

1. The CEO of a large metropolitan health-care facility would like to assess the effect of the recent implementation of the Six Sigma management approach on customer satisfaction. A random sample of 100 patients is selected from a list of patients who were at the facility the past week and also a year ago:

SATISFIED LAST YEAR	SATISFIED NOW		Total
	Yes	No	
Yes	67	5	72
No	20	8	28
Total	87	13	100

What is the p -value associated with the test statistic for testing $H_0 : \pi_1 \geq \pi_2$

$H_1 : \pi_1 < \pi_2$?

- 0.0014
- 0.9986
- 0.14
- 0.86
- 0.05

2. A fast food restaurant that sells burritos is concerned about the variability in the amount of filling that different employees place in the burritos. To achieve product consistency, they need the variance to be no more than 1.7 ounces². A sample of $n=18$ burritos showed a sample variance of 2.89 ounces². Using a 0.1 level of significance, the calculated test statistic and critical value, respectively, are

- 28.9 and 24.769
- 2.17 and 24.769
- 28.9 and 1.7
- 30.5 and 26.769
- 26.769 and 2.17

3. At the 0.025 level of significance and to test that satisfaction was lower last year, prior to introduction of Six Sigma management, the test statistic is a

- Z-test
- T-test
- χ^2 -test
- F-test
- Durbin-Watson test

4. Two candidates for governor participated in a televised debate. A political pollster recorded the preferences of 600 registered voters in a random sample prior to and after the debate:

Preference prior to debate	Preference After debate		Total
	Candidate A	Candidate B	
Candidate A	246	36	282
Candidate B	12	306	318
Total	258	342	600

To determine if there is evidence of difference in the proportion of voters who favored candidate A prior to and after the debate, the p-value associated with the test statistic is

- 0.0005.
- 0.2000
- 0.0300
- 0.1000
- 0.0001

5. A sample of eight earnings per share estimates is shown below:

Company	AT&T	Caterpillar	Kodak	Exxon	hp	IBM	McDonalds	Wal-Mart
Estimated Earnings per Share	2.92	4.65	4.27	3.09	3.57	7.04	2.64	1.74

Using the above information, answer the following:

Assuming $\alpha = 0.05$ and to test that the standard deviation, in the earnings per share estimates, exceeds 1.5, the critical value of the test is

- 14.0671
- 16.0128
- 12.0170
- 15.5073
- 17.5345

6. The least squares method minimizes which of the following?

- a. Error Sum of Squares
- b. Total Sum of Squares
- c. Regression Sum of Squares
- d. Total variation
- e. The regression equation

7. We are interested to fit a line for predicting the score of students using the study time. In theory, which way should the line slope?

- a. The slope of the line should be positive because there is a direct relationship between the two variables.
- b. The slope of the line should be negative because there is a direct relationship between the two variables.
- c. The slope of the line should be positive because there is an inverse relationship between the two variables.
- d. The slope of the line should be negative because there is an inverse relationship between the two variables.
- e. The slope of the line should be negative because there is strong relationship between the two variables

8. We are interested to fit a line for predicting the weight of adults using the time they spend in gym exercising. The data are available on weight and time spent in gym exercising for a sample of 17 adults. Suppose that the sample correlation coefficient was calculated and found to be -0.921 . Which of the following statements best explains this value?

- a. There is a strong inverse linear relationship between the two variables.
- b. There is a strong inverse relationship between the two variables.
- c. There is a weak direct relationship between the two variables.
- d. There is no relationship between the two variables.
- e. -92.1% of the variation in time spent in gym is explained by weight

An agent for a residential real estate company in a large city would like to be able to predict the monthly rental cost for apartments, based on the size of an apartment, as defined by square footage. The agent selects a sample of 25 apartments in a particular residential neighborhood. Summary data are:

$$\sum_{i=1}^{25} y_i = 34660, \quad \sum_{i=1}^{25} x_i = 28383, \quad \sum_{i=1}^{25} S_{xx} = 1999747, \quad \sum_{i=1}^{25} S_{yy} = 3139726, \quad \sum_{i=1}^{25} S_{xy} = 2130019$$

J. Use the above information to answer the following five questions:

9. What is the predicted monthly rent for an apartment that has 1,000 square feet
 - a. \$1242.10
 - b. \$842.50
 - c. \$1438.40
 - d. 8. \$1524
 - e. 9. \$2560.90

