Final Exam  
STAT 1222  
December 2021

1. A researcher wishes to estimate the average amount spent per person by visitors to a museum. She takes a random sample of forty-five visitors and obtains an average of $26 per person. The population of interest is
(a) the forty-five selected visitors.
(b) the museum researchers.
(c) visitors to the museum.
(d) the amount of money spent per person.
(e) $26 per person.

2. A university dean wants to know how many credits students are taking. They randomly sample some students and put the results in the histogram below. How many students from the sample are taking between 3 and 9 credits?

(a) 3
(b) 11
(c) 3
(d) 2
(e) 2

3. How would you find the mean of data in the range B1 to B6 in Excel?
(a) =MEAN(B1:B6)
(b) =MEAN.S(B1:B6)
(c) =AVERAGE(B1:B6)
(d) =AVERAGE(B1:B6)
(e) =MEAN(B1:B6)

4. Using Excel, how would you find the standard deviation for a sample of data located in cells A1 to A10?
(a) =STDEV.S(A1:A10)
(b) =STDEV.P(A1:A10)
(c) =STDEV(A1:A10)
(d) =STDEV.P(A1:A10)
(e) =STDEV.S(A1:A10)

5. Which of the following would NOT be an appropriate measure of central tendency when looking at incomes of a middle class neighborhood when one neighbor just sold their startup idea to Google for $4,000,000?
(a) Mean
(b) Median
(c) Mode
(d) All other answer choices are acceptable measures of central tendency
(e) No other answer choice was an acceptable measure of central tendency

6. Seven bear tracks were found in the woods. Their length was measured in inches and recorded in the table below:

|        | 9.25 | 9.25 | 9.4 | 9.5 | 9.6 | 9.6 | 9.6 |

If another bear track was found that was 8 in, what would happen to the mean?
(a) It would increase.
(b) Impossible to tell without further information.
(c) It wouldn't change.
(d) It would decrease.

7. Mean, median, and mode are all measures of the ___________ of a data set, while variance and standard deviation measure the ___________ of a data set.
(a) randomness, center
(b) center, spread
(c) spread, randomness
(d) spread, center
(e) center, randomness

8. The score made by a particular student on a national standardized exam is the 75th percentile. This means that
(a) He got about 75% of the answers correct.
(b) About 75% of all scores on the exam were higher than his.
(c) About 75% of all scores on the exam were lower than his.
(d) His score is 75% of the average score.
9. The mean score on a standardized math exam is 77.2; the standard deviation is 9.5. Zack is told that the z-score of his exam score is 1.3. Zack's score is
(a) above average
(b) below average
(c) exactly average
(d) impossible to say without seeing the other scores

10. A company conducts a survey of 1000 randomly selected individuals to get their overall impressions of the past year. Results are shown below. What is the probability that the next person surveyed has a positive impression?

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>410</td>
</tr>
<tr>
<td>Negative</td>
<td>550</td>
</tr>
<tr>
<td>Don't Know</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td>1000</td>
</tr>
</tbody>
</table>

\[
P(\text{Positive}) = \frac{410}{1000} = 0.41
\]

11. The probability distribution of a discrete random variable \(X\) is given by

<table>
<thead>
<tr>
<th>(x)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p(x))</td>
<td>0.34</td>
<td>0.16</td>
<td>0.10</td>
<td>0.02</td>
<td></td>
</tr>
</tbody>
</table>

Find the value of \(p\), the missing probability in the table.

\[
(a) 0.43 \quad (b) 0.49 \quad (c) 0.17 \quad (d) 0.21 \quad (e) 0.28
\]

12. The probability distribution of a discrete random variable \(X\) is shown in the following table:

<table>
<thead>
<tr>
<th>(x)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P(x))</td>
<td>0.24</td>
<td>0.20</td>
<td>0.37</td>
<td>0.05</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Find \(P(x \geq 2)\).

\[
(b) 0.76 \quad (c) 0.56 \quad (d) 0.24 \quad (e) 0.20 \quad (f) 0.20
\]

Use this information for the next two questions: The following two-way contingency table gives the breakdown of the population of adults in a particular locale according to highest level of education and whether or not the individual regularly takes dietary supplements:

<table>
<thead>
<tr>
<th>Education</th>
<th>Takes Supplements</th>
<th>Does Not Take Supplements</th>
</tr>
</thead>
<tbody>
<tr>
<td>No High School Diploma</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>High School Diploma</td>
<td>0.06</td>
<td>0.44</td>
</tr>
<tr>
<td>Undergraduate Degree</td>
<td>0.09</td>
<td>0.20</td>
</tr>
<tr>
<td>Graduate Degree</td>
<td>0.02</td>
<td>0.01</td>
</tr>
</tbody>
</table>

13. An adult is selected at random. The probability that the person's highest level of education is an undergraduate degree is

\[
(a) 0.09 \quad (b) 0.23 \quad (c) 0.37 \quad (d) 0.16 \quad (e) 0.44
\]

