MULTIPLE CHOICE. Provide an appropriate response.

1) The data given were collected to determine whether the life span of randomly selected mammals, in years, can be predicted from their corresponding gestation periods.

<table>
<thead>
<tr>
<th>Gestation</th>
<th>8</th>
<th>2.1</th>
<th>1.3</th>
<th>1</th>
<th>11.5</th>
<th>5.3</th>
<th>3.8</th>
<th>24.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life span</td>
<td>30</td>
<td>13</td>
<td>8</td>
<td>4</td>
<td>28</td>
<td>11</td>
<td>12</td>
<td>42</td>
</tr>
</tbody>
</table>

Identify the explanatory variable
A) mammal breed
B) months
C) life span
D) climate
E) gestation

2) The data given were collected to determine whether the life span of randomly selected mammals, in years, can be predicted from their corresponding gestation periods.

<table>
<thead>
<tr>
<th>Gestation</th>
<th>8</th>
<th>2.1</th>
<th>1.3</th>
<th>1</th>
<th>11.5</th>
<th>5.3</th>
<th>3.8</th>
<th>24.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life span</td>
<td>30</td>
<td>13</td>
<td>8</td>
<td>4</td>
<td>28</td>
<td>11</td>
<td>12</td>
<td>42</td>
</tr>
</tbody>
</table>

Identify the response variable
A) life span
B) climate
C) length of coat
D) gestation
E) breed of mammal

Determine whether the samples are independent or dependent.

3) The effectiveness of a new headache medicine is tested by measuring the amount of time before the headache is cured for patients who use the medicine and another group of patients who use a placebo drug.

A) Dependent samples
B) Independent samples
C) Cannot be determined from the information given

4) The effectiveness of a headache medicine is tested by measuring the intensity of a headache in patients before and after drug treatment. The data consist of before and after intensities for each patient.

A) Independent samples
B) Dependent samples
C) Cannot be determined from the information given

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Fill in the blank.

5) Dependent samples result when the data are ____________________ pairs.
MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

From the sample statistics, find the value of \( \hat{p}_1 - \hat{p}_2 \), the point estimate of the difference of proportions.

6) \( n_1 = 216 \quad n_2 = 186 \)  
   \( x_1 = 76 \quad x_2 = 99 \)
   A) none of these  
   B) 0.57  
   C) -0.57  
   D) -0.180  
   E) 0.180

Select the most appropriate answer.

7) The number of standard errors that the sample estimate \( (\hat{p}_1 - \hat{p}_2) \) of \( (p_1 - p_2) \) falls from its null hypothesis value of zero is called
   A) the level of significance.  
   B) the margin of error.  
   C) the test statistic.  
   D) none of these  
   E) the P-value.

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Provide an appropriate response.

8) When confidence intervals for \( p_1 - p_2 \) do not contain zero, how can we tell which population proportion is predicted to be larger?

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

9) A 95% confidence interval for the difference in means for a collection of paired sample data is \( (0, 3.4) \). Based on the same sample, a traditional significance test fails to support the claim of \( \mu_d > 0 \). What can you conclude about the significance level \( \alpha (\alpha = 1 - .95) \) of the hypothesis test?
   A) \( \alpha = 0.01 \)  
   B) \( \alpha = 0.05 \)  
   C) \( \alpha = 0.95 \)  
   D) \( \alpha > 0.05 \)  
   E) \( \alpha < 0.05 \)
Interpret the test statistic.

10) In 2006, the General Social Survey asked respondents "When a person has a disease that cannot be cured, do you think doctors should be allowed by law to end the patient's life by some painless means if the patient and his family request it?". In testing $H_0: p_M = p_F$ versus $H_a: p_M > p_F$, where $p_M$ is the proportion of males who responded "yes" and $p_F$ is the proportion of females who responded "yes", suppose the $z$ statistic is 4.48. What is your conclusion using a significance level of 0.05?

A) Since the $p$-value is approximately 1, we reject the null hypothesis and conclude that the percentage of males holding this belief is greater than the percentage of females.
B) Since the $p$-value is approximately 0, we reject the null hypothesis and conclude that the percentage of females holding this belief is greater than the percentage of males.
C) Since the $p$-value is approximately 0, we reject the null hypothesis and conclude that the percentage of males holding this belief is greater than the percentage of females.
D) Since the $p$-value is approximately 1, we are unable to reject the null hypothesis, there is insufficient evidence to conclude that the percentage of males holding this belief is greater than the percentage of females.
E) Since the $p$-value is approximately 0, we are unable to reject the null hypothesis, there is insufficient evidence to conclude that the percentage of males holding this belief is greater than the percentage of females.

