

Unit 3 : Week 3 - Quiz 3

[Save Answers](#)[Submit for Grading](#)

1. One purpose of regression is to understand the relationship between variables.
(Points : 1)

- ☐ True
☐ False

2. One purpose of regression is to predict the value of one dependent variable based one or more independent variables.
(Points : 1)

- ☐ True
☐ False

3. A scatter diagram is a graphical depiction of the relationship between the dependent and independent variables.
(Points : 1)

- ☐ True
☐ False

4. There is a strong relationship between variables if the data points lie in a straight line.
(Points : 1)

- ☐ True
☐ False

5.

A Random error that can be predicted in regression analysis.

(Points : 1)

- ☐ True
☐ False

6. The sum of the squared errors is maximized in the regression line.
(Points : 1)

- ☐ True
☐ False

7. The SSR indicates how much of the total variability in the independent variable is explained by the regression model.
(Points : 1)

- ☐ True
- ☐ False

8. The coefficient of determination takes on values between -1 and + 1.
(Points : 1)

- ☐ True
- ☐ False

9. When we look at the residuals, we are also looking at the errors. (Points : 1)

- ☐ True
- ☐ False

10.

The regression model assumes the errors are normally distributed.

(Points : 1)

- ☐ True
- ☐ False

11. Exponential smoothing is always a superior forecasting method to regression, so exponential smoothing should be used whenever the appropriate software is available.
(Points : 1)

- ☐ True
- ☐ False

12.

Time-series models attempt to predict the future by using historical data.

(Points : 1)

- ☐ True
- ☐ False

13. The most common quantitative causal model is regression analysis. (Points : 1)

- ☐ True
- ☐ False

14. Time-series models enable the forecaster to include specific representations of various qualitative and quantitative factors. (Points : 1)

- ☐ True
- ☐ False

15. Four components of time series are trend, moving average, exponential smoothing, and seasonality. (Points : 1)

- ☐ True
☐ False

16. A scatter diagram for a time series may be plotted on a two-dimensional graph with the vertical axis representing the variable to be forecast (such as sales). (Points : 1)

- ☐ True
☐ False

17.

Exponential smoothing cannot be used for data with a trend.

(Points : 1)

- ☐ True
☐ False

18. A seasonal index must be between -1 and +1.

(Points : 1)

- ☐ True
☐ False

19. The process of isolating linear trend and seasonal factors to develop a more accurate forecast is called regression. (Points : 1)

- ☐ True
☐ False

20. Adaptive smoothing is analogous to exponential smoothing where the coefficients α and β are periodically updated to improve the forecast. (Points : 1)

- ☐ True
☐ False

21. Which of the following is not classified as a qualitative forecasting model?

- A) exponential smoothing
B) Delphi method
C) jury of executive opinion
D) sales force composite
E) consumer market survey

(Points : 1)

- ☐ A)
☐ B)
☐ C)
☐ D)
☐ E)

22. Which of the following statements about scatter diagrams is true?

- A) Time is always plotted on the y-axis.
- B) It can depict the relationship among three variables simultaneously.
- C) It is helpful when forecasting with qualitative data.
- D) The variable to be forecasted is placed on the y-axis.
- E) It is not a good tool for understanding time-series data.

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

23. Enrollment in a particular class for the last four semesters has been 115, 120, 110, and 130. Suppose a one-semester moving average was used to forecast enrollment (this is sometimes referred to as a naïve forecast). Thus, the forecast for the second semester would be 115, for the third semester it would be 120, and for the last semester it would be 110. What would the MSE be for this situation?

- A) 196.00
- B) 230.67
- C) 175.00
- D) 11.60
- E) None of the above

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

24. When is the exponential smoothing model equivalent to the naïve forecasting model?

- A) $\alpha = 0$
- B) $\alpha = 0.5$
- C) $\alpha = 1$
- D) during the first period in which it is used
- E) never

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

25. Assume that you have tried three different forecasting models. For the first, the MAD = 2.5, for the second, the MSE = 10.5, and for the third, the MAPE = 2.7. We can then say:

- A) the third method is the best.
- B) the second method is the best.
- C) methods one and three are preferable to method two.
- D) method two is least preferred.
- E) None of the above

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

26. As one increases the number of periods used in the calculation of a moving average,

- A) greater emphasis is placed on more recent data.
- B) less emphasis is placed on more recent data.
- C) the emphasis placed on more recent data remains the same.
- D) it requires a computer to automate the calculations.
- E) one is usually looking for a long-term prediction.

