Part I - Multiple Choice (Questions 1-10) – Circle the letter of the answer of your choice.

1. An SRS of size 100 is taken from a population having proportion 0.8 successes. An independent SRS of size 400 is taken from a population having proportion 0.5 successes. The sampling distribution for the difference in sample proportions has what standard deviation?
   (a) 1.3
   (b) 0.40
   (c) 0.047
   (d) 0.0002
   (e) None of the above.

2. What is the best reason for performing a paired experiment rather than a two-independent sample experiment in this case?
   (a) It is easier to do since we need fewer experimental units and each unit receives more than one treatment.
   (b) It allows us to remove variation in the results caused by other factors since we can compare both treatments within the same experimental unit.
   (c) The computer program is more accurate since we work only with the differences.
   (d) It requires fewer assumptions since we are interested only in the difference between treatments.
   (e) It allows us to do more experiments since we use each experimental unit twice.

The next three questions refer to the following situation.
An experiment was conducted to assess the efficacy of spraying oats with malathion (at 0.25 lb/acre) to control the cereal leaf beetle. A sample of 10 farms was selected at random from southwest Manitoba. Each farm was assigned at random to either the control group (no spray) or the treatment group (spray). At the conclusion of the experiment, a plot on each farm was selected and the number of larvae per stem was measured. Here are two possible computer outputs from (only one of which is correct; some output hidden):

<table>
<thead>
<tr>
<th>t-Tests: separate estimates of ( \mu_1, \mu_2 )</th>
<th>t-Test, paired samples: not spray - spray</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test ( H_0: \mu(\text{not spray})-\mu(\text{spray}) = 0 )</td>
<td>Test ( H_0: \mu = 0 ) vs ( H_a: \mu &gt; 0 )</td>
</tr>
<tr>
<td>vs ( H_a: \mu(\text{not spray})-\mu(\text{spray}) &gt; 0 )</td>
<td>Sample mean = 1.0440</td>
</tr>
<tr>
<td>Sample mean (not spray) = 4.0947</td>
<td>t-statistic = 1.896 with * d.f.</td>
</tr>
<tr>
<td>Sample mean (spray) = 3.0508</td>
<td></td>
</tr>
</tbody>
</table>

3. The appropriate test statistic and P-value are
   (a) 1.896, 0.033  (b) 1.896, 0.131  (c) 1.896, 0.065  (d) 1.887, 0.059  (e) 1.887, 0.118

4. A Type II error would occur if
   (a) we conclude malathion is ineffective when in fact it was effective.
   (b) we conclude malathion is effective when in fact it is ineffective.
   (c) we conclude malathion is effective when in fact it is effective.
   (d) we conclude malathion is ineffective when in fact it is ineffective.
   (e) we conclude malathion is neither ineffective nor effective.

5. Power refers to
   (a) the ability to detect an effect of malathion when in fact there is no effect.
   (b) the ability to not detect an effect of malathion when in fact there is no effect.
   (c) the ability to detect an effect of malathion when in fact there is an effect.
   (d) the ability to not detect an effect of malathion when in fact there is an effect.
   (e) the ability to make a correct decision regardless of whether malathion has an effect or not.
6. The Excellent Drug Company claims its aspirin tablets will relieve headaches faster than any other aspirin on the market. To determine whether Excellent's claim is valid, random samples of size 15 are chosen from aspirins made by Excellent and the Simple Drug Company. An aspirin is given to each of the 30 randomly selected persons suffering from headaches and the number of minutes required for each to recover from the headache is recorded. The sample results are:

<table>
<thead>
<tr>
<th></th>
<th>( \bar{x} )</th>
<th>( s^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent (E)</td>
<td>8.4</td>
<td>4.2</td>
</tr>
<tr>
<td>Simple (S)</td>
<td>8.9</td>
<td>4.6</td>
</tr>
</tbody>
</table>

A 5% significance level test is performed to determine whether Excellent's aspirin cures headaches significantly faster than Simple's aspirin. The appropriate hypotheses to be tested are

(a) \( H_0: \mu_E - \mu_S = 0 \); \( H_a: \mu_E - \mu_S > 0 \)
(b) \( H_0: \mu_E - \mu_S = 0 \); \( H_a: \mu_E - \mu_S \neq 0 \)
(c) \( H_0: \mu_E - \mu_S = 0 \); \( H_a: \mu_E - \mu_S < 0 \)
(d) \( H_0: \mu_E - \mu_S < 0 \); \( H_a: \mu_E - \mu_S = 0 \)
(e) \( H_0: \mu_E - \mu_S > 0 \); \( H_a: \mu_E - \mu_S = 0 \)

7. 42 of 65 randomly selected people at a baseball game report owning an iPod. 34 of 52 randomly selected people at a rock concert occurring at the same time across town reported owning an iPod. A researcher wants to test the claim that the proportion of iPod owners at the two venues is not the same. A 90% confidence interval for the difference in population proportions is \((-0.154, 0.138)\). Which of the following gives the correct outcome of the researchers' test of the claim?

(a) Since the confidence interval includes 0, the researcher can conclude that the proportion of iPod owners at the two venues is the same.
(b) Since the confidence interval includes 0, the researcher can conclude that the proportion of iPod owners at the two venues may be the same.
(c) Since the confidence interval includes 0, the researcher can conclude that the proportion of iPod owners at the two venues is different.
(d) Since the confidence interval includes more positive than negative values, we can conclude that a higher proportion of people at the baseball game own iPods than at the rock concert.
(e) We cannot draw a conclusion about a claim without performing a significance test.

8. A researcher wants to see if birds that build larger nests lay larger eggs. She selects two random samples of nests: one of small nests and the other of large nests. She weighs one egg from each nest. The data are summarized below.

<table>
<thead>
<tr>
<th></th>
<th>Small nests</th>
<th>Large nests</th>
</tr>
</thead>
<tbody>
<tr>
<td>sample size</td>
<td>60</td>
<td>159</td>
</tr>
<tr>
<td>sample mean (g)</td>
<td>37.2</td>
<td>35.6</td>
</tr>
<tr>
<td>sample variance</td>
<td>24.7</td>
<td>39.0</td>
</tr>
</tbody>
</table>

A 95% confidence interval for the difference between the average mass of eggs in small and large nests is:

(a) \((37.2 - 35.6) \pm 2.000 \sqrt{\frac{24.7^2}{60} + \frac{39.0^2}{159}}\)
(b) \((37.2 - 35.6) \pm 2.009 \sqrt{\frac{24.7^2}{59} + \frac{39.0^2}{158}}\)
(c) \((37.2 - 35.6) \pm \sqrt{\frac{24.7^2}{59} + \frac{39.0^2}{158}}\)
(d) \((37.2 - 35.6) \pm \sqrt{\frac{24.7^2 + 39.0^2}{59 + 158}}\)
(e) None of these

9. To use the two-sample \( t \) procedure to perform a significance test on the difference between two means, we assume that

(a) the populations' standard deviations are known.
(b) the samples from each population are independent.
(c) the distributions are exactly Normal in each population.
(d) the sample sizes are large.
(e) all of the above
10. The following are percents of fat found in 5 samples of each of two brands of ice cream:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th></th>
<th></th>
<th></th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5.7</td>
<td>4.5</td>
<td>6.2</td>
<td>6.3</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>6.3</td>
<td>5.7</td>
<td>5.9</td>
<td>6.4</td>
<td>5.1</td>
</tr>
</tbody>
</table>

Which of the following procedures is appropriate to test the hypothesis of equal average fat content in the two types of ice cream?
(a) Paired t test with 5 df.
(b) Two-sample t test with 4 df.
(c) Paired t test with 4 df.
(d) Two-sample t test with 9 df.
(e) Two-proportion z test

Part II - Free Response (Question 11-13) - Show your work and explain your results clearly.

