Sai Parthish Mandumula
1002022847

DAMT

Assignment - 2

11.1. given data is
The transmitted files : 30.
The average size of = 126 Kbytes := $\bar{S}$
and standard deviation = 35 k bytes = S_s
The average transmittance time was = 0.04 secs
and the standard deviation = 0.01 seconds
The correlation between time and the size
was ~~80~~ 0.86 = λ
from the given information $T - \bar{T}$ =
$$\lambda \frac{S_T}{S_S}(S - \bar{S})$$

By. Substituting the values
$$T - 0.04 = 0.86 \frac{0.01}{35}(S - 126)$$

and $$T = 0.86 \frac{0.01}{35}(S - 126) + 0.04 \quad (1)$$

By the above equation, we can come to know to the time that transmit a 400 kbyte
To S = 400 kbytes
$$T = 0.86 \frac{0.01}{35}(400 - 126) + 0.04$$
$$= 0.0673 + 0.04$$
$$T = 0.1074 \text{ seconds.}$$

Sai Parthish Mandumula
1002022847.

11.2 The following statistics were obtained for a sample of size $n = 75$:
the predictor variable x has mean $32.2, = \bar{x}$
Variance 6.4; $v(x)$, $v(y) = 2.8$, $cov(x,y) = 3.6$

a) Linear regression equation predicting Y based on x is expressed as.

$$y - \bar{y} = \frac{cov(x,y)}{\sigma x^2}(x - \bar{x})$$

$$\Rightarrow y - 8.4 = \frac{3.6}{6.4}(x - 32.2)$$

$$\Rightarrow y = 0.56x - 18.11 + 8.4$$

$$y = 0.56x - 9.71$$

∴ The linear regression equation is a follow
$y = 0.5628x - 9.71.$

(b) form the parameters that are given and calculate from making ANOVA

$$S.S_{tot} = (n-1)Sy^2$$

$$= (75-1) \times 2.8$$

$$= 207.2$$

$$SS_{reg} = (n-1)\,b_1^2 \times Sx^2$$

$$(75-1) \times (0.5625)^2 \times 6.4$$

$$= 149.85$$

$$SS_{ERR} = SS_{tot} - S_{reg}$$

$$= \cancel{207}\ 207 - 149 = 57.35$$

Sai Parthish Mandumula
1002022847

ANOVA table for above model is given by

Source	Sum of squares	Degree of freedom	mean Square	F
model	149.85	1	149.85	190.748
Error	57.35	73	0.7856	
Total	207.2	74	2.8	

Total variation estimated by R-Square

$$R^2 = \frac{SSreg}{SS\,tot} = \frac{149.85}{207.2} = 0.723$$

∴ total variation of Y in X is 0.723.

ⓒ The given $[1-\alpha]100\%$ confidence interval to the slope

$$\left[b_1 - t_{\alpha/2}\frac{s}{\sqrt{S_{xx}}}, b_1 + t_{\alpha/2}\frac{s}{\sqrt{S_{xx}}}\right]$$

Here the $s = \sqrt{0.7856}$

$= \underline{0.886}$

Sai Parthish Mandhemula
1002022842.

for the 99% confidence interval $t_{0.005,74}$ =

$$S_{xx} = \Sigma(x_i - \bar{x})^2$$

$$= n s_x^2 - 76 \times 64$$

$$= 480$$

Here we have 99% for the regression slope is calculated as

$$C_{99\%} = \left[b_1 - t_{\alpha/2} \frac{s}{\sqrt{S_{xx}}},\ b_1 + t_{\alpha/2} \frac{s}{\sqrt{S_{xx}}} \right)$$

$$= \frac{0.5625 \times \sqrt{480}}{0.886} = 13.91$$

∴ we have the T value which is greater than the theory value of t-distribution at 99% level of significance & 74 degrees of freedom regression slope so the significant at 99% level.

11.5)

Sai Parthish Mandumula
1002022842

11.9 from the sample 11.10. on p. 390
The given information:

$n = 7$, $\bar{x}_2 = 9.57$, $s^2_{x_2} = 63.29$, $\bar{y} = 35$

$s^2_y = 242$, $r_{x_2 y} = 0.758$

Compute the least Squares Estimates.

$$b_1 = r\sqrt{\frac{s^2_y}{s^2_{x_2}}} = (0.758)\sqrt{\frac{242}{63.29}} = 1.48$$

$$b_0 = \bar{y} - b_1\bar{x}_2 = 35 - (1.48)(9.57) = 20.84.$$

The fitted regression line has the equation:

$$\hat{y} = 20.84 + 1.48x_2$$

The coefficient of determination is

$$R^2 = r^2 = 0.525$$

Showing that 52.5% of the total variation of the number of processed requests is Explained by the number of tables only.

$$SS_{TOT} = (n-1)s^2_y = (6)(242) = 1452$$
$$(n-1 = 6 \text{ d.f})$$

$$SS_{REG} = R^2 SS_{TOT} = (0.525)(1452)$$
$$= 834.9 \text{ (1 d.f)}$$

$$SS_{ERR} = SS_{TOT} - SS_{REG} = 1452 - 834.9$$
$$= 617.1 \ (n-2 = 5 \text{ d.f})$$

Sai Parthish Mandumula
1002022847.

