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 in 1990 ① $H_A$: The current market shares is DIFFERENT from those of 1990

The assumption is: Each expected frequency is at least 5. ($e_i \geq 5 \ \forall i = 1,2,3,4,5$) ①

The test statistic: ②

| class(i) | $O_i$ | $p_i$ | $e_i$ | $(O_i-e_i)^2/e_i$ |
|---|---|---|---|---|
| GM | 391 | 0.36 | 360 | 2.669 |
| Japan | 202 | 0.26 | 260 | 12.938 |
| Ford | 275 | 0.21 | 210 | 20.119 |
| Chrys. | 53 | 0.09 | 90 | 15.211 |
| Other | 79 | 0.08 | 80 | 0.0125 |
| Total | 1000 | 1 | 1000 | 50.951 |

$$\chi_0^2 = \sum_1^5 \frac{(O_i-e_i)^2}{e_i} = \boxed{50.951} = \chi^2_{cal}$$ ②

The critical value: $\chi^2_{k-1,0.1} = \chi^2_{4,0.1} = 7.7794$ ①

The decision rule: If $\chi_0^2 > \chi_\alpha^2 \Rightarrow$ Reject $H_0$ ①

Since $\chi_0^2 = 50.951 > 7.7794 = \chi_\alpha^2 \Rightarrow$ Reject $H_0$

The conclusion: The current shares DIFFER from those of 1990. ①

Type of error: Type-I error ①

If no Combining = -3



**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:

| | Delivery time | | | |
|---|---|---|---|---|
| **Computer Ordering** | Below average | Equal to average | Above average | |
| No | 4 $\boxed{8.4}$ | ~~12~~ 20 $\boxed{15.6}$ | ~~8~~ 6 | 24 |
| Yes | 10 $\boxed{5.6}$ | ~~4~~ 6 $\boxed{10.4}$ | ~~2~~ 4 | 16 |
| | 14 | 26 | 10 | $\boxed{40}$ |

(2)

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: $H_o$: The delivery time is INDEP. from Computer ordering $H_A$: They are NOT independent. (1)

The assumption is: The expected frequency of each cell is at least 5 ($e_{ij} \geq 5 \ \forall i,j = 1,2$) (1)

The test statistic value:

$$\chi^2_0 = \chi^2_{cal} = \sum_1^2 \sum_1^2 \frac{(O_{ij}-e_{ij})^2}{e_{ij}} = \frac{(4-8.4)^2}{8.4} + \frac{(20-15.6)^2}{15.6} + \frac{(10-5.6)^2}{5.6} + \frac{(6-10.4)^2}{10.4}$$ (3)

$$= 2.305 + 1.241 + 3.457 + 1.862$$

$$= \boxed{8.8645}$$ (2) ~~8.9286~~

The critical value: $\chi^2_{(r-1)(c-1),\alpha} = \chi^2_{1,0.05} = \boxed{3.8415}$ (1) ~~5.9915~~

The decision rule: If $\chi^2_{cal} > \chi^2_\alpha \Rightarrow$ Reject $H_o$

Since $\chi^2_{cal} = 8.8645 > 3.8415 = \chi^2_\alpha \Rightarrow$ Reject $H_o$ (1)

The conclusion: The delivery time and the computer assisted ordering are NOT indep. (Dependent). (1)

The last Column should be Combined with the 2nd Column or 3 marks will be DEDUCED

**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:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Copiers serviced (X)** | 4 | 2 | 5 | 7 | 1 | 3 | 4 | 5 | 2 | 4 | 6 |
| **Minutes required (Y)** | 140 | 68 | 103 | 145 | 60 | 51 | 103 | 134 | 110 | 90 | 112 |

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.

