STT 231 – 001
FINAL EXAMINATION REVIEW SHEET III
FALL 2015

PART I
HYPOTHESES TESTING FOR ONE PROPORTION
HYPOTHESES TESTING FOR THE DIFFERENCE BETWEEN TWO PROPORTIONS

1. A two-sided or two-tailed hypothesis test is one in which
   A. the null hypothesis includes values in either direction from a specific standard.
   B. the null hypothesis includes values in one direction from a specific standard.
   C. the alternative hypothesis includes values in one direction from a specific standard
   D. the alternative hypothesis includes values in either direction from a specific standard

   KEY: D

2. Null and alternative hypotheses are statements about
   A. population parameters.
   B. sample parameters.
   C. sample statistics.
   D. it depends - sometimes population parameters and sometimes sample statistics.

   KEY: A

3. Which statement is correct about a \( p \)-value?
   A. The smaller the \( p \)-value the stronger the evidence in favor of the alternative hypothesis.
   B. The smaller the \( p \)-value the stronger the evidence in favor the null hypothesis
   C. Whether a small \( p \)-value provides evidence in favor of the null hypothesis depends on whether the test is one-sided or two-sided.
   D. Whether a small \( p \)-value provides evidence in favor of the alternative hypothesis depends on whether the test is one-sided or two-sided.

   KEY: A

4. A hypothesis test gives a \( p \)-value of 0.03. If the significance level \( \alpha = 0.05 \), the results are said to be
   A. not statistically significant because the \( p \)-value \( \leq \alpha \).
   B. statistically significant because the \( p \)-value \( \leq \alpha \).
   C. practically significant because the \( p \)-value \( \leq \alpha \).
   D. not practically significant because the \( p \)-value \( \leq \alpha \).

   KEY: B

5. A hypothesis test gives a \( p \)-value of 0.050. If the significance level \( \alpha = 0.05 \), the results are said to be
   A. not statistically significant because the \( p \)-value is not smaller than \( \alpha \).
   B. statistically significant because the \( p \)-value \( \leq \alpha \).
   C. practically significant because the \( p \)-value is the same as \( \alpha \).
   D. inconclusive because the \( p \)-value is not smaller nor larger than \( \alpha \).

   KEY: B

6. The likelihood that a statistic would be as extreme or more extreme than what was observed is called a
   A. statistically significant result.
   B. test statistic.
   C. significance level.
   D. \( p \)-value.

   KEY: D

239
7. The data summary used to decide between the null hypothesis and the alternative hypothesis is called a
   A. statistically significant result.
   B. test statistic.
   C. significance level.
   D. p-value.
   KEY: B

8. The designated level (typically set at 0.05) to which the p-value is compared to, in order to decide whether the
   alternative hypothesis is accepted or not is called a
   A. statistically significant result.
   B. test statistic.
   C. significance level.
   D. none of the above.
   KEY: C

9. When the p-value is less than or equal to the designated level of 0.05, the result is called a
   A. statistically significant result.
   B. test statistic.
   C. significance level.
   D. none of the above.
   KEY: A

10. Which of the following conclusions is not equivalent to rejecting the null hypothesis?
    A. The results are statistically significant.
    B. The results are not statistically significant.
    C. The alternative hypothesis is accepted.
    D. The p-value ≤ α (the significance level)
    KEY: B

11. If the result of a hypothesis test for a proportion is statistically significant, then
    A. the null hypothesis is rejected.
    B. the alternative hypothesis is rejected.
    C. the population proportion must equal the null value.
    D. None of the above.
    KEY: A

12. The smaller the p-value, the
    A. stronger the evidence against the alternative hypothesis.
    B. stronger the evidence for the null hypothesis.
    C. stronger the evidence against the null hypothesis.
    D. None of the above.
    KEY: C

13. Which of the following is not a valid conclusion for a hypothesis test?
    A. Reject the null hypothesis.
    B. Do not reject the null hypothesis.
    C. We have proven the null hypothesis is true.
    D. We have proven the alternative hypothesis is true.
    KEY: C and D
14. In hypothesis testing for one proportion, the "null value" is used in which of the following?
   A. The null hypothesis.
   B. The alternative hypothesis.
   C. The computation of the test statistic.
   D. All of the above.
   KEY: D

15. A result is called statistically significant whenever
   A. the null hypothesis is true.
   B. the alternative hypothesis is true.
   C. the $p$-value is less than or equal to the significance level.
   D. the $p$-value is larger than the significance level.
   KEY: C

16. Which one of the following is not true about hypothesis tests?
   A. Hypothesis tests are only valid when the sample is representative of the population for the question of interest.
   B. Hypotheses are statements about the population represented by the samples.
   C. Hypotheses are statements about the sample (or samples) from the population.
   D. Conclusions are statements about the population represented by the samples.
   KEY: C

17. In a hypothesis test which of the following can (and should) be determined before collecting data?
   A. The null and alternative hypotheses.
   B. The value of the test statistic.
   C. The $p$-value.
   D. Whether the test statistic will be positive or negative.
   KEY: A

18. The level of significance (usually .05) associated with a significance test is the probability
   A. that the null hypothesis is true.
   B. that the alternative hypothesis is true.
   C. of not rejecting a true null hypothesis.
   D. of rejecting a true null hypothesis.
   KEY: D

Questions 19 to 22: Suppose the significance level for a hypothesis test is $\alpha = 0.05$.

19. If the $p$-value is 0.001, the conclusion is to
   A. reject the null hypothesis.
   B. accept the null hypothesis.
   C. not reject the null hypothesis.
   D. None of the above.
   KEY: A

20. If the $p$-value is 0.049, the conclusion is to
   A. reject the null hypothesis.
   B. accept the null hypothesis.
   C. not reject the null hypothesis.
   D. None of the above.
   KEY: A
21. If the $p$-value is 0.05, the conclusion is to
   A. reject the null hypothesis.
   B. accept the alternative hypothesis.
   C. not reject the null hypothesis.
   D. None of the above.
   KEY: C

22. If the $p$-value is 0.999, the conclusion is to
   A. reject the null hypothesis.
   B. accept the alternative hypothesis.
   C. not reject the null hypothesis.
   D. None of the above.
   KEY: C

23. In hypothesis testing, a Type 1 error occurs when
   A. the null hypothesis is not rejected when the null hypothesis is true.
   B. the null hypothesis is rejected when the null hypothesis is true.
   C. the null hypothesis is not rejected when the alternative hypothesis is true.
   D. the null hypothesis is rejected when the alternative hypothesis is true.
   KEY: B

24. In hypothesis testing, a Type 2 error occurs when
   A. the null hypothesis is not rejected when the null hypothesis is true.
   B. the null hypothesis is rejected when the null hypothesis is true.
   C. the null hypothesis is not rejected when the alternative hypothesis is true.
   D. the null hypothesis is rejected when the alternative hypothesis is true.
   KEY: C

