# KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
## DEPARTMENT OF MATHEMATICS & STATISTICS
## DHAHRAN, SAUDI ARABIA

STAT 212: BUSINESS STATISTICS I I

Semester 063

Second Major *SOLUTIONS*

Saturday Aug 8, 2007

7:00 - 8:30PM

Please **circle** your:

**Instructor's name** & **section number**

Mohammad F. Saleh — Sec 1: (9:20 – 10:20) Sec 2 : (10:30 – 11:30)

Marwan M. Almomani — Sec 3 : (10:30 – 11:30)

Name: Student ID#: Serial #:

**Directions:**

1) You must **show all work** to obtain full credit for questions on this exam.
2) DO NOT round your answers at each step. Round answers only if necessary at **your final step to 4 decimal places**.

| Question No | Full Marks | Marks Obtained |
|---|---|---|
| Q1 | 14 | |
| Q2 | 14 | |
| Q3 | 14 | |
| Q4 | 23 | |
| Total | 65 | |

 *SOLUTIONS* Al-Momani, M.

**Question One (2+1+6+1+2+1+1 = 14 pts):**

Last rating period, the percentages of the viewers watching several channels between 11:00 P.M and 11:30 PM in a major TV market were as follows:

| | WDUX (News) | WWTY (News) | WACO (Cheers Reruns) | WTJW (News) | Other |
|---|---|---|---|---|---|
| $p_i$ = | 10% | 27% | 19% | 9% | 35% |

And in the current rating period, a survey of 2,000 viewers gives the following:

| | WDUX (News) | WWTY (News) | WACO (Cheers Reruns) | WTJW (News) | Other |
|---|---|---|---|---|---|
| | 182 | 536 | 354 | 151 | 777 |
| $e_i = n*p_i \to$ | 200 | 540 | 380 | 180 | 700 |

2.5 pts

Do you think that the viewing shares in the current rating period differ from those in the last rating period at 0.10 level of significance?

a. The test hypotheses are:
   $H_0$: The viewing shares in the current rating period is the same as those in the last rating period.
   $H_A$: The viewing shares in the current rating period differ from those in the last rating period.
   2 pts

b. The assumption is:
   $e_i \geq 5$ } 1 pt

c. The test statistic is:

$$\chi^2_c = \sum \frac{(o_i - e_i)^2}{e_i} = \frac{(182-200)^2}{200} + \frac{(536-540)^2}{540} + \frac{(354-380)^2}{380} + \frac{(151-180)^2}{180} + \frac{(777-700)^2}{700}$$

2.5 pts

$$= 1.62 + 0.0296 + 1.7789 + 4.6722 + 8.47$$

$$= 16.5707 \quad \} \text{ 1 pt}$$

d. The critical value is: $\chi^2_{\alpha, k-1} = \chi^2_{.10, 4} = 7.7794$ 1 pt

e. The decision rule and the decision are:

Reject $H_0$ if $\chi^2_c > \chi^2_{\alpha, k-1}$

$16.5707 > 7.7794$ $\therefore$ Reject $H_0$ } 2 pts

f. Your conclusion is:
The viewing shares in the current rating period differ from those in the last rating period 1 pt

g. Based on your decision, what type of error you might have committed?

Type I error. } 1 pt

## Question Two (2+7+1+2+1+1 = 14 pts):

Marketers know that tastes differ in various regions of the country. In the rental car business, an industry expert has given the opinion that there are strong regional preferences of size of car and quotes. The following data in support of the view:

(3) pts

| Preferred Car Type | | Region of Country | | | | |
|---|---|---|---|---|---|---|
| | | Northeast | Southeast | Northwest | Southwest | Total |
| | Full – size | 100 (100) | 118 (100) | 108 (100) | 74 (100) | 400 |
| | Intermediate | 122 (125) | 105 (125) | 125 (125) | 148 (125) | 500 |
| | All other | 28 (25) | 27 (25) | 17 (25) | 28 (25) | 100 |
| | Total | 250 | 250 | 250 | 250 | 1000 |

Do the data support the expert's opinion at the 0.05 significance level?

a. The test hypotheses are:

$H_0$: Region of Country and preferred car type are independent (2) pts

$H_A$: " " " " " " " " not independent.

b. The test statistic is:

$$\chi_c^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(o_{ij}-e_{ij})^2}{e_{ij}}$$

$$= \frac{(100-100)^2}{100} + \frac{(118-100)^2}{100} + \frac{(108-100)^2}{100} + \frac{(74-100)^2}{100} +$$

$$\frac{(122-125)^2}{125} + \frac{(105-125)^2}{125} + \frac{(125-125)^2}{125} + \frac{(148-125)^2}{125} + \quad (3) \text{ pts}$$

$$\frac{(28-25)^2}{25} + \frac{(27-25)^2}{25} + \frac{(17-25)^2}{25} + \frac{(28-25)^2}{25}$$

$$= 0 + 3.24 + 0.64 + 6.76 + 0.072 + 3.2 + 0 + 4.232 + .36 + .16 + 2.56 + 0.36 = 21.584 \quad \} (1) pt$$

c. The critical value is:

$$\chi^2_{\alpha,(r-1)(c-1)} = \chi^2_{.05,6} = 12.5916 \quad \} (1) pt$$

d. The decision rule and the decision are:

Reject $H_0$ if $\chi_c^2 > \chi^2_{\alpha,(r-1)(c-1)}$

$21.584 > 12.5916$ $\therefore$ Reject $H_0$ } (2) pts

e. The conclusion is:

The region country and preferred car type are NOT independent } (1) pt

f. What are other assumptions required to perform the test?