11. Europeans have been more skeptical than Americans about the use of genetic engineering to improve foods. A sample survey gathered responses from random samples of 863 Americans and 12,178 Europeans. The European sample was larger because Europe is divided into several nations. Subjects were asked to consider using modern biotechnology in the production of foods, for example, to make them higher in protein, keep longer, or change in taste. They were asked if they considered this "risky for society." In all, 52% of Americans and 64% of Europeans thought the application was risky.

(a) Is there convincing evidence that more than half of all adult Americans consider applying biotechnology to the production of foods risky? Use a significance test to support your answer.

\[
p_A = 0.52
\]

\[
H_0: p_A = 0.5
\]

\[
H_a: p_A > 0.5
\]

\[
\alpha = 0.05
\]

\[
z = \frac{0.52 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{863}}} = 1.19
\]

\[
p\text{value} = 0.1167
\]

Since our p-value is greater than alpha, we fail to reject the null hypothesis. We do not have reason to believe that more than half of Americans consider this risky.

(b) Construct and interpret a 99% confidence interval for the percentage difference between Europe and the United States.

\[
\hat{p}_A = 0.52 \quad \hat{p}_E = 0.64
\]

\[
\hat{p}_E - \hat{p}_A = \pm z^* (SE)
\]

\[
0.64 - 0.52 \pm 2.326 \left( \frac{0.52 \times 0.48}{863} + \frac{0.64 \times 0.36}{12178} \right)
\]

\[
(0.075, 0.165)
\]

We are 99% confident that the true difference is between 7% and 16%.
Researchers studying the learning of speech often compare measurements made on the recorded speech of adults and children. One variable of interest is called the voice onset time (VOT). Here are the results for 6-year-old children and adults asked to pronounce the word “bees.” The VOT is measured in milliseconds and can be either positive or negative.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>T</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children</td>
<td>10</td>
<td>-3.67</td>
<td>33.89</td>
</tr>
<tr>
<td>Adults</td>
<td>20</td>
<td>-23.17</td>
<td>50.74</td>
</tr>
</tbody>
</table>

(a) Do the results give the researchers evidence that the VOT distinguishes adults from children?

\[ H_0: \mu_A - \mu_C = 0 \]
\[ H_a: \mu_A - \mu_C \neq 0 \]
\[ \alpha = 0.05 \]

Conditions
- Assume for both groups
- Independence 100% all boys olds
- 200% all adults
- Normality - small samples - proceed with caution

\[ t = \frac{-23.17 + 3.67}{\sqrt{\frac{50.74^2}{20} + \frac{33.89^2}{10}}} = -1.249 \quad df = 25.38 \]

p-value = 0.223

Since our p-value is greater than alpha, we fail to reject the null hypothesis. We do not have reason to believe children and adults test differently.

(b) The researchers in the study looked at VOTs for adults and children pronouncing several different words. Explain why they should not perform a separate two-sample t test for each word and conclude that the words with a significant difference (say, \( P < 0.05 \)) distinguish children from adults.

Children and adults pronounce some words the same so should compare the group of words to distinguish adults from children. The word might be confounding.
Part I - Multiple Choice (Questions 1-10) – Circle the letter of the answer of your choice.

1. An SRS of size 100 is taken from a population having proportion 0.8 successes. An independent SRS of size 400 is taken from a population having proportion 0.5 successes. The sampling distribution for the difference in sample proportions has what standard deviation?
   (a) 1.3
   (b) 0.40
   (c) 0.047
   (d) 0.0002
   (e) None of the above.

2. What is the best reason for performing a paired experiment rather than a two-independent sample experiment in this case?
   (a) It is easier to do since we need fewer experimental units and each unit receives more than one treatment.
   (b) It allows us to remove variation in the results caused by other factors since we can compare both treatments within the same experimental unit.
   (c) The computer program is more accurate since we work only with the differences.
   (d) It requires fewer assumptions since we are interested only in the difference between treatments.
   (e) It allows us to do more experiments since we use each experimental unit twice.

The next three questions refer to the following situation.
An experiment was conducted to assess the efficacy of spraying oats with malathion (at 0.25 lb/acre) to control the cereal leaf beetle. A sample of 10 farms was selected at random from southwest Manitoba. Each farm was assigned at random to either the control group (no spray) or the treatment group (spray). At the conclusion of the experiment, a plot on each farm was selected and the number of larvae per stem was measured. Here are two possible computer outputs from (only one of which is correct; some output hidden):

<table>
<thead>
<tr>
<th>Test, separate estimates of ( \mu_1, \mu_2 )</th>
<th>Test, paired samples: not spray - spray</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test ( H_0: \mu(\text{not spray}) - \mu(\text{spray}) = 0 ) vs ( H_A: \mu(\text{not spray}) - \mu(\text{spray}) &gt; 0 )</td>
<td>Test ( H_0: \mu = 0 ) vs ( H_A: \mu &gt; 0 )</td>
</tr>
<tr>
<td>Sample mean(\text{not spray}) = 4.0947</td>
<td>Sample mean = 1.0440</td>
</tr>
<tr>
<td>Sample mean(\text{spray}) = 3.0508</td>
<td></td>
</tr>
<tr>
<td>( t )-statistic = 1.896 with * d.f.</td>
<td>( t )-statistic = 1.887 with * d.f.</td>
</tr>
</tbody>
</table>

3. The appropriate test statistic and \( P \)-value are
   (a) 1.896, 0.033  (b) 1.896, 0.131  (c) 1.896, 0.065  (d) 1.887, 0.059  (e) 1.887, 0.118

4. A Type II error would occur if
   (a) we conclude malathion is ineffective when in fact it was effective.
   (b) we conclude malathion is effective when in fact it is ineffective.
   (c) we conclude malathion is effective when in fact it is effective.
   (d) we conclude malathion is ineffective when in fact it is ineffective.
   (e) we conclude malathion is neither ineffective nor effective.

