

Linear regression interview questions

This document provides questions and answers related to concepts of linear regression.

1. In simple linear regression, there is _____ dependent variable and _____ independent variable(s)
 1. One, multiple
 2. Multiple, one
 3. One, one
 4. Multiple, multiple

2. In multi-linear regression, there is _____ dependent variable and _____ independent variable(s)
 1. Multiple, one
 2. One, multiple
 3. Multiple, multiple
 4. One, one

3. It is OK to add independent variables to a multi-linear regression model as it increases the explained variance of the model and makes the model more efficient
 1. True
 2. False

4. Linear or multilinear regression helps in predicting _____
 1. Continuous valued output
 2. Discrete valued output

5. Linear regression analysis helps in studying _____ relationship between variables.
 1. Deterministic
 2. Statistical

6. Linear regression analysis helps in doing which of the following?
 1. Causal analysis
 2. Effects in forecasting
 3. Forecasting trends
 4. All of the above

7. The best fit line is achieved by finding values of the parameters which minimizes the sum of _____
1. Sum of squared regression (SSR)
 2. Sum of squared residuals/errors (SSE)
 3. Sum of squares total (SST)
8. Best fit line is also termed as _____
1. The maximum squares regression line
 2. Least squares regression line
9. Which of the following can be used to understand the statistical relationship between dependent and independent variables in linear regression?
1. Coefficient of determination
 2. Correlation coefficient
 3. Both of the above
 4. None of the above
10. It is absolutely OK to state that correlation does imply causation
1. True
 2. False
11. The value of coefficient of determination, R-squared, is _____
1. Less than 0
 2. Greater than 1
 3. Between 0 and 1
12. Which of the following can be used to understand the positive or negative relationship between dependent and independent variables
1. Coefficient of determination
 2. Pearson correlation coefficient
 3. Both of the above
 4. None of the above
13. The goal of the regression model is to achieve the R-squared value _____
1. Closer to 0
 2. Closer to 1
 3. More than 1

14. Pearson correlation coefficient does always have a positive value
1. True
 2. False
15. Value of Pearson correlation coefficient near to zero represents the fact there is a stronger relationship between dependent and independent variables
1. True
 2. False
16. Population correlation coefficient and sample correlation coefficient are one and the same
1. True
 2. False
17. The value of the Pearson correlation coefficient falls in the range of _____
1. 0 and 1
 2. 0 and -1
 3. -1 and 1
 4. 1 and 2
18. The large value of R-squared can be safely interpreted as the fact that the estimated regression line fits the data well.
1. True
 2. False
19. The value of R-squared does not depend upon the data points; Rather it only depends upon the value of parameters
1. True
 2. False
20. The value of correlation coefficient and coefficient of determination is used to study the strength of the relationship in _____
1. Samples only
 2. Both Samples and Population
 3. Population only

21. Which of the following tests can be used to determine whether a linear association exists between the dependent and independent variables in a simple linear regression model?
1. T-test
 2. ANOVA F-test
 3. Both of the above
 4. None of the above
22. In order to estimate the population parameter, the null hypothesis is that the population parameter is _____ to zero?
1. Equal
 2. Not equal
23. Which of the following can be used for learning the value of parameters for the regression model for population and not just the samples?
1. Hypothesis testing
 2. Confidence intervals
 3. Both of the above
 4. None of the above
24. The value of R-Squared _____ with the addition of every new independent variable?
1. May increase or decrease
 2. Always increases
 3. Always decreases
25. In order to reject the null hypothesis while estimating the population parameter, p-value has to be _____ given 0.05 is set as the significance level
1. More than 0.05
 2. Less than 0.05
26. The value of _____ may increase or decrease based on whether a predictor variable enhances the model or not
1. R-squared
 2. Adjusted R-squared
27. The value of Adjusted R-squared _____ if the predictor variable enhances the model less than what is predicted by chance?
1. Increases
 2. Decreases

28. In regression model t-tests, the value of t-test statistics is equal to _____?

1. Coefficient divided by Standard error of the coefficient
2. Standard error of coefficient divided by coefficient
3. Coefficient plus standard error of the coefficient

29. In ANOVA test for regression, degrees of freedom (regression) is _____

1. Equal to the number of parameters being estimated
2. One more than the number of parameters being estimated
3. One less than the number of parameters being estimated

30. In ANOVA test for regression, degrees of freedom (regression) is _____

1. Equal to the number of predictor variables
2. One more than the number of predictor variables
3. One less than the number of predictor variables

31. For SST as the sum of squares total, SSE as the sum of squared errors, and SSR as the sum of squares regression, which of the following is correct?

1. $SST = SSR - SSE$
2. $SST = SSR + SSE$
3. $SST = SSR/SSE$

32. The value of the R-squared or coefficient of determination is which of the following?

1. SSR / SST
2. SSE / SST

33. Mean squared error can be calculated as _____

1. Sum of squares residuals or error/degrees of freedom
2. Sum of squares regression/ degrees of freedom
3. Sum of squares total/ degrees of freedom

34. Sum of Squares Regression (SSR) is _____

1. Sum of Squares of predicted value minus the average value of the dependent variable
2. Sum of Squares of Actual value minus predicted value
3. Sum of Squares of Actual value minus the average value of the dependent variable

35. Sum of Squares Error (SSE) is _____

1. Sum of Squares of predicted value minus the average value of the dependent variable
2. Sum of Squares of Actual value minus predicted value
3. Sum of Squares of Actual value minus the average value of the dependent variable

36. Sum of Squares Total (SST) is _____

1. Sum of Squares of predicted value minus the average value of the dependent variable
2. Sum of Squares of Actual value minus predicted value
3. Sum of Squares of Actual value minus the average value of the dependent variable

37. _____ the value of the sum of squares regression (SSR), better the regression model

1. Greater
2. Lesser

38. The objective for regression model is to minimize _____ and maximize _____

1. SSR, SSE
2. SSE, SSR
3. SSR, SST
4. SSE, SST

39. Which of the following can be used to test the hypothesis that there exists a linear regression model with at least one predictor variable?

1. F-test
2. T-test

40. Which of the following is the ratio of explained variance and unexplained variance in relation to doing hypothesis testing with regression model?

1. T-statistics
2. F-statistics

Answers: 1.3, 2.2, 3.2, 4.1, 5.2, 6.4, 7.2, 8.2, 9.3, 10.2, 11.3, 12.3, 13.2, 14.2, 15.2, 16.2, 17.3, 18.1, 19.1, 20.2, 21.3, 22.1, 23.3, 24.2, 25.2, 26.2, 27.2, 28.1, 29.3, 30.3, 31.2, 32.1, 33.1, 34.1, 35.2, 36.3, 37.1, 38.2, 39.1, 40.2