Identify the null hypothesis, alternative hypothesis, test statistic, P-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim.

1) A supplier of 3.5” disks claims that no more than 1% of the disks are defective. In a random sample of 600 disks, it is found that 3% are defective, but the supplier claims that this is only a sample fluctuation. At the 0.01 level of significance, test the supplier’s claim that no more than 1% are defective.

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

Find the P-value for the indicated hypothesis test.

2) In a sample of 47 adults selected randomly from one town, it is found that 9 of them have been exposed to a particular strain of the flu. Find the P-value for a test of the claim that the proportion of all adults in the town that have been exposed to this strain of the flu is 8%.
   A) 0.0024    B) 0.0524    C) 0.0048    D) 0.0262

Identify the null hypothesis, alternative hypothesis, test statistic, P-value, conclusion about the null hypothesis, and final conclusion that addresses the original claim.

3) A random sample of 100 pumpkins is obtained and the mean circumference is found to be 40.5 cm. Assuming that the population standard deviation is known to be 1.6 cm, use a 0.05 significance level to test the claim that the mean circumference of all pumpkins is equal to 39.9 cm.

Test the given claim using the traditional method of hypothesis testing. Assume that the sample has been randomly selected from a population with a normal distribution.

4) A public bus company official claims that the mean waiting time for bus number 14 during peak hours is less than 10 minutes. Karen took bus number 14 during peak hours on 18 different occasions. Her mean waiting time was 7 minutes with a standard deviation of 2.3 minutes. At the 0.01 significance level, test the claim that the mean is less than 10 minutes.

Use the traditional method to test the given hypothesis. Assume that the population is normally distributed and that the sample has been randomly selected.

5) A manufacturer uses a new production method to produce steel rods. A random sample of 17 steel rods resulted in lengths with a standard deviation of 2.4 cm. At the 0.10 significance level, test the claim that the new production method has lengths with a standard deviation different from 3.5 cm, which was the standard deviation for the old method.
MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Solve the problem.

6) For large numbers of degrees of freedom, the critical $\chi^2$ values can be approximated as follows:

$$\chi^2 = \frac{1}{2} (z + \sqrt{2k - 1})^2,$$

where $k$ is the number of degrees of freedom and $z$ is the critical value. To find the lower critical value, the negative $z$-value is used, to find the upper critical value, the positive $z$-value is used.

Use this approximation to estimate the critical value of $\chi^2$ in a right-tailed hypothesis test with $n = 125$ and $\alpha = 0.01$.

A) $\chi^2 \approx 162.833$    B) $\chi^2 \approx 167.285$    C) $\chi^2 \approx 168.448$    D) $\chi^2 \approx 163.981$

Construct a scatter diagram for the given data.

7) Construct a scatter diagram for the given data.

<table>
<thead>
<tr>
<th>$x$</th>
<th>0.17</th>
<th>0.07</th>
<th>0.16</th>
<th>0.34</th>
<th>-0.16</th>
<th>0.3</th>
<th>0.51</th>
<th>0.08</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>0.5</td>
<td>0.85</td>
<td>0.35</td>
<td>0.46</td>
<td>0.01</td>
<td>0.53</td>
<td>0.22</td>
<td>-0.03</td>
</tr>
</tbody>
</table>

A) ![Scatter Diagram A]  B) ![Scatter Diagram B]
Find the value of the linear correlation coefficient r.

8) \( x \) 62 53 64 52 52 54 58
   \( y \) 158 176 151 164 164 174 162
   A) 0.754  B) -0.775  C) -0.081  D) 0

Use the given data to find the equation of the regression line. Round the final values to three significant digits, if necessary.

9) \( x \) 6 8 20 28 36
   \( y \) 2 4 13 20 30
   A) \( y = -3.79 + 0.801x \)  B) \( y = -2.79 + 0.897x \)
   C) \( y = -3.79 + 0.897x \)  D) \( y = -2.79 + 0.950x \)

Solve the problem.

10) For the data below, determine the value of the linear correlation coefficient r between y and \( x^2 \). Test whether the correlation is significant. Use a significance level of 0.01.

   \( x \) 1.2 2.7 4.4 6.6 9.5
   \( y \) 1.6 4.7 9.9 24.5 39.0
   A) 0.873  B) 0.985  C) 0.990  D) 0.913

Find the total variation for the paired data.

11) The equation of the regression line for the paired data below is \( y = 6.18286 + 4.33937x \). Find the explained variation, the unexplained variation, the total variation, and the coefficient of determination.