5. Power refers to
   (a) the ability to detect an effect of malathion when in fact there is no effect.
   (b) the ability to not detect an effect of malathion when in fact there is no effect.
   (c) the ability to detect an effect of malathion when in fact there is an effect.
   (d) the ability to not detect an effect of malathion when in fact there is an effect.
   (e) the ability to make a correct decision regardless of whether malathion has an effect or not.
6. The Excellent Drug Company claims its aspirin tablets will relieve headaches faster than any other aspirin on the market. To determine whether Excellent's claim is valid, random samples of size 15 are chosen from aspirins made by Excellent and the Simple Drug Company. An aspirin is given to each of the 30 randomly selected persons suffering from headaches and the number of minutes required for each to recover from the headache is recorded. The sample results are:

<table>
<thead>
<tr>
<th></th>
<th>( \bar{x} )</th>
<th>( s^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent (E)</td>
<td>8.4</td>
<td>4.2</td>
</tr>
<tr>
<td>Simple (S)</td>
<td>8.9</td>
<td>4.6</td>
</tr>
</tbody>
</table>

A 5% significance level test is performed to determine whether Excellent's aspirin cures headaches significantly faster than Simple's aspirin. The appropriate hypotheses to be tested are

(a) \( H_0: \mu_E - \mu_S = 0 ; H_a: \mu_E - \mu_S > 0 \)
(b) \( H_0: \mu_E - \mu_S = 0 ; H_a: \mu_E - \mu_S \neq 0 \)
(c) \( H_0: \mu_E - \mu_S = 0 ; H_a: \mu_E - \mu_S < 0 \)
(d) \( H_0: \mu_E - \mu_S < 0 ; H_a: \mu_E - \mu_S = 0 \)
(e) \( H_0: \mu_E - \mu_S > 0 ; H_a: \mu_E - \mu_S = 0 \)

7. 42 of 65 randomly selected people at a baseball game report owning an iPod. 34 of 52 randomly selected people at a rock concert occurring at the same time across town reported owning an iPod. A researcher wants to test the claim that the proportion of iPod owners at the two venues is not the same. A 90% confidence interval for the difference in population proportions is \((-0.154, 0.138)\). Which of the following gives the correct outcome of the researchers' test of the claim?

(a) Since the confidence interval includes 0, the researcher can conclude that the proportion of iPod owners at the two venues is the same.
(b) Since the confidence interval includes 0, the researcher can conclude that the proportion of iPod owners at the two venues may be the same.
(c) Since the confidence interval includes 0, the researcher can conclude that the proportion of iPod owners at the two venues is different.
(d) Since the confidence interval includes more positive than negative values, we can conclude that a higher proportion of people at the baseball game own iPods than at the rock concert.
(e) We cannot draw a conclusion about a claim without performing a significance test.

8. A researcher wants to see if birds that build larger nests lay larger eggs. She selects two random samples of nests: one of small nests and the other of large nests. She weighs one egg from each nest. The data are summarized below.

<table>
<thead>
<tr>
<th></th>
<th>Small nests</th>
<th>Large nests</th>
</tr>
</thead>
<tbody>
<tr>
<td>sample size</td>
<td>60</td>
<td>159</td>
</tr>
<tr>
<td>sample mean (g)</td>
<td>37.2</td>
<td>35.6</td>
</tr>
<tr>
<td>sample variance</td>
<td>24.7</td>
<td>39.0</td>
</tr>
</tbody>
</table>

A 95% confidence interval for the difference between the average mass of eggs in small and large nests is:

(a) \((37.2 - 35.6) \pm \frac{2.000 \sqrt{24.7^2 + 39.0^2}}{60 + 159}\)
(b) \((37.2 - 35.6) \pm \frac{2.009 \sqrt{24.7^2 + 39.0^2}}{59 + 158}\)
(c) \((37.2 - 35.6) \pm \frac{24.7}{59} + \frac{39.0}{158}\)
(d) \((37.2 - 35.6) \pm \frac{24.7 + 39.0}{59 + 158}\)
(e) None of these

9. To use the two-sample \(t\) procedure to perform a significance test on the difference between two means, we assume that
(a) the populations' standard deviations are known.
(b) the samples from each population are independent.
(c) the distributions are exactly Normal in each population.
(d) the sample sizes are large.
(e) all of the above
10. The following are percents of fat found in 5 samples of each of two brands of ice cream:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.7</td>
<td>6.3</td>
<td>6.2</td>
</tr>
<tr>
<td>4.5</td>
<td>5.7</td>
<td>6.3</td>
</tr>
<tr>
<td>6.3</td>
<td>5.9</td>
<td>6.4</td>
</tr>
<tr>
<td>6.2</td>
<td>6.4</td>
<td>5.1</td>
</tr>
</tbody>
</table>

Which of the following procedures is appropriate to test the hypothesis of equal average fat content in the two types of ice cream?
(a) Paired t test with 5 df.
(b) Two-sample t test with 4 df.
(c) Paired t test with 4 df.
(d) Two-sample t test with 9 df.
(e) Two-proportion z test

Part II – Free Response (Question 11-13) – Show your work and explain your results clearly.

11. Europeans have been more skeptical than Americans about the use of genetic engineering to improve foods. A sample survey gathered responses from random samples of 863 Americans and 12,178 Europeans. (The European sample was larger because Europe is divided into several nations.) Subjects were asked to consider using modern biotechnology in the production of foods, for example, to make them higher in protein, keep longer, or change in taste. They were asked if they considered this "risky for society." In all, 52% of Americans and 64% of Europeans thought the application was risky.

(a) Is there convincing evidence that more than half of all adult Americans consider applying biotechnology to the production of foods risky? Use a significance test to support your answer.

(b) Construct and interpret a 99% confidence interval for the percentage difference between Europe and the United States.
Researchers studying the learning of speech often compare measurements made on the recorded speech of adults and children. One variable of interest is called the voice onset time (VOT). Here are the results for 6-year-old children and adults asked to pronounce the word "bees." The VOT is measured in milliseconds and can be either positive or negative.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>$\bar{x}$</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children</td>
<td>10</td>
<td>-3.67</td>
<td>33.89</td>
</tr>
<tr>
<td>Adults</td>
<td>20</td>
<td>-23.17</td>
<td>50.74</td>
</tr>
</tbody>
</table>

(a) Do the results give the researchers evidence that the VOT distinguishes adults from children?

(b) The researchers in the study looked at VOTs for adults and children pronouncing several different words. Explain why they should not perform a separate two-sample $t$ test for each word and conclude that the words with a significant difference (say, $P < 0.05$) distinguish children from adults.
Part I - Multiple Choice (Questions 1-10) – Circle the letter of the answer of your choice.

Questions 1 to 3 relate to the following situation.
A well-known chewing gum maker wants to determine if people who chew gum have a preference of flavors. A random sample of sales is selected. Here are the results:

<table>
<thead>
<tr>
<th>Flavor</th>
<th>Peppermint</th>
<th>Cinnamon</th>
<th>Wintergreen</th>
<th>Spearmint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number sold</td>
<td>25</td>
<td>19</td>
<td>22</td>
<td>14</td>
</tr>
</tbody>
</table>

1. An appropriate null hypothesis for a significance test would be
   (a) $\mu = 20$
   (b) There is a flavor preference.
   (c) The cell counts (above) are independent of flavor.
   (d) At least one of the cell counts (above) is different from the other three.
   (e) There is no flavor preference.

2. Which of the following are conditions for a chi-square test?
   I. There have to be at least 800 gum chewers in the population.
   II. If $p =$ proportion of gum chewers, then $np \geq 10$ and $n(1 - p) \geq 10$.
   III. All of the cell counts (above) have to be at least 5.
   IV. The sample has to be random.

