1. (12 Marks) The following table gives the gasoline mileages for 50 randomly selected cars from a certain brand. It is claimed that the gasoline mileages for this brand is normally distributed with mean 31 and standard deviation 0.8 miles/gallon.

Mileage Class	Number of cars	P_i	e_i
< 30.95	13	0.4751	23.755
30.95 < 32.15	25	0.4496	22.480
> 32.15	12	0.0753	3.765
Total	**50**	1	50.000

Test this claim at 1% significance level.

(1) H_0: The mileage is Normally distributed.

H_A: " " " NOT " " .

$P_1 = P(X < 30.95) = P\left(Z < \frac{30.95-31}{0.8}\right)$

$= P(Z < -0.06) = 0.5 - 0.0249 = \boxed{0.4751}$ (1)

$P_3 = P(X > 32.15) = P\left(Z > \frac{32.15-31}{0.8}\right)$

$= P(Z > 1.44) = 0.5 - 0.4247 = \boxed{0.0753}$ (1)

$e_i = nP_i \Rightarrow e_1 = nP_1 = 50(0.4751) = \boxed{23.755}$ (1)

$e_2 = nP_2 = 50(0.4496) = \boxed{22.48}$ (1), $e_3 = \boxed{3.765}$ (1)

$\chi^2 = \sum_1^3 \frac{(O_i - e_i)^2}{e_i} = \frac{(13-23.755)^2}{23.755} + \frac{(25-22.48)^2}{22.48} + \frac{(12-3.765)^2}{3.765}$ (1)

$= 4.8693 + 0.2825 + 18.0120 = \boxed{23.164}$ (1)

$\chi^2_{0.05,2} = \boxed{5.9915}$

(1) If $\chi^2_{cal} > \chi^2_{tab} \Rightarrow$ Reject H_0

∴ (1) Since $\chi^2_{cal} = 23.164 > 5.9915 = \chi^2_{tab}$, Reject H_0

(1) That is, the mileage is NOT normally distributed.

(1) Assumptions: (1) Sample size ≥ 30.

(2) $e_i \geq 5$.

2. (8 Marks) A survey was recently conducted in which males and females were surveyed and asked whether they owned a laptop personal computer. The following data were observed:

	Males	Females	Total
Have Laptop	120 107.67	70 82.33	190
No Laptop	50 62.33	60 47.67	110
Total	170	130	300

Given this information, is whether having a laptop is independent of gender? Explain.

(1) H_0: Gender is Indep. of having a laptop.

H_A: " " NOT indep. " " " "

$$e_{ij} = \frac{(n_{i.})(n_{.j})}{n} \Rightarrow n_{11} = \frac{190 \times 170}{300} = 107.67 \quad (1)$$

$$\chi^2_{cal} = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(O_{ij} - e_{ij})^2}{e_{ij}} = \frac{(107.67-120)^2}{107.67} + \cdots + \frac{(60-47.67)^2}{47.67} =$$

$$= \boxed{8.887}\ (1), \quad \chi^2_{(2-1)(2-1);\,0.05} = \chi^2_{1,0.05} = 3.8415\ (1)$$

(2) Since $\chi^2_{cal} > \chi^2_{tab} \Rightarrow$ Rej. H_0.

(1) The two factors are NOT independent.

Assumptions: (1) sample size ≥ 30

(1) (2) $e_{ij} \geq 5$

3. (8 Marks) A manufacturing company is interested in predicting the number of defects that will be produced each hour on the assembly line. The managers believe that there is a relationship between the defect rate and the production rate per hour. The *managers believe that they can use production rate to predict the number of defects.* The following data were collected for 10 randomly selected hours.

Defects (X)	20	30	10	20	30	25	30	20	10	40
Production rate (Y)	400	450	350	375	400	400	450	300	300	300

Given the following

$\Sigma x = 235$, $\Sigma y = 3725$, $\Sigma x^2 = 6325$, $\Sigma y^2 = 1418125$ and $\Sigma xy = 89000$.

a. Find the correlation coefficient between the two variables.
b. Is there a significant relationship between the two variables? Explain.