   \( x \) 9 7 2 3 4 22 17
   \( y \) 43 35 16 21 23 102 81
   A) 6,531.37  B) 13.479  C) 6,693.27  D) 6,544.86

Construct the indicated prediction interval for an individual \( y \).

12) The equation of the regression line for the paired data below is . Find the 99% prediction interval of \( y \) for \( x = 12 \).

   \( x \) 9 7 2 3 4 22 17
   \( y \) 43 35 16 21 23 102 81
   A) 53.6 < \( y < 71.3 \)  B) 51.1 < \( y < 65.4 \)  C) 56.4 < \( y < 68.5 \)  D) 59.9 < \( y < 64.9 \)
Solutions

Problem 1
Let $p$ be the population proportion of defective disks.

1. The null hypothesis is $H_0 : p = 0.01$.
2. The alternative hypothesis is $H_1 : p > 0.01$.
3. The test is one-tailed.
4. The test statistics (page 408) is
   
   
   $$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \approx \frac{0.03 - 0.01}{\sqrt{0.01 \cdot 0.99/600}} = 4.92.$$

5. From Table A-2 we see that the critical value for the one-tailed test at significance level 0.01 is 2.33.
6. Because the critical value is smaller than the test statistics we reject the null-hypothesis.
7. From Table A-2 we see that the P-value is smaller than 0.0001 (Indeed, it can be computed that the P-value is about to 0.0000004 or four over ten million).
8. We conclude that the claim of the disks’ supplier is not supported by the sample evidence and moreover is extremely unlikely to be valid.

Problem 2
1. The null-hypothesis is $H_0 = 0.08$.
2. The alternative hypothesis is $H_1 : p \neq 0.08$
3. The test is two-tailed.
4. The test statistics is
   
   $$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \approx \frac{9/47 - 0.08}{\sqrt{0.08 \cdot 0.92/47}} = 2.82.$$

5. From Table A-2 we find that P-value is $0.0024 \times 2 = 0.0048$. The correct answer is “C”.

Problem 3
1. The null-hypothesis is $H_0 : \mu = 39.9$.
2. The alternative hypothesis is $H_1 : \mu \neq 39.9$.
3. The test is two-tailed.
4. The test statistics is (page 419)
   
   $$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{40.5 - 39.9}{1.6/\sqrt{100}} = 3.75.$$

5. From Table A-2 we see that the critical value for two-tailed test at significance level 0.05 is 1.96.
6. Because the critical value is smaller than the test statistics we reject the null hypothesis.
7. From Table A-2 we see that the P-value is smaller than 0.0002 (Indeed, the P-value is approximately 0.000177).
8. The claim is not supported by the sample evidence and it is very unlikely that the claim is valid.

Problem 4
1. The null hypothesis is \( H_0 : \mu = 10 \).
2. The alternative hypothesis is \( H_1 : \mu > 10 \).
3. The test is one-tailed.
4. The test statistics is (page 426)
   \[
   t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{7 - 10}{2.3/\sqrt{14}} \approx -0.3486.
   \]
5. Because the population distribution is unknown we use the Student distribution with 13 degrees of freedom. From Table A-3 we see that the critical value for one-tailed test at the 0.01 significance level is 2.650.
6. We have no evidence to reject the null-hypothesis.
7. The sample evidence supports the claim.

Problem 5
1. The null-hypothesis is \( H_0 : \sigma = 3.5 \).
2. The alternative hypothesis is \( H_1 : \sigma \neq 3.5 \).
3. The test is two-tailed.
4. The test statistics is (page 436)
   \[
   \chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{16 \cdot 2.4^2}{3.5^2} \approx 7.523.
   \]
5. From Table A-4 we find that the critical value for 16 degrees of freedom and the area in the left tail 0.05 (one half of the significance level) is 7.962.
6. Because the test statistics is smaller than the critical value we reject the null-hypothesis.
7. The sample evidence supports the claim at the significance level 0.10.

Problem 6
Because the test is right-tailed we use the positive value \( z_{0.01} = 2.33 \). Number of degrees of freedom is \( k = n - 1 = 124 \). Plugging in these numbers into the formula given in the problem we get
\[
\chi^2 = \frac{1}{2} (2.33 + \sqrt{247})^2 \approx 162.833.
\]
The correct answer is “A”.
Problem 7

Above we see the scatterplot constructed by Statdisk.
The closest answer is “C”.

Problem 8
We use the formula 10-1 on page 520.

\[ r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}} \]

\[ = \frac{7 \cdot 64636 - 395 \cdot 1149}{\sqrt{7 \cdot 22437 - 395^2} \sqrt{7 \cdot 189053 - 1149^2}} \approx -0.775. \]

The correct answer is “B”.

