1) A company has two manufacturing plants, and company officials want to determine whether there is a difference in the average age of workers at the two locations. The ages of five randomly selected workers at each plant were recorded. For the confidence interval procedure to work properly, what additional assumptions must be satisfied?

A) Both samples were randomly and independently selected from their respective populations.
B) The population variances are equal.
C) Both populations have approximate normal distributions.
D) All of the above are necessary.

2) When testing for a difference between the means of a treatment group and a placebo group, the computer display below is obtained. Using a 0.02 significance level, is there sufficient evidence to support the claim that the treatment group (variable 1) comes from a population with a mean that is larger than the mean for the placebo population? Explain.

<table>
<thead>
<tr>
<th>t-Test: Two Sample for Means</th>
<th>Variable 1</th>
<th>Variable 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mean</td>
<td>171.6392</td>
<td>168.7718</td>
</tr>
<tr>
<td>3 Known Variance</td>
<td>47.51672</td>
<td>41.08293</td>
</tr>
<tr>
<td>4 Observations</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>5 Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6 t</td>
<td>2.154057</td>
<td></td>
</tr>
<tr>
<td>7 P(t=1) one-tail</td>
<td>0.0158</td>
<td></td>
</tr>
<tr>
<td>8 T Critical one-tail</td>
<td>1.644833</td>
<td></td>
</tr>
<tr>
<td>9 P(t=1) two-tail</td>
<td>0.0316</td>
<td></td>
</tr>
<tr>
<td>10 T Critical two-tail</td>
<td>1.959961</td>
<td></td>
</tr>
</tbody>
</table>

A) There is not enough information to decide.
B) No, the P-value for a one-tail test is 0.0158, which is greater than the significance level of 0.02. There is not sufficient evidence to support the claim that the two population means are different.
C) Yes, the P-value for a two-tail test is 0.0316, which is greater than the significance level of 0.02. There is sufficient evidence to support the claim that the two population means are different.
D) No, the P-value for a two-tail test is 0.0316, which is greater than the significance level of 0.02. There is not sufficient evidence to support the claim that the two population means are different.
E) Yes, the P-value for a one-tail test is 0.0158, which is less than the significance level of 0.02. There is sufficient evidence to support the claim to support the claim that the treatment group has a mean that is larger than the mean for the placebo population.

3) In a random sample of 500 people in their 30's, 18% were smokers. In a random sample of 450 people in their 40's, 15% were smokers. The 95% confidence interval for the difference between the population proportions $P_{30's} - P_{40's}$ turned out to be:

\[-0.0183 < p_1 - p_2 < 0.0761.\]

What does this imply?

A) There is not a significant difference between the smoking habits of people in their 30's and people in their 40's, but a higher percentage of 30-somethings sampled smoked.
B) A slightly higher proportion of people in their 40's smoke than people in their 30's.
C) 1.83% of people in their 30's smoke while 7.61% of people in their 40's smoke.
D) There is a significant difference between the smoking habits of people in their 30's and people in their 40's, a higher percentage of 30-somethings smoke.
E) A higher proportion of people in their 40's smoke than people in their 30's.
4) What kind of statistical procedure should be used to analyze data from the following study:

We are interested in determining whether men's and women's attitudes on the environment differ. To do so we sampled 100 men and 100 women and asked: "Do you think the environment is a major concern?" Of those sampled, 67 women and 53 men responded that they believed it is.

A) Small sample independent t-test using a pooled variance
B) Matched pairs dependent t-test
C) Large sample independent sample z-test
D) Z-test to compare two different proportions

Solve the problem.

5) A researcher is investigating which of two newly developed automobile engine oils is better at prolonging the life of an engine. Since there are a variety of automobile engines, 20 different engine types were randomly selected and were tested using each of the two engine oils. The number of hours of continuous use before engine breakdown was recorded for each engine oil. Based on the information provided, what type of analysis will yield the most useful information?