$R^2 adj$?

Source	Sum of Square	degree of freedom	mean Square	F
model	834	1	8.34	6.76
Error	617	5	123	
Total	1454	6	242	

So, the adjusted R^2

$$R^2 adj = 1 - \frac{Mss\ err}{Mss\ tot} = 1 - \frac{123}{242}$$

So the $R^2 adj = 1 - 0.51$
$= \underline{0.49}$

(b) mean Square and F-ratio

Source	Sum of Square	Degree of freedom	Mean Square	F ratio
model	834.9	1	834.9	6.78
Error	617.1	5	123.42	
Total	1452	6		

$$\text{mean Square model} = \frac{\text{Sum of Square model}}{\text{degree of freedom model}}$$

$$= \frac{834.9}{1} = \underline{834.9}$$

$$\text{mean Square Error} = \frac{\text{Sum of Square Error}}{\text{degree of freedom Error}}$$

Sai Parthish Mandumula
1002022847.

$$\Rightarrow \frac{617.1}{5}$$

$$\Rightarrow 123.42$$

$$F\text{-ratio} = \frac{\text{mean Square model}}{\text{mean Square Error}}$$

$$= \frac{834.9}{123.42}$$

$$= \underline{6.764}$$

11.5) From the above problem we have the values.

Year, x	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	13
Investment, Y	17	23	31	29	33	39	39	40	41	44	4[illegible]
Reporting profit z.	NO	NO	Yes	NO	Yes	Yes	Yes	Yes	NO	Yes	Y[illegible]

Prediction of matrix x and response vector x

$$x = \begin{array}{ccc} 1 & 2003 & 0 \\ 1 & 2004 & 0 \\ 1 & 2005 & 1 \\ 1 & 2006 & 0 \\ 1 & 2007 & 1 \\ 1 & 2008 & 1 \\ 1 & 2009 & 1 \end{array}$$

Sai Parthish Yandamula
100202284A

1 2010 1
1 2011 0
1 2012 1
1 2013 1

X r = 17 23 31 39 33 39 40 41 44 47

Compute $X^T X = \begin{pmatrix} 11 & 58 & 7 \\ 58 & 814 & 64 \\ 7 & 64 & 7 \end{pmatrix}$ $X^T y = \begin{pmatrix} 383 \\ 3356 \\ 273 \end{pmatrix}$

$b = (X^T X)^{-1} (X^T y) = \begin{pmatrix} 13.36 \\ 2.36 \\ 4.09 \end{pmatrix}$

The regression equation is represented as

$y = 13.36 + 2.63x + 4.09z$ — (1)

(b) Substituting x = 15 and z = 1 in the eq (1)

$y = 13.36 + 2.36x + 4.09z$

$13.36 + 2.36(15) + 4.09(1)$

52.85

Hence, the predicted investment amount is .52, 8501.

(c) Substituting x = 15 & z = 0 in eq (1)

$y = 13.36 + 2.36x + 4.09z$

$13.36 + 2.36(15) + 4.09(0)$

$= 48.76$

Sai Parthish Mandumula
1002022847

Hence, the predicted investment amount is $ 48,760
The investment reduce is $ 40,900

(d) multivariant ANOVA table.

$$\hat{y} = Xb$$

$$\hat{y}^T = \begin{bmatrix} 20.4 & 22.8 & 29.2 & 27.5 & 34 & 36.3 & 38.7 & 41 & 39.3 & 45.7 & 48 \end{bmatrix}$$

$$SS_{REG} = (\hat{y} - \bar{y})^T (\hat{y} - \bar{y}) = 808.02$$

$$SS_{Err} = (y - \hat{y})^T (y - \hat{y}) = 33.62$$

ANOVA table is represented as follows.

Source	Sum of Square	Degree of freedom	Mean Square	F
model	808.02	2	404	96.1
Error	33.62	8	4.2	
Total	841.62			

(e)

$$R^2 = \frac{SS_{REG}}{SS_{TOT}}$$

$$R^2 = \frac{808.02}{841.6}$$

$$R^2 = 0.96$$

Sai Parthish Mandumula
1002022847.

Since, it is more than the previous value of R^2

$\therefore$ The new variable 2 explains 3.9% of the total variation in y.