   (a) I and III only
   (b) II and IV only
   (c) IV only
   (d) I and IV only
   (e) III and IV only

3. An appropriate interpretation of this inference procedure is
   (a) gum chewers prefer peppermint flavor.
   (b) one or more of the conditions for inference is violated, so we can’t conclude anything from this study.
   (c) the $P$-value is so large that $H_0$ can be rejected. There is no gum preference.
   (d) there is insufficient evidence that gum chewers have a preference among these four flavors.
   (e) there is sufficient evidence of a gum preference.

4. Are all employees equally prone to having accidents? To investigate this hypothesis, a researcher looked at a light manufacturing plant and classified the accidents by type and by age of the employee.

<table>
<thead>
<tr>
<th>Age</th>
<th>Sprain</th>
<th>Burn</th>
<th>Cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 25</td>
<td>9</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>25 or over</td>
<td>61</td>
<td>13</td>
<td>12</td>
</tr>
</tbody>
</table>

A chi-square test gave a test statistic of 20.78. If we test at $\alpha = 0.05$
   (a) there appears to be no association between accident type and age.
   (b) age seems to be independent of accident type.
   (c) accident type does not seem to be independent of age.
   (d) there appears to be a 20.78% correlation between accident type and age.
   (e) the proportion of sprain, cuts, and burns seems to be similar for both age classes.

5. A survey was conducted to investigate whether alcohol consumption and smoking are related. The following information was compiled for 600 individuals:

<table>
<thead>
<tr>
<th>Smoker</th>
<th>Nonsmoker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drinker</td>
<td>193</td>
</tr>
<tr>
<td>Nondrinker</td>
<td>89</td>
</tr>
</tbody>
</table>

Which of the following statements is true?
(a) The appropriate alternative hypothesis is $H_a$: Smoking and alcohol consumption are independent.
(b) The appropriate null hypothesis is $H_0$: Smoking and alcohol consumption are not independent.
(c) The calculated value of the test statistic is 3.84.
(d) The calculated value of the test statistic is 7.86.
(e) At level 0.01 we conclude that smoking and alcohol consumption are related.
Questions 6 to 9 refer to the following situation.

In the paper "Color Association of Male and Female Fourth-Grade School Children" (Journal of Psychology, 1988, 383-388), reported on a study in which children were asked to indicate what emotion they associated with the color red. The response and the sex of the child are noted and summarized below. The first number in each cell is the count; the second number is the row percent.

<table>
<thead>
<tr>
<th></th>
<th>Anger</th>
<th>Happy</th>
<th>Love</th>
<th>Pain</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>27</td>
<td>19</td>
<td>39</td>
<td>17</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>26.47</td>
<td>18.63</td>
<td>38.24</td>
<td>16.67</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>34</td>
<td>12</td>
<td>38</td>
<td>28</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>30.36</td>
<td>10.71</td>
<td>33.93</td>
<td>25.00</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>61</td>
<td>31</td>
<td>77</td>
<td>45</td>
<td>214</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Statistic</th>
<th>DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>*</td>
<td>4.629</td>
<td>*****</td>
</tr>
</tbody>
</table>

6. The null hypothesis is
   (a) emotional association with red is independent of gender.
   (b) gender is dependent upon the emotional association with red.
   (c) the probability of associating a specific emotion with red is related to gender.
   (d) male and female children's emotional reaction to the color red are homogenous
   (e) the color red is independent of the emotion associated with it and with gender.

7. Under a suitable null hypothesis, the expected frequency for the cell corresponding to Anger and Males is
   (a) 15.9.
   (b) 55.7.
   (c) 30.4.
   (d) 31.9.
   (e) 29.1.

8. The null hypothesis will be rejected at \( \alpha = 0.05 \) if the test statistic exceeds
   (a) 3.84.
   (b) 5.99.
   (c) 7.81.
   (d) 9.49.
   (e) 14.07.

9. A Type I error would be committed if
   (a) we conclude that the sex of the child and the emotion associated with red are independent when in fact they are not independent.
   (b) we conclude that the sex of the child and the emotion associated with red are not independent when in fact they are not independent.
   (c) we conclude that the proportion of emotions associated with red differs between males and females when in fact they are the same.
   (d) we conclude that the proportion of emotions associated with red is the same for males and females when in fact they are the same.
   (e) we fail to find any association between the color red and emotions for either sex.

10. Each person in a random sample of 50 was asked to state his/her sex and preferred color. The resulting frequencies are shown below.

<table>
<thead>
<tr>
<th></th>
<th>Color</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>Male</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Female</td>
<td>15</td>
<td>6</td>
</tr>
</tbody>
</table>

A chi-square test is used to test the null hypothesis that sex and preferred color are independent. Which of the following statements is a correct decision about the null hypothesis?
   (a) Reject at the 0.005 level.
   (b) Reject at the 0.01 level but not at the 0.005 level.
   (c) Reject at the 0.025 level but not at the 0.01 level.
   (d) Reject at the 0.05 level but not at the 0.025 level.
   (e) Accept at the 0.05 level.
Part II – Free Response (Question 11-13) – Show your work and explain your results clearly.

11. An experiment in chicken breeding results in offspring having very curly, slightly curly, or normal feathers. If this is the result of a single gene system, then the proportions of offspring in the three phenotypes should be 0.25, 0.50, and 0.25 respectively. In one such experiment, 93 chickens were born. 20 had normal feathers, 50 had slightly curly feathers, and 23 had very curly feathers. Carry out a test to determine whether the genetic model seems to hold in this setting.

- feathers: VC SC N
- obs: 23 50 20 93
- exp: 23.25 46.5 23.25

\[ X^2 \text{ GOF test} \quad \alpha = 0.05 \]

Ho: the proportions of types of feathers match the proportions of a single gene system.
Ha: the proportions are different than the single gene system.

\[ X^2 = \sum \frac{(\text{obs} - \text{exp})^2}{\text{exp}} = 0.7204 \quad df = 2 \]

P-value = 0.6975

Our p-value is greater than our alpha level so we fail to reject the null hypothesis. We have reason to believe the genetic model holds.

12. Describe the differences in the three types of chi-squared tests.

\[ X^2 \text{ GOF} = \text{one sample, one variable, see if proportions match a set proportion.} \]

\[ X^2 \text{ homogeneity} = \text{two samples - one from each of two populations, see if populations are the same.} \]

\[ X^2 \text{ independence} = \text{one population, two variables, see if two variables are independent.} \]
13. Your school administration wants to install new soft drink machines in the gym and cafeteria. Their market analysts want to know if flavor preference depends on the location. A random sample of 180 students was selected and interviewed. Their location and soft drink preference are given in the table.