A) Matched pairs comparison of population means.
B) Independent samples comparison of population proportions.
C) Independent samples comparison of population means.
D) Matched pairs comparison of population proportions.

6) A marketing study was conducted to compare the mean age of male and female purchasers of a certain product. Random and independent samples were selected for both male and female purchasers of the product. What type of analysis should be used to compare the mean age of male and female purchasers?

A) An independent samples comparison of population proportions.
B) A paired difference comparison of population means.
C) An independent samples comparison of population means.
D) A test of a single population mean.

7) When blood levels are low at an area hospital, a call goes out to local residents to give blood. The blood center is interested in determining which sex—males or females—is more likely to respond. Random, independent samples of 60 females and 100 males were each asked if they would be willing to give blood when called by a local hospital. A success is defined as a person who responds to the call and donates blood. The goal is to compare the percentage of the successes between the male and female responses. What type of analysis should be used?

A) An independent samples comparison of population means.
B) A paired difference comparison of population means.
C) An independent samples comparison of population proportions.
D) A test of a single population proportion.
8) We are interested in comparing the average supermarket prices of two leading colas. Our sample was taken by randomly selecting eight supermarkets and recording the price of a six-pack of each brand of cola at each supermarket. The data are shown in the following table:

<table>
<thead>
<tr>
<th>Supermarket</th>
<th>Brand 1</th>
<th>Brand 2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2.25</td>
<td>$2.30</td>
<td>-$0.05</td>
</tr>
<tr>
<td>2</td>
<td>2.47</td>
<td>2.45</td>
<td>0.02</td>
</tr>
<tr>
<td>3</td>
<td>2.38</td>
<td>2.44</td>
<td>-0.06</td>
</tr>
<tr>
<td>4</td>
<td>2.27</td>
<td>2.29</td>
<td>-0.02</td>
</tr>
<tr>
<td>5</td>
<td>2.15</td>
<td>2.25</td>
<td>-0.10</td>
</tr>
<tr>
<td>6</td>
<td>2.25</td>
<td>2.25</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>2.36</td>
<td>2.42</td>
<td>-0.06</td>
</tr>
<tr>
<td>8</td>
<td>2.37</td>
<td>2.40</td>
<td>-0.03</td>
</tr>
</tbody>
</table>

\[ x_1 = 2.3125 \quad x_2 = 2.3500 \quad d = -0.0375 \]
\[ s_1 = 0.1007 \quad s_2 = 0.0859 \quad s_d = 0.0381 \]

What assumptions are needed for a confidence interval for the mean difference to be valid?

A) The population of paired differences has an approximately normal distribution.
B) The population variances are equal.
C) The samples were independently selected from each population.
D) All of the above are needed.

9) We sampled 100 men and 100 women and asked: “Do you think the environment is a major concern?” Of those sampled, 67 women and 53 men responded that they believed it is. For the confidence interval procedure to work properly, what additional assumptions must be satisfied?

A) The population variances are equal.
B) Both samples were randomly and independently selected from their respective populations.
C) Both populations have approximate normal distributions.
D) All of the above are necessary.

10) A certain manufacturer is interested in evaluating two alternative manufacturing plans consisting of different machine layouts. Because of union rules, hours of operation vary greatly for this particular manufacturer from one day to the next. Twenty-eight random working days were selected and each plan was monitored and the number of items produced each day was recorded. Some of the collected data is shown below:

<table>
<thead>
<tr>
<th>DAY</th>
<th>PLAN 1 OUTPUT</th>
<th>PLAN 2 OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1234 units</td>
<td>1311 units</td>
</tr>
<tr>
<td>2</td>
<td>1355 units</td>
<td>1366 units</td>
</tr>
<tr>
<td>3</td>
<td>1300 units</td>
<td>1289 units</td>
</tr>
</tbody>
</table>

What assumptions are necessary for the above test to be valid?

