MATU9D2 : Practical Statistics Spring 2017

Weekly 8 : Hand Calculations : Solutions

(i) See graph paper

(ii) $n = 10$ $\quad \Sigma x = 5181$ $\quad \Sigma y = 246$

$\Sigma x^2 = 2707469$ $\quad \Sigma y^2 = 6510$ $\quad \Sigma xy = 130547$

$$S_{xy} = \Sigma xy - \frac{\Sigma x \Sigma y}{n} = 130547 - \frac{5181 \times 246}{10} = 3094.4$$

$$S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{n} = 2707469 - \frac{5181^2}{10} = 23192.9$$

$$S_{yy} = \Sigma y^2 - \frac{(\Sigma y)^2}{n} = 6510 - \frac{246^2}{10} = 458.4$$

$$r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{3094.4}{\sqrt{23192.9 \times 458.4}}$$

$$= \frac{3094.4}{3260.617} = 0.949 //$$

The correlation coefficient between number of registrations and number of manatees killed is 0.949.

$R^2 = 0.949^2 = 0.901$ ie 90.1% of the variability in the number of manatees killed is explained by the linear relationship with the number of registrations.

(iii) $H_0 : \rho = 0$ $\quad$ where ρ = population correlation between

$H_1 : \rho \neq 0$ $\quad$ no. of manatees killed & no. of regns.

Significance Level 0.05

Test Statistic $\quad t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t(n-2)$ under H_0.

Observed Test Statistic

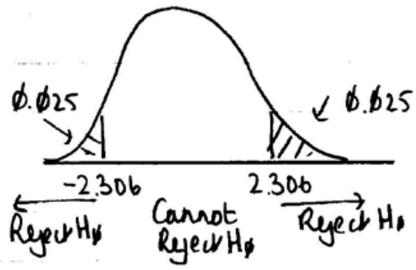
$n = 10 \qquad r = 0.949$

$$t = \frac{0.949\sqrt{8}}{\sqrt{1-0.949^2}} = \frac{2.6842}{0.3153} = 8.513$$

Rejection Region

2 tailed ; 0.05 ; $t(8)$

Critical Values $= t(8;\, 0.025) = 2.306$



p value

$= 2 \times P(\,t(8) > 8.513) < 2 \times 0.0005 = 0.001$

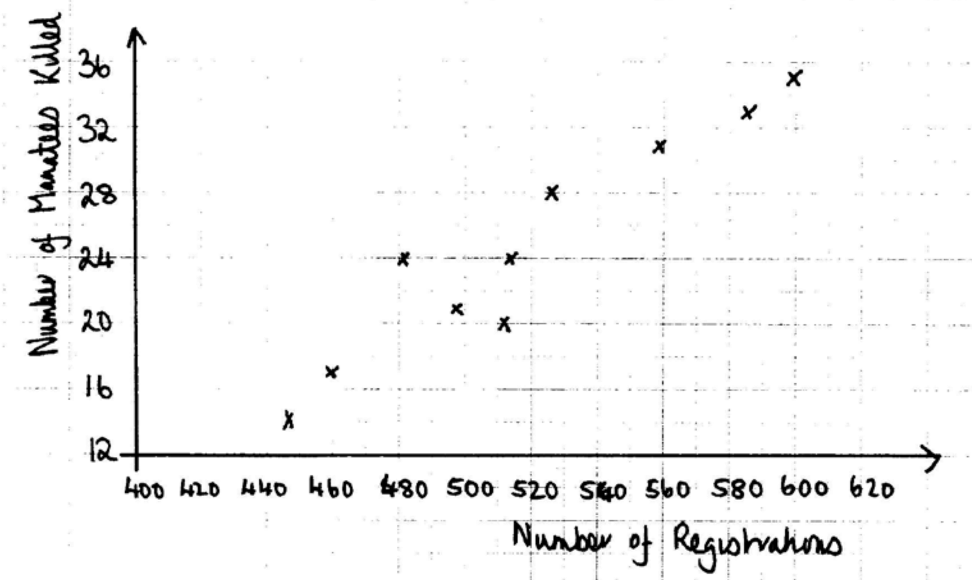
$P(t(8) > 5.041) = 0.0005$

Conclusion

1. Observed Test Statistic is in the Rejection Region $(8.513 > 2.306)$ so we can reject H_0 in favour of H_1 at 5% level & conclude that there is sufficient evidence, at 5% level, to suggest a significant correlation between number of manatees killed & no. of registrations