25. In a hypothesis test the decision was made to not reject the null hypothesis. Which type of mistake could have been made?
   A. Type 1.
   B. Type 2.
   C. Type 1 if it's a one-sided test and Type 2 if it's a two-sided test.
   D. Type 2 if it's a one-sided test and Type 1 if it's a two-sided test.
   KEY: B

26. If, in a hypothesis test, the null hypothesis is actually true, which type of mistake can be made?
   A. Type 1.
   B. Type 2.
   C. Type 1 if it's a one-sided test and Type 2 if it's a two-sided test.
   D. Type 2 if it's a one-sided test and Type 1 if it's a two-sided test.
   KEY: A

27. In an American criminal trial, the null hypothesis is that the defendant is innocent and the alternative hypothesis is that the defendant is guilty. Which of the following describes a Type 2 error for a criminal trial?
   A. A guilty verdict for a person who is innocent.
   B. A guilty verdict for a person who is not innocent.
   C. A not guilty verdict for a person who is guilty
   D. A not guilty verdict for a person who is innocent
   KEY: C
28. A Washington Post poll shows that concerns about housing payments have spiked despite some improvements in the overall economy. In all, 53 percent of the 900 American adults surveyed said they are "very concerned" or "somewhat concerned" about having the money to make their monthly payment. Let $p$ represent the population proportion of all American adults who are "very concerned" or "somewhat concerned" about having the money to make their monthly payment. Which are the appropriate hypotheses to assess if a majority of American adults are worried about making their mortgage or rent payments?

A. $H_0: p = 0.50$ versus $H_a: p > 0.50$
B. $H_0: p \geq 0.50$ versus $H_a: p < 0.50$
C. $H_0: p = 0.53$ versus $H_a: p > 0.53$
D. $H_0: p = 0.50$ versus $H_a: p = 0.53$

KEY: A

Questions 29 to 32: A hypothesis test for a population proportion $p$ is given below:

- $H_0: p = 0.10$
- $H_a: p \neq 0.10$

For each sample size $n$ and sample proportion $\hat{p}$ compute the value of the $z$-statistic.

29. Sample size $n = 100$ and sample proportion $\hat{p} = 0.10$. $z$-statistic = ?
   A. -1.00
   B. 0.00
   C. 0.10
   D. 1.00

KEY: B

30. Sample size $n = 100$ and sample proportion $\hat{p} = 0.15$. $z$-statistic = ?
   A. -1.12
   B. 0.04
   C. 1.12
   D. 1.67

KEY: D

31. Sample size $n = 500$ and sample proportion $\hat{p} = 0.04$. $z$-statistic = ?
   A. -6.84
   B. -4.47
   C. 4.47
   D. 6.84

KEY: B

32. Sample size $n = 500$ and sample proportion $\hat{p} = 0.20$. $z$-statistic = ?
   A. -7.45
   B. -5.59
   C. 5.59
   D. 7.45

KEY: D
Questions 33 to 37: A sample of \( n = 200 \) college students is asked if they believe in extraterrestrial life and 120 of these students say that they do. The data are used to test \( H_0: p = 0.5 \) versus \( H_1: p > 0.5 \), where \( p \) is the population proportion of college students who say they believe in extraterrestrial life. The following Minitab output was obtained:

<table>
<thead>
<tr>
<th>Sample</th>
<th>X</th>
<th>N</th>
<th>Sample p</th>
<th>95.0 % CI</th>
<th>Z-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>120</td>
<td>200</td>
<td>0.600000</td>
<td>(0.532105, 0.667895)</td>
<td>2.83</td>
<td>0.002</td>
</tr>
</tbody>
</table>

33. What is the correct description of the area that equals the \( p \)-value for this problem?
   A. The area to the right of 0.60 under a standard normal curve.
   B. The area to the right of 2.83 under a standard normal curve.
   C. The area to the right of -2.83 under a standard normal curve.
   D. The area between 0.532105 and 0.667895 under a standard normal curve.
   KEY: B

34. Suppose that the alternative hypothesis had been \( H_1: p \neq 0.5 \). What would have been the \( p \)-value of the test?
   A. 0.002
   B. 0.001
   C. 0.004
   D. 0.5
   KEY: C

35. Using a 5% significance level, what is the correct decision for this significance test?
   A. Fail to reject the null hypothesis because the \( p \)-value is greater than 0.05.
   B. Fail to reject the null hypothesis because the \( p \)-value is less than 0.05.
   C. Reject the null hypothesis because the \( p \)-value is greater than 0.05.
   D. Reject the null hypothesis because the \( p \)-value is less than 0.05.
   KEY: D

36. Using a 5% significance level, what is the correct conclusion for this significance test?
   A. The proportion of college students who say they believe in extraterrestrial life is equal to 50%.
   B. The proportion of college students who say they believe in extraterrestrial life is not equal to 50%.
   C. The proportion of college students who say they believe in extraterrestrial life seems to be greater than 50%.
   D. The proportion of college students who say they believe in extraterrestrial life seems to be equal to 60%.
   KEY: C

37. Based on the decision made in question 36, what mistake could have been made?
   A. Type 1.
   B. Type 2.
   C. Neither one; the \( p \)-value is so small that no mistake could have been made.
   KEY: A

38. About 90% of the general population is right-handed. A researcher speculates that artists are less likely to be right-handed than the general population. In a random sample of 100 artists, 83 are right-handed. Which of the following best describes the \( p \)-value for this situation?
   A. The probability that the population proportion of artists who are right-handed is 0.90.
   B. The probability that the population proportion of artists who are right-handed is 0.83.
   C. The probability the sample proportion would be as small as 0.83, or even smaller, if the population proportion of artists who are right-handed is actually 0.90.
   D. The probability that the population proportion of artists who are right-handed is less than 0.90, given that the sample proportion is 0.83.
   KEY: C
39. Consider testing the alternative hypothesis that the proportion of adult Canadians opposed to same-sex marriage in Canada is less than 0.5. The test was conducted based on a poll of $n = 1003$ adults and it had a $p$-value of 0.102. Which of the following describes the probability represented by the $p$-value for this test?

A. It is the probability that fewer than half of all adults in Canada that year were opposed to same-sex marriage.
B. It is the probability that more than half of all adult Canadians that year were opposed to same-sex marriage.
C. It is the probability that a sample of 1003 adults in Canada that year would result in 48% or fewer saying they are opposed to same-sex marriage, given that a majority (over 50%) of Canadian adults actually were opposed that year.
D. It is the probability that a sample of 1003 adults in Canada that year would result in 48% or fewer saying they are opposed to same-sex marriage, given that 50% of Canadian adults actually were opposed that year.

KEY: D

Questions 40 to 44: An airport official wants to prove that the $p_1$ = proportion of delayed flights after a storm for Airline 1 was different from $p_2$ = the proportion of delayed flights for Airline 2. Random samples from the two airlines after a storm showed that 50 out of 100 of Airline 1’s flights were delayed, and 70 out of 200 of Airline 2’s flights were delayed.