<table>
<thead>
<tr>
<th>Soft drink</th>
<th>Gym</th>
<th>Cafeteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coca-Cola</td>
<td>33</td>
<td>57</td>
</tr>
<tr>
<td>Pepsi</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>Sprite</td>
<td>5</td>
<td>18</td>
</tr>
</tbody>
</table>

Determine if there is evidence that flavor preference depends on the location.

\[ \chi^2 \text{ test for independence} \quad \alpha = 0.05 \]

\[ H_0: \text{location and soft drink are independent} \]
\[ H_a: \text{location and soft drink preference are dependent} \]

Conditions
- \( SEs \) given
- All expected counts > 75

\[ \chi^2 = \sum \frac{(O - E)^2}{E} = 21.425 \quad df = 2 \]

\[ p-value = 0.0002 \]

Since our \( p-value \) < \( \alpha \), we reject the null hypothesis. We have reason to believe that location and soft drink preference are dependent.
Part I - Multiple Choice (Questions 1-10) – Circle the letter of the answer of your choice.

Questions 1 to 3 relate to the following situation.
A well-known chewing gum maker wants to determine if people who chew gum have a preference of flavors. A random sample of sales is selected. Here are the results:

<table>
<thead>
<tr>
<th>Flavor</th>
<th>Peppermint</th>
<th>Cinnamon</th>
<th>Wintergreen</th>
<th>Spearmint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number sold</td>
<td>25</td>
<td>19</td>
<td>22</td>
<td>14</td>
</tr>
</tbody>
</table>

1. An appropriate null hypothesis for a significance test would be
(a) \( \mu = 20 \)
(b) There is a flavor preference.
(c) The cell counts (above) are independent of flavor.
(d) At least one of the cell counts (above) is different from the other three.
(e) There is no flavor preference.

2. Which of the following are conditions for a chi-square test?
I. There have to be at least 800 gum chewers in the population.
II. If \( p \) = proportion of gum chewers, then \( np \geq 10 \) and \( n(1 - p) \geq 10 \).
III. All of the cell counts (above) have to be at least 5.
IV. The sample has to be random.
(a) I and III only
(b) II and IV only
(c) IV only
(d) I and IV only
(e) III and IV only

3. An appropriate interpretation of this inference procedure is
(a) gum chewers prefer peppermint flavor.
(b) one or more of the conditions for inference is violated, so we can’t conclude anything from this study.
(c) the \( P \)-value is so large that \( H_0 \) can be rejected. There is no gum preference.
(d) there is insufficient evidence that gum chewers have a preference among these four flavors.
(e) there is sufficient evidence of a gum preference.

4. Are all employees equally prone to having accidents? To investigate this hypothesis, a researcher looked at a light manufacturing plant and classified the accidents by type and by age of the employee.

<table>
<thead>
<tr>
<th>Age</th>
<th>Sprain</th>
<th>Burn</th>
<th>Cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 25</td>
<td>9</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>25 or over</td>
<td>61</td>
<td>13</td>
<td>12</td>
</tr>
</tbody>
</table>

A chi-square test gave a test statistic of 20.78. If we test at \( \alpha = 0.05 \)
(a) there appears to be no association between accident type and age.
(b) age seems to be independent of accident type.
(c) accident type does not seem to be independent of age.
(d) there appears to be a 20.78% correlation between accident type and age.
(e) the proportion of sprain, cuts, and burns seems to be similar for both age classes.

5. A survey was conducted to investigate whether alcohol consumption and smoking are related. The following information was compiled for 600 individuals:

<table>
<thead>
<tr>
<th>Smoker</th>
<th>Nonsmoker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drinker</td>
<td>193</td>
</tr>
<tr>
<td>Nondrinker</td>
<td>89</td>
</tr>
</tbody>
</table>

Which of the following statements is true?
(a) The appropriate alternative hypothesis is \( H_a: \) Smoking and alcohol consumption are independent.
(b) The appropriate null hypothesis is \( H_0: \) Smoking and alcohol consumption are not independent.
(c) The calculated value of the test statistic is 3.84.
(d) The calculated value of the test statistic is 7.86.
(e) At level 0.01 we conclude that smoking and alcohol consumption are related.
Questions 6 to 9 refer to the following situation.
In the paper "Color Association of Male and Female Fourth-Grade School Children" (Journal of Psychology, 1988, 383-388), reported on a study in which children were asked to indicate what emotion they associated with the color red. The response and the sex of the child are noted and summarized below. The first number in each cell is the count; the second number is the row percent.

<table>
<thead>
<tr>
<th></th>
<th>Anger</th>
<th>Happy</th>
<th>Love</th>
<th>Pain</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>27</td>
<td>19</td>
<td>39</td>
<td>17</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>26.47</td>
<td>18.63</td>
<td>38.24</td>
<td>16.67</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>34</td>
<td>12</td>
<td>38</td>
<td>28</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>30.36</td>
<td>10.71</td>
<td>33.93</td>
<td>25.00</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>61</td>
<td>31</td>
<td>77</td>
<td>45</td>
<td>214</td>
</tr>
</tbody>
</table>

Statistic
Pearson Chi-Square
DFValueProb
*
4.629*****

6. The null hypothesis is
(a) emotional association with red is independent of gender.
(b) gender is dependent upon the emotional association with red.
(c) the probability of associating a specific emotion with red is related to gender.
(d) male and female children's emotional reaction to the color red are homogenous
(e) the color red is independent of the emotion associated with it and with gender.

7. Under a suitable null hypothesis, the expected frequency for the cell corresponding to Anger and Males is
(a) 15.9.
(b) 55.7.
(c) 30.4.
(d) 31.9.
(e) 29.1.

8. The null hypothesis will be rejected at \( \alpha = 0.05 \) if the test statistic exceeds
(a) 3.84.
(b) 5.99.
(c) 7.81.
(d) 9.49.
(e) 14.07.

9. A Type I error would be committed if
(a) we conclude that the sex of the child and the emotion associated with red are independent when in fact they are not independent.
(b) we conclude that the sex of the child and the emotion associated with red are not independent when in fact they are not independent.
(c) we conclude that the proportion of emotions associated with red differs between males and females when in fact they are the same.
(d) we conclude that the proportion of emotions associated with red is the same for males and females when in fact they are the same.
(e) we fail to find any association between the color red and emotions for either sex.

10. Each person in a random sample of 50 was asked to state his/her sex and preferred color. The resulting frequencies are shown below.

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>Blue</th>
<th>Green</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>5</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>Female</td>
<td>15</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

A chi-square test is used to test the null hypothesis that sex and preferred color are independent. Which of the following statements is a correct decision about the null hypothesis?
(a) Reject at the 0.005 level.
(b) Reject at the 0.01 level but not at the 0.005 level.
(c) Reject at the 0.025 level but not at the 0.01 level.
(d) Reject at the 0.05 level but not at the 0.025 level.
(e) Accept at the 0.05 level.
Part II – Free Response (Question 11-13) – Show your work and explain your results clearly.

11. An experiment in chicken breeding results in offspring having very curly, slightly curly, or normal feathers. If this is the result of a single gene system, then the proportions of offspring in the three phenotypes should be 0.25, 0.50, and 0.25 respectively. In one such experiment, 93 chickens were born. 20 had normal feathers, 50 had slightly curly feathers, and 23 had very curly feathers. Carry out a test to determine whether the genetic model seems to hold in this setting.

12. Describe the differences in the three types of chi-squared tests.
13. Your school administration wants to install new soft drink machines in the gym and cafeteria. Their market analysts want to know if flavor preference depends on the location. A random sample of 180 students was selected and interviewed. Their location and soft drink preference are given in the table.

