

Practice questions from Chapter 8 (Simple Regression Analysis) (Solutions)

1. The coefficient of determination (R-square) varies between -1 and + 1. **False** Coeff. of correlation varies between -1 and +1. Coefficient of determination varies between 0 and 1.
2. A value of correlation close to zero implies a weak relationship between two variables. **True**
3. A negative value of the covariance between two variables implies a negative relationship between them. **True**
4. In simple regression there is one dependent variable whereas in multiple regression there are many dependent variables. **False**. The number of dependent variables is always 1
5. The best fit line for scatterplot is also called the least squares line. **True**
6. $y = a + bx$ is a population regression model. **False**. This is the sample regression model. In the population regression model, we will use α and β for the coefficients and we will have an error term (ϵ) and we will not use y -hat as the dependent variable.
7. Each year a nationally recognized publication conducts its "Survey of America's Best Graduate and Professional Schools."

An academic advisor wants to predict the typical starting salary of a graduate at a top business school using the GMAT score of the school as a predictor variable. A simple linear regression of SALARY versus GMAT using 25 data points is shown below.

$$a = -92040 \quad b = 228 \quad s = 3213 \quad r^2 = .66 \quad r = .81 \quad df = 23 \quad t = 6.67$$

Give a practical interpretation of $r = .81$.

- A) 81% of the sample variation in SALARY can be explained by using GMAT in a straight -line model.
 - B) There appears to be a positive correlation between SALARY and GMAT.**
 - C) We estimate SALARY to increase 81% for every 1-point increase in GMAT.
 - D) We can predict SALARY correctly 81% of the time using GMAT in a straight -line model.
8. To investigate the relationship between yield of potatoes, y , and level of fertilizer application, x , a researcher divides a field into eight plots of equal size and applies differing amounts of fertilizer to each. The yield of potatoes (in pounds) and the fertilizer application (in pounds) are recorded for each plot. The data are as follows:

x	1	1.5	2	2.5	3	3.5	4	4.5
y	25	31	27	28	36	35	32	34

Summary statistics yield $SS(Total) = 112$ and $SS(Residual) = 52.476$.

 - a. Calculate the coefficient of determination (or R-square value).
 $1 - 52.476/112 = 1 - 0.468536 = 0.531464$
 - b. Calculate the coefficient of correlation. **Sqrt(0.531464)= 0.729**
 - c. What percent of the sample variation in y can be explained by the simple linear model? **53.146%**

9. Following is a regression output using Excel

Regression Statistics	
Multiple R	0.870807013
R Square	0.758304853
Adjusted R Square	0.697881066
Standard Error	1.450779757
Observations	6

ANOVA				
	df	SS	MS	F
Regression	1	26.41428571	26.41428571	12.54977376
Residual	4	8.419047619	2.104761905	
Total	5	34.83333333		

	Coefficients	Standard Error	t Stat	P-value
Intercept	3.638095238	1.66922253	2.179514818	0.094795901
X	1.228571429	0.346802694	3.542565985	0.023958073

- A. What is a? **3.638**
- B. What is b? **1.2285**
- C. What is the equation for the predicted value y? **$y = 3.638 + 1.2285X$**
- D. Interpret b. **As X increases by one unit, Y increases by average of 1.2285 units**
- E. What is the estimated standard error of the regression model? In other words, what is s?

1.45

- F. What is the estimated standard error of the least square slope b? **0.3468**

10. The following data was collected between a dependent variable y and an independent variable x

x_i	y_i
6	3
7	1
4	4
5	2
3	5

- a) What is the equation for the least square line? **$y = 7.5 - 0.9X$**
- b) What is y for $x = 4.5$? **3.45**
- c) What is y for $x = 2.5$? **5.7**
- d) What is y for $x = 2$? **5.25**
- e) What is the estimated standard error of the regression model? (s ?) **0.7958**
- f) What is the estimated standard error of the slope beta? **0.251661**
- g) What is the 95% CI for beta? **-1.7009 to -0.0991**
- h) What is the coefficient of correlation? **0.9**
- i) What is the coefficient of determination? **0.81**
- j) What is the t value for the coefficient b? **-3.57624**