11) In 2006, the General Social Survey asked respondents "Should divorce in this country be easier or more difficult to obtain than it is now?". In testing $H_0: p_M = p_F$ versus $H_a: p_M > p_F$, where $p_M$ is the proportion of males who responded "more difficult" and $p_F$ is the proportion of females who responded "more difficult", suppose the $z$ statistic is 0.33. What is your conclusion using a significance level of 0.05?

A) Since $z < 1.645$, we reject the null hypothesis and conclude that the proportion of males who responded "more difficult" is greater than the proportion of females who responded "more difficult".
B) Since $z < 1.645$, we reject the null hypothesis and conclude that the proportion of females who responded "more difficult" is greater than the proportion of males who responded "more difficult".
C) Since $z = 0.33$ is greater than 0.05, we reject the null hypothesis and conclude that the proportion of males who responded "more difficult" is greater than the proportion of females who responded "more difficult".
D) Since $z < 1.645$, we are unable to reject the null hypothesis. There is insufficient evidence to show that the proportion of males who responded "more difficult" is greater than the proportion of females who responded "more difficult".
E) Since $z = 0.32$ is greater than 0.05, we reject the null hypothesis and conclude that the proportion of females who responded "more difficult" is greater than the proportion of males who responded "more difficult".
Interpret the given confidence interval.

12) For many of the years of its existence, the General Social Survey has asked respondents whether marijuana should be made legal. A 95% confidence interval for $p_{2006} - p_{1980}$ is given by $(0.08, 0.14)$ where $p_{2006}$ is the proportion of respondents who answered "legal" in 2006 and $p_{1980}$ is the proportion of respondents who responded "legal" in 1978. Based on the 95% confidence interval, what can we conclude about the percentage of respondents who favor legalization in 1980 versus 2006?

A) Since all of the values in the confidence interval are less than 1, we are unable to conclude that there is a significant difference between the percentage in favor of legalization in 1980 and 2006.
B) Since all of the values in the confidence interval are greater than 0, we can conclude that the percentage in favor of legalization was greater in 1980 than it is in 2006.
C) Since all of the values in the confidence interval are less than 1, we can conclude that there is a significant difference between the percentage in favor of legalization in 1980 and 2006.
D) Since all of the values in the confidence interval are greater than 0, we can conclude that the percentage in favor of legalization is greater in 2006 than it was in 1980.

A two-sided significance test for two population proportions is to be performed using the $P$-value approach. The null hypothesis is $H_0: p_1 = p_2$ and the alternative is $H_a: p_1 \neq p_2$. Use the given sample data to find the $P$-value for the significance test. Give an interpretation of the $P$-value.

13) $n_1 = 200 \quad n_2 = 100$
   $x_1 = 11 \quad x_2 = 8$

A) $P$-value $= 0.2$; The probability that the proportions are equal is 0.2.
B) $P$-value $= 0.4$; If there is no difference in the proportions, there is about a 40% chance of seeing the observed difference or larger due to natural sampling variation.
C) $P$-value $= 0.4$; The probability that the proportions are equal is 0.4.
D) $P$-value $= 0.025$; We are 95% confident that the difference in proportions is $0.025$.
E) $P$-value $= 0.025$; We are 95% confident that the difference in proportions is $0.025$.

14) $n_1 = 50 \quad n_2 = 75$
   $x_1 = 20 \quad x_2 = 15$

A) $P$-value $= 0.015$; The probability that the proportions are unequal is 0.015.
B) $P$-value $= 0.015$; The probability that the proportions are equal is 0.015.
C) $P$-value $= 0.007$; If there is no difference in the proportions, there is about a 0.7% chance of seeing the observed difference or larger due to natural sampling variation.
D) $P$-value $= 0.015$; If there is no difference in the proportions, there is about a 1.5% chance of seeing the observed difference or larger due to natural sampling variation.
E) $P$-value $= 0.2$; The difference in proportions is 0.2.
SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Provide an appropriate response.

