1. The shares of the U.S. automobile market held in 1990 by General Motors, Japanese manufacturers, Ford, Chrysler, and other manufacturers were respectively 36%, 26%, 21%, 9%, and 8%. Suppose that a new survey of 1,000 new-car buyers shows the following purchase frequencies: GM(391), Japanese(202), Ford(275), Chrysler(53), and Other(79). Test at 10% significance level to determine whether the current market shares differ from those of 1990. What type of error you might have committed in your decision?

The hypotheses are: $H_0$: The current market shares is as 1990

$H_a$: The current market shares is different from those 1990.

(1 point)

The assumption is: Each expected frequency ($e_i$) is at least 5 $e_i \geq 5$ for all $i = 1, 2, ..., k$.

(1 point)

The test statistic:

<table>
<thead>
<tr>
<th>Class(i)</th>
<th>$o_i$</th>
<th>$e_i$</th>
<th>$e_i-o_i$</th>
<th>$(o_i-e_i)^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM</td>
<td>391</td>
<td>360</td>
<td>31</td>
<td>2.669</td>
</tr>
<tr>
<td>Japanese</td>
<td>202</td>
<td>260</td>
<td>-58</td>
<td>12.938</td>
</tr>
<tr>
<td>Ford</td>
<td>275</td>
<td>210</td>
<td>65</td>
<td>20.119</td>
</tr>
<tr>
<td>Chrysler</td>
<td>53</td>
<td>90</td>
<td>-37</td>
<td>15.211</td>
</tr>
<tr>
<td>Other</td>
<td>79</td>
<td>80</td>
<td>-1</td>
<td>0.0125</td>
</tr>
<tr>
<td>Total</td>
<td>1000</td>
<td>1000</td>
<td>0</td>
<td>50.950</td>
</tr>
</tbody>
</table>

(2 points)

The critical value: $\chi^2_{0.05, 5} = 7.7794$ (1 point)

The decision rule: $H_0$ is rejected if $\chi^2 > \chi^2_{0.05, 5}$

$\Rightarrow 50.950 > 7.7794$

$\Rightarrow$ Reject $H_0$.

(1 point)

The conclusion: The current market shares differ from those 1990.

(1 point)

Type of error: $H_0$ was rejected $\Rightarrow$ Type I error.

(1 point)
2. A book marketing research study about the relationship between delivery time and computer-assisted ordering was conducted. A sample of 40 firms shows that 16 use computer-assisted ordering, while 24 do not. Furthermore, past data are used to categorize each firm’s delivery times as below the industry average, equal to the industry average, or above the industry average as given in the table below:

<table>
<thead>
<tr>
<th>Delivery time</th>
<th>Below average</th>
<th>Equal to average</th>
<th>Above average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer Ordering</td>
<td>4</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>No</td>
<td>4</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>Yes</td>
<td>10</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>16</td>
<td>10</td>
</tr>
</tbody>
</table>

Using the above table what do you conclude about the relationship between delivery time and computer-assisted ordering? Use 5% significance level.

The hypotheses are: $H_0$: The delivery time and computer-assisted ordering are independent. $H_a$: The delivery time and computer-assisted ordering are not independent.

The assumption is: The expected frequency for each cell is at least 5 ($e_{ij} = n \cdot p_{ij} \geq 5$ for all $i = 1, ..., k$).

The test statistic:

$$
X^2 = \sum_{i=1}^{k} \sum_{j=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}
$$

Combine last column with the previous ones to get:

<table>
<thead>
<tr>
<th>Below average</th>
<th>Equal or above average</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>4</td>
</tr>
<tr>
<td>Yes</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
</tr>
</tbody>
</table>

$$
X^2 = \frac{24^2}{8.4} + \frac{20^2}{15.6} + \frac{6^2}{5.6} + \frac{4^2}{10.4} = 2.3048 + 1.2410 + 3.6711 + 1.8615 = 8.8415
$$

The critical value:

$$
X^2_{k,(r-1)(c-1)} = X^2_{0.05,(2-1)(2-1)} = X^2_{0.05,1} = 3.8415
$$

The decision rule:

Reject $H_0$ if $X^2 > X^2_{k,(r-1)(c-1)}$

$\Rightarrow 8.8415 > 3.8415 \Rightarrow$ Reject $H_0$

The conclusion:

The delivery time and computer-assisted ordering are not independent (Dependent).

Without combining:

$$
X^2 = \sum_{i=1}^{k} \sum_{j=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(4-8.4)^2}{8.4} + \frac{(12-9.6)^2}{9.6} + \frac{(8-6)^2}{6} + \frac{(10-5.6)^2}{5.6} + \frac{(4-10.4)^2}{10.4}
$$

$$
= 2.3048 + 0.6000 + 0.6667 + 3.4571 + 0.9000 + 1.0000
$$

$$
= 8.9286
$$

$$
X^2_{k,(r-1)(c-1)} = X^2_{0.05,(2-1)(2-1)} = X^2_{0.05,1} = 5.9915
$$

Reject $H_0$ if $X^2 > X^2_{k,(r-1)(c-1)} = 8.9286 > 5.9915 \Rightarrow$ Reject $H_0$
3. Accu-Copiers, Inc., sells and services the Accu-500 copying machine. As part of its standard service contract, the company agrees to perform routine service on this copier. To obtain information about the time it takes to perform routine service, Accu-Copiers has collected data for 11 service calls. The data are as follows:

<table>
<thead>
<tr>
<th>Copiers serviced (X)</th>
<th>4</th>
<th>2</th>
<th>5</th>
<th>7</th>
<th>1</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>2</th>
<th>4</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minutes required (Y)</td>
<td>140</td>
<td>68</td>
<td>103</td>
<td>145</td>
<td>60</td>
<td>51</td>
<td>103</td>
<td>134</td>
<td>110</td>
<td>90</td>
<td>112</td>
</tr>
</tbody>
</table>

