MATU9D2 : Practical Statistics   Spring 2017

Practical 8 : Solutions

These calculations were completed in Practical 7

1. (a) See attached graph paper

(b) $n = 10$ $\quad \sum x = 190$ $\quad \sum x^2 = 3940$

$\sum y = 173$ $\quad \sum y^2 = 3269$ $\quad \sum xy = 2988$

$$S_{xy} = \sum xy - \frac{\sum x \sum y}{n} = 2988 - \frac{190 \times 173}{10} = -299$$

neg. since neg. relation.

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n} = 3940 - \frac{190^2}{10} = 330$$

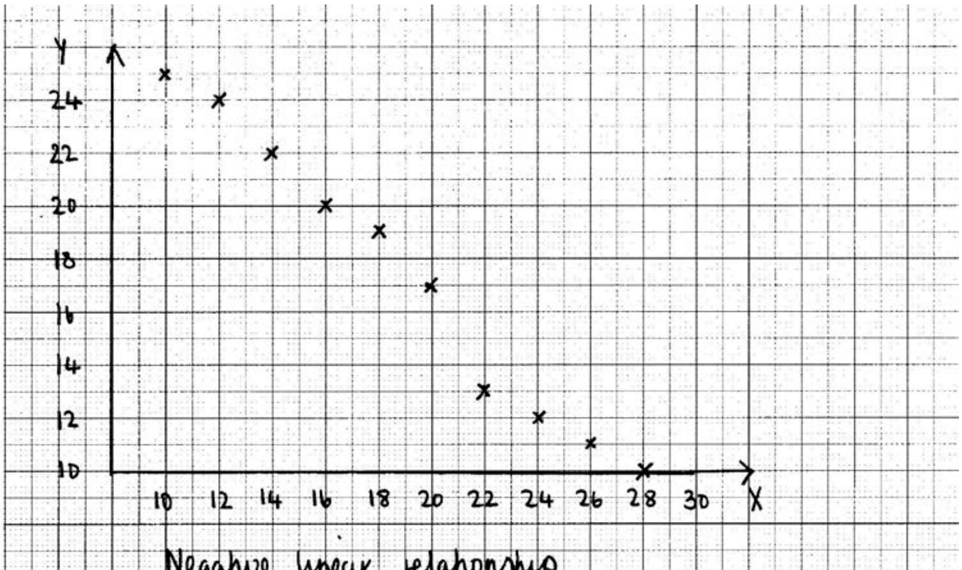
$$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n} = 3269 - \frac{173^2}{10} = 276.1$$

MUST always be pos.

$$r = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}} = \frac{-299}{\sqrt{330 \times 276.1}} = -0.991$$

$R^2 = (-0.991)^2 = 0.982$

ie. 98.2% of the variability in y is explained by linear relationship with x.



Negative linear relationship
so correlation should be negative
& since very close to all points on a
straight line correlation should
be close to -1

Question 1 (a)

$H_\emptyset : \rho = 0 \qquad H_1 : \rho \neq 0$

Significance level $\quad \emptyset.\emptyset 5$

Test Statistic $\quad t = r\sqrt{\frac{n-2}{1-r^2}} \quad \sim t(n-2)$ under $H_\emptyset$

Observed Test Statistic

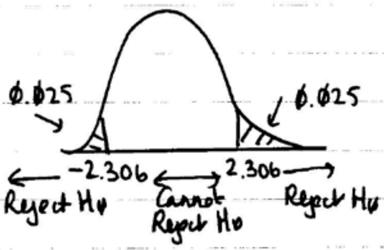
$n = 10 \qquad r = -0.991$

$$t = -0.991\sqrt{\frac{10-2}{1-(-0.991)^2}} = -0.991\sqrt{\frac{8}{0.0179}}$$

$$= -20.95 \; //$$

Rejection Region $\quad \emptyset.\emptyset 5$; 2 tailed ; $t(n-2)$

Critical Value $= t(8; \emptyset.\emptyset 25) = 2.306$



p value $= 2 \times P(t(8) > 20.95) < 2 \times 0.0005 = 0.001$

since $P(t(8) > 5.041) = 0.0005$

Conclusion Observed Test Statistic is in the Rejection Region so can reject $H_\emptyset$ in favour of H_1 at 5% level. In fact, $p < 0.01$ so can also reject $H_\emptyset$ in favour of H_1 at 1% & conclude that correlation is significantly different to zero

Question 1 (b) Using Summary Calculations from Practical 7

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} = \frac{-299}{330} = -0.906 //$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 17.3 - (-0.906) \times 19 = 34.51 //$$

So Regression Line is $y = 34.51 - 0.906x$

$$\hat{\sigma}^2 = \frac{S_{yy} - \frac{S_{xy}^2}{S_{xx}}}{n-2} = \frac{276.1 - \frac{(-299)^2}{330}}{8} = \frac{5.18788}{8}$$

$$= 0.6485 //$$

Error (Residual) Variance $= 0.6485$

Question 1 (c)

When $x = 10$, $y = 34.51 - 0.906 \times 10 = 25.45$

$x = 26$, $y = 34.51 - 0.906 \times 26 = 10.95$

Add these points to the graph & join up

– See graph paper

Question 1 (d)

$H_0: \beta = 0$

$H_1: \beta \neq 0$

<u>Significance Level</u> 0.05

<u>Test Statistic</u> $t = \dfrac{\hat{\beta}}{\sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}}} \sim t(n-2)$ under H_0