<table>
<thead>
<tr>
<th>Soft drink</th>
<th>Gym</th>
<th>Cafeteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coca-Cola</td>
<td>33</td>
<td>57</td>
</tr>
<tr>
<td>Pepsi</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>Sprite</td>
<td>5</td>
<td>35</td>
</tr>
</tbody>
</table>

Determine if there is evidence that flavor preference depends on the location.
AP Statistics
4/13/09 Wood/Myers
Test #14 (Chapter 15)

I promise that the only resources that I have used are my textbook and/or my notes. I have not used any
living organisms, which includes electronic communication.

Honor Pledge

Part I - Multiple Choice (Questions 1-10) – Circle the letter of the answer of your choice.

1. Which of (a) through (d) is NOT one of the basic assumptions that must be satisfied in order to perform inference for regression of $y$ on $x$?
   (a) For each value of $x$, the corresponding population of $y$-values is Normally distributed.
   (b) The standard deviation $\sigma$ of the population of $y$-values corresponding to a particular value of $x$ is always the same regardless of the specific value of $x$.
   (c) The sample size (the number of paired observations $(x, y)$ in the sample data) exceeds 30.
   (d) There exists a straight line $y = \alpha + \beta x$ such that, for each value of $x$, the mean $\mu_y$ of the corresponding population of $y$-values lies on that straight line.
   (e) All of (a) through (d) are required assumptions.

2. If the assumptions for regression inference are met, then a Normal probability plot of the residuals should be
   (a) bell-shaped.
   (b) a group of randomly scattered points.
   (c) roughly linear.
   (d) clearly curved.
   (e) "S"-shaped.

3. Inference for regression on the population regression slope $\beta$ is based on which of the following distributions?
   (a) The $t$ distribution with $n - 1$ degrees of freedom
   (b) The standard Normal distribution
   (c) The chi-square distribution with $n - 1$ degrees of freedom
   (d) The $t$ distribution with $n - 2$ degrees of freedom
   (e) The Normal distribution with mean $\mu$ and standard deviation $\sigma$

Questions 4 and 5 refer to the following situation:
One concern about the depletion of the ozone layer is that the increase in ultraviolet light will decrease crop yields. An experiment was conducted in a greenhouse where soybean plants were exposed to varying levels of UV rays—measured in Dobson units. At the end of the experiment the yield (kg) was measured. A regression analysis was performed; here is some output:

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>uv</td>
</tr>
</tbody>
</table>

4. Which of the following is correct?
   (a) If the UV reading is increased by 1 Dobson unit, the yield is expected to increase by 0.0463 kg.
   (b) If the yield increases by 1 kg, the UV reading is expected to decline by 0.0463 Dobson units.
   (c) The estimated yield is 3.98 kg when the UV reading is 0 Dobson units.
   (d) The predicted yield is 4.3 kg when the UV reading is 20 Dobson units.
   (e) The $t$ ratios are used to test if the estimated slope is different from zero.

5. The null and alternative hypotheses for a test of the slope, the test statistic, and the $P$-value are
   (a) $H_0: \beta = 0; H_a: \beta \neq 0; t^* = -4.31; P$-value = 0.0008.
   (b) $H_0: \alpha = 0; H_a: \alpha < 0; t^* = -74.01; P$-value < 0.0001.
   (c) $H_0: \beta = 0; H_a: \beta < 0; t^* = -4.31; P$-value = 0.0004.
   (d) $H_0: \hat{\alpha} = 0; H_a: \hat{\alpha} > 0; t^* = -4.31; P$-value = 0.0004.
   (e) $H_0: \hat{\alpha} = 0; H_a: \hat{\alpha} \neq 0; t^* = -4.31; P$-value = 0.0008.
6. If a test of hypotheses rejects $H_0: \beta = 0$ in favor of the alternative hypothesis $H_a: \beta > 0$, where $\beta$ is the population regression slope, then the least-squares regression line
(a) is useful for predicting $y$, given $x$ (within the limits of $x$-values covered by the data).
(b) slopes downward and to the right when plotted on the scatterplot of paired observations $(x, y)$.
(c) can be extrapolated beyond the limits of the $x$-values covered by the data to predict $y$ at any possible $x$.
(d) is not useful for predicting $y$, given $x$.
(e) has an intercept that is greater than zero.

**The following information is used in Questions 7 and 8.**

A random sample of 80 companies from the Forbes 500 list was selected and the relationship between sales (in hundreds of thousands of dollars) and profits (in hundreds of thousands of dollars) was investigated by regression. A least-squares regression line was fitted to the data using statistical software, with sales as the explanatory variable and profits as the response variable. Here is the output from the software:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>s.e. of Coefficient</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-176.644</td>
<td>61.16</td>
<td>0.0050</td>
</tr>
<tr>
<td>Sales</td>
<td>0.092498</td>
<td>0.0075</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

7. Using the above data, approximately what is a 90% confidence interval for the slope of the least-squares regression line?
(a) $0.0925 \pm 0.0075$
(b) $0.0925 \pm 0.012$
(c) $-0.0925 \pm 0.0075$
(d) $-0.0925 \pm 0.012$
(e) None of the above.

8. Using the above data, is there strong evidence (and if so, why) of a straight-line relationship between sales and profits?
(a) Yes, because the slope of the least-squares line is positive.
(b) Yes, because the P-value for testing if the slope is 0 is quite small.
(c) No, because the value of the square of the correlation is relatively small.
(d) It is impossible to say because we are not given the actual value of the correlation.
(e) None of the above.

**The following information is used in Questions 9 and 10.**

A marine biologist wants to test the effect of water temperature on the average dive duration for sea otters. Several otters are available for an experiment. The biologist collects the following data:

<table>
<thead>
<tr>
<th>Water temp.°C</th>
<th>Dive duration (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Otter</td>
<td>$x$</td>
</tr>
<tr>
<td>J2</td>
<td>4</td>
</tr>
<tr>
<td>J1</td>
<td>8</td>
</tr>
<tr>
<td>B7</td>
<td>8</td>
</tr>
<tr>
<td>B9</td>
<td>12</td>
</tr>
<tr>
<td>M3</td>
<td>12</td>
</tr>
<tr>
<td>D4</td>
<td>16</td>
</tr>
<tr>
<td>B8</td>
<td>20</td>
</tr>
</tbody>
</table>

We want to determine if water temperature is useful in predicting dive duration. Here is output from Minitab for these data:

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>Stdev</th>
<th>t-ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>52.789</td>
<td>5.257</td>
<td>10.04</td>
<td>0.000</td>
</tr>
<tr>
<td>H2Otemp</td>
<td>3.3684</td>
<td>0.4216</td>
<td>***</td>
<td>***</td>
</tr>
</tbody>
</table>

$s = 5.557$  
R-sq = 92.7%  
R-sq(adj) = 91.3%

9. The $t$ statistic for testing $H_0$ has been left out. From the output, the $t$-statistic has the value
(a) 7.99.  
(b) 10.04.  
(c) 0.124.  
(d) 0.927.  
(e) 15.67.