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

$$= \frac{10(89000) - (235)(3725)}{\sqrt{[63250 - (235)^2][~~14181250~~ - (3725)^2]}} \quad ①$$

$$= \boxed{0.295311} \quad ①$$

b. ① $H_0: \rho = 0$ vs. $H_A: \rho \neq 0$

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.295311}{\sqrt{\frac{1-0.087}{8}}} = \frac{0.295311}{0.338} =$$

$$= \boxed{0.874} \ ①$$

$$t_{\frac{\alpha}{2}, n-2} = t_{0.025, 8} = 2.306 \ ①$$

① Since $|t_{call}| = 0.874 \not> 2.306 = |t_{tab}| \Rightarrow$ Do NOT rej. H_0

① There is No significant relationship between the two variables.

Assumptions: ① Data are quantitative. ①
① Have bivariate Normal.

4. (12 Marks) A study was done in which the high daily temperature and the number of traffic accidents within the city were recorded. These sample data are shown as follows:

Temperature (X)	91	56	75	68	50	39	98
# of Accidents (Y)	7	4	9	11	3	5	8

$H_0: \beta_1 = 0$

$H_A: \beta_1 \neq 0$ 1

Given the following

$\Sigma x = 477,\ \Sigma y = 47,\ \Sigma x^2 = 35291,\ \Sigma y^2 = 365,\ \Sigma xy = 3413.$

$\sum(x-\bar{x})^2 = 2786.86,\ \sum(y-\bar{y})^2 = 49.4286$ and $\sum(x-\bar{x})(y-\bar{y}) = 210.286.$

a. Find the equation for predicting the number of accidents using the temperature.
b. Construct a 95% confidence interval for the slope of the regression line.
c. Use (b) to test weather there is a linear relationship between the two variables.

a. $b_1 = \frac{S_{xy}}{S_{xx}} = \frac{210.286}{2786.86} = \boxed{0.0755}$ 1

$b_0 = \bar{y} - b_1\bar{x} = \frac{47}{7} - (0.0755)(\frac{477}{7})$

$= \boxed{1.5695}$ 1

1 $\hat{y} = (1.5695) + (0.0755)X$

b. A $100(1-\alpha)\%$ C.I. for β_1 is $b_1 \pm t_{\frac{\alpha}{2}, n-2} S_{b_1}$

$S_{b_1} = \frac{S_\epsilon}{\sqrt{\Sigma(x-\bar{x})^2}} = \frac{\sqrt{SSE/n-2}}{\sqrt{\Sigma(x-\bar{x})^2}} = \sqrt{\frac{(SST-SSR)/n-2}{\Sigma(x-\bar{x})^2}}$ 1

$SST = 49.4286$ 1 & $SSR = b_1 S_{xy} = (0.0755)(210.286)$
$= 15.8766$ 1

$SSE = 49.4286 - 15.8766 = \boxed{33.552}$ 1

So $S_{b_1} = \sqrt{\frac{33.552/5}{2786.86}} = \sqrt{\frac{6.7104}{2786.86}} = \sqrt{0.00241} = \boxed{0.0491}$ 1

$t_{0.025,5} = \boxed{2.5706}$

A 95% CI for β_1 is $0.0755 \pm (2.5706)(0.0491)$

$(0.0755 \pm 0.1261) = \boxed{[-0.0506,\ 0.2016]}$ 1

c. Since $0 \in$ C.I. $\Rightarrow$ we do NOT rej. H_0
There is No relationship between X & Y 1 1

5. (10 Marks) A study was recently performed to determine how much tip income waiters should make based on the size of the bill at each table. A random sample of bills and resulting tips were collected and analyzed using Minitab as shown in the next page, *use the Minitab output to answer the following questions:*
   a. Find the correlation coefficient between the bill and the tip and comment on your finding.
   b. Test weather the value of the slope is more than 0.1. Give your conclusion.
   c. Find a 95% confidence interval for a tip from a bill with $100.

a. $r = \sqrt{R^2} = \sqrt{0.877} = 0.936$ (1)

⇒ There is a STRONG POSITIVE linear (1) relationship between the tip and the bill

c. $(11.74, 29.29)$ (1)

b. $H_0: \beta_1 \le 0.1$ (1)
$H_A: \beta_1 > 0.1$ ⇒ $t = \dfrac{b_1 - \beta_{10}}{S_{b_1}}$ ⇒

$t = \dfrac{0.2128 - 0.1}{0.02814} = 4.01$ (1) (1) (1)

$t_{n-2;\alpha} = t_{8;0.05} = \boxed{1.8595}$

(2) Since $t_{cal} = 4.01 > 1.8595 = t_{tab}$ ⇒ Reject H_0 which means that

(1) The slope is GREATER than 0.1