

8.S: Regression (Solution)

1.d

2. A measure of the degree to which variation of one variable is related to variation in one or more other variables. The most commonly used correlation coefficient indicates the degree to which variation in one variable is described by a straight line relation with another variable.

Suppose that sample information is available on family income and Years of schooling of the head of the household. A correlation coefficient = 0 would indicate no linear association at all between these two variables. A correlation of 1 would indicate perfect linear association (where all variation in family income could be associated with schooling and vice versa).

3. a. 81% of the variation in the money spent for repairs is explained by the age of the auto

4. b. 16

5. The coefficient of determination is r^2 with $0 \leq r^2 \leq 1$, since $-1 \leq r \leq 1$.

6. True

7. d. on a scale from -1 to +1, the degree of linear relationship between the two variables is +.10

8. d. there exists no linear relationship between X and Y

9. Approximately 0.9

10. d. neither of the above changes will affect r .

11. Definition: A t test is obtained by dividing a regression coefficient by its standard error and then comparing the result to critical values for Students' t with Error df . It provides a test of the claim that $\beta_i = 0$ when all other variables have been included in the relevant regression model.

Example: Suppose that 4 variables are suspected of influencing some response. Suppose that the results of fitting $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + e_i$ include:

Table 8.S. 1

Variable	Regression coefficient	Standard error of regular coefficient
.5	1	-3
.4	2	+2
.02	3	+1
.6	4	-.5

t calculated for variables 1, 2, and 3 would be 5 or larger in absolute value while that for variable 4 would be less than 1. For most significance levels, the hypothesis $\beta_1 = 0$ would be rejected. But, notice that this is for the case when X_2 , X_3 , and X_4 have been included in the regression. For most significance levels, the hypothesis $\beta_4 = 0$ would be continued (retained) for the case where X_1 , X_2 , and X_3 are in the regression. Often this pattern of results will result in computing another regression involving only X_1 , X_2 , X_3 , and examination of the t ratios produced for that case.

12. c. those who score low on one test tend to score low on the other.

13. False. Since $H_0 : \beta = -1$ would not be rejected at $\alpha = 0.05$, it would not be rejected at $\alpha = 0.01$.

14. True

15. d

16. Some variables seem to be related, so that knowing one variable's status allows us to predict the status of the other. This relationship can be measured and is called correlation. However, a high correlation between two variables in no way proves that a cause-and-effect relation exists between them. It is entirely possible that a third factor causes both variables to vary together.

17. True

18. $Y_j = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_3 \cdot X_3 + b_4 \cdot X_4 + b_5 \cdot X_6 + e_j$

19. d. there is a perfect negative relationship between Y and X in the sample.

20. b. low

21. The precision of the estimate of the Y variable depends on the range of the independent (X) variable explored. If we explore a very small range of the X variable, we won't be able to make much use of the regression. Also, extrapolation is not recommended.

22. $\hat{y} = -3.6 + (3.1 \cdot 7) = 18.1$

23. Most simply, since -5 is included in the confidence interval for the slope, we can conclude that the evidence is consistent with the claim at the 95% confidence level.

Using a t test: $H_0 : B_1 = -5$ $H_A : B_1 \neq -5$ $t_{\text{calculated}} = \frac{-5 - (-4)}{1} = -1$ $t_{\text{critical}} = -1.96$.

Since $t_{\text{calc}} < t_{\text{crit}}$ we retain the null hypothesis that $B_1 = -5$.

24. True.

$t_{(\text{critical}, df=23, \text{two-tailed}, \alpha=.02)} = \pm 2.5$

$t_{\text{critical}, df=23, \text{two-tailed}, \alpha=.01} = \pm 2.8$

25.

1. $80 + 1.5 \cdot 4 = 86$

2. No. Most business statisticians would not want to extrapolate that far. If someone did, the estimate would be 110, but some other factors probably come into play with 20 years.

26. d. one quarter

27. b. $r = -.77$

28.

1. $-.72, .32$

2. the t value

3. the t value

29.

1. The population value for β_2 , the change that occurs in Y with a unit change in X_2 , when the other variables are held constant.

2. The population value for the standard error of the distribution of estimates of β_2 .

3. $.8, .1, 16 = 20 - 4$.

This page titled [8.S: Regression \(Solution\)](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [OpenStax](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.

- [13.12: Solutions](#) by OpenStax is licensed [CC BY 4.0](#). Original source: <https://openstax.org/details/books/introductory-business-statistics>.