10. The $P$-value is
(a) less than 0.001.  
(b) between 0.001 and 0.01.  
(c) between 0.01 and 0.05.  
(d) between 0.05 and 0.10.  
(e) greater than 0.10.
Part II – Free Response (Question 11-12) – Show your work and explain your results clearly.

11. A teacher asked her 8 introductory statistics students to record the total amount of time they spent studying for a particular test. The amounts of study time \( x \) (in hours) and the resulting test grades \( y \) are given below.

<table>
<thead>
<tr>
<th>( x )</th>
<th>2</th>
<th>1</th>
<th>1.5</th>
<th>0.5</th>
<th>1</th>
<th>3</th>
<th>0</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y )</td>
<td>92</td>
<td>81</td>
<td>84</td>
<td>68</td>
<td>85</td>
<td>96</td>
<td>48</td>
<td>74</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot of the data.

(b) Use your calculator to obtain the equation of the least-squares regression line and the correlation.

\[
y = 60.71 + 12.94x \quad r = 0.81
\]

(c) Explain in words what the slope \( b \) of the least-squares line says about hours studied and grade awarded.

On average, for every additional hour studied, the grade increases by 12.94%.

(d) What is the estimate of \( \beta \) from the data? What is your estimate of the intercept \( \alpha \) of the true regression line?

\[
\hat{\beta} \approx 12.94 \quad \hat{\alpha} \approx 60.71
\]

(e) Use your calculator to calculate the residuals. Report the sum of the residuals and the sum of the squares of the residuals. Then use these results to estimate the standard deviation \( \sigma \) in the regression model. Interpret this value in context.

\[
\sum (y - \hat{y}) = 0 \\
\sum (y - \hat{y})^2 = 560.35
\]

(f) Calculate and interpret \( SE_b \).

\[
\sum (x - \bar{x})^2 = 6.375 \\
SE_b = \frac{9.66}{\sqrt{6.375}} = 3.83
\]

(g) Do we have evidence that the number of hours studied helps predict grade awarded on this statistics test?

\[
H_0: \beta = 0 \\
H_a: \beta > 0 \\
\alpha = 0.05
\]

Conditions
1. true relationship linear – assume
2. observations independent – yes
3. all \( y \)'s, for every \( x \), \( \sim N \) – assume
4. all \( y \)'s, for every \( x \), approx same std. dev. – assume

\[
t = \frac{12.94 - 0}{3.83} = 3.38
\]

\( df = 6 \)

\( p \)-value = 0.007

Since our \( p \)-value is less than our \( \alpha \), we reject the null. We have reason to believe that there is a linear relationship, therefore study hours does help predict test grade.
12. A mathematics professor wishes to analyze the relationship between the number of papers (in hundreds) graded by his department's student homework graders and the total amount of money paid to the graders. He collects data for 12 randomly chosen graders and uses MINITAB to do regression analysis. Below is a portion of the Minitab output. (Here, COST = amount paid (dollars), PAPERS = number of papers in hundreds, and the intervals listed at the bottom are computed for 1600 papers.)

The regression equation is
\[
\text{COST} = 35.8 + 12.1 \text{ PAPERS}
\]

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>Stdev</th>
<th>t-ratio</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>35.80</td>
<td>17.06</td>
<td>2.10</td>
<td>0.062</td>
</tr>
<tr>
<td>PAPERS</td>
<td>12.0835</td>
<td>0.9738</td>
<td>12.41</td>
<td>0.000</td>
</tr>
</tbody>
</table>

\[s = 6.526 \quad \text{R-sq} = 93.9\% \quad \text{R-sq (adj)} = 93.3\%

(a) What is the least-squares regression equation?

\[
\text{COST} = 35.8 + 12.1 \text{ PAPERS}
\]

(c) What is the standard deviation \( s \) in the regression model? Interpret this value in context.

\[s = 6.526 \quad \text{The average distance from actual cost and predicted cost is \$6.526.}
\]

(d) Interpret the slope of the least-squares regression line in the context of this problem.

\[\text{as the number of papers increases by 100, on average the graders get paid \$12.10 more.}
\]

(e) The model for regression inference has three parameters: \( \alpha \), \( \beta \), and \( \sigma \). Estimate these parameters from the data.

\[\alpha \approx 35.8 \quad \beta \approx 12.0835 \quad \sigma \approx 6.526
\]

(f) Calculate and interpret a 95% confidence interval on the slope of the least-squares regression line

Conditions
1. true relationship is linear - assume
2. independent observations - yes
3. for every \( x \), \( y \sim N \) - assume
4. for every \( x \), \( y \)'s have same sd - assume

\[t^* = 2.23 \quad df = 10
\]

\[12.0835 \pm 2.23 (0.9738)
\]

\[(9.91, 14.25)
\]

We are 95% confident that the true slope of the regression line between number of papers and cost is

\[9.91 \text{ to } 14.25\]
Part I - Multiple Choice (Questions 1-10) – Circle the letter of the answer of your choice.

1. Which of (a) through (d) is NOT one of the basic assumptions that must be satisfied in order to perform inference for regression of $y$ on $x$?
   (a) For each value of $x$, the corresponding population of $y$-values is Normally distributed.
   (b) The standard deviation $\sigma$ of the population of $y$-values corresponding to a particular value of $x$ is always the same regardless of the specific value of $x$.
   (c) The sample size (the number of paired observations $(x, y)$ in the sample data) exceeds 30.
   (d) There exists a straight line $y = \alpha + \beta x$ such that, for each value of $x$, the mean $\mu_y$ of the corresponding population of $y$-values lies on that straight line.
   (e) All of (a) through (d) are required assumptions.

2. If the assumptions for regression inference are met, then a Normal probability plot of the residuals should be
   (a) bell-shaped.
   (b) a group of randomly scattered points.
   (c) roughly linear.
   (d) clearly curved.
   (e) "S"-shaped.

3. Inference for regression on the population regression slope $\beta$ is based on which of the following distributions?
   (a) The $t$ distribution with $n - 1$ degrees of freedom
   (b) The standard Normal distribution
   (c) The chi-square distribution with $n - 1$ degrees of freedom
   (d) The $t$ distribution with $n - 2$ degrees of freedom
   (e) The Normal distribution with mean $\mu$ and standard deviation $\sigma$

Questions 4 and 5 refer to the following situation:
One concern about the depletion of the ozone layer is that the increase in ultraviolet light will decrease crop yields. An experiment was conducted in a greenhouse where soybean plants were exposed to varying levels of UV rays—measured in Dobson units. At the end of the experiment the yield (kg) was measured. A regression analysis was performed; here is some output:

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>$uv$</td>
</tr>
</tbody>
</table>

4. Which of the following is correct?
   (a) If the UV reading is increased by 1 Dobson unit, the yield is expected to increase by 0.0463 kg.
   (b) If the yield increases by 1 kg, the UV reading is expected to decline by 0.0463 Dobson units.
   (c) The estimated yield is 3.98 kg when the UV reading is 0 Dobson units.
   (d) The predicted yield is 4.3 kg when the UV reading is 20 Dobson units.
   (e) The $t$ ratios are used to test if the estimated slope is different from zero.

