. This information and the gender of the student was recorded as follows:



Is the use of a safety belt independent of gender?

1. For problem, consider that 1% of a population has a particular disease. A new test for identifying the disease has been developed. If the person has the disease, the test is positive 94% of the time. If the person does not have the disease, the test is positive 2% of the time.

What is the probability that a randomly selected person from this population tests positive?

1. For problem, consider that 1% of a population has a particular disease. A new test for identifying the disease has been developed. If the person has the disease, the test is positive 94% of the time. If the person does not have the disease, the test is positive 2% of the time.

A person is randomly selected from this population and tested. She tests positive. Which of the following best represents the probability that she has the disease?

1. Use the following information for problem. On any given day, the probability that Megan will be late for work is 0.2.Whether or not she is late to work is independent from day to day.Megan was late to work today. What is the probability that she will NOT be late to work tomorrow?
2. Use the following information for problem. On any given day, the probability that Megan will be late for work is 0.2.Whether or not she is late to work is independent from day to day.Which of the following is closest to the probability that Megan will be late to work at least one of the five days next week?
3. Use the following information for problem. On any given day, the probability that Megan will be late for work is 0.2.Whether or not she is late to work is independent from day to day.What is the probability that Megan will be on time exactly three days and then be late on the fourth one?
4. Let X be a normal random variable with mean 20 and standard deviation 5.What is the probability that a randomly selected value of X is between 15 and 25?
5. A random sample of size 25 is selected from a population that is normally distributed with a mean of 15 and a standard deviation of 4.What is the sampling distribution of the sample mean?
6. Find t\* such that the probability that a randomly selected observation from a t-distribution with 16 degrees of freedom is less than t\* is 0.1.
7. Let z be a standard normal random variable. Find the probability that a randomly selected value of z is between –2.1 and 0.4.
8. A researcher set a 95% confidence interval on the mean length of fish in a recreational lake and found it to be from 6.2 to 8.7 inches. Which of the following is a proper interpretation of this interval?
9. A large university wanted to know whether toilet paper in the campus restrooms should be hung so that the sheets rolled off over the top or under the bottom of the roll. Two hundred students were randomly selected to participate in a survey. Each selected student was asked his or her preference on how to hang the toilet paper. Researchers found 66% preferred that the sheets roll over the top of the roll. Which of the following would be used to set a 95% confidence interval on the proportion of this university's student population favoring the sheets to roll off the top?
10. An owner of a swimming pool wants to know whether or not she needs to add chlorine to the water. Because of costs and the fact that too much chlorine is unpleasant for swimmers, she wants to be sure that chlorine is needed before adding it. What is the swimming pool owner's null hypothesis and what would be a type I error?
11. Suppose the scores on an examination are normally distributed with a mean of 50 and a standard deviation of 10. What is the probability that the score of a student will be higher than 56.5?

1. Suppose the scores on an examination are normally distributed with a mean of 50 and a standard deviation of 10. What proportion of the students score below 45?

1. Suppose the scores on an examination are normally distributed with a mean of 50 and a standard deviation of 10.what is the z score that corresponds to the score 44?

1. Suppose the scores on an examination are normally distributed with a mean of 50 and a standard deviation of 10.What is the raw score that corresponds to z = 1.5?
2. Suppose the scores on an examination are normally distributed with a mean of 50 and a standard deviation of 10.If repeated samples of size n = 25 is taken from the scores, what is the standard deviation of the distribution of the sample mean?

1. A consumer advocate claims that more than 10% of the bolts from supplier A are defective. To test this claim, the correct alternative hypothesis is

1. A consumer claims that car model of type 1 has a lower average miles per gallon than car model of type 2. Let μ1 and μ2 represent the average miles per gallon for types 1 and 2 respectively. Which of the following is the correct null hypothesis?
2. Suppose the coefficient of correlation between the two variables x and y was found to be 0.96, we can say that
3. Eleven cars of a certain model, between one and seven years of age, were randomly selected from the classified ads. The following summary statistics on their ages (x in years) and prices (y in 1000 dollars) were used to obtain the regression equation



Which of the following statements is correct?

1. Eleven cars of a certain model, between one and seven years of age, were randomly selected from the classified ads. The following summary statistics on their ages (x in years) and prices (y in 1000 dollars) were used to obtain the regression equation



 The predicted price for a 5-year-old car is

1. A sample consists of four observations: {1, 3, 5, 7}. What is the standard deviation?
2. Suppose a researcher conducts an experiment to test a hypothesis. If she doubles her sample size, which of the following will increase?

I. The power of the hypothesis test.
II. The effect size of the hypothesis test.
III. The probability of making a Type II error.

1. In hypothesis testing, which of the following statements is always true?

I. The P-value is greater than the significance level.
II. The P-value is computed from the significance level.
III. The P-value is the parameter in the null hypothesis.
IV. The P-value is a test statistic.
V. The P-value is a probability.

1. Which of the following would be a reason to use a one-sample t-test instead of a one-sample z-test?

I. The standard deviation of the population is unknown.
II. The null hypothesis involves a continuous variable.
III. The sample size is large (greater than 40).

1. Which one of these statistics is unaffected by outliers?
2. Which of the following would indicate that a dataset is not bell-shaped (Normal distribution curve)?
3. In a truly normal frequency distribution
4. How might the standard deviation (S) of a normal distribution be greater  than the mean?
5. In a class of 100, the mean on a certain exam was 50, the standard  deviation, 0. This means

1. Suppose the mean on the final exam is 24 (of 40), with a standard deviation of 1.5.  If you get a 21, how well do you do (relative to the rest of the class)?

You can claim that there is a significant difference between scores from two  groups if

1. A statistician conducts a test of the following set of hypotheses: *H*0: μ = 29 versus the alternative μ < 29. Based on a random sample of 38, he found the value of the test statistic to be 1.85. What is the *p*-value associated with the test?
2. A botanist believes that she has developed a new fertilizer that promotes growth in flowers better than the standard fertilizer. She randomly assigns half of her flower plants to the new fertilizer and half to the standard one. The plants are each properly cared for throughout the growing season, and the growth of each plant is recorded at the end of the season. Statistically, what does the researcher want to do?
3. A statistician was testing the following set of hypotheses: *H*0: *p* = 0.1 versus *H*a: *p* ≠ 0.1. Using a random sample of size 230, she found *zT* = 1.37.What is the *p*-value associated with this test?
4. The mean value of homes in a large city was reported to be $157,000. The local chamber of commerce thought the mean value of homes was higher than this reported value. Let μ be the mean value of home in the city, and let be the mean value of homes in a randomly selected sample of homes in the city. What is the appropriate set of hypotheses for the chamber of commerce to test?
5. A homeowner thinks she may need a new roof. Because of costs, she wants to be sure a new roof is really needed before having the current one replaced. What is the homeowner's null hypothesis and what would be a type I error?