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

for all $i,j$.

**Question Three (3+3+8 = 14pts):**

The following table shows how many weeks a sample of 6 persons have worked at an automobile inspection station and the number of cars each one inspected between noon and 2 P.M. on a given day:

| | | | | | | |
|---|---|---|---|---|---|---|
| Number of cars inspected (y) | 14 | 13 | 15 | 21 | 23 | 21 |
| Number of weeks employed (x) | 1 | 3 | 4 | 7 | 8 | 13 |

You have calculated some of the necessary summary information to carry out the analyses as follows:

$$\sum x = 36,\ \sum x^2 = 308,\ \sum y = 107,\ \sum y^2 = 2001 \text{ and } \sum xy = 717$$

a. Obtain the correlation coefficient

$$r = \frac{717 - \frac{(36)(107)}{6}}{\sqrt{\left[308 - \frac{(36)^2}{6}\right]\left[2001 - \frac{(107)^2}{6}\right]}} = \underbrace{\frac{75}{\sqrt{(92)(92.8333)}} = \frac{75}{92.41573}}_{②\ \text{pts}} = 0.8116 \quad ①\text{pt}$$

b. Interpret the value of the linear correlation coefficient in terms of the linear relationship between the two variables.

There is a positive strong linear relationship between (Y) number of cars inspected and (X) number of weeks employed. ③ pts

c. (2+2+1+2+1 = 8 pts) At 1% level of significance, do the data provide sufficient evidence to conclude that the number of weeks employed (x) and the number of cars inspected (y) are negatively linear correlated?

I. State the hypotheses:

$H_0$: $\rho \geq 0$ vs. $H_A$: $\rho < 0$ } ② pts

II. The test statistic is:

$$t_c = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.8116}{\sqrt{\frac{1-(.8116)^2}{6-2}}} = 2.7784 \quad \} \ ②\text{ pts}$$

III. The critical value is:

$t_{\alpha, n-2} = t_{.01, 4} = 3.7469$ } ① pt

IV. The decision rule and the decision are:

Reject $H_0$ if $t_c < -t_{\alpha,n-2}$

$2.7784 \not< -3.7469$ } ② pts

$\therefore$ Do NOT reject $H_0$

V. The conclusion is:

The data do not provide a sufficient evidence that the number of weeks employed (X) and number of cars inspected (Y) are negatively linearly correlated. ① pt

**Question Four (2+6+1+4+1+3+1+1+4 = 23 pts):**

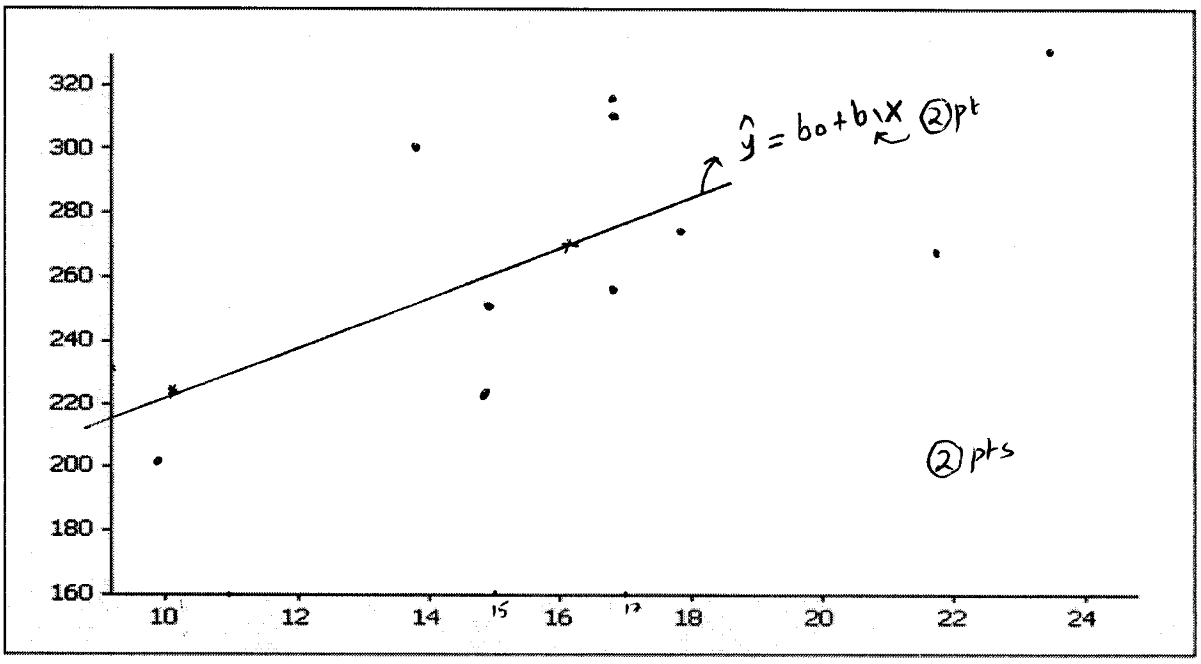
A manufacturing company is interested in predicting the cost of certain product. The manager believes that there is a relationship between the cost (in Dollars) of the product and the size (in millimeter) of the product. The manager believes that he can use production size to predict the cost of the product. The following data were collected randomly.