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

27. Sales for boxes of Girl Scout cookies over a 4-month period were forecasted as follows: 100, 120, 115, and 123. The actual results over the 4-month period were as follows: 110, 114, 119, 115. What was the MAD of the 4-month forecast?

- A) 0
- B) 5
- C) 7
- D) 108
- E) None of the above

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

28. Which of the following is not considered to be one of the components of a time series?

- A) trend
- B) seasonality
- C) variance
- D) cycles
- E) random variations

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

29. Demand for soccer balls at a new sporting goods store is forecasted using the following regression equation: $Y = 98 + 2.2X$ where X is the number of months that the store has been in existence. Let April be represented by $X = 4$. April is assumed to have a seasonality index of 1.15. What is the forecast for soccer ball demand for the month of April (rounded to the nearest integer)?

- A) 123
- B) 101
- C) 107
- D) 115
- E) None of the above

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

30. A tracking signal was calculated for a particular set of demand forecasts. This tracking signal was positive. This would indicate that

- A) demand is greater than the forecast.
- B) demand is less than the forecast.
- C) demand is equal to the forecast.
- D) the MAD is negative.
- E) None of the above

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

31. Which of the following statements is true regarding a scatter diagram?

- A) It provides very little information about the relationship between the regression variables.
- B) It is a plot of the independent and dependent variables.
- C) It is a line chart of the independent and dependent variables.
- D) It has a value between -1 and +1.
- E) It gives the percent of variation in the dependent variable that is explained by the independent variable.

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

32. If computing a causal linear regression model of $Y = a + bX$ and the resultant r^2 is very near zero, then one would be able to conclude that

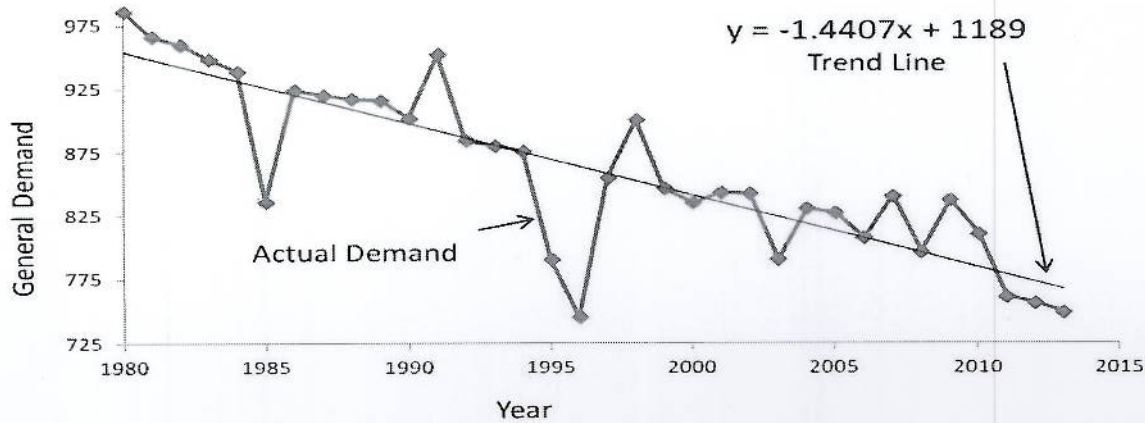
- A) $Y = a + bX$ is a good forecasting method.
- B) $Y = a + bX$ is not a good forecasting method.
- C) a multiple linear regression model is a good forecasting method for the data.
- D) a multiple linear regression model is not a good forecasting method for the data.
- E) None of the above

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

33.

The diagram below illustrates data with a



- A) negative correlation coefficient.
- B) zero correlation coefficient.
- C) positive correlation coefficient.
- D) correlation coefficient equal to +1.
- E) None of the above

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

34. Which of the following is not an assumption of the regression model?

- A) The errors are independent.
- B) The errors are normally distributed.
- C) The errors have constant variance.
- D) The mean of the errors is zero.
- E) The errors should have a standard deviation equal to one.

(Points : 1)

- ☐ A

- ☐ B
- ☐ C
- ☐ D
- ☐ E

35. The residual plot in a good regression model shows

- A) a cone pattern.
- B) a random pattern.
- C) an arched pattern.
- D) an increasing pattern.
- E) a decreasing pattern.