A) Both populations must be approximately normally distributed.
B) The population of paired differences must be approximately normally distributed.
C) None of these listed, since the Central Limit Theorem can be applied.
D) The population variances must be approximately equal.
11) We sampled 100 men and 100 women and asked: “Do you think the environment is a major concern?” Of those sampled, 67 women and 53 men responded that they believed it is. For the confidence interval procedure to work properly, what additional assumptions must be satisfied?

A) The population variances are equal.
B) Both samples were randomly and independently selected from their respective populations.
C) Both populations have approximate normal distributions.
D) All of the above are necessary.

12) University administrators are trying to decide where to build a new parking garage on campus. The administrators have determined that the parking garage will be built either by the college of engineering or by the college of business. To help make the final decision, the university has randomly and independently asked students from each of the two colleges to estimate how long they usually take to find a parking spot on campus (in minutes). Based on their sample, the following 95% confidence interval (for \( \mu_e - \mu_b \)) was created: \((4.20, 10.20)\). What conclusion can the university make about the population mean parking times based on this confidence interval?

A) They are 95% confident that the mean parking time of all business students is less than the mean parking time of all engineering students.
B) They are 95% confident that the mean parking time of all business students exceeds the mean parking time of all engineering students.
C) They are 95% confident that the mean parking time of all business students equals the mean parking time of all engineering students.

13) A researcher wants to determine if gender differences exist in the area of domestic violence. The researcher surveyed 1000 men and 1000 women and asked each if they felt that domestic violence was a major concern in today’s society. Two hundred fifty of the men and 400 of the women responded by saying they did feel domestic violence was a major concern. What criterion is used to assess whether the Central Limit Theorem can be applied to this problem?

A) The samples are independently selected.
B) Both sample sizes are at least 30.
C) \( np > 15 \) and \( nq > 15 \)
D) The population proportions are equal.

14) A researcher is investigating which of two newly developed automobile engine oils is better at prolonging the life of an engine. Since there are a variety of automobile engines, 20 different engine types were randomly selected and were tested using each of the two engine oils. The number of hours of continuous use before engine breakdown was recorded for each engine oil. Based on the information provided, what type of analysis will yield the most useful information?

A) Matched pairs comparison of population means.
B) Independent samples comparison of population proportions.
C) Independent samples comparison of population means.
D) Matched pairs comparison of population proportions.

15) A researcher is investigating which of two newly developed automobile engine oils is better at prolonging the life of an engine. Since there are a variety of automobile engines, 20 different engine types were randomly selected and were tested using each of the two engine oils. The number of hours of continuous use before engine breakdown was recorded for each engine oil. Suppose the following 95% confidence interval for \( \mu_A - \mu_B \) was calculated: \((100, 2500)\). Which of the following inferences is correct?

A) We are 95% confident that no significant differences exists in the mean number of hours of continuous use before breakdown of engines using oils A and B.
B) We are 95% confident that an engine using oil B has a higher mean number of hours of continuous use before breakdown than does an engine using oil A.
C) We are 95% confident that the mean number of hours of continuous use of an engine using oil A is between 100 and 2500 hours.
D) We are 95% confident that an engine using oil A has a higher mean number of hours of continuous use before breakdown than does an engine using oil B.
16) Which supermarket has the lowest prices in town? All claim to be cheaper, but an independent agency recently was asked to investigate this question. The agency randomly selected 100 items common to each of two supermarkets (labeled A and B) and recorded the prices charged by each supermarket. The summary results are provided below:

\[ \bar{x}_A = 2.09 \quad \bar{x}_B = 1.99 \quad \bar{d} = .10 \]
\[ s_A = 0.22 \quad s_B = 0.19 \quad s_d = .03 \]

Which of the following assumptions is necessary for a confidence interval for the mean difference to be valid?

A) The population of paired differences has an approximate normal distribution.
B) None of these assumptions are necessary.
C) The population variances must be equal.
D) The samples are randomly and independently selected.