5. The null and alternative hypotheses for a test of the slope, the test statistic, and the $P$-value are
   (a) $H_0: \beta = 0; H_a: \beta \neq 0; t* = -4.31; P$-value = 0.0008.
   (b) $H_0: \alpha = 0; H_a: \alpha < 0; t* = -74.01; P$-value < 0.0001.
   (c) $H_0: \beta = 0; H_a: \beta < 0; t* = -4.31; P$-value = 0.0004.
   (d) $H_0: \alpha = 0; H_a: \alpha > 0; t* = -4.31; P$-value = 0.0004.
   (e) $H_0: \alpha = 0; H_a: \alpha \neq 0; t* = -4.31; P$-value = 0.0008.
6. If a test of hypotheses rejects $H_0: \beta = 0$ in favor of the alternative hypothesis $H_1: \beta > 0$, where $\beta$ is the population regression slope, then the least-squares regression line
   (a) is useful for predicting $y$, given $x$ (within the limits of $x$-values covered by the data).
   (b) slopes downward and to the right when plotted on the scatterplot of paired observations $(x, y)$.
   (c) can be extrapolated beyond the limits of the $x$-values covered by the data to predict $y$ at any possible $x$.
   (d) is not useful for predicting $y$, given $x$.
   (e) has an intercept that is greater than zero.

The following information is used in Questions 7 and 8.

A random sample of 80 companies from the Forbes 500 list was selected and the relationship between sales (in hundreds of thousands of dollars) and profits (in hundreds of thousands of dollars) was investigated by regression. A least-squares regression line was fitted to the data using statistical software, with sales as the explanatory variable and profits as the response variable. Here is the output from the software:

```
Dependent variable is Profits
R squares = 66.2%
s = 466.2 with 80 - 2 = 78 degrees of freedom

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>s.e. of Coefficient</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-176.644</td>
<td>61.15</td>
<td>0.0050</td>
</tr>
<tr>
<td>Sales</td>
<td>0.092498</td>
<td>0.0075</td>
<td>*0.0001</td>
</tr>
</tbody>
</table>
```

7. Using the above data, approximately what is a 90% confidence interval for the slope of the least-squares regression line?
   (a) 0.0925 ± 0.0075
   (b) 0.0925 ± 0.012
   (c) -0.0925 ± 0.0075
   (d) -0.0925 ± 0.012
   (e) None of the above.

8. Using the above data, is there strong evidence (and if so, why) of a straight-line relationship between sales and profits?
   (a) Yes, because the slope of the least-squares line is positive.
   (b) Yes, because the P-value for testing if the slope is 0 is quite small.
   (c) No, because the value of the square of the correlation is relatively small.
   (d) It is impossible to say because we are not given the actual value of the correlation.
   (e) None of the above.

The following information is used in Questions 9 and 10.

A marine biologist wants to test the effect of water temperature on the average dive duration for sea otters. Several otters are available for an experiment. The biologist collects the following data:

<table>
<thead>
<tr>
<th>Otter</th>
<th>Water temp.°(C)</th>
<th>Dive duration (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J2</td>
<td>4</td>
<td>63</td>
</tr>
<tr>
<td>J1</td>
<td>8</td>
<td>75</td>
</tr>
<tr>
<td>B7</td>
<td>8</td>
<td>84</td>
</tr>
<tr>
<td>B9</td>
<td>12</td>
<td>91</td>
</tr>
<tr>
<td>M3</td>
<td>12</td>
<td>101</td>
</tr>
<tr>
<td>D4</td>
<td>16</td>
<td>110</td>
</tr>
<tr>
<td>B8</td>
<td>20</td>
<td>115</td>
</tr>
</tbody>
</table>

We want to determine if water temperature is useful in predicting dive duration. Here is output from Minitab for these data:

```
R-sq = 92.7% R-sq(adj) = 91.3%

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>Stdev</th>
<th>t-ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>52.789</td>
<td>5.257</td>
<td>10.04</td>
<td>0.000</td>
</tr>
<tr>
<td>H2Otemp</td>
<td>3.3684</td>
<td>0.4216</td>
<td>***</td>
<td>***</td>
</tr>
</tbody>
</table>

s = 5.557    R-sq = 92.7%    R-sq(adj) = 91.3%
```

9. The $t$ statistic for testing $H_0$ has been left out. From the output, the $t$-statistic has the value
   (a) 7.99.  (b) 10.04.  (c) 0.124.  (d) 0.927.  (e) 15.67.

10. The $P$-value is
    (a) less than 0.001.
    (b) between 0.001 and 0.01.
    (c) between 0.01 and 0.05.
    (d) between 0.05 and 0.10.
    (e) greater than 0.10.
11. A teacher asked her 8 introductory statistics students to record the total amount of time they spent studying for a particular test. The amounts of study time $x$ (in hours) and the resulting test grades $y$ are given below.

<table>
<thead>
<tr>
<th>$x$</th>
<th>2</th>
<th>1</th>
<th>1.5</th>
<th>0.5</th>
<th>1</th>
<th>3</th>
<th>0</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>92</td>
<td>81</td>
<td>84</td>
<td>68</td>
<td>85</td>
<td>96</td>
<td>48</td>
<td>74</td>
</tr>
</tbody>
</table>

(a) Make a scatterplot of the data.

(b) Use your calculator to obtain the equation of the least-squares regression line and the correlation.

(c) Explain in words what the slope $b$ of the least-squares line says about hours studied and grade awarded.

(d) What is the estimate of $\beta$ from the data? What is your estimate of the intercept $\alpha$ of the true regression line?

(e) Use your calculator to calculate the residuals. Report the sum of the residuals and the sum of the squares of the residuals. Then use these results to estimate the standard deviation $\sigma$ in the regression model. Interpret this value in context.

(f) Calculate and interpret $SE_b$.

(g) Do we have evidence that the number of hours studied helps predict grade awarded on this statistics test?
12. A mathematics professor wishes to analyze the relationship between the number of papers (in hundreds) graded by his department's student homework graders and the total amount of money paid to the graders. He collects data for 12 randomly chosen graders and uses MINITAB to do regression analysis. Below is a portion of the Minitab output. (Here, COST = amount paid (dollars), PAPERS = number of papers in hundreds, and the intervals listed at the bottom are computed for 1600 papers.)

The regression equation is
COST = 35.8 + 12.1 PAPERS

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>Stdev</th>
<th>t-ratio</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>35.8</td>
<td>17.06</td>
<td>2.10</td>
<td>0.062</td>
</tr>
<tr>
<td>PAPERS</td>
<td>12.0835</td>
<td>0.9738</td>
<td>12.41</td>
<td>0.000</td>
</tr>
</tbody>
</table>

s = 6.526 R-sq = 93.9% R-sq (adj) = 93.3%

(a) What is the least-squares regression equation?

(c) What is the standard deviation $s$ in the regression model? Interpret this value in context.

(d) Interpret the slope of the least-squares regression line in the context of this problem.

(e) The model for regression inference has three parameters: $\alpha$, $\beta$, and $\sigma$. Estimate these parameters from the data.

(f) Calculate and interpret a 95% confidence interval on the slope of the least-squares regression line.