Simple Linear Regression

11.5 Confidence Interval for the Slope β1

1. Researchers looked at the link between the popularity of elementary school aged children (8 – 11 years old) and their ability to recognized emotion in facial expressions. The 11 participant students were each assessed by their peers for popularity in an initial phase of the experiment. The mean popularity score given to the students by their peers is reported below in the first row. The score was based on a five-point scale with five indicating they are well liked and a score of one indicating they are disliked. In the second row, the number of correct answers each student had on an emotion-recognition test is recorded. The test involved inferring the emotional state of people in photographs. Use a 98% confidence level to form a confidence interval for the true slope for the linear relationship between popularity and scores on the emotional recognition exam. Do not forget to interpret the results.
(Note: $SS_{xx} = 10.1964, SS_{yy} = 546.1818, SS_{xy} = 72.3545$)

Subject	1	2	3	4	5	6	7	8	9	10	11
Popularity	1.3	2.4	1.8	4.3	3.1	3.9	3.2	2.1	4.1	1.9	2.5
Emotional Recognition	10	16	14	32	25	25	26	16	29	12	18

2. Psychologists wanted to know if the order of questions in a survey of happiness could have an effect on the correlation between responses. They ran two studies on college students where they asked about happiness and recent dating history. In the first study, they first asked ten single students to rate their happiness on a visual-analog scale, and then they asked them how many dates they had gone on in the last six months. Use the results and a 95% confidence level to form a confidence interval for the true slope of the linear relationship between happiness and number of dates when the number of dates is asked about after the happiness question. Interpret the results.
(Note: $SS_{xx} = 99.6, SS_{yy} = 19.06, SS_{xy} = -3.9$)

Subject	1	2	3	4	5	6	7	8	9	10
Number of dates	10	5	6	2	1	0	8	3	2	1
Happiness	5.1	4.2	7.2	7.9	4.6	5.2	3.8	3.4	4.1	5.5

3. The psychologists from problem 2 surveyed another ten students and asked them the same questions as in question 2, except they switched the order of the questions. This meant they asked about the number of dates before asking the students to assess their happiness. Form a 95% confidence interval for the true slope of the least squares line

used to model the linear relationship between happiness and number of dates when the number of dates is asked about first. Interpret the results.
(Note: $SS_{xx} = 93.6, SS_{yy} = 14.945, SS_{xy} = 35.6$)

Subject	1	2	3	4	5	6	7	8	9	10
Number of dates	9	7	5	2	3	1	9	2	0	4
Happiness	7.1	7.2	6.2	4.2	5.4	4.3	6.9	4.4	3.9	5.9

Answers:

1. The preliminary calculations and final answer:

   $SS_{xx} = 10.1964, SS_{yy} = 546.1818, SS_{xy} = 72.3545$

   $\hat{\beta}_1 = 7.0961$

   $SSE = 32.7483$

   $S^2 = 3.6387$

   $S_{\beta_1} = 0.59736$

   Critical value: $t_{\alpha/2,n-2} = t_{0.01,9} = 2.821$

   Margin of Error: $E = t_{\alpha/2} * S_{\hat{\beta}_1} = 2.821 * 0.59736 = 1.6852$

   Interval: $\left(\hat{\beta}_1 - E, \hat{\beta}_1 + E\right) = \left(7.0961 - 1.6852, 7.0961 + 1.6852\right) = \left(5.41, 8.78\right)$

   The sample data produced a positive interval which indicates a positive relationship exists between the two variables. This means that a child who recognizes emotional states in others easily tends to be more popular with his/her peers and vice versa. It also can be said that a unit change in a student's popularity score produces an increase on the emotional recognition test somewhere between 5.41 and 8.78 points (the standard definition of the slope of a line is: a unit change in X produces a $\hat{\beta}_1$ change in Y).

2. The preliminary calculations and final answer:

$SS_{xx} = 99.6, SS_{yy} = 19.06, SS_{xy} = -3.9$

$\hat{\beta}_1 = -0.03916$

$SSE = 18.9073$

$S^2 = 2.3634$

$S_{\beta_1} = 0.1540$

Critical value: $t_{\alpha/2,n-2} = t_{0.025,8} = 2.306$

Margin of Error: $E = t_{\alpha/2} * S_{\hat{\beta}_1} = 2.306 * 0.1540 = 0.355124$

Interval:

$\left(\hat{\beta}_1 - E, \hat{\beta}_1 + E\right) = \left(-0.03916 - 0.355124, -0.03916 + 0.355124\right) = \left(-0.394, 0.316\right)$

The fact that the interval includes zero leads us to conclude that there is not a linear relationship between these two variables (note: having no linear relationship does not necessarily mean there is no relationship. There could be another kind of relationship). This means that there does not seem to be a correlation between number of dates and happiness.

3. The preliminary calculations and final answer:

$SS_{xx} = 93.6, SS_{yy} = 14.945, SS_{xy} = 35.6$

$\hat{\beta}_1 = 0.3803$

$SSE = 1.4048$

$S^2 = 0.1756$

$S_{\beta_1} = 0.04331$

Critical value: $t_{\alpha/2,n-2} = t_{0.025,8} = 2.306$

Margin of Error: $E = t_{\alpha/2} * S_{\hat{\beta}_1} = 2.306 * 0.04331 = 0.099873$

Interval: $\left(\hat{\beta}_1 - E, \hat{\beta}_1 + E\right) = \left(0.3803 - 0.099873, 0.3803 + 0.099873\right) = \left(0.2804, 0.4802\right)$

Since both of the interval limits are positive, there seems to be a positive linear relation between number of dates and happiness. This is different from what we saw in problem 2, but in this study, they asked the questions in a different order. Finally, the interval says that the true slope is between 0.28 and 0.48, which means that a unit change in the number of dates produces between a 0.28 and 0.48 change on the happiness scale.