15) *Newsweek* magazine (September 19, 1983) reported results of a poll to investigate whether there was a difference between the percentage of women and the percentage of men who approved of the way Ronald Reagan was handling his job as president. Of 511 women interviewed, 40% indicated approval. Of 507 men, 49% indicated approval.

a. Find a 90% confidence interval for the difference between the proportion of men and the proportion of women who approved of Reagan's performance. What assumption do you make to use this method?

b. Interpret the interval obtained in part (a). Can you conclude that there is a gender gap?

16) A recent General Social Survey asked subjects their opinions about government spending on defense and government spending on the environment. For each, should it increase, or should it decrease? The results are given in the table.

<table>
<thead>
<tr>
<th>Defense Spending</th>
<th>Environment Spending</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase</td>
<td>Increase</td>
</tr>
<tr>
<td>Decrease</td>
<td>Decrease</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Increase</th>
<th>Decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase</td>
<td>130</td>
<td>53</td>
</tr>
<tr>
<td>Decrease</td>
<td>256</td>
<td>51</td>
</tr>
</tbody>
</table>

a. Find the sample proportion who favor an increase for defense spending and the sample proportion who favor an increase for spending on the environment.

b. Give all steps of a test of the hypothesis that the corresponding population proportions are equal.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

17) When we fail to reject the null hypothesis, we
A) have obtained a t-value greater than our critical t-value.
B) have committed a Type I error.
C) conclude that sampling variability is responsible for our obtained difference.
D) prove that there is no significant difference between the two groups
E) have committed a Type II error.

18) A researcher is interested in the academic performance differences between individuals using an optimistic versus a pessimistic approach to their studies. If the researcher fails to find a significant difference, when in fact one exists in the population:
A) the null hypothesis was correctly accepted.
B) a Type I error has been made.
C) a Type II error has been made.
D) the research hypothesis was correctly accepted.
E) the null hypothesis was correctly rejected.
In the recent past, a cola manufacturer invited consumers to take a blind taste test. Consumers were asked to determine which of two sodas (Brand 1 - the manufacturer's own soda, or Brand 2 - the competitor's soda) they preferred. The manufacturer was interested in determining which factors played a role in people's taste preferences. One of the factors studied was the gender of the consumer. Let $p_1$ be the true proportion of men who prefer Brand 1 and let $p_2$ be the true proportion of women who prefer Brand 1. Below is a computer printout comparing the taste preferences of the men and women surveyed.

$$H_0: p_1 = p_2$$

SAMPLES SELECTED FROM soda(brand1, brand2)

<table>
<thead>
<tr>
<th>sex</th>
<th>samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>males</td>
<td>(sex=0, males)</td>
</tr>
<tr>
<td>females</td>
<td>(sex=1, females)</td>
</tr>
</tbody>
</table>

$$\hat{p}_1 = 0.422018$$

sample size of $X = 109$

$$\hat{p}_2 = 0.25$$

sample size of $Y = 52$

$$\hat{p}_1 - \hat{p}_2 = 0.172018$$

$$z = 2.11825$$

To determine if a difference exists in the taste preferences of men and women, which alternative hypothesis would the manufacturer use?

A) $H_a: \mu_2 - \mu_1 = 0$
B) $H_a: \mu_2 - \mu_1 \neq 0$
C) $H_0: \mu_2 - \mu_1 \neq 0$
D) $H_a: p_1 - p_2 \neq 0$
E) $H_a: p_1 - p_2 = 0$
SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

20) General Social Surveys have often asked, "Do you agree or disagree with this statement? Women should take care of running their homes and leave running the country up to men." The results from 1974 (sample 1) and 1998 (sample 2) are shown in the table, where X is the number who agreed. Is there a change in opinion from 1974 to 1998 (the last year this was asked)? Answer by reporting from the printout:

a. the 95% confidence interval, giving your interpretation.
b. the hypotheses, test statistic, and P-value of a significance test, giving your interpretation.
c. Are the methods used in parts (a) and (b) valid? Explain.