| | |
|---|---|
| The correlation coefficient = | $r = \dfrac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} = \dfrac{4773 - \frac{(43)(1116)}{11}}{\sqrt{\left(201 - \frac{(43)^2}{11}\right)\left(123368 - \frac{(1116)^2}{11}\right)}}$ (2) $= 0.71$ (1) |
| The assumptions are: a. Two variables are Quantitative | b. Two variables have the Bivariate Normal distn. (2) |
| The hypotheses are: $H_0$: $\rho \leq 0$ | $H_A$: $\rho > 0$ (1) |
| The test statistic value: | $t_0 = t_{cal} = \dfrac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \dfrac{0.71}{\sqrt{\frac{1-(0.71)^2}{9}}} = \dfrac{0.71}{\sqrt{0.0551}} = \dfrac{0.71}{0.235} = \boxed{3.025}$ (1) (1) |
| The critical value: | $t_{\alpha, n-2} = t_{0.025, 9} = \boxed{2.2622}$ (2) 2.8214 X |
| Decision Rule: If $t_0 > t_\alpha \Rightarrow$ Reject $H_0$ | Since $t_{cal} = 3.025 > 2.2622 = t_\alpha \Rightarrow$ Reject $H_0$ (1) |
| Conclusion: | Yes, the relationship is direct (positive). (1) |

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 100,000's of bottles)
The following sums were obtained,

$n = 30$, $\sum x = 112.05$, $\sum x^2 = 418.742$, $\sum y = 251.48$, $\sum y^2 = 2121.53$, $\sum xy = 938.442$, and $SSE = 10.495$

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

1. The assumptions are: a. The residuals are indep b. The residules are normally dist. c. The residules have Constant variance d. The relationship between X & Y is linear (2)

2. The fitted regression equation is: $\hat{Y} =$ 21.653 − 3.553 $X$ (1)

$$b_1 = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} = \frac{938.442 - \frac{(112.05)(251.48)}{30}}{418.742 - \frac{(112.05)^2}{30}} \rightarrow \frac{-0.8358}{0.23525} = \boxed{-3.553}$$ (1)

$$b_0 = \bar{y} - b_1\bar{x} = \frac{251.48}{30} + (3.553)\left(\frac{112.05}{30}\right) = \boxed{21.653}$$ (1)

3. The standard error of the regression model is:

$$S_e = \sqrt{MSE} = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{10.495}{28}} = \sqrt{0.375} = \boxed{0.61223}$$ (1)

4. The predicted value of the *demand* if the *price* was \$4.00 is:

$$\hat{y}_{4.00} = 21.653 - (3.553)(4) = 7.441 \text{ H of T bottles} = 744,100 \text{ bottles}$$ (1)

5. A 99% C.I. for the *demand* if the *price* of a bottle was \$4.00 is:

$x_p = 4.00$, $\bar{x} = 3.735$, $\sum(x-\bar{x})^2 = 0.23525$, $\hat{y}_{4.00} = 7.441$

$t_{\frac{0.01}{2}, n-2} = t_{0.005, 28} = \boxed{2.7633}$ (1)

A 99% C.I is $\hat{y} \pm t_{\alpha/2} S_e \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum(x-\bar{x})^2}} = 7.441 \pm (2.763)(0.6122)\sqrt{1 + \frac{1}{30} + \frac{(4-3.735)^2}{0.23525}}$ (1)

$\Rightarrow$ C.I. is $[7.441 \pm 1.969] = [5.472, 9.41]$ (1)

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.

| | |
|---|---|
| The hypotheses are: $H_o$: $\beta_1 \leq 3$     $H_A$: $\beta_1 > 3$   1 | |
| The test statistic: | $t_o = \dfrac{b_1 - \beta_{1_o}}{s_{b_1}} = \dfrac{-3.553 - 3}{s_e/\sqrt{0.23525}} = \dfrac{-6.553}{0.61223/0.485}$ 1 <br> $= \boxed{-5.1912}$ 1 |
| The critical value: | $t_{\alpha, n-2} = t_{0.1, 28} = \boxed{1.3125}$ 1 |
| The decision rule: | If $t_o = t_{cal} > t_{\alpha, n-2} \Rightarrow$ Reject $H_o$ <br> Since $t_o = -5.1912 \not> 1.3125 = t_\alpha \Rightarrow$ Do NOT reject $H_o$ 1 |
| The conclusion: | The demand will ~~do~~ increase by at most 300,000 bottles. 1 |

*With Our Best Wishes*