14. An adult is selected at random. The probability that the person's highest level of education is a high school diploma OR that the person takes dietary supplements regularly is

\[
(a) 0.06 \quad (b) 0.44 \quad (c) 0.20 \quad (d) 0.05 \quad (e) 0.70
\]

15. 16% of victims of financial fraud know the perpetrator of the fraud personally. Let \(X\) be the number of people in a random sample of 24 victims of financial fraud who knew the perpetrator personally. Then \(X\) is binomial with:

\[
(a) n = 16 \quad p = 0.18
\]

\[
(b) n = 24 \quad p = 0.18
\]

\[
(c) n = 16 \quad p = 0.24
\]

\[
(d) n = 24 \quad p = 0.24
\]

\[
(e) n = 24 \quad p = 0.16
\]

16. A survey shows that people use cell phones an average of 1.6 years with a standard deviation of 0.3 years. A user is randomly selected. If cell phone use is normally distributed, we can use Excel to calculate the probability that the randomly selected user uses their phone for less than 1 year with the function:

\[
(a) =1 - \text{NORM.DIST}(1, 1.6, 0.3, \text{TRUE})
\]

\[
(b) =1 - \text{NORM.INV}(1, 1.6, 0.3)
\]

\[
(c) =\text{NORM.DIST}(1, 1.6, 0.3, \text{TRUE})
\]

\[
(d) =1 - \text{NORM.INV}(1, 1.6, 0.3, \text{FALSE})
\]

17. A survey shows that people use cell phones an average of 1.6 years with a standard deviation of 0.3 years. A user is randomly selected. If cell phone use is normally distributed, we can use Excel to calculate the probability that the randomly selected user uses their phone for more than 1 year with the function:

\[
(a) =1 - \text{NORM.DIST}(1, 1.6, 0.3, \text{TRUE})
\]

\[
(b) =\text{NORM.DIST}(1, 1.6, 0.3, \text{TRUE})
\]

\[
(c) =\text{NORM.DIST}(1, 1.6, 0.3, \text{TRUE})
\]

\[
(d) =\text{NORM.DIST}(1, 1.6, 0.3, \text{FALSE})
\]

\[
(e) =1 - \text{NORM.INV}(1, 1.6, 0.3)
\]
16. A randomly selected sample of size 17 is taken. Assume the population is normally distributed. Which function finds the probability that \( x \) is less than 1.47?
(a) \( t \text{DIST}(1.4, 17, \text{TRUE}) \)
(b) \( t \text{DIST}(1.4, 16, \text{TRUE}) \)
(c) \( t = 1 - t \text{DIST}(1.4, 16, \text{TRUE}) \)
(d) \( t \text{DIST}(1.4, 0.1, \text{TRUE}) \)
(e) \( t \text{DIST}(1.4, 16, \text{TRUE}) \)

19. Which of the following is NOT a property of the t distribution?
(a) The total area under the curve is equal to 1
(b) Bell-shaped
(c) Centered at 0 with standard deviation of 1
(d) The shape is determined by degrees of freedom
(e) Symmetric about the mean

20. Cell phone bills for a city's residents have a mean of $64 and a standard deviation of $14. Random samples of 100 bills are drawn from this population, and the mean of each sample is found. What is the mean of the sampling distribution?
(a) 26.04 (b) 6.44 (c) 14 (d) 6 (e) 0.14

21. Cell phone bills for a city's residents have a mean of $64 and a standard deviation of $14. Random samples of 100 bills are drawn from this population, and the mean of each sample is found. What is the standard deviation of the mean?
(a) 14 (b) 0.14 (c) 64 (d) 6 (e) 0.34

22. In a random sample of 55 people, the mean body mass index (BMI) was 25.7 and the standard deviation was 9.12. Assume the body mass indexes are normally distributed. Which distribution will you use to calculate a confidence interval for BMI?
(a) Cannot be determined
(b) \( z \) distribution
(c) \( p \) distribution
(d) \( t \) distribution
(e) \( c \) distribution

23. In a random sample of 14 people, the mean body mass index (BMI) was 25.7 and the standard deviation was 9.12. Assume the body mass indexes are normally distributed. Which distribution will you use to calculate a confidence interval for BMI?
(a) Cannot be determined
(b) \( z \) distribution
(c) \( p \) distribution
(d) \( t \) distribution
(e) \( c \) distribution

24. A voter wants to know how much time governors spend on average at golf courses per year. He finds data on 12 governors. Assuming the population is normal, what equation should he use to make a confidence interval for the population mean time?
(a) \( 2 \cdot \text{SE} \) distribution
(b) \( z \) distribution
(c) \( p \) distribution
(d) \( \text{t} \) distribution
(e) \( c \) distribution