17) A marketing study was conducted to compare the mean age of male and female purchasers of a certain product. Random and independent samples were selected for both male and female purchasers of the product. It was desired to test to determine if the mean age of all female purchasers exceeds the mean age of all male purchasers. The sample data is shown here:

Female: \( n = 10, \) sample mean = 50.30, sample standard deviation = 13.215
Male: \( n = 10, \) sample mean = 39.80, sample standard deviation = 10.040

Which of the following assumptions must be true for the pooled test of hypothesis to be valid?

I. Both the male and female populations of ages must possess approximately normal probability distributions.
II. Both the male and female populations of ages must possess population variances that are equal.
III. Both samples of ages must have been randomly and independently selected from their respective populations.

A) III only \hspace{1cm} B) II only \hspace{1cm} C) I only \hspace{1cm} D) I, II, and III

18) The newspaper recently ran an article indicating differences in perception of sexual harassment on the job between men and women. The article claimed that women perceived the problem to be much more prevalent than did the men. One question asked of both men and women was "Do you think sexual harassment is a major problem in the American workplace?" 24% of the men and 62% of the women responded "Yes." The newspaper created a 99.2% confidence interval for the true difference in proportions and reported it to be -.28 to -.48. What can be said about the proportions with 99.2% reliability?

A) The true proportion of all men exceeds the true proportion of all women who believe that sexual harassment is a major problem in the American workplace.
B) Nothing can be said as negative proportions are meaningless.
C) The true proportion of all women equals the true proportion of all men who believe that sexual harassment is a major problem in the American workplace.
D) The true proportion of all women exceeds the true proportion of all men who believe that sexual harassment is a major problem in the American workplace.

19) The owners of an industrial plant want to determine which of two types of fuel (gas or electricity) will produce more useful energy at a lower cost. In an analysis of the difference of means of the two samples, the owners were not able to reject \( H_0 \) in the test \( H_0: (\mu_E - \mu_G) = 0 \) vs. \( H_{\alpha}: (\mu_E - \mu_G) > 0 \). What is our best interpretation of the result?

A) The mean cost for electricity is not different from the mean cost for gas.
B) The mean cost for electricity is less than the mean cost for gas.
C) The mean cost for electricity is greater than the mean cost for gas.
D) The mean cost for electricity is different from the mean cost for gas.
20) What kind of statistical procedure should be used to analyze data from the following study:

The accuracy of verbal responses is tested in an experiment in which subjects report their weight and they are then weighed on a physician's scale. The data consists of the reported weight and the measured weight for each subject.

A) Z-test to compare two different proportions
B) Matched pairs dependent t-test
C) Large sample independent sample z-test
D) Small sample independent t-test using a pooled variance

21) A researcher wishes to determine whether men perform worse in Statistics than women. Use the computer output below to test the claim at a 10% significance level that women perform better than men in Statistics (Note: we made the comparison: \( \mu_m - \mu_w \)).

<table>
<thead>
<tr>
<th>test stat</th>
<th>df</th>
<th>p-value</th>
<th>mean</th>
<th>st. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.61</td>
<td>44</td>
<td>0.273</td>
<td>-3.61</td>
<td>5.92</td>
</tr>
</tbody>
</table>

A) Do not reject the null; The test does not show a significant difference between men and women.
B) Reject the null; The test does not show a significant difference between men and women.
C) Reject the null; The test shows that women perform better than men in Statistics.
D) Do not reject the null; The test shows that men perform better than women in Statistics.
E) Reject the null; The test shows that men perform worse than women in Statistics.

22) In the context of a hypothesis test for two proportions, which of the following statements about the pooled sample proportion, \( \bar{p} \), is/are true?
I. It estimates the common value of \( p_1 \) and \( p_2 \) under the assumption of equal proportions.
II. It is obtained by averaging the two sample proportions \( \hat{p}_1 \) and \( \hat{p}_2 \).
III. It is equal to the proportion of successes in both samples combined.

A) I and II  B) I and III  C) III only  D) I, II, and III

Provide an appropriate response.