40. What are the appropriate null and alternative hypotheses?

A. $H_0: p_1 - p_2 = 0$ and $H_a: p_1 - p_2 \neq 0$
B. $H_0: p_1 - p_2 \neq 0$ and $H_a: p_1 - p_2 = 0$
C. $H_0: p_1 - p_2 = 0$ and $H_a: p_1 - p_2 < 0$
D. $H_0: p_1 - p_2 = 0$ and $H_a: p_1 - p_2 > 0$

KEY: A

41. What is the value of the test statistic?

A. 0.79
B. 2.00
C. 2.50
D. None of the above

KEY: C

42. What is the $p$-value? (computer or calculator for normal distribution required)

A. $p$-value = 0.0062
B. $p$-value = 0.0124
C. $p$-value = 0.0456
D. $p$-value = 0.2148

KEY: B

43. For a significance level of $\alpha = 0.05$, are the results statistically significant?

A. No, the results are not statistically significant because the $p$-value < 0.05.
B. Yes, the results are statistically significant because the $p$-value < 0.05.
C. No, the results are not statistically significant because the $p$-value > 0.05.
D. Yes, the results are statistically significant because the $p$-value > 0.05.

KEY: B

44. Report your conclusion in terms of the two airlines.

A. The proportion of delayed flights after a storm for Airline 1 seems to be different than the proportion of delayed flights after a storm for Airline 2.
B. The results are not statistically significant; there is not enough evidence to conclude there is a difference between the two proportions.
C. The difference in proportions of delayed flights is at least 15%.
D. The proportion of delayed flights after a storm for Airline 1 seems to be the same as the proportion of delayed flights after a storm for Airline 2.

KEY: A
Questions 45 to 49: An airport official wants to prove that the \( p_1 \) = proportion of delayed flights after a storm for Airline A is less than \( p_2 \) = the proportion of delayed flights for Airline B. Random samples from two airlines after a storm showed that 51 out of 200 of Airline A’s flights were delayed, and 60 out of 200 of Airline B’s flights were delayed.

45. What are the appropriate null and alternative hypotheses?
   A. \( H_0: p_1 - p_2 = 0 \) and \( H_a: p_1 - p_2 \neq 0 \)
   B. \( H_0: p_1 - p_2 \leq 0 \) and \( H_a: p_1 - p_2 = 0 \)
   C. \( H_0: p_1 - p_2 = 0 \) and \( H_a: p_1 - p_2 < 0 \)
   D. \( H_0: p_1 - p_2 = 0 \) and \( H_a: p_1 - p_2 > 0 \)

   KEY: C

46. What is the value of the test statistic?
   A. -1.00
   B. -1.50
   C. -2.00
   D. None of the above

   KEY: A

47. What is the \( p \)-value? (computer or calculator for normal distribution required)
   A. \( p \)-value = 0.0228
   B. \( p \)-value = 0.0668
   C. \( p \)-value = 0.1587
   D. \( p \)-value = 0.3174

   KEY: C

48. For a significance level of \( \alpha = 0.05 \), are the results statistically significant?
   A. No, the results are not statistically significant because the \( p \)-value < 0.05.
   B. Yes, the results are statistically significant because the \( p \)-value < 0.05.
   C. No, the results are not statistically significant because the \( p \)-value > 0.05
   D. Yes, the results are statistically significant because the \( p \)-value > 0.05.

   KEY: C

49. Report your conclusion in terms of the two airlines.
   A. The proportion of delayed flights for Airline A seems to be less than the proportion of delayed flights for Airline B.
   B. The results are not statistically significant: there is not enough evidence to conclude that the proportion of delayed flights for Airline A is less than the proportion for Airline B.
   C. The difference in proportions of delayed flights is at least 4%.
   D. The proportion of delayed flights after a storm for Airline A seems to be greater than the proportion of delayed flights after a storm for Airline B.

   KEY: B
PART II
HYPOTHESES TESTING FOR ONE MEAN
HYPOTHESES TESTING FOR THE DIFFERENCE BETWEEN TWO MEANS

50. Which statement is not true about hypothesis tests?
   A. Hypothesis tests are only valid when the sample is representative of the population for the
      question of interest.
   B. Hypotheses are statements about the population represented by the samples.
   C. Hypotheses are statements about the sample (or samples) from the population.
   D. Conclusions are statements about the population represented by the samples.
   KEY: C

51. The primary purpose of a significance test is to
   A. estimate the p-value of a sample.
   B. estimate the p-value of a population.
   C. decide whether there is enough evidence to support a research hypothesis about a sample.
   D. decide whether there is enough evidence to support a research hypothesis about a population.
   KEY: D

52. The level of significance associated with a significance test is the probability
   A. of rejecting a true null hypothesis.
   B. of not rejecting a true null hypothesis.
   C. that the null hypothesis is true.
   D. that the alternative hypothesis is true.
   KEY: A

53. A result is called statistically significant whenever
   A. the null hypothesis is true.
   B. the alternative hypothesis is true.
   C. the p-value is less or equal to the significance level.
   D. the p-value is larger than the significance level.
   KEY: C

54. Which of the following is not one of the steps for hypothesis testing?
   A. Determine the null and alternative hypotheses.
   B. Verify data conditions and calculate a test statistic.
   C. Assuming the null hypothesis is true, find the p-value.
   D. Assuming the alternative hypothesis is true, find the p-value.
   KEY: D

55. Which of the following is not a correct way to state a null hypothesis?
   A. \( H_0: \)
   B. \( H_0: \mu = 10 \)
   C. \( H_0: \mu = 0 \)
   D. \( H_0: \mu = 0.5 \)
   KEY: A
56. In hypothesis testing for one mean, the "null value" is not used in which of the following?
   A. The null hypothesis.
   B. The alternative hypothesis.
   C. The computation of the test statistic.
   D. The (null) standard error.

   KEY: D

57. A null hypothesis is that the average pulse rate of adults is 70. For a sample of 64 adults, the
    average pulse rate is 71.8. A significance test is done and the p-value is 0.02. What is the most
    appropriate conclusion?
   A. Conclude that the population average is 70.
   B. Conclude that the population average is 71.8.
   C. Reject the hypothesis that the population average is 70.
   D. Reject the hypothesis that the sample average is 70.

   KEY: C

58. The p-value for a one-sided test for a mean was 0.04. The p-value for the corresponding two-sided
    test would be:
   A. 0.02
   B. 0.04
   C. 0.06
   D. 0.08

   KEY: D

59. A null hypothesis is that the mean cholesterol level is 200 in a certain age group. The alternative is
    that the mean is not 200. Which of the following is the most significant evidence against the null
    and in favor of the alternative?
   A. For a sample of n = 25, the sample mean is 220.
   B. For a sample of n = 10, the sample mean is 220.
   C. For a sample of n = 50, the sample mean is 180.
   D. For a sample of n = 20, the sample mean is 180.