<u>Observed Test Statistic</u>

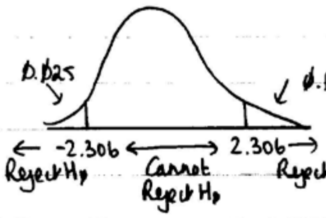
$\hat{\beta} = -0.906 \qquad \hat{\sigma}^2 = 0.6485 \qquad S_{xx} = 330$

$$t = \frac{-0.906}{\sqrt{\frac{0.6485}{330}}} = \frac{-0.906}{0.04433} = -20.44$$

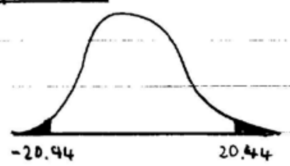
<u>Rejection Region</u> 0.05 ; 2tailed ; $t(8)$

<u>Critical Values</u> $= t(8; 0.025)$

$= 2.306$



<u>p value</u> $= 2 \times P(t(8) > 20.44) < 2 \times 0.0005 = 0.001$



p value = shaded region

<u>Conclusion</u>

Same conclusion from Rejection Region or pvalue. Can reject H_0 in favour of H_1 at 5% level (Obs TS in RR & $p < 0.05$) So evidence to conclude slope is significantly different to zero.

Question 2

Let x = no. of applications & y = yield in tons per acre

(a) See plot – not linear but not a smooth curve so cannot transform

(b) $n = 7$

$\Sigma x = 36$ $\quad \Sigma y = 45$

$\Sigma x^2 = 246$ $\quad \Sigma y^2 = 371$ $\quad \Sigma xy = 281$

$S_{xy} = 49.5714$ $\quad S_{xx} = 60.8571$ $\quad S_{yy} = 81.7143$

So $\hat{\beta} = \dfrac{S_{xy}}{S_{xx}} = \dfrac{49.5714}{60.8571} = 0.8146$ //

$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = \left(\dfrac{45}{7}\right) - 0.8146 \times \left(\dfrac{36}{7}\right) = 2.2392$ //

So regression line is $y = 2.239 + 0.815x$

(c) Predict y when $x = 7$ ie $y = 2.239 + 0.815 \times 7$
$= 7.944$

ie Predict yield of 7.944 tons per acre for 7 fertiliser apps.

(d) 95% CI for mean yield for 7 fertiliser apps. ($x_0 = 7$)

$$\hat{\alpha} + \hat{\beta}x_0 \pm t(n-2; 0.025)\sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}\right)}$$

$\hat{\alpha} = 2.239$ $\quad \hat{\beta} = 0.815$ $\quad x_0 = 7$

$t(5; 0.025) = 2.571$ $\quad \hat{\sigma}^2 = 8.267$ $\quad n = 7$

$\bar{x} = \Sigma x / n = 36/7 = 5.1428$ $\quad S_{xx} = 60.8571$

$$\hat{\sigma}^2 = \frac{S_{yy} - \frac{S_{xy}^2}{S_{xx}}}{n-2} = \frac{81.7143 - \frac{49.5714^2}{60.8571}}{5}$$

$$= \frac{41.3357}{5} = 8.267 \ //$$

$$(2.239 + 0.815 \times 7) \pm 2.571 \times \sqrt{8.267\left(\frac{1}{7} + \frac{(7 - 5.1428)^2}{60.8571}\right)}$$

$$7.944 \pm 2.571\sqrt{8.267 \times 0.1995}$$

$$7.944 \pm 2.571 \times 1.2843$$

$$7.944 \pm 3.302$$

$$(4.642, \ 11.246)$$

(e) 95% PI for yield for a single future set of 7 applications

$$\hat{\alpha} + \hat{\beta}x_0 \pm t(n-2; 0.025)\sqrt{\hat{\sigma}^2\left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right)}$$

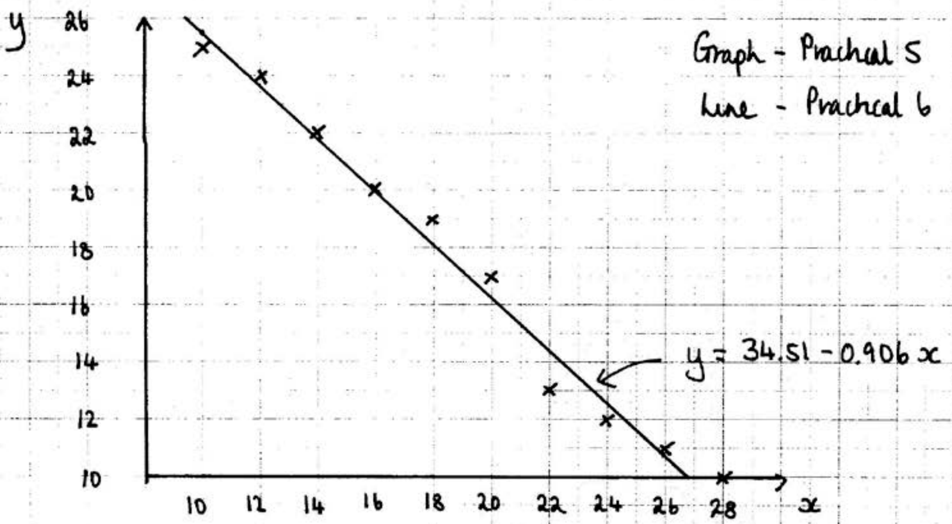
$$7.944 \pm 2.571\sqrt{8.267 \times 1.1995}$$

$$7.944 \pm 2.571 \times 3.149$$

$$7.944 \pm 8.096$$

$$(-0.15, \ 16.04)$$

Question 1



Graph - Practical 5
Line - Practical 6

Question 2