

STAT 2223-002

HOMEWORK 3

Fall 2018

Due: Tuesday, September 25, 2018

Please show all your steps, answers without clear and logical support work will be graded as 0.

Name \_\_\_\_\_

Class ID \_\_\_\_\_

Homework 3 contains 11 questions, with a total of 24 points.

Just a friendly reminder: Question 1, 8, 9 10 will be graded based on accuracy. Question 2-7 and 11 will be graded by completion. Homework 3 is a good preparation/review of your upcoming test.

1. (12 points) Seven employees at a large corporation were randomly selected and their years of education ( $x$ ) and annual salaries (in  $y$  thousand dollars) were collected and yielded the following

$$\sum x = 122, \quad \sum x^2 = 2210, \quad \sum y = 390, \quad \sum y^2 = 24826, \quad \sum xy = 7153.$$

- a. Compute  $SS_{xx}$ ,  $SS_{xy}$  and  $SS_{yy}$  (3 points)

- b. Calculate the sample correlation coefficient  $r$  between  $x$  and  $y$ . (1 point)

- c. Find the equation of the estimated regression line relating  $y$  to  $x$ . (2 points)

d. Test whether  $x$  is useful for predicting  $y$ , at  $\alpha = 0.05$ . (4 points)

1)  $H_0$ :                       $H_a$ :

2) Identify the test statistic and calculate its value

3) Construct the rejection region

4) Your decision

e. Construct 95% confidence interval for the population slope  $\beta_1$ . (2 points)

**Use the following information to answer Questions 2-6**

A career counselor wishes to estimate, based on the following sample data, the mean increase in the annual salary of people (in thousands of dollars) per additional year of education pursued.

Years, $x$	12	13	14	15	16	17
Salary, $y$	25	31	43	52	58	60

2. Which of the following is the null and the alternative hypotheses to test whether the number of years of education pursued is useful for predicting the annual salary?(1 point)

- a.  $H_0: \beta_1 = 0$  vs.  $H_a: \beta_1 \leq 0$
- b.  $H_0: \beta_1 = 0$  vs.  $H_a: \beta_1 \neq 0$
- c.  $H_0: \beta_1 = 0$  vs.  $H_a: \beta_1 > 0$
- d.  $H_0: \beta_1 = 0$  vs.  $H_a: \beta_1 \geq 0$
- e.  $H_0: \beta_1 = 0$  vs.  $H_a: \beta_1 < 0$

3. The value of the test statistic is: (1 point)

- a. -0.099.
- b. 9.064.
- c. -9.064.
- d. 10.064.
- e. 0.099.

4. The rejection region at the 1% significance level is: (1 point)

- a.  $[4.604, \infty)$
- b.  $(-\infty, -3.747] \cup [3.747, \infty)$
- c.  $(-\infty, -4.604] \cup [4.604, \infty)$
- d.  $(-\infty, -3.747]$
- e.  $[3.747, \infty)$

5. Which of the following is the correct decision regarding the rejection of the null hypothesis? (1 point)

- a. The decision is not to reject  $H_0$ , because the test statistic falls in the rejection region  $(-\infty, -3.747] \cup [3.747, \infty)$ .
- b. The decision is to reject  $H_0$ , because the test statistic falls in the rejection region  $[4.604, \infty)$ .
- c. The decision is not to reject  $H_0$ , because the test statistic falls in the rejection region  $[3.747, \infty)$ .

- d. The decision is to reject  $H_0$ , because the test statistic falls in the rejection region  $(-\infty, -3.747]$ .
- e. The decision is to reject  $H_0$ , because the test statistic falls in the rejection region  $(-\infty, -4.604] \cup [4.604, \infty)$ .

6. Identify the correct interpretation of the results obtained in the previous question. (1 point)

- a. The number of years of education pursued is useful for predicting the annual salary at 1% significance level but not at 5% significance level.
- b. The number of years of education pursued is not useful for predicting the annual salary at 1% significance level.
- c. The number of years of education pursued is useful for predicting the annual salary at 1% significance level.
- d. The number of years of education pursued is not useful for predicting the annual salary at any significance level.
- e. The number of years of education pursued is useful for predicting the annual salary at 1% and 5% significance levels but not at 10% significance level.

7. The coefficient of determination measures \_\_\_\_\_. (1 point)

- a. the proportion of the variability in  $y$  that is not accounted for by the linear relationship between  $x$  and  $y$ .
- b. the linear relationship between  $x$  and  $y$ .
- c. the proportion of the variability in  $y$  that is accounted for by the linear relationship between  $x$  and  $y$ .
- d. the nonlinear relationship between  $x$  and  $y$ .
- e. the expected change in  $y$  brought about by a unit increase in  $x$ .

### Use the following information to answer Questions 8-9

A career counselor wishes to estimate, based on the following sample data, the mean increase in the annual salary of people (in thousands of dollars) per additional year of education pursued.

Years, $x$	12	13	14	15	16	17
Salary, $y$	25	31	43	52	58	60

8. Using any of the three formulas, which of the following is the coefficient of determination? (1 point)

- a. 0.9808
- b. 0.0379
- c. 0.9788
- d. 0.9620
- e. 0.0212

9. Which of the following is the correct interpretation based on the coefficient of determination in the previous problem? (1 point)
- a. 98.08% of the variability in the annual salary is explained by the number of years of education pursued.
  - b. 96.2% of the variability in the annual salary is explained by the number of years of education pursued.
  - c. 3.79% of the variability in the annual salary is explained by the number of years of education pursued.
  - d. 97.88% of the variability in the annual salary is explained by the number of years of education pursued.
  - e. 2.12% of the variability in the annual salary is explained by the number of years of education pursued.

10. Find the 95% confidence interval for the mean value of  $y$  at  $x = 15$  (2 points)

11. Find the 95% prediction interval for the new observation of  $y$  at  $x = 15$  (2 points)