10. Suppose that the observed monthly rent for an apartment of size 1,000 square feet is \$1,425, the corresponding residual is
 - a. \$182.9
 - b. \$399.6
 - c. \$196.3
 - d. \$281.9
 - e. \$1318.8

11. The standard error of the estimate is
 - a. 194.6
 - b. 1048.7
 - c. 356.7
 - d. 256.6
 - e. 136.2

12. A 99% confidence interval estimate for the slope of the regression line is
 - a. [0.679, 1.451]
 - b. [0.831, 1.299]
 - c. [0.582, 1.792]
 - d. [0.679, 1.299]
 - e. [0.582, 1.451]

13. To test whether the monthly rental cost for apartments, is positively related to the size of an apartment, the measured value of the test statistic is

- a. 7.74
- b. 0.85
- c. 0.72
- d. 5.00
- e. 4.78

The summary of data on birth rate (births per 10,000 population) and GNP (gross national product per capita cross times 10 to the power of negative 3 end exponent) of 14 developing countries is given below:

$$\sum x = 17.05, \sum x^2 = 41.22, \sum y = 55, \sum xy = 52.37$$

Use the above information to answer the following two questions

14. Fit a linear regression line to predict the birth rate. The **slope** of the estimated regression line is found to be

- a. -0.714
- b. -0.627
- c. 0.627
- d. 0.860
- e. 4.779

15. Fit a linear regression line to predict the birth rate. The **intercept** of the estimated regression line is found to be

- a. 4.799
- b. -0.714
- c. -0.627
- d. 0.627
- e. 0.860

16. For the following regression equation, $\hat{y} = 100 - 12X_1 + 5X_2 - 4X_1X_2$, a unit increase in X_1 ,

while holding X_2 constant at a value of 2, decreases the value of y on average by:

- a. 20
- b. 90
- c. 11
- d. 100
- e. 89

The following regression output was obtained from a study of architectural firms. The dependent variable is the total amount of fees in millions of dollars

Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	3710.00	742	12.89	0.000
Error	46	2647.38	57.55		
Total	51	6357.38			

Coefficients				
Term	Coef	SE Coef	T-Value	P-Value
Constant	7.987	2.967	2.690	0.010
x1	0.122	0.031	3.920	0.000
x2	-1.220	0.053	-2.270	0.028
x3	-0.063	0.039	-1.610	0.114
x4	0.523	0.142	3.690	0.001
x5	-0.065	0.040	-1.620	0.112

- X_1 is the number of architects employed by the company.
 X_2 is the number of engineers employed by the company.
 X_3 is the number of years involved with health care projects.
 X_4 is the number of states in which the firm operates.
 X_5 is the percent of the firm's work that is health care-related.

Based on this information, answer the next four questions

17. Write out the regression equation.

- $\hat{y} = 7.987 + 0.122X_1 - 1.22X_2 - 0.063X_3 + 0.523X_4 - 0.065X_5$
- $\hat{y} = 0.122 - 1.22X_1 - 1.42X_2 - 0.063X_3 + 0.523X_4 - 0.065X_5$
- $\hat{y} = 7.987 + 0.122X_1 - 1.22X_2 - 0.063X_3 + 0.523X_4 + 0.065X_5$
- $\hat{y} = 7.987 + 0.122X_1 - 1.22X_2 + 0.063X_3 + 0.523X_4 - 0.065X_5$
- $\hat{y} = 7.987 - 0.122X_1 + 1.22X_2 - 0.063X_3 + 0.523X_4 - 0.065X_5$

18. How large is the sample and how many independent variables are there?

- Sample size is 52 and there are 5 independent variables.
- Sample size is 52 and there are 6 independent variables.
- Sample size is 51 and there are 5 independent variables.
- Sample size is 46 and there are 5 independent variables.
- Sample size is 46 and there are 6 independent variables.

19. Conduct a test of hypothesis to see if any of the set of regression coefficients could be different from 0. Use the .05 significance level. What is your conclusion?