   KEY: C

60. A test of $H_0: \mu = 0$ versus $H_a: \mu > 0$ is conducted on the same population independently by two
    different researchers. They both use the same sample size and the same value of $\alpha = 0.05$. Which of
    the following will be the same for both researchers?
   A. The p-value of the test.
   B. The power of the test if the true $\mu = 6$.
   C. The value of the test statistic.
   D. The decision about whether or not to reject the null hypothesis.

   KEY: B

61. Which of the following is not true about hypothesis testing?
   A. The null hypothesis defines a specific value of a population parameter, called the null value.
   B. A relevant statistic is calculated from population information and summarized into a "test
      statistic."
   C. A p-value is computed on the basis of the standardized "test statistic."
   D. On the basis of the p-value, we either reject or fail to reject the null hypothesis.

   KEY: B
Questions 62 to 66: An investigator wants to assess whether the mean \( \mu = 185 \) pounds is the average weight of passengers flying on small planes exceeds the FAA guideline of average total weight of 185 pounds (passenger weight including shoes, clothes, and carry-on). Suppose that a random sample of 51 passengers showed an average total weight of 200 pounds with a sample standard deviation of 59.5 pounds. Assume that passenger total weights are normally distributed.

62. What are the appropriate null and alternative hypotheses?
   A. \( H_0: \mu = 185 \) and \( H_a: \mu < 185 \)
   B. \( H_0: \mu = 185 \) and \( H_a: \mu > 185 \)
   C. \( H_0: \mu = 185 \) and \( H_a: \mu \neq 185 \)
   D. \( H_0: \mu \neq 185 \) and \( H_a: \mu = 185 \)

KEY: B

63. What is the value of the test statistic?
   A. \( t = 1.50 \)
   B. \( t = 1.65 \)
   C. \( t = 1.80 \)
   D. None of the above

KEY: C

64. What is the \( p \)-value?
   A. \( p \)-value = 0.039
   B. \( p \)-value = 0.053
   C. \( p \)-value = 0.070
   D. None of the above

KEY: A

65. For a significance level of \( \alpha = 0.05 \), are the results statistically significant?
   A. No, the results are not statistically significant because the \( p \)-value < 0.05.
   B. Yes, the results are statistically significant because the \( p \)-value < 0.05.
   C. No, the results are not statistically significant because the \( p \)-value > 0.05.
   D. Yes, the results are statistically significant because the \( p \)-value > 0.05.

KEY: B

66. Which of the following is an appropriate conclusion?
   A. The results are statistically significant so the average total weight of all passengers appears to be greater than 185 pounds.
   B. The results are statistically significant so the average total weight of all passengers appears to be less than 185 pounds.
   C. The results are not statistically significant so there is not enough evidence to conclude the average total weight of all passengers is greater than 185 pounds.
   D. None of the above.

KEY: A
67. Spatial perception is measured on a scale from 0 to 10. A group of 9th grade students are tested for spatial perception. SPSS was used to obtain descriptive statistics of the spatial perception scores in the sample.

<table>
<thead>
<tr>
<th>One-Sample Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>N Mean Std. Deviation</td>
</tr>
<tr>
<td>Spatial Perception</td>
</tr>
<tr>
<td>18 6.00 1.815</td>
</tr>
<tr>
<td>Std. Error Mean</td>
</tr>
<tr>
<td>.428</td>
</tr>
</tbody>
</table>

Using α = 0.01, is the mean score significantly different from 5?
A. Yes, because $t = 2.34$ and this is greater than $t^* = 2.11$.
B. No, because $t = 2.34$ and this is smaller than $t^* = 3.97$.
C. No, because $t = 0.55$ and this is smaller than $t^* = 2.11$.
D. No, because $t = 2.34$ and this is smaller than $t^* = 2.90$.

KEY: D

68. The amount of time the husband and the wife spend on housework is measured for 15 women and their 15 husbands. For the wives the mean was 7 hours/week and for the husbands the mean was 4.5 hours/week. The standard deviation of the differences in time spent on housework was 2.85.

What is the value of the test statistic for testing the difference in mean time spent on housework between husbands and wives?
A. 0.88
B. 2.40
C. 3.40
D. 4.80

KEY: C

69. The head circumference is measured for 25 girls and their younger twin sisters. The mean of the older twin girls was 50.23 cm and the mean of the younger twins was 49.96 cm. The standard deviation of the differences was 1 cm. Is this difference significant at a significance level of 5%?
A. Yes, because $t = 1.35$ and the $p$-value < 0.10.
B. Yes, because $t = 6.25$ and the $p$-value < 0.10.
C. No, because $t = 1.35$ and the $p$-value > 0.10.
D. No, because $t = 0.27$ and the $p$-value > 0.10.

KEY: C

70. An experiment is conducted with 15 seniors who are taking Spanish at Oak View High School. A randomly selected group of eight students is first tested with a written test and a day later with an oral exam. To avoid order effects, the other seven students are tested in reverse order. The instructor is interested in the difference in grades between the two testing methods. SPSS is used to obtain descriptive statistics for the grades of the two tests.

<table>
<thead>
<tr>
<th>Test</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talk</td>
<td>15</td>
<td>1.523</td>
<td>1.530</td>
</tr>
<tr>
<td>Write</td>
<td>15</td>
<td>4.168</td>
<td>2.047</td>
</tr>
<tr>
<td>Talk - Write</td>
<td>15</td>
<td>-2.643</td>
<td>2.182</td>
</tr>
</tbody>
</table>
Is there a significant difference between the mean grades using the two different testing methods? Use α = 0.05.

A. Yes, because t = -4.69 and the p-value < 0.05.
B. No, because t = -1.21 and the p-value > 0.05.
C. Yes, because t = -6.63 and the p-value < 0.05.
D. No, because t = -1.48 and the p-value > 0.05.

KEY: A

Questions 71 to 74: It is known that for right-handed people, the dominant (right) hand tends to be stronger. For left-handed people who live in a world designed for right-handed people, the same may not be true. To test this, muscle strength was measured on the right and left hands of a random sample of 15 left-handed men and the difference (left - right) was found. The alternative hypothesis is one-sided (left hand stronger). The resulting t-statistic was 1.80.

71. This is an example of
   A. a two-sample t-test.
   B. a paired t-test.
   C. a pooled t-test.
   D. an unpooled t-test.

KEY: B

72. Which of the following is true about the conditions necessary to carry out the t-test in this situation?
   A. Because the sample size is small (15), the population of differences must be assumed to be approximately normal, but no assumption about the variances is required.
   B. Because the sample size is not small (15 + 15 = 30), the population of differences need not be assumed to be approximately normal, and no assumption about the variances is required.
   C. Because the sample size is small (15), the population of differences must be assumed to be approximately normal, and the variances of the right and left hand strengths must be assumed to be equal.
   D. Because the sample size is not small (15 + 15 = 30), the population of differences need not be assumed to be approximately normal, but the variances of the right and left hand strengths must be assumed to be equal.