Do the data provide sufficient evidence to conclude that there is a direct relationship between the number of copiers serviced and the time it takes to be serviced? Use a significance level of 0.025.

\[
\text{The correlation coefficient } r = \frac{\sum xy - \left( \frac{\sum x \sum y}{n} \right)}{\sqrt{\left( \sum x^2 - \left( \frac{\sum x^2}{n} \right)^2 \right) \left( \sum y^2 - \left( \frac{\sum y^2}{n} \right)^2 \right)}} = \frac{4773 - \left( \frac{(43)(116)}{11} \right)}{\sqrt{\left( 12368 - \left( \frac{(43)^2}{11} \right) \right) \left( 116^2 \right)}}
\]

\[
= 0.7103 \quad \text{(1 point)}
\]

The assumptions are:

a. The data are interval or ratio level of measurement.

b. The two variables are distributed as a bivariate normal distribution.

The hypotheses are: $H_0$: $\rho = 0$, $H_A$: $\rho \neq 0$ (1 point)

The test statistic value:

\[
\hat{r} = \frac{r}{\sqrt{1-r^2}} = \frac{0.7103}{\sqrt{1-0.5072}}
\]

\[
= 3.027 \quad \text{(1 point)}
\]

The critical value:

\[
\tau_{\alpha, n-2} = 2.8214 \quad \text{(1 point)}
\]

Decision Rule:

Reject $H_0$ if $|\hat{r}| > \tau_{\alpha, n-2} \Rightarrow 3.027 > 2.8214$

\[
\text{Reject } H_0 \quad \text{(1 point)}
\]

Conclusion: There is a significant linear (positive) relationship between the two variables. (1 point)
4. Enterprise Industries produces FRESH, a brand of liquid laundry detergent. In order to study the relationship between the price and demand for FRESH, the company has gathered data concerning demand for FRESH over the last 30 sales periods where,

X: The price (in dollars) per bottle of FRESH and Y: The demand for FRESH (in 100,000s of bottles)

The following sums were obtained,

\[ n = 30, \sum x = 112.05, \sum x^2 = 418.742, \sum y = 251.48, \sum y^2 = 2121.53, \sum xy = 938.442, \text{ and } SSE = 10.495 \]

Assuming that X is the independent variable and Y is the dependent variable then

1. The assumptions are:
   a. Error values are independent.
   b. Error values are normally distributed.
   c. Error variance is constant.
   d. The relationship between X and Y is linear.

2. The fitted regression equation is:
   \[ \hat{Y} = 21.6524 - 3.5528X \]

   \[ b_1 = \frac{\sum xy - (\frac{\sum x \cdot \sum y}{n})}{\sum x^2 - (\frac{(\sum x)^2}{n})} = \frac{938.442 - (\frac{112.05 \cdot 251.48}{30})}{418.742 - (\frac{(112.05)^2}{30})} = \frac{-0.9358}{0.23525} = -3.5528 \]

   \[ b_0 = \bar{y} - b_1 \bar{x} = (\frac{251.48}{30}) - (-3.5528)(\frac{112.05}{30}) = 21.6524 \]

3. The standard error of the regression model is:
   \[ s_c = \sqrt{\frac{MSE}{n-k-1}} = \sqrt{\frac{10.495}{30-1-1}} = 0.6122 \]

4. The predicted value of the demand if the price was $4.00 is:
   \[ \hat{Y}_4 = 21.6524 - 3.5528(4) = 7.4412 \]

5. A 99% C.I. for the demand if the price of a bottle was $4.00 is:
   \[ x_p = 4.00, \bar{x} = \frac{112.05}{30} = 3.7350, \frac{\sum (x - \bar{x})^2}{n} = 418.742 - (\frac{(112.05)^2}{30}) = 2.7633 \]

   \[ L_{x_p, n-2} = t_{0.005}, 28 = 2.052, 28 = 2.7633 \]

   A 99% C.I. is:
   \[ \hat{Y} \pm t_{0.005} \cdot s_c \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum (x - \bar{x})^2}} \]

   \[ \Rightarrow 7.4412 \pm 1.959523 \Rightarrow [5.4889, 9.3935] \]
Do you think that the demand will increase by at most 300,000 bottles if the price was decreased by $1? Justify your answer using 10% significance level.

<table>
<thead>
<tr>
<th>The hypotheses are: $H_0$: $\beta_1 \leq 3$</th>
<th>$H_A$: $\beta_1 &gt; 3$</th>
<th>1 point</th>
</tr>
</thead>
<tbody>
<tr>
<td>The test statistic: $t_c = \frac{b_1 - \beta_{10}}{S_{b_1}} = \frac{-3.5528 - 3}{0.6122/\sqrt{0.2533}}$</td>
<td>$S_{b_1} = \frac{S_e}{\sqrt{(x-x)}^2}$</td>
<td>1 point</td>
</tr>
<tr>
<td></td>
<td>$= \frac{-5.1921}{1.2621}$</td>
<td>1 point</td>
</tr>
<tr>
<td>The critical value: $t_{\alpha, n-2} = t_{0.10, 28} = 1.3125$</td>
<td>0 point</td>
<td></td>
</tr>
<tr>
<td>The decision rule: $t_c &gt; t_{\alpha, n-2}$ ⇒ $-5.1921 &gt; 1.3125$</td>
<td>$\Rightarrow$ Do not reject $H_0$</td>
<td>1 point</td>
</tr>
<tr>
<td>The conclusion: The demand will increase by at most 300,000 bottles</td>
<td>1 point</td>
<td></td>
</tr>
</tbody>
</table>

With Our Best Wishes