Linear Regression Interview Questions

These questions can be found as practice tests on our website, https://vitalflux.com, on this page, 40 Linear Regression Interview Questions for Data Scientists.

1. In _____ regression, there is _____ dependent variable and _____ independent variable(s)
    - Simple linear, one, multiple
    - Multiple, multiple, one
    - Simple linear, one, one
    - Multiple, one, multiple
2. In _____ regression, there is _____ dependent variable and _____ independent variable(s)
    - Simple linear, multiple, one
    - Simple linear, one, multiple
    - Multiple, one, multiple
    - Multiple, multiple, multiple
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 model more effcient
    - True
    - False
4. Linear or multilinear regression helps in predicting _____
    - Continuous valued output
    - Discrete valued output
5. Regression analysis helps in studying _____ relationship between variables.
    - Deterministic
    - Statistical
6. Regression analysis helps in doing which of the following?
    - Causal analysis
    - Effects in forecasting
    - Forecasting trends
    - All of the above
7. The best fit line is achieved by finding values of the parameters which minimizes the sum of _____
    - Prediction errors
    - Squared prediction errors
8. Best fit line is also termed as _____
    - Maximum squares regression line
    - 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?
    - Coefficient of determination
    - Correlation coefficient
    - Both of the above
    - None of the above

10. It is absolutely OK to state that correlation does imply causation
    - True
    - False
11. The value of coefficient of determination, R-squared, is _____
    - Less than 0
    - Greater than 1
    - Between 0 and 1
12. Which of the following can be used to understand the positive or negative relationship between dependent and independent variables
    - Coefficient of determination
    - Pearson correlation coefficient
13. The goal of the regression model is to achieve the R-squared value _____
    - Closer to 0
    - Closer to 1
    - More than 1
    - Less than 1
14. Pearson correlation coefficient is _____ to coefficient of determination
    - Directly proportional
    - Inversely proportional
15. Pearson correlation coefficient does always have positive value
    - True
    - False
16. Value of Pearson correlation coefficient near to zero represents the fact there is a stronger relationship between dependent and independent variables
    - True
    - False
17. Population correlation coefficient and sample correlation coefficient are one and the same
    - True
    - False
18. The value of Pearson correlation coefficient falls in the range of _____
    - 0 and 1
    - 0 and -1
    - -1 and 1
    - 1 and 2
19. The value of correlation coefficient and R-squared remains same for all samples of data
    - True
    - False
20. The large value of R-squared can be safely interpreted as the fact that estimated regression line fits the data well.
    - True
    - False
21. The value of R-squared does not depend upon the data points; Rather it only depends upon the value of parameters
    - True
    - False
22. The value of correlation coefficient and coefficient of determination is used to study the strength of relationship in _____
    - Samples only

    - Both Samples and Population
    - Population only
23. 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?
    - T-test
    - ANOVA F-test
    - Both of the abov
    - None of the abovee
24. In order to estimate population parameter, the null hypothesis is that the population parameter is _____ to zero?
    - Equal
    - Not equal
25. Which of the following can be used for learning the value of parameters for regression model for population and not just the samples?
    - Hypothesis testing
    - Confidence intervals
    - Both of the above
    - None of the above
26. The value of R-Squared _____ with addition of every new independent variable?
    - May increase or decrease
    - Always increases
    - Always decreases
27. In order to reject the null hypothesis while estimating population parameter, p-value has to be _____
    - More than 0.05
    - Less than 0.05
28. The value of _____ may increase or decrease based on whether a predictor variable enhances the model or not
    - R-squared
    - Adjusted R-squared
29. The value of Adjusted R-squared _____ if the predictor variable enhances the model less than what is predicted by chance?
    - Increases
    - Decreases
30. In regression model t-tests, the value of t-test statistics is equal to _____?
    - Coefficient divided by Standard error of coefficient
    - Standard error of coefficient divided by coefficient
    - Coefficient plus standard error of coefficient
31. In ANOVA test for regression, degrees of freedom (regression) is _____
    - Equal to number of parameters being estimated
    - One more than the number of parameters being estimated
    - One less than the number of parameters being estimated
32. In ANOVA test for regression, degrees of freedom (regression) is _____
    - Equal to number of predictor variables
    - One more than the number of predictor variables
    - One less than the number of predictor variables
33. For SST as sum of squares total, SSE as sum of squared errors and SSR as sum of squares regression, which of the following is correct?

- SST = SSR – SSE
- SST = SSR + SSE
- SST = SSR/SSE

34. The value of coefficient of determination is which of the following?
    - SSR / SST
    - SSE / SST
35. Mean squared error can be calculated as _____
    - Sum of squares error / degrees of freedom
    - Sum of squares regression/ degrees of freedom
    - Sum of squares total/ degrees of freedom
36. Sum of Squares Regression (SSR) is _____
    - Sum of Squares of predicted value minus average value of dependent variable
    - Sum of Squares of Actual value minus predicted value
    - Sum of Squares of Actual value minus average value of dependent variable
37. Sum of Squares Error (SSE) is _____
    - Sum of Squares of predicted value minus average value of dependent variable
    - Sum of Squares of Actual value minus predicted value
    - Sum of Squares of Actual value minus average value of dependent variable
38. Sum of Squares Total (SST) is _____
    - Sum of Squares of predicted value minus average value of dependent variable
    - Sum of Squares of Actual value minus predicted value
    - Sum of Squares of Actual value minus average value of dependent variable
39. _____ the value of sum of squares regression (SSR), better the regression model
    - Greater
    - Lesser
40. The objective for regression model is to minimize _____ and maximize _____
    - SSR, SSE
    - SSE, SSR
    - SSR, SST
    - SSE, SST