KEY: A

73. Assuming the conditions are met, based on the t-statistic of 1.80 the appropriate decision for this test using α = 0.05 is:
   A. df = 14, so p-value < 0.05 and the null hypothesis can be rejected.
   B. df = 14, so p-value > 0.05 and the null hypothesis cannot be rejected.
   C. df = 28, so p-value < 0.05 and the null hypothesis can be rejected.
   D. df = 28, so p-value > 0.05 and the null hypothesis cannot be rejected.

KEY: A

74. Which of the following is an appropriate conclusion?
   A. The results are statistically significant so the left hand appears to be stronger.
   B. The results are statistically significant so the left hand does not appear to be stronger.
   C. The results are not statistically significant so there is not enough evidence to conclude the left hand appears to be stronger.
   D. None of the above.
KEY: A

75. Suppose that a difference between two groups is examined. In the language of statistics, the alternative hypothesis is a statement that there is
A. no difference between the sample means.
B. a difference between the sample means.
C. no difference between the population means.
D. a difference between the population means.

KEY: D

76. The maximum distance at which a highway sign can be read is determined for a sample of young people and a sample of older people. The mean distance is computed for each age group. What's the most appropriate null hypothesis about the means of the two groups?
A. The population means are different.
B. The sample means are different.
C. The population means are the same.
D. The sample means are the same.

KEY: C

77. Researchers want to see if men have a higher blood pressure than women do. A study is planned in which the blood pressures of 50 men and 50 women will be measured. What's the most appropriate alternative hypothesis about the means of the men and women?
A. The sample means are the same.
B. The sample mean will be higher for men.
C. The population means are the same.
D. The population mean is higher for men than for women.

KEY: D

78. When comparing two means, the situation most likely to lead to a result that is statistically significant but of little practical importance is
A. when the actual difference is large and the sample sizes are large.
B. when the actual difference is large and the sample sizes are small.
C. when the actual difference is small and the sample sizes are large.
D. when the actual difference is small and the sample sizes are small.

KEY: C

79. When comparing two means, which situation is most likely to lead to a result that is statistically significant? (Consider all other factors equal, such as significance level and standard deviations.)
A. $\bar{x}_1 = 10$, $\bar{x}_2 = 20$ and the sample sizes are $n_1 = 25$ and $n_2 = 25$
B. $\bar{x}_1 = 10$, $\bar{x}_2 = 25$ and the sample sizes are $n_1 = 25$ and $n_2 = 25$
C. $\bar{x}_1 = 10$, $\bar{x}_2 = 20$ and the sample sizes are $n_1 = 15$ and $n_2 = 20$
D. $\bar{x}_1 = 10$, $\bar{x}_2 = 25$ and the sample sizes are $n_1 = 25$ and $n_2 = 35$

KEY: D
Questions 80 and 81: A null hypothesis is that the mean nose lengths of men and women are the same. The alternative hypothesis is that men have a longer mean nose length than women.

80. Which of the following is the correct way to state the null hypothesis?
A. $H_0: \mu = 0.50$
B. $H_0: \bar{x}_1 - \bar{x}_2 = 0$
C. $H_0: \mu_1 - \mu_2 = 0$
D. $H_0: \mu_1 - \mu_2 = 0$

KEY: D

81. A statistical test is performed for assessing if men have a longer mean nose length than women. The $p$-value is 0.225. Which of the following is the most appropriate way to state the conclusion?
A. The mean nose lengths of the populations of men and women are identical.
B. There is not enough evidence to say that the populations of men and women have different mean nose lengths.
C. Men have a greater mean nose length than women.
D. The probability is 0.225 that men and women have the same mean nose length.

KEY: B

Questions 82 to 86: An airport official wants to assess if the flights from one airline (Airline 1) are less delayed than flights from another airline (Airline 2). Let $\mu_1 =$ average delay for Airline 1 and $\mu_2 =$ average delay for Airline 2. A random sample of 10 flights for Airline 1 shows an average of 9.5 minutes delay with a standard deviation of 3 minutes. A random sample of 10 flights for Airline 2 shows an average of 12.63 minutes delay with a standard deviation of 3 minutes. Assume delay times are normally distributed, but do not assume the population variances are equal. Use the conservative “by hand” estimate for the degrees of freedom.

82. What are the appropriate null and alternative hypotheses?
A. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 \neq 0$
B. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 = 0$
C. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 < 0$
D. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 > 0$

KEY: C

83. What is the value of the test statistic?
A. $t = -1.80$
B. $t = -2.00$
C. $t = -2.33$
D. None of the above

KEY: C

84. What is the $p$-value?
A. $p$-value = 0.021
B. $p$-value = 0.022
C. $p$-value = 0.038
D. None of the above

KEY: B
85. For a significance level of $\alpha = 0.05$, are the results statistically significant?
   A. No, results are not statistically significant because the $p$-value < 0.05.
   B. Yes, results are statistically significant because the $p$-value < 0.05.
   C. No, results are not statistically significant because the $p$-value > 0.05
   D. Yes, results are statistically significant because the $p$-value > 0.05.
   KEY: B

86. Which of the following is an appropriate conclusion?
   A. The average delay for Airline 1 does appear to be less than the average delay for Airline 2.
   B. The results are not statistically significant so there is not enough evidence to conclude the average delay for Airline 1 is less than the average delay for Airline 2.
   C. The average delay for Airline 1 is at least 3 minutes less than Airline 2.
   D. None of the above.
   KEY: A

Questions 87 to 91: A bank official wants to assess if the average delays for two airlines are different. Let $\mu_1 = \text{average delay for Airline 1}$ and $\mu_2 = \text{average delay for Airline 2}$. A random sample of 10 flights for Airline 1 shows an average of 6 minutes delay with a standard deviation of 5 minutes. A random sample of 10 flights for Airline 2 shows an average of 12 minutes delay with a standard deviation of 5 minutes. Assume delay times are normally distributed, but do not assume the population variances are equal. Use the conservative “by hand” estimate for the degrees of freedom.

87. What are the appropriate null and alternative hypotheses?
   A. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 \neq 0$
   B. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 = 0$
   C. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 < 0$
   D. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 > 0$
   KEY: A

88. What is the value of the test statistic?
   A. $t = -2.68$
   B. $t = -1.50$
   C. $t = 1.50$
   D. $t = 2.68$
   KEY: A

89. For a significance level of 0.05, what is the critical value?
   A. The critical value = 1.83.
   B. The critical value = 2.10.
   C. The critical value = 2.26.
   D. None of the above.
   KEY: C

90. Are the results statistically significant?
   A. No, results are not statistically significant because $|t| < 1.83$.
   B. Yes, results are statistically significant because $|t| > 2.26$.
   C. No, results are not statistically significant because $|t| > 2.26$.
   D. None of the above.
   KEY: B

254
91. Which of the following is an appropriate conclusion?
   A. The average delay for Airline 1 does appear to be different than the average delay for Airline 2.
   B. The results are not statistically significant so there is not enough evidence to conclude that the average delays are different.
   C. The average delay for Airline 1 is at least 3 minutes less than for Airline 2.
   D. None of the above.