<table>
<thead>
<tr>
<th>Sample</th>
<th>X</th>
<th>N</th>
<th>Sample p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>509</td>
<td>1431</td>
<td>0.355695</td>
</tr>
<tr>
<td>2</td>
<td>280</td>
<td>1814</td>
<td>0.154355</td>
</tr>
</tbody>
</table>

Difference = \( p_1 - p_2 \)
Estimate for difference: 0.201340
95% CI for difference: (0.171480, 0.231201)
Test for difference = 0 (vs. not =0): Z = 13.27 P-Value = 0.000

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

21) From the sample statistics, find the value of the pooled estimate \( \hat{p} \) used.

\[ n_1 = 100 \quad n_2 = 100 \]
\[ x_1 = 34 \quad x_2 = 30 \]

A) 0.04
B) none of these
C) 0.64
D) 0.35
E) 0.32

22) \( n_1 = 216 \quad n_2 = 186 \)
\[ x_1 = 76 \quad x_2 = 99 \]

A) 0.18
B) 0.88
C) 0.44
D) none of these
E) 0.35

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Fill in the blank.

23) The proportion \( \hat{p} \) used in calculating the z test statistic when comparing two population proportions is called an \( \underline{\underline{\underline{\text{pooled estimate}}}} \), because it estimates the common value of \( p_1 \) and \( p_2 \) presuming \( H_0 \) is true.
MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Provide an appropriate response.

24) The General Social Survey asked respondents whether marijuana should be made legal in many of its surveys. Calculate a 95% confidence interval for \( p_{2006} - p_{1978} \) where \( p_{2006} \) is the proportion of respondents who answered "legal" in 2006 and \( p_{1978} \) is the proportion of respondents who responded "legal" in 1978 using the information given in the table below.

<table>
<thead>
<tr>
<th></th>
<th>1978</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal</td>
<td>447</td>
<td>672</td>
</tr>
<tr>
<td>n</td>
<td>1464</td>
<td>1828</td>
</tr>
</tbody>
</table>

A) (-0.09, -0.03)
B) none of these
C) (0.64, 0.71)
D) (0.03, 0.09)
E) (-0.64, -0.71)
Answer Key
Testname: 10_1

1) E
2) A
3) B
4) B
5) matched
6) D
7) C
8) If the confidence interval contains all positive values, \( p_1 > p_2 \); however, if the confidence interval contains all negative values, \( p_2 > p_1 \).
9) E
10) C
11) D
12) D
13) B
14) D
15) a. Let \( p_1 \) = population proportion of men who approved of Reagan’s performance and \( p_2 \) = population proportion of women who approved of Reagan’s performance. Based on the sample information given, the 90% CI for \( p_1 - p_2 \) is (0.039, 0.141). The assumptions made to use this method are: 1) independent random samples for the two groups; 2) large enough sample sizes \( n_1 \) and \( n_2 \) so that, in each sample, there are at least 10 "successes" and at least 10 "failures";
   b. We can be 90% confident that the population proportion of people who approved of Reagan’s performance is between 0.039 higher and 0.141 higher for men than for women. Since both endpoints of the confidence interval for \( p_1 - p_2 \) are positive, we can conclude that there is a gender gap.
16) a. 0.373, 0.788; b. The assumptions for the two-sided test are 1) categorical response variable for two groups; 2) independent random samples; 3) there are at least 5 "successes" and 5 "failures" in each sample. \( H_0: p_1 = p_2 \), \( H_a: p_1 \neq p_2 \), \( z = -13.17 \), P-value = 0.0000; Since the P-value is below 0.05, there is extremely strong evidence against the null hypothesis that the population proportions favoring an increase in government spending for defense and the environment are the same at \( \alpha = 0.05 \).
17) D
18) B
19) D
20) a. The 95% confidence interval for \( p_1 - p_2 \) is (0.171, 0.231). We can be 95% confident that the population proportion who agreed with the statement is between 0.171 higher and 0.231 higher in 1974 than in 1998. Since both endpoints of the confidence interval for \( p_1 - p_2 \) are positive, we infer that \( p_1 - p_2 \) is positive. That is, it seems that \( p_1 \) is greater than \( p_2 \). Therefore, we can conclude that there is a change in opinion from 1974 to 1998; b. \( H_0: p_1 = p_2 \), \( H_a: p_1 \neq p_2 \), \( z = 13.27 \), P-value = 0.0000; Since the P-value is below 0.05, there is extremely strong evidence against the null hypothesis that the population proportions agreeing with the statement were the same in 1974 as they are in 1998 at \( \alpha = 0.05 \); c. The methods used in parts (a) and (b) are valid, because 1) categorical response variable for two groups; 2) independent random samples; and 3) each sample has at least 10 "successes" and 10 "failures."
21) E
22) C
23) pooled
24) D