

Problem Set 9

Question 1

Data pairs: (1, 4), (2, 3), (3, 5), (4, 10)

Part A

$$\bar{y} = \frac{4 + 3 + 5 + 10}{4} = 5.5$$

$$\bar{x} = \frac{1 + 2 + 3 + 4}{4} = 2.5$$

$$\beta_1 = \frac{\sum_{i=1}^4 (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^4 (x_i - \bar{x})^2}$$

$$\beta_1 = \frac{(1 - 2.5)(4 - 5.5) + (2 - 2.5)(3 - 5.5) + (3 - 2.5)(5 - 5.5) + (4 - 2.5)(10 - 5.5)}{(1 - 2.5)^2 + (2 - 2.5)^2 + (3 - 2.5)^2 + (4 - 2.5)^2} = 2$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

$$\beta_0 = 5.5 - 2 \cdot 2.5 = .5$$

$$y = .5 + 2x$$

Part B

$$R^2 = 1 - \frac{SSE^*}{SST} = 1 - \frac{\sum_{i=1}^4 [(y_i - \bar{y}) - \beta_1(x_i - \bar{x})]^2}{\sum_{i=1}^4 (y_i - \bar{y})^2}$$

$$R^2 = 1 - \frac{(4 - (2 \cdot 1) - .5)^2 + (3 - (2 \cdot 2) - .5)^2 + (5 - (2 \cdot 3) - .5)^2 + (10 - (2 \cdot 4) - .5)^2}{(4 - 5.5)^2 + (3 - 5.5)^2 + (5 - 5.5)^2 + (10 - 5.5)^2} = .69$$

Part C

$$95\% \text{ Confidence Interval} = [\beta_1 - t_{\alpha/2, n-2} \frac{\sigma}{\sqrt{s_{xx}}}, \beta_1 + t_{\alpha/2, n-2} \frac{\sigma}{\sqrt{s_{xx}}}]$$

$$\sigma = \sqrt{\frac{SSE^*}{n-2}} = \sqrt{\frac{9}{2}} = 2.12$$

$$s_{xx} = \sum_{i=1}^4 (x_i - \bar{x})^2 = (1 - 2.5)^2 + (2 - 2.5)^2 + (3 - 2.5)^2 + (4 - 2.5)^2 = 5$$

$$95\% \text{ Confidence Interval} = [2 - 4.303 \cdot \frac{2.12}{\sqrt{5}}, 2 + 4.303 \cdot \frac{2.12}{\sqrt{5}}] = [-2.08, 6.08]$$

Part D

Hypothesis test H_0 is $\beta_1 = 0$ and H_1 is $\beta_1 \neq 0$

$$t = \frac{\beta_1}{\sigma/\sqrt{s_x x}} = \frac{2}{2.12/\sqrt{5}} = 2.11$$

$$p = P(T_2 < -2.11) + P(T_2 > 2.11) = .169$$

```
1  p <- 1 - pt(2.11, df = 2) + pt(-2.11, df = 2)
```

We fail to reject the null hypothesis since $.169 > .05$.

Question 2

Part A

```
1 quartet <- read.csv("Data/Quartet.csv")
2
3 Fit_1 = lm(quartet$y1 ~ quartet$x1)
4 Fit_2 = lm(quartet$y2 ~ quartet$x2)
5 Fit_3 = lm(quartet$y3 ~ quartet$x3)
6 Fit_4 = lm(quartet$y4 ~ quartet$x4)
```

The four fit lines are:

Series 1: $y = 3.0001 + .5001x$

Series 2: $y = 3.001 + .5000x$

Series 3: $y = 3.0025 + .4997x$

Series 4: $y = 3.0017 + .4999x$

Part B

```
1 summary(Fit_1)
2 summary(Fit_2)
3 summary(Fit_3)
4 summary(Fit_4)
```

Series 1: $R^2 = .6665$

Series 2: $R^2 = .6662$

Series 3: $R^2 = .6663$

Series 4: $R^2 = .6667$