2. $p = 0.001$ ie probability of observing this data if H_0 is true is 0.001 ie 0.1% chance so can reject H_0 in favour of H_1 at 5% & 1% level.

(i)



Number of Manatees Killed
36
32
28
24
20
16
12
400 420 440 460 480 500 520 540 560 580 600 620
Number of Registrations

(iv) $n = 10 \qquad \Sigma x = 5181 \qquad \Sigma y = 246$

$\Sigma x^2 = 2707469 \qquad \Sigma y^2 = 6510 \qquad \Sigma xy = 130547$

$$S_{xy} = \Sigma xy - \frac{\Sigma x \Sigma y}{n} = 130547 - \frac{5181 \times 246}{10} = 3094.4$$

$$S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{n} = 270469 - \frac{(5181)^2}{10} = 23192.9$$

$$S_{yy} = \Sigma y^2 - \frac{(\Sigma y)^2}{n} = 6510 - \frac{(246)^2}{10} = 458.4$$

$$\beta = \frac{S_{xy}}{S_{xx}} = \frac{3094.4}{23192.9} = 0.1334 \,//$$

$$\alpha = \bar{y} - \beta \bar{x} = \frac{246}{10} - 0.1334 \times \frac{5181}{10}$$

$$= -44.515 \,//$$

Regression line is $y = -44.515 + 0.1334x$

Residual Variance is $\sigma^2 = \dfrac{S_{yy} - \dfrac{S_{xy}^2}{S_{xx}}}{n-2}$

$$= \frac{458.4 - \dfrac{3094.4^2}{23192.9}}{8}$$

$$= \frac{45.5447}{8} = 5.693 \,//$$

Add fitted line to graph

when $x = 440$ $\quad y = -44.515 + 0.1334 \times 440 = 14.181$

when $x = 600$ $\quad y = -44.515 + 0.1334 \times 600 = 35.525$

Add points to graph & join by a straight line
- see graph paper

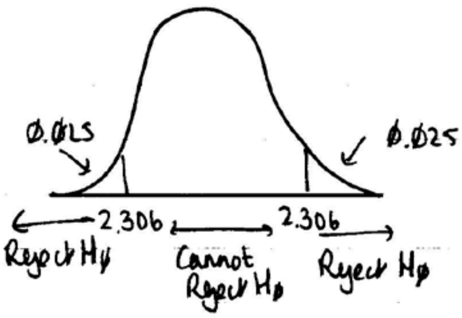
(v) $H_\emptyset : \beta = 0$
$H_1 : \beta \neq 0$
Significance level $\emptyset.\emptyset5$

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

<u>Observed Test Statistic</u>
$\hat{\beta} = 0.1334 \qquad \hat{\sigma}^2 = 5.693 \qquad S_{xx} = 23192.9$

$$t = \frac{0.1334}{\sqrt{\dfrac{5.693}{23192.9}}} = \frac{0.1334}{0.01567} = 8.515$$

<u>Rejection Region</u> $\quad$ 2 tailed ; $\emptyset.\emptyset5$; $t(8)$
Critical Values $= t(8; \emptyset.\emptyset25)$
$= 2.306$



$\underline{\text{p value}} = 2 \times P(t(8) > 8.515) < 2 \times 0.0005 = 0.001$

$\uparrow P(t(8) > 5.041) = 0.0005$

<u>Conclusion</u>

1. Observed test statistic is in the Rejection Region so can reject H_0 in favour of H_1 at 5% level. Conclude evidence that the slope is significantly different to zero.