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

36. A prediction equation for the number of people (in 1,000s) and the annual waste total (in metric tons) was performed using simple linear regression. In the regression printout shown below, what can be said about the level of significance for the overall model?

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.715257514
R Square	0.511593311
Adjusted R Square	0.496330602
Standard Error	50.43909112
Observations	34

ANOVA

	df	SS	MS	F	Significance F
Regression	1	85276.17996	85276.17996	33.51917	1.99773E-06
Residual	32	81411.26122	2544.101913		
Total	33	166687.4412			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%
Intercept	612.3757625	43.18367091	14.18072502	2.39E-15	524.4135038
# People	1.128638579	0.194943283	5.789574096	2E-06	0.731552108

- A) Population size is not a good predictor for the amount of waste.
 B) The significance level for the intercept indicates the model is not valid.
 C) The significance level for population indicates the slope is equal to zero.
 D) The significance level for population indicates the slope is not equal to zero.
 E) None of the above

(Points : 1)

- ☐ A
☐ B
☐ C
☐ D
☐ E

37. A healthcare executive is using regression to predict total revenues. She has decided to include both patient length of stay and insurance type in her model. Insurance type can be grouped into the following categories: Medicare, Medicaid, Managed Care, Self-Pay, and Charity. Which of the following is true?

- A) Insurance type will be represented in the regression model by five binary variables.
 B) Insurance type will be represented in the regression model by six dummy variables.
 C) Insurance type will be represented in the regression model by five dummy variables.
 D) Insurance type will be represented in the regression model by four binary variables.
 E) Neither binary nor dummy variables are necessary for the regression model.

(Points : 1)

- ☐ A
☐ B

- ☐ C
- ☐ D
- ☐ E

38. A hospital administrator is using regression to predict total revenues for the hospital she is on the board of director's for. She is deciding whether or not to include both patient length of stay and insurance type in her model. Her first regression model only included patient length of stay. The resulting r^2 was .83, with an adjusted r^2 of .82 and her level of significance was .003. In the second model, she included both patient length of stay and insurance type. The r^2 was .84 and the adjusted r^2 was .80 for the second model and the level of significance did not change. Which of the following statements is true?

- A) The second model is a better model.
- B) The first model is a better model.
- C) The r^2 increased when additional variables were added because these variables significantly contribute to the prediction of total revenues.
- D) The adjusted r^2 always increases when additional variables are added to the model.
- E) None of the above statements are true.

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

39. Which of the following is not a common pitfall of regression?

- A) If the assumptions are not met, the statistical tests may not be valid.
- B) Nonlinear relationships cannot be incorporated.
- C) Two variables may be highly correlated to one another but one is not causing the other to change.
- D) Concluding that a statistically significant relationship implies practical value.
- E) Using a regression equation beyond the range of X is very questionable.

(Points : 1)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

40. Which of the following statements (are) is not true about regression models? (More than one answer may apply)

- A) Estimates of the slope are found from sample data.
- B) The regression line minimizes the sum of the squared errors.
- C) The error is found by subtracting the actual data value from the predicted data value.
- D) The dependent variable is the explanatory variable.
- E) The intercept coefficient is not typically interpreted.

(Points : 2)

- ☐ A
- ☐ B
- ☐ C
- ☐ D
- ☐ E

41. Describe the purpose and structure of a scatter diagram. (Points : 4)

Spellchecker

42. Explain the purposes of regression models. (Points : 3)

Spellchecker

43. What is the difference between simple linear regression models and multiple regression models? (Points : 2)

Spellchecker

44. Explain what an error is in simple linear regression (Points : 1)

Spellchecker

45. Demand for a particular type of battery fluctuates from one week to the next. A study of the last six weeks provides the following demands (in dozens): 4, 5, 3, 2, 8, 10 (last week).

Forecast demand for the next week using a two-week moving average. (Round to nearest hundredth) (Points : 1)

46. Demand for a particular type of battery fluctuates from one week to the next. A study of the last six weeks provides the following demands (in dozens): 4, 5, 3, 2, 8, 10 (last week).

Forecast demand for the next week using a three-week moving average. (Round to nearest hundredth) (Points : 1)

47.

Use simple exponential smoothing with $\alpha = 0.3$ to forecast battery sales for February through May. Assume

that the forecast for January was for 22 batteries.

Month	Automobile Battery Sales	Forecasted Battery Sales
January	42	-----
February	33	
March	28	
April	59	

(Points : 5)

Spellchecker

48.