Problem 9
We use formula 10-2 on page 542 to compute the slope of the regression line.

\[ b_1 = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} = \frac{5 \cdot 1944 - 98 \cdot 69}{5 \cdot 2580 - 98^2} \approx 0.897 \]

To find the y-intercept of the regression line we use formula 10-3.

\[ b_0 = \bar{y} - b_1 \bar{x} = \frac{69 - 0.897 \cdot 98}{5} \approx -3.78 \]

The correct answer is “C”.
**Problem 10**

First we compute the correlation coefficient. Because we use values $x^2$ instead of $x$ formula 10-1 takes the form

$$
 r = \frac{n \sum x^2 y - (\sum x^2)(\sum y)}{\sqrt{n(\sum x^2)^2 - (\sum x^2)^2} \sqrt{n(\sum y^2)^2 - (\sum y)^2}} = \frac{5 \cdot 4815.201 - 161.9 \cdot 79.7}{\sqrt{5 \cdot 10472.5634 - 161.9^2 \cdot 79.7^2}} \approx 0.990.
$$

The correct answer to the first part of the problem is “C”.

Next we follow the procedure described on page 527.

The null-hypothesis $H_0: \rho = 0$ (There is no linear correlation).

The alternative hypothesis $H_1: \rho \neq 0$ (There is linear correlation).

The test statistics

$$
 t = \frac{r}{\sqrt{1 - r^2)/(n-2)} = \frac{0.990}{\sqrt{0.0199/3}} \approx 12.1554.
$$

From table A-3 we see that the critical value for two-tailed test at a significance level 0.01 is 5.841. Because the test statistics is larger than the critical value we reject the null-hypothesis. There is a significant correlation between values of $x^2$ and $y$.

**Problem 11**

We will use formula 10-4 on page 559. First we need to compute the predicted values according to the formula

$$
 \hat{y} = 6.18286 + 4.33937x
$$

The results of computations are presented in the table below.

<table>
<thead>
<tr>
<th>$x$</th>
<th>9</th>
<th>7</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>22</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>43</td>
<td>35</td>
<td>16</td>
<td>21</td>
<td>23</td>
<td>102</td>
<td>81</td>
</tr>
<tr>
<td>$\hat{y}$</td>
<td>45.23719</td>
<td>36.55845</td>
<td>14.8616</td>
<td>19.20097</td>
<td>23.54034</td>
<td>101.649</td>
<td>79.95215</td>
</tr>
</tbody>
</table>

The mean of $y$-values is $\bar{y} = 45.85714$.

The explained variation is $\sum (\hat{y} - \bar{y})^2 \approx 6531.36578$.

The unexplained variation is $\sum (y - \hat{y})^2 \approx 13.47940$.

The total variation is $6531.36578 + 13.47940 = 6544.84519$. The closest answer is “D”.

The coefficient of determination is

$$
 r^2 = \frac{\text{explained variation}}{\text{total variation}} = \frac{6531.36578}{6544.84519} \approx 0.99794.
$$
Problem 12
We use the formulas for prediction interval on page 561.
\[
\hat{y} - E < y < \hat{y} + E,
\]
\[
E = t_{\alpha \frac{2}{2}} s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \overline{x})^2}{n(\sum x^2) - (\sum x)^2}}, \quad \text{and}
\]
\[
s_e = \sqrt{\frac{\sum (y - \hat{y})^2}{n - 2}}.
\]
First we compute the standard error of estimate. From the previous problem we know that the unexplained variation \(\sum (y - \hat{y})^2\) is 13.47940, whence
\[
s_e = \sqrt{\frac{13.47940}{5}} \approx 1.64191
\]
Next from Table A-3 keeping in mind that the number of degrees of freedom is \(n - 2 = 5\) we find that \(t_{0.005} = 4.032\).
Finally
\[
x_0 = 12, \quad \overline{x} = 9.14286, \quad \sum x = 64, \quad \sum x^2 = 932, \quad \text{and} \quad \hat{y} = 6.1829 + 4.3394 \cdot 12 \approx 58.3
\]
Plugging in these numbers we get that the margin of error is
\[
E = 4.032 \cdot 1.64191 \cdot \sqrt{1 + \frac{1}{7} + \frac{7(12 - 9.14286)^2}{7 \cdot 932 - 64^2}} \approx 7.1
\]
The prediction interval is \(51.1 < y < 65.4\). The correct answer is “B”.