| Cost (Dollars) (y) | 245 | 312 | 279 | 308 | 201 | 219 | 270 | 324 | 300 | 255 |
|---|---|---|---|---|---|---|---|---|---|---|
| Lot Size (x) | 15 | 17 | 18 | 17 | 10 | 15 | 22 | 24 | 14 | 17 |

Also, the following summary statistics is obtained by the manager to predict the cost of the product using production size.

$\sum x = 169$, $\sum x^2 = 2997$, $\sum y = 2713$, $\sum y^2 = 751337$, $\sum xy = 46783$, and $SSE = 9118$

a. Draw a scatter diagram to verify the assumption that the relationship is linear.




b. Fit a straight line to these data by the method of least squares, and draw its graph on the diagram obtained in part (a).

$$b_1 = \frac{46783 - \frac{(169)(2713)}{10}}{2997 - \frac{(169)^2}{10}} = \frac{933.3}{140.9} = 6.62385 \quad \} \text{ ② pts}$$

$$b_0 = \bar{y} - b_1 \bar{x} = 271.3 - (6.62385)(16.9) = 159.3570 \quad \} \text{ ① pt}$$

$\hat{Y} =$ 159.3570 + 6.62385 $X$ } ① pt.

The line passes $(\bar{x}, \bar{y}) = (16.9, 271.3)$

when $x = 10$, $y = 225.6$

c. Interpret the slope of the regression line. $b_1 = 6.62385$

As the lot size (X) increases by 1 millimeter, then the cost (Y) will increase by \$6.62 ① pt

d. What is the percentage of the variation of the cost that explained by the variation of the size?

$R^2 = \frac{SSR}{SST}$ ; $SST = \Sigma(y-\bar{y})^2 = \Sigma y^2 - n\bar{y}^2 = 751337 - (10)(271.3)^2$
$= 15300.1$ ① pt

$SSR = SST - SSE = 15300.1 - 9118 = 6182.1$ } ① pt

$\therefore R^2 = \frac{6182.1}{15300.1} = 0.4041$ ① pt $\Rightarrow$ The percentage $= 40.41\%$ } ① pt

e. The standard error of the regression slope is:

$S_\varepsilon = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{9118}{10-2}} = \sqrt{1139.75} = 33.7602$ ① pt

$\therefore S_{b_1} = \frac{S_\varepsilon}{\sqrt{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}} = \frac{33.7602}{\sqrt{2997 - \frac{(169)^2}{10}}} = \frac{33.7602}{\sqrt{140.9}} = 2.8441$ } ① pt

f. Construct a 95% confidence interval for the true regression slope and interpret this interval estimate.

$1-\alpha = .95 \Rightarrow \alpha = .05$ $\therefore t_{\alpha/2, n-2} = t_{.025, 8} = 2.3060$ } ① pt

① pt A 95% C.I. for $\beta_1$ is: $b_1 \pm t_{\alpha/2, n-2} \cdot S_{b_1}$

$6.62385 \pm (2.3060)(2.8441)$

$6.62385 \pm 6.5585$

$0.0654 \le \beta_1 \le 13.1823$ } ① pt

Interpretation: As the lot size (X) increases by 1 millimeter, then the cost (Y) increases by on average between .06 and 13.18 dollars.

g. Based on the C.I that you obtained in part (f), do you think that there is a significant relationship between the size and the cost of the product? Explain.

The zero value does not belong to the C.I., so the linear relationship between (X) and (Y) is significant ① pt

h. Use the regression equation that you obtained in part (b) to predict the cost of a product if the size of the lot is 19 millimeter.

$\hat{y}(19) = 159.3570 + 6.62385(19)$
$= 285.21$ } ① pt

i. Find the 95% confidence interval for the average of a lot of size of 19 millimeter.

A 95% C.I. for $E(Y|X_p = 19)$ is:

$\hat{y} \pm t_{\alpha/2, n-2} \, S_\varepsilon \sqrt{\frac{1}{n} + \frac{(X_p - \bar{x})^2}{\Sigma(x-\bar{x})^2}}$

$285.210 \pm (2.3060)(33.7602)\sqrt{\frac{1}{10} + \frac{(19-16.9)^2}{140.9}}$ } ① pt

$285.210 \pm 28.8700$

$[256.34, 314.08]$ } ② pts