The following table represents the number of applicants at popular private college in the last four years.

Month	New members
2007	10,067
2008	10,940
2009	11,116
2010	10,999

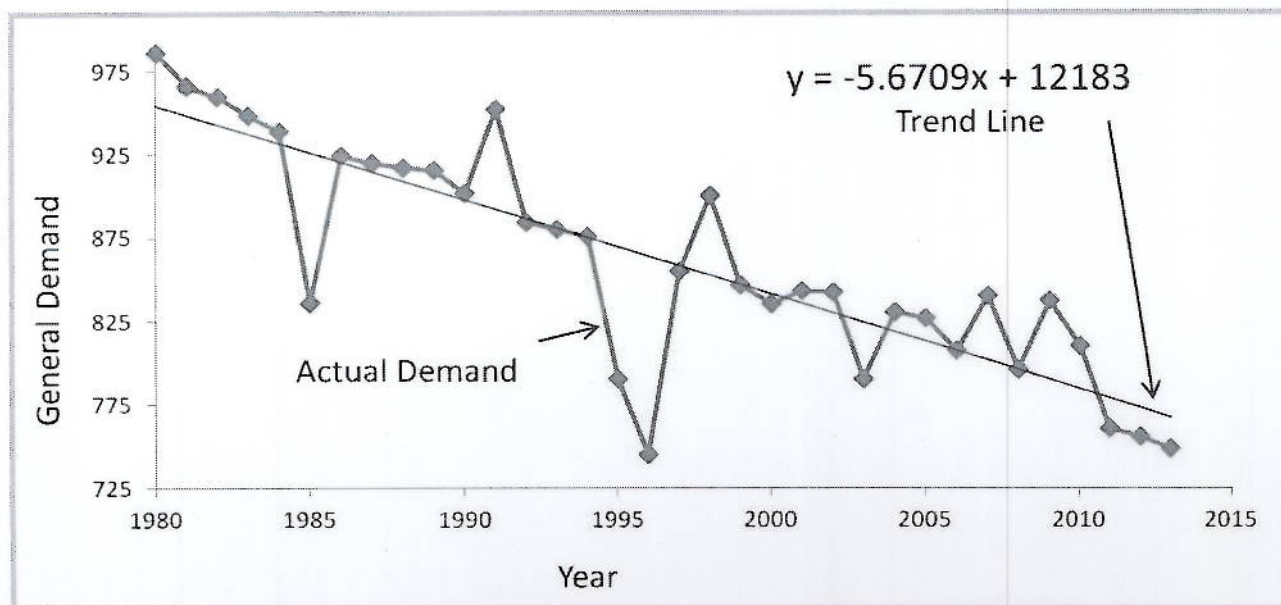
Assuming $\alpha = 0.2$, $\beta = 0.3$, an initial forecast of 10,000 for 2007, and an initial trend adjustment of 0 for 2007, use exponential smoothing with trend adjustment to come up with a forecast for 2011 on the number of applicants.

Forecast for 2011 = _____ (Fill in the blank with correct answer)

(Points : 3)

49.

Using the diagram below. The predicted demand of 2014 using the regression model is ____ (Round answer to nearest whole number)



(Points : 1)

50.

George Crowder is conducting research on monthly expenses for a waste management company. His *dependent variable is annual waste in metric tons* while his *independent variable is number of people (in thousands)* that live in 25 selected cities. Below is his Excel output.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.616
R Square	0.379
Adjusted R Square	0.352
Standard Error	60.349
Observations	25

ANOVA					
	df	SS	MS	F	Significance F
Regression	1	51193.075	51193.075	14.056	0.001
Residual	23	83764.925	3641.953		
Total	24	134958			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%
Intercept	559.173	73.500	7.608	0.000	407.126
# People	1.379	0.368	3.749	0.001	0.618

- (a) What is the prediction equation for the annual metric tons of waste based on the population?
- (b) Based on his model, each additional (1000) increase in people the predicted waste in metric tons is increased by how much?
- (c) Based on the significance F-test, is this model a good prediction equation?
- (d) What percent of the variation in waste is explained by the size of the population?
- (e) Can the null hypothesis that the slope is zero be rejected? Why or why not?
- (f) What is the value of the correlation coefficient and what conclusions can you draw about the model based on the correlation coefficient.

(Points : 12)

Spellchecker

Save Answers

Submit for Grading