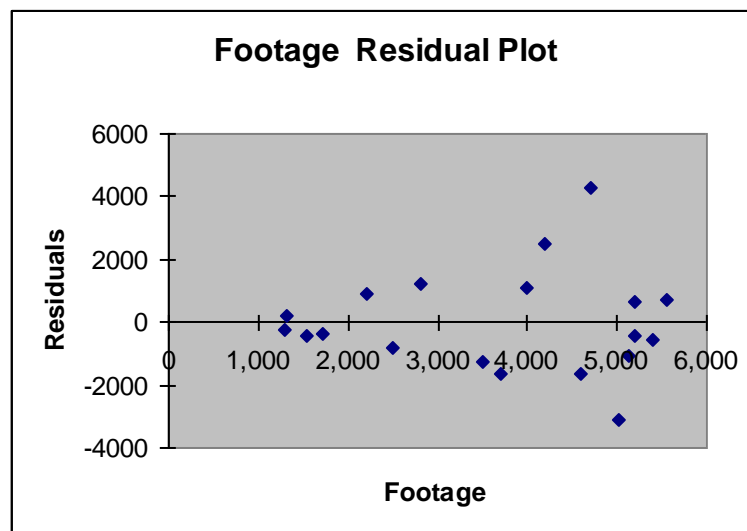
- We reject H_0 and conclude that at least one of the regression coefficients is different from 0.
- We do not reject H_0 and conclude that at least one of the regression coefficients is different from 0.

- c. We do not reject H_0 and conclude that none of the regression coefficients is different from 0.
- d. We reject H_1 and conclude that at least one of the regression coefficients is different from 0.
- e. We cannot tell what our decision will be from the information given.

20. Conduct a test of hypothesis for each independent variable. Use the .05 significance level.

- a. Variables x_1 , x_2 and x_4 are significantly affecting y because their respective p -values are less than the significance level 0.05. On contrary, x_3 and x_5 are insignificant.
- b. Variables x_1 , x_3 and x_4 are significantly affecting y because their respective p -values are less than the significance level 0.05. On contrary, x_2 and x_5 are insignificant
- c. Variables x_1 , x_2 and x_5 are significantly affecting y because their respective p -values are less than the significance level 0.05. On contrary, x_3 and x_4 are insignificant
- d. Variables x_2 , x_3 and x_4 are significantly affecting y because their respective p -values are less than the significance level 0.05. On contrary, x_1 and x_5 are insignificant
- e. We cannot tell what our decision will be from the information given

21. Based on the residual plot below, you will conclude that there might be a violation of which of the following assumptions.



- a. Homoscedasticity
- b. Linearity of the relationship
- c. Normality of errors
- d. Independence of errors
- e. None of the above

An economist is interested in relating the consumption for an economy (in \$ billions) to the gross domestic product (GDP) (in \$ billions) and aggregate Price (consumer price index). The Microsoft Excel output of the regression analysis is partially reproduced below:

SUMMARY OUTPUT					
Regression Statistics					
Multiple R	0.991				
R Square	0.982				
Adjusted R Square	0.976				
Standard Error	0.299				
Observations	10				
ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Signif F</i>
Regression	2	33.4163	16.7082	186.325	0.0001
Residual	7	0.6277	0.0897		
Total	9	34.0440			
	<i>Coeff</i>	<i>StdError</i>	<i>t Stat</i>	<i>P-value</i>	
Intercept	-0.0861	0.5674	-0.152	0.8837	
GDP	0.7654	0.0574	13.340	0.0001	
Price	-0.0006	0.0028	-0.219	0.8330	

Based on the above information, answer the next four questions.

22. The p-value for the regression model as a whole is

- 0.0001
- 0.8837
- 0.8330
- 0.0897
- 0.0574

23. The estimated mean consumption level for an economy with GDP equal to \$4 billion and an aggregate Price index of 150, is

- 2.8855
- 114.72
- 28.855
- 11.472
- 124.77

24. One economy in the sample had an aggregate consumption level of \$4 billion, a GDP of \$6 billion, and an aggregate Price level of 200. What is the residual for this data point?

- a. -0.386
- b. 2.8855
- c. 3.806
- d. -28.85
- e. 0.836

25. To test for the significance of the coefficient on aggregate price, the value of the relevant t-statistic is

- a. -0.219
- b. 13.340
- c. 0.5674
- d. -0.152
- e. 0.0574