23) In a random sample of 500 people aged 20–24, 22% were smokers. In a random sample of 450 people aged 25–29, 14% were smokers. A 95% confidence interval for the difference between the proportion of 20–24 year olds and the proportion of 25–29 year olds who are smokers is \( 0.032 < p_1 - p_2 < 0.128 \).

Which of the following statements give a correct interpretation of this confidence interval?
I. We can be 95% confident that the interval 0.032 to 0.128 contains the true difference between the two population proportions.
II. There is a 95% chance that the true difference between the two population proportions lies between 0.032 and 0.128.
III. If the process were repeated many times, each time selecting random samples of 500 people aged 20–24 and 450 people aged 25–29 and each time constructing a confidence interval for \( p_1 - p_2 \), 95% of the time the true difference between the two population proportions will lie between 0.032 and 0.128.
IV. If the process were repeated many times, each time selecting random samples of 500 people aged 20–24 and 450 people aged 25–29 and each time constructing a confidence interval for \( p_1 - p_2 \), 95% of the time the confidence interval limits will contain the true difference between the two population proportions.

A) I and III  B) I and IV  C) II and IV  D) II and III
24) Ten different obese patients are tested for the number of calories they eat per day before and after learning about portion control. A 95% confidence interval for the mean of the differences (before – after) yielded the following result:

\[ 212 < \mu_d < 325. \]

Interpret the results.

A) Before the patients consumed 212 calories per day, after 325 calories.
B) There is a small reduction in their calorie intake, but it is not significant.
C) It is estimated that, patients will typically consume between 212 and 325 calories more per day after they learn about portion control.
D) Patients will typically consume between 212 and 325 calories less per day after the course.
E) There is no difference between the patient’s calorie intake before and after the course.

25) What kind of statistical procedure should be used to analyze data from the following study: Ten athletes perform a maximum squat test before being given growth hormone. They are then tested a week later to determine if they have improved their maximum squat.

A) Matched pairs dependent t-test 
B) Large sample independent sample z-test 
C) Small sample independent t-test using a pooled variance 
D) The F-test to compare population variances. 
E) Z-test to compare two different proportions

26) What kind of statistical procedure should be used to analyze data from the following study: An extermination firm is testing two brands of pesticide spray, for the effectiveness against ants. Under controlled conditions, each spray is used on 10 randomly selected ants. The measurements are the number of seconds until all 10 ants are dead.

A) Matched pairs dependent t-test 
B) Small sample independent t-test using a pooled variance 
C) Z-test to compare two different proportions 
D) Large sample independent sample z-test

27) What kind of statistical procedure should be used to analyze data from the following study: A survey is conducted of teens from inner city schools and suburban schools to compare the typical volume of soda consumed by the two groups over the past month. The summary data is given as: Inner city, \( n = 28, \mu = 360 \text{ oz}, s = 48 \text{ oz} \) and Suburban, \( n = 29, \mu = 336 \text{ oz}, s = 45 \text{ oz} \) (assume equal variances)

A) Z-test to compare two different proportions 
B) Matched pairs dependent t-test 
C) The F-test to compare population variances. 
D) Small sample independent t-test using a pooled variance 
E) Large sample independent sample z-test
28) A researcher was interested in comparing the amount of time spent watching television by women and by men. Independent simple random samples of 14 women and 17 men were selected, and each person was asked how many hours he or she had watched television during the previous week. The summary statistics are as follows.

<table>
<thead>
<tr>
<th></th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>11.7</td>
<td>14.2</td>
</tr>
<tr>
<td>$s_1$</td>
<td>3.9</td>
<td>5.2</td>
</tr>
<tr>
<td>$n_1$</td>
<td>14</td>
<td>17</td>
</tr>
</tbody>
</table>

The following 99% confidence interval was obtained for $\mu_W - \mu_m$, the difference between the mean amount of time spent watching television for women and for men:

$$-7.43 \text{ hrs} < \mu_W - \mu_m < 2.43 \text{ hrs}.$$ 

What does the confidence interval suggest about the population means?

A) The confidence interval limits include 0 which suggests that the two population means are unlikely to be equal. There appears to be a significant difference between the mean amount of time spent watching television for women and for men.