   KEY: A

Questions 92 to 96: A researcher wants to assess if the average age when women first marry has increased from 1960 to 1990. Let $\mu_1 =$ average age of first marriage for women in 1990 and $\mu_2 =$ average age of first marriage for women in 1960. A random sample of 10 women married in 1990 showed an average age at marriage of 24.95 years, with a sample standard deviation of 2 years. A random sample of 20 women married in 1960 showed an average age at marriage of 23.1 years, with a sample standard deviation of 1.5 years. Assume that age of first marriage for women is normally distributed, but do not assume the population variances are equal. Use the conservative “by hand” estimate for the degrees of freedom.

92. What are the appropriate null and alternative hypotheses?
   A. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 \neq 0$
   B. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 = 0$
   C. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 < 0$
   D. $H_0: \mu_1 - \mu_2 = 0$ and $H_a: \mu_1 - \mu_2 > 0$

   KEY: D

93. What is the value of the test statistic?
   A. $t = 1.80$
   B. $t = 2.00$
   C. $t = 2.33$
   D. $t = 2.58$

   KEY: D

94. What is the $p$-value?
   A. $p$-value $= 0.015$
   B. $p$-value $= 0.022$
   C. $p$-value $= 0.038$
   D. None of the above

   KEY: A

95. For a significance level of $\alpha = 0.05$, are the results statistically significant?
   A. No, results are not statistically significant because the $p$-value $< 0.05$.  
   B. Yes, results are statistically significant because the $p$-value $< 0.05$. 
   C. No, results are not statistically significant because the $p$-value $> 0.05$  
   D. Yes, results are statistically significant because the $p$-value $> 0.05$.

   KEY: B

96. Which of the following is an appropriate conclusion?
   A. The average age of first marriage for women in 1990 does appear to be greater than the average age in 1960.
   B. The results are not statistically significant so there is not enough evidence to conclude that the average age in 1990 is greater than the average age in 1960.
   C. The average age of marriage is at least 23 years old for both 1960 and 1990.

   KEY: B

255
97. In each of the following cases, we wish to test the null hypothesis $H_0: \mu = 10$ vs. $H_a: \mu \neq 10$. In which case can this null hypothesis not be rejected at a significance level $\alpha = 0.05$?
A. 90% confidence interval for $\mu$ is (9 to 12).
B. 95% confidence interval for $\mu$ is (11 to 21).
C. 98% confidence interval for $\mu$ is (4.4 to 9.4).
D. 99% confidence interval for $\mu$ is (5 to 9).
KEY: A

98. In each of the following cases, we wish to test the null hypothesis $H_0: \mu = 18$. In which case can this null hypothesis be rejected at a significance level $\alpha = 0.05$?
A. 95% confidence interval for $\mu$ is (16 to 21), $H_a: \mu \neq 18$.
B. 90% confidence interval for $\mu$ is (15 to 20), $H_a: \mu < 18$.
C. 90% confidence interval for $\mu$ is (19 to 26), $H_a: \mu > 18$.
D. 90% confidence interval for $\mu$ is (15 to 22), $H_a: \mu \neq 18$.
KEY: C

99. Each of the following presents a two-sided confidence interval and the alternative hypothesis of a corresponding hypothesis test. In which case can we not use the given confidence interval to make a decision at a significance level $\alpha = 0.05$?
A. 95% confidence interval for $p_1 - p_2$ is (-0.15 to 0.07), $H_a: p_1 - p_2 \neq 0$.
B. 95% confidence interval for $p$ is (.12 to .28), $H_a: p > 0.10$.
C. 90% confidence interval for $\mu$ is (101 to 105), $H_a: \mu \neq 100$.
D. 90% confidence interval for $\mu_1 - \mu_2$ is (3 to 15), $H_a: \mu_1 - \mu_2 > 0$.
KEY: C

100. As Internet usage flourishes, so do questions of security and confidentiality of personal information. A survey of U.S. adults resulted in a 95% confidence interval for the proportion of all U.S. adults who would never give personal information to a company of (0.22, 0.30). Based on this interval, which one of the null hypotheses below (versus a two-sided alternative) can be rejected?
A. $H_0: p = 2/7$
B. $H_0: p = 1/3$
C. $H_0: p = 1/4$
D. $H_0: p = 0.26$
KEY: B
Questions 101 to 106: A small bakery is trying to predict how many loaves of bread to bake daily. They randomly sample daily sales records from the past year. In these days, the bakery sold between 40 and 80 loaves per day with a 95% confidence interval for the population mean daily loaf demand given by (51, 54).

101. What was the mean number of loaves sold daily for the sample of 50 days?
   A. 50  
   B. 60  
   C. 52.5  
   D. Cannot be determined.
   KEY: C

102. How would a 99% confidence interval compare to the 95% confidence interval?
   A. It would be wider.  
   B. It would be narrower.  
   C. It would be the same width.  
   D. Cannot be determined.
   KEY: A

103. Based on this interval, what is the decision for testing \( H_0: \mu = 50 \) versus \( H_a: \mu \neq 50 \) at the 5% significance level?
   A. Reject the null hypothesis.  
   B. Fail to reject the null hypothesis.  
   C. Cannot be determined.
   KEY: A

104. Based on this interval, what is the decision for testing \( H_0: \mu = 50 \) versus \( H_a: \mu \neq 50 \) at the 10% significance level?
   A. Reject the null hypothesis.  
   B. Fail to reject the null hypothesis.  
   C. Cannot be determined.
   KEY: A

105. Based on this interval, what is the decision for testing \( H_0: \mu = 52 \) versus \( H_a: \mu \neq 52 \) at the 5% significance level?
   A. Reject the null hypothesis.  
   B. Fail to reject the null hypothesis.  
   C. Cannot be determined.
   KEY: B

106. Based on this interval, what is the decision for testing \( H_0: \mu = 52 \) versus \( H_a: \mu \neq 52 \) at the 10% significance level?
   A. Reject the null hypothesis.  
   B. Fail to reject the null hypothesis.  
   C. Cannot be determined.
   KEY: C
107. For what problem is a one-sample $t$-test used?
   A. To test a hypothesis about a proportion.
   B. To test a hypothesis about a mean.
   C. To test a hypothesis about the difference between two means for independent samples.
   D. To test a hypothesis about the difference between two proportions for independent samples.
KEY: B

108. For what problem is a two-sample $t$-test used?
   A. To test a hypothesis about a mean.
   B. To test a hypothesis about the mean difference for paired data.
   C. To test a hypothesis about the difference between two means for independent samples.
   D. To test a hypothesis about the difference between two proportions for independent samples.
KEY: C

MORE! MORE! MORE! PROBLEMS

14. A clean standard requires that vehicle exhaust not exceed specified limits for various pollutants. Many states require that cars be tested annually to be sure they meet these standards. Suppose state regulators double check a random sample of cars that a suspect repair shop has certified as okay. They will revoke the shop’s license if they find significant evidence that the shop is certifying vehicle that do not meet standards.