2. $p \doteq 0.001$ ie probability of observing this data if H_0 is true is 0.001 ie 0.1% chance
ie can reject H_0 in favour of H_1 at 5% & 1% levels.

(vi) 95% Confidence Interval for Mean y when $x = x_0$ ($x_0 = 460$)

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

$\hat{\alpha} = -44.515$ $\quad \hat{\beta} = 0.1334$ $\quad t(8; 0.025) = 2.306$

$\hat{\sigma}^2 = 5.693$ $\quad n = 10$ $\quad \bar{x} = \frac{\Sigma x}{n} = \frac{5181}{10} = 518.1$

$S_{xx} = 23192.9$

$x_0 = 460$

$$(-44.515 + 0.1334 \times 460) \pm 2.306\sqrt{5.693\left(\frac{1}{10} + \frac{(460 - 518.1)^2}{23192.9}\right)}$$

$$16.849 \pm 2.306\sqrt{5.693 \times 0.2455}$$

$$16.849 \pm 2.306 \times 1.1823$$

$$16.849 \pm 2.726$$

$$(14.123, 19.575)$$

(vii) 95% Prediction Interval for y when single future x_0

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

$\hat{\alpha} = -44.515$ $\quad \hat{\beta} = 0.1334$ $\quad t(8; 0.025) = 2.306$

$\hat{\sigma}^2 = 5.693$ $\quad n = 10$ $\quad \bar{x} = 518.1$ $\quad$ same as (iv)

$S_{xx} = 23192.9$ $\quad x_0 = 460$

$$16.849 \pm 2.306\sqrt{5.693\left(1 + \frac{1}{10} + \frac{(460-518.1)^2}{23192.9}\right)}$$

$$16.849 \pm 2.306\sqrt{5.693 \times (1 + 0.2455)}$$

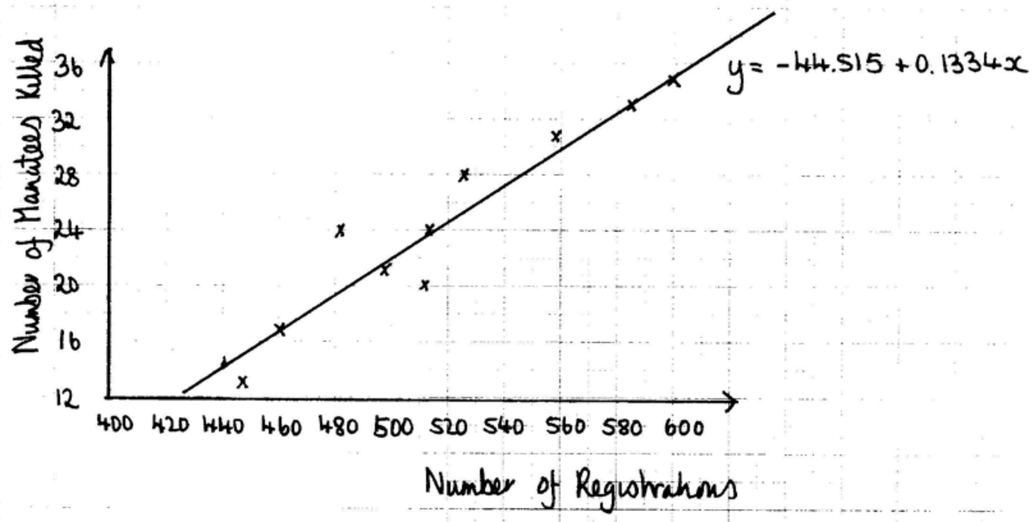
from (iv)

$$16.849 \pm 2.306 \times 2.6628$$

$$16.849 \pm 6.140$$

$$(10.709, 22.989)$$

(iv)



Number of Manatees Killed
36
32
28
24
20
16
12
400
420
440
460
480
500
520
540
560
580
600
Number of Registrations
y= -44.515 + 0.1334x