25. Which of the following statements would correspond to a null hypothesis?
(a) The population proportion is more than 0.26
(b) The population mean is less than 24
(c) The population mean is not 13
(d) The population proportion is at most 0.75

26. Which of the following statements would correspond to an alternative hypothesis?
(a) The population mean is not 22
(b) The population proportion is at least 0.9
(c) The population proportion is less than or equal to 0.90
(d) The population mean is no more than 17

27. The target temperature for a hot beverage the moment it is dispensed from a vending machine is 168°F. A consumer organization believes that a particular brand of machine dispenses beverages whose temperature exceeds 168°F. In this scenario, a Type I error would be:
(a) Deciding that the beverages do not exceed 168°F when they do.
(b) Deciding that the beverages exceed 168°F when they don't.
28. A diabetic claims that the average cost of insulin per year for a Type I diabetic is $5,605. She takes a sample of 100 Type I diabetics and finds their average cost is $5,700 with a standard deviation of $570. Use \( n = 0.05 \) to test the claim.

What are the hypotheses?

- \( H_0: \mu = 5605 \) (claim)
- \( H_a: \mu \neq 5605 \)

(b) \( H_0: \mu \geq 5605 \)
- \( H_a: \mu < 5605 \) (claim)
- \( H_a: \mu \leq 5605 \) (claim)
- \( H_a: \mu > 5605 \) (claim)
- \( H_a: \mu = 5605 \) (claim)

(c) \( H_0: \mu = 5605 \)
- \( H_a: \mu < 5605 \) (claim)
- \( H_a: \mu > 5605 \) (claim)
- \( H_a: \mu \neq 5605 \)

Which function in Excel finds the test statistic?

(a) \( \text{NORMDIST}(5605, 5700, 570, \text{TRUE}) \)
(b) \( \text{NORMDIST}(5700, 5605, 570, \text{TRUE}) \)
(c) \( (5605 - 5700)^2 / 570^2 / 100 \)
(d) \( \text{NORMINV}(5700, 5605, 570^2 / 100) \)

30. A diabetic claims that the average cost of insulin per year for a Type I diabetic is $5,605. She takes a sample of 100 Type I diabetics. She also finds that the p-value for the hypothesis test is 0.0724. Using \( n = 0.05 \), what is the decision?

(a) Reject \( H_0 \)
(b) Reject \( H_a \)
(c) Fail to reject \( H_0 \)
(d) Fail to reject \( H_a \)

31. A diabetic claims that the average cost of insulin per year for a Type I diabetic is $5,605. She takes a sample of 100 Type I diabetics. She performs a hypothesis test and finds that the test statistic is \(-5.5, 10 \). What is the interpretation for the test statistic?

(a) There is enough evidence to reject the claim.
(b) There is enough evidence to support the claim.
(c) There is not enough evidence to support the claim.
(d) There is not enough evidence to reject the claim.

32. A nutritionist claims that the average amount of sugar in a 16 oz soda is at least 54 g. He randomly samples 13 sodas and finds they contain an average of 50 g of sugar with a standard deviation of 4 g. Assume the population is normally distributed and use \( n = 0.05 \) to test the claim.

Which function in Excel finds the test statistic?

(a) \( t(150, 13) \)
(b) \( t(150, 12, \text{TRUE}) \)
(c) \( \text{NORMDIST}(50, 54, 4, \text{TRUE}) \)
(d) \( \text{NORMDIST}(54, 50, 4, \text{TRUE}) \)

33. A nutritionist claims that the average amount of sugar in a 16 oz soda is at least 54 g. He randomly samples 13 sodas. Assume the population is normally distributed and suppose the test statistic is -1.19. Which function in Excel finds the p-value?

(a) \( t(11, -1.19, 12, \text{TRUE}) \)
(b) \( 1 - \text{NORMDIST}(1.19, 0, 1, \text{TRUE}) \)
(c) \( 1 - \text{NORMDIST}(1.19, 13, \text{TRUE}) \)
(d) \( 1 - \text{NORMDIST}(1.19, 10, \text{TRUE}) \)

34. A nutritionist claims that the average amount of sugar in a 16 oz soda is at least 54 g. He randomly samples 13 sodas. Assume the population is normally distributed and suppose the p-value is 0.0054. Using \( n = 0.10 \), what is the decision?

(a) Reject \( H_0 \)
(b) Fail to reject \( H_0 \)
(c) Fail to reject \( H_a \)
(d) Fail to support \( H_0 \)

35. A nutritionist claims that the average amount of sugar in a 16 oz soda is at least 54 g. He randomly samples 13 sodas. Assume the population is normally distributed and suppose the p-value is 0.0054. Using \( n = 0.10 \), what is the decision?

(a) There is enough evidence to reject the claim.
(b) There is enough evidence to support the claim.
(c) There is not enough evidence to support the claim.
(d) There is not enough evidence to reject the claim.