The appropriate hypotheses are:

A. Null hypothesis: The regulators decide that the shop is meeting standards.
   Alternative hypothesis: The regulators decide that the shop is not meeting standards.
B. Null hypothesis: The shop’s emission standards are higher than for other shops.
   Alternative hypothesis: The shop’s emission standards are not higher than for other shops.
C. Null hypothesis: The shop is meeting the emissions standards.
   Alternative hypothesis: The shop is not meeting the emission standards.
D. Null hypothesis: The repair shop’s license is revoked.
   Alternative hypothesis: The repair shop’s license is not revoked.
QUESTIONS 15 - 17
The National center for Education Statistics monitors many aspects of elementary and secondary education nationwide. Their 1996 numbers are often used as a baseline to assess changes. In 1996 31% of students reported that their mothers had graduated from college. In 2000, responses from 8368 students found that this figure had grown to 32%. Is this evidence that there has been an increase in education level among mothers?

15. The appropriate hypotheses are:

A. $H_0$: $p = 0.31; H_A$: $p > 0.31$  
B. $H_0$: $p > 0.31; H_A$: $p < 0.31$

C. $H_0$: $p = 0.32; H_A$: $p \neq 0.32$  
D. $H_0$: $p < 0.31; H_A$: $p > 0.32$

16. Let all necessary conditions be satisfied. Find the test statistic ($z$), and the P-value.

A. $z = 1.978$, P-value = 0.0047  
B. $z = 1.978$, P-value = 0.024

C. $z = 0.0051$, P-value = -1.978  
D. $z = 0.048$, P-value = 1.978

17. State your conclusion.

A. With a P-value of 1.978, we fail to reject the null hypothesis.
B. No conclusion can be drawn because the P-value is too big to be considered in this context.

C. With a P-value of 0.024, we reject the null hypothesis. There is evidence to suggest that the percentage of students whose mothers are college graduates has changed since 1996. In fact, the evidence suggests that the percentage has increased.
D. The confidence level was not given, so no conclusion can be arrived at.

18. National data in the 1960s showed that about 44% of the adult population had never smoked cigarettes. In 1995 a national health survey interviewed a random sample of 881 adults and found that 52% had never been smokers.

(a) Create a 95% confidence interval for the proportion of adults (in 1995) who had never been smokers.

(b) Does this provide evidence of a change in behavior among Americans? Using your confidence interval, test an appropriate hypothesis and state your conclusion.
ANSWERS
(a) (0.487, 0.553) (b) Null Hypo: p = 0.44; Alter Hypo: p not equal to 0.44. Since 44% is not in the 95% CI, we will reject the Null Hypo. There is strong evidence that, in 1995, the percentage of adults who have never smoked was not 44%.

19. A company hopes to improve customer satisfaction, setting as a goal no more than 5% negative comments. A random survey of 350 customers found only 10 with complaints.
(a) Create a 95% confidence interval for the true level of dissatisfaction among customers.
(b) Does this provide evidence that the company has reached its goal? Using your confidence interval, test an appropriate hypothesis and state your conclusion.

ANSWERS
(a) (0.011, 0.046) (b) Null Hypo: p = 0.05; Alter. Hypo: p < 0.05. Since 5% is not in the 95% CI, we will reject the Null Hypo. There is strong evidence that less that 5% of customers have complaints.

20. A company with a fleet of 150 cars found that the emissions systems of 7 out of 22 they tested failed to meet pollution control guidelines. Is this strong evidence that more than 20% of the fleet might be out of compliance? Test an appropriate hypothesis and state your conclusion. Be sure the appropriate assumptions and conditions are satisfied before you proceed.

ANSWER
Null Hypo: p = 0.20; Alter. Hypo: p > 0.20. Two conditions are not satisfied (verify!!!!!!!) We cannot proceed with a hypothesis test.

21. In a rural area only about 30% of the wells that are drilled find adequate water at a depth of 100 feet or less. A local man claims to be able to find water by “dowsing” - using a forked stick to indicate where the well should be drilled. You check with 80 of his customers and find that 27 have wells less than 100 feet deep. What do you conclude about this claim? (We consider a P – value of around 5% to represent strong evidence)

(a) Write appropriate hypotheses.
(b) Check the necessary assumptions.
(c) Perform the mechanics of the test. What is the P – value?
(d) Explain carefully what the P – value means in this context.
(e) What’s your conclusion?
FOR SOLUTION TO PROBLEM 21, SEE CLASS NOTES

22. In the 1980s it was generally believed that congenital abnormalities affected about 5% of the nation's children. Some people believe that the increase in the number of chemicals in the environment has led to an increase in the incidence of abnormalities. A recent study examined 384 children and found that 46 of them showed signs of an abnormality. Is this strong evidence that the risk has increased? (We consider a P-value of around 5% to represent strong evidence)

(a) Write appropriate hypotheses.
(b) Check the necessary assumptions.
(c) Perform the mechanics of the test. What is the P-value?
(d) Explain carefully what the P-value means in this context.
(e) What's your conclusion?
(f) Do environmental chemicals cause congenital abnormalities?

FOR SOLUTION TO PROBLEM 22, SEE CLASS NOTES

QUESTIONS 23 - 25

The National Center for Education Statistics monitors many aspects of elementary and secondary education nationwide. Their 1996 numbers are often used as a baseline to assess changes. In 1996, 34% of students had not been absent from school even once during the previous month. In the 2000 survey, responses from 8302 students showed that this figure had slipped to 33%. Officials would of course be concerned if student attendance were declining. Do these figures give evidence of a decrease in student attendance?

23. Write appropriate hypotheses.

A. \( H_0 : p = 0.33 \); \( H_A : p < 0.33 \)
B. \( H_0 : p > 0.34 \); \( H_A : p < 0.34 \)
C. \( H_0 : p = 0.34 \); \( H_A : p < 0.34 \)
D. \( H_0 : p < 0.33 \); \( H_A : p = 0.34 \)
E. \( H_0 : p = 0.34 \); \( H_A : p = 0.33 \)
24. Perform the test by finding the Test statistics and the P-value.