B) The confidence interval limits include 0 which suggests that the two population means might be equal. There does not appear to be a significant difference between the mean amount of time spent watching television for women and for men.

C) The confidence interval concludes that women watch 7.43 hours and men watch 2.43 hrs of television, so women watch more TV than men.

D) The confidence interval strongly suggests that the mean amount of time spent watching television for women is smaller than for men.

E) The confidence interval strongly suggests that the mean amount of time spent watching television for women is larger than for men.

29) Consider the following set of salary data:

<table>
<thead>
<tr>
<th></th>
<th>Men (1)</th>
<th>Women (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>100</td>
<td>80</td>
</tr>
<tr>
<td>Mean</td>
<td>$12,850</td>
<td>$13,000</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>$345</td>
<td>$500</td>
</tr>
</tbody>
</table>

What assumptions are necessary to perform a test for the difference in population means?

A) The two samples were independently selected from the populations of men and women.
B) The population variances of salaries for men and women are equal.
C) Both of the target populations have approximately normal distributions.
D) All of the above are necessary.
30) The table below shows the weights, in pounds, of seven subjects before and after following a particular diet for two
months.

<table>
<thead>
<tr>
<th>Subject</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>178</td>
<td>160</td>
<td>192</td>
<td>167</td>
<td>196</td>
<td>176</td>
<td>150</td>
</tr>
<tr>
<td>After</td>
<td>171</td>
<td>151</td>
<td>190</td>
<td>172</td>
<td>182</td>
<td>178</td>
<td>138</td>
</tr>
</tbody>
</table>

\[ \bar{X}_d = 5.286, \ \bar{X}_b = 174.1429, \ \bar{X}_a = 168.8571, \ S^2_p = 303.1429, \ S_d = 7.1581 \]

A hypothesis test at the 1% significance level is conducted to determine if the diet is effective. The critical value is 3.143. Do the data provide sufficient evidence to conclude that the diet is effective in reducing weight?

A) The test stat is \( t = 0.57 \). At the 1% significance level, the data do not provide sufficient evidence to conclude that the diet is effective in reducing weight.
B) There results are inconclusive.
C) The test stat is \( t = 1.954 \). At the 1% significance level, the data does provide sufficient evidence to conclude that the diet is effective in reducing weight.
D) The test stat is \( t = 0.57 \). At the 1% significance level, the data does provide sufficient evidence to conclude that the diet is effective in reducing weight.
E) The test stat is \( t = 1.954 \). At the 1% significance level, the data do not provide sufficient evidence to conclude that the diet is effective in reducing weight.

31) A researcher wishes to determine whether Toyota Corollas use less gas than Honda Civics after 10,000 miles of use. Use the computer output below to test the claim at a 3% significance level that Corollas use less gas on average than Civics (Note: we made the comparison: \( \mu_{\text{corolla}} - \mu_{\text{civic}} \)).

<table>
<thead>
<tr>
<th>test stat</th>
<th>df</th>
<th>p-value</th>
<th>mean dif</th>
<th>st. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.23</td>
<td>44</td>
<td>0.013</td>
<td>-4.28</td>
<td>1.92</td>
</tr>
</tbody>
</table>

A) Reject the null; The test shows the Corolla consumes less gas.
B) Reject the null; The test does not show a significant difference between the two cars.
C) Do not reject the null; The test does not show a significant difference between the two cars.
D) Reject the null; The test shows the Civic consumes less gas.
E) Do not reject the null; The test shows he Civic use more gas than the Corolla.
Answer Key
Testname:

1) D  
2) E  
3) A  
4) D  
5) A  
6) C  
7) C  
8) A  
9) B  
10) B  
11) B  
12) A  
13) C  
14) A  
15) D  
16) B  
17) D  
18) D  
19) A  
20) B  
21) A  
22) B  
23) B  
24) D  
25) A  
26) B  
27) D  
28) B  
29) A  
30) E  
31) A