A. \( Z_o = 0.0052; P \text{- value} = 0.027 \)
B. \( Z_o = -1.923; P \text{- value} = 0.027 \)
C. \( Z_o = 0.048; P \text{- value} = 0.0052 \)
D. \( Z_o = 0.01; P \text{- value} = 0.048 \)
E. \( Z_o = 1.923; P \text{- value} = 0.027 \)

25. State your conclusion.

A. With a P-value of 0.027, we fail to reject the null hypothesis. There is no evidence to suggest that the percentage of students with perfect attendance in the previous month has decreased in 2000.
B. With a P-value of 0.0052, we reject the null hypothesis. There is evidence to suggest that the percentage of students with perfect attendance in the previous month has decreased in 2000.
C. With a P-value of 0.048, essentially 0.05, equal to the significance level, we retain the null hypothesis. There is no evidence to suggest that the percentage of students with perfect attendance in the previous month has decreased in 2000.
D. With a P-value of 0.027, we reject the null hypothesis. There is evidence to suggest that the percentage of students with perfect attendance in the previous month has decreased in 2000.
E. With a P-value of 0.027, we retain the null hypothesis. There is evidence to suggest that the percentage of students with perfect attendance in the previous month has decreased in 2000.

26. A governor is concerned about his “negatives” – the percentage of state residents who express disapproval of his job performance. His political committee pays for a series of TV ads, hoping that they can keep the negatives below 30%. They will use follow-up polling to assess the ad’s effectiveness. After the political campaign, the pollsters check the governor’s negatives. They test the hypothesis that the ads produced no change against the alternative that the negatives are now below 30%, and find a P-value of 0.22. What conclusion is appropriate?

A. There is a 22% chance that the ads worked.
B. There is a 78% chance that the ads worked.
C. There is a 22% chance that the poll they conducted is correct.
D. There is a 22% chance that natural sampling variation could produce poll results like these if there is really no change in public opinion.
E. None of the above is appropriate.
27. The seller of a loaded die claims that it will favor the outcome 6. We don’t believe that claim, and roll the die 200 times to test an appropriate hypothesis. Our P-value turns out to be 0.03. Which conclusion is appropriate?

A. There’s a 3% chance that the die is fair.
B. There’s a 97% chance that the die is fair.
C. There’s a 3% chance that a loaded die could randomly produce the results we observed, so it’s reasonable to conclude that the die is fair.
D. There’s a 3% chance that a fair die could randomly produce the results we observed, so it’s reasonable to conclude that the die is loaded.
E. None of the above.

QUESTIONS 28 - 31

In 1960, census results indicated that the age at which American men first married had a mean of 23.3 years. It is widely suspected that young people today are waiting longer to get married. We want to find out if the mean age of first marriage has increased during the past 40 years. We plan to test our hypothesis by selecting a random sample of 40 men who married for the first time last year. Do you think the necessary assumptions for inference are satisfied? Explain.

28. Write appropriate hypotheses.

A. \( H_0 : \mu = 23.3 \); \( H_A : \mu < 23.3 \)
B. \( H_0 : \mu > 23.3 \); \( H_A : \mu = 23.3 \)
C. \( H_0 : \mu = 23.3 \); \( H_A : \mu < 23.3 \)
D. \( H_0 : \mu = 23.3 \); \( H_A : \mu > 23.3 \)
E. None of the above

29. Describe the approximate sampling distribution model for the mean age in such samples.

A. \( N \left( 23.3, \frac{\sigma}{40} \right) \)
B. \( t_{39} \left( 23.3, \frac{\sigma}{\sqrt{40}} \right) \)
C. \( N \left( 23.3, \frac{s}{\sqrt{40}} \right) \)
D. \( t_{39} \left( 23.3, \frac{s}{\sqrt{40}} \right) \)
E. None of the above
30. The men in our sample married at an average age of 24.2 years, with a standard deviation of 5.3 years. What is the P-value for this result? Explain in context what this P-value means.

A. P-value = 0.8544; If the mean age at first marriage is still 23.3 years, there is a 0.8544 chance of getting a sample mean of 24.2 years or older simply from natural sampling variation.
B. P-value = 1.07; If the mean age at first marriage is still 23.3 years, then, the P-value is what we observe from data due to natural sampling variation.
C. P-value = 0.1456; If the mean age at first marriage is still 23.3 years, there is a 0.1456 chance of getting a sample mean of 24.2 years or older simply from natural sampling variation.
D. P-value = 0.1456; If the mean age at first marriage is still 23.3 years, there is a 0.1456 chance of getting a sample mean of 24.2 years or younger simply from natural sampling variation.
E. P-value = 1.07; If the mean age at first marriage is still 23.3 years, then, the P-value is what we observe in a sample of size n = 40, with mean 24.2 years or older.

31. What is your conclusion?

A. We fail to reject the null hypothesis. There is no evidence to suggest that the mean age at first marriage has changed from 23.3 years, the mean in 1960, to 24.2 years.
B. We fail to reject the null hypothesis. There is strong evidence to suggest that the mean age at first marriage has changed from 23.3 years, the mean in 1960, to 24.2 years.
C. We reject the null hypothesis. There is strong evidence to suggest that the mean age at first marriage has changed from 23.3 years, the mean in 1960, to 24.2 years.
D. We reject the null hypothesis. There is no strong evidence to suggest that the mean age at first marriage has changed from 23.3 years, the mean in 1960, to 24.2 years.
E. We cannot draw any conclusion in this case because the P-value = 1.07 is exceedingly big.

32. Hungry? Researchers investigated how the size of a bowl affects how much ice cream people tend to scoop when serving themselves. At an "ice cream social," people were randomly given either a 17 oz or a 34 oz bowl (both large enough that they would not be filled to capacity). They were then invited to scoop as much ice cream as they liked. Did the bowl size change the selected portion size? Here are the summaries:
<table>
<thead>
<tr>
<th>Small Bowl</th>
<th>Large Bowl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Test an appropriate hypothesis and state your conclusions. For assumptions and conditions that you cannot test, you may assume that they are sufficiently satisfied to proceed.

**ANSWER:** Reject $H_0$. The mean amount of ice cream people put into a bowl is related to the size of the bowl. $t_0 = 2.104$; $P$-value $= 0.04087$

33. **Thirsty?** Researchers randomly assigned participants either a tall, thin “highball” glass or a short, wide “tumbler,” each of which held 355 ml. Participants were asked to pour a shot (1.5 oz = 44.3 ml) into their glass. Did the shape make a difference in how much liquid they poured? Here are the summaries:

<table>
<thead>
<tr>
<th>highball</th>
<th>tumbler</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Test an appropriate hypothesis and state your conclusion. For assumptions and conditions that you cannot test, you may assume that they are sufficiently satisfied to proceed.

**ANSWER:** Reject the null hypothesis. The mean amount of liquid people pour into a glass is related to the shape of the glass. People tend to pour, on average, into a small, wide tumbler than into a tall, narrow highball glass. $t_0 = 7.71$; $P$-value $= 0$
PROBLEMS ON TYPE I AND TYPE II ERRORS