1. In a linear equation, what is the difference between a dependent variable and an independent variable?

If x and y are two variables in an algebraic equation and every value of x is linked with any other value of y, then 'y' value is said to be a function of x value known as an independent variable, and 'y' value is known as a dependent variable.

2. What is the concept of simple linear regression? Give a specific example.

Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

3. In a linear regression, define the slope.

The slope of a regression line (b) represents the rate of change in y as x changes.

4. Determine the graph's slope, where the lower point on the line is represented as (3, 2) and the higher point is represented as (2, 2).

5. In linear regression, what are the conditions for a positive slope?

Slope shows both steepness and direction. With positive slope the line moves upward when going from left to right.

6. In linear regression, what are the conditions for a negative slope?

If the slope is negative, y decreases as x increases and the function runs downhill.

7. What is multiple linear regression and how does it work?

Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable.

8. In multiple linear regression, define the number of squares due to error.

9. In multiple linear regression, define the number of squares due to regression.

Regression sum of squares (also known as the sum of squares due to regression or explained sum of squares) The regression sum of squares describes how well a regression model represents the modeled data.

In a regression equation, what is multicollinearity?

Multicollinearity happens when independent variables in the regression model are highly correlated to each other. It makes it hard to interpret of model and also creates an overfitting problem. It is a common assumption that people test before selecting the variables into the regression model.

11. What is heteroskedasticity, and what does it mean?

Heteroscedasticity means unequal scatter. In regression analysis, we talk about heteroscedasticity in the context of the residuals or error term. Specifically, heteroscedasticity is a systematic change in the spread of the residuals over the range of measured values.

12. Describe the concept of ridge regression.

Ridge regression is a model tuning method that is used to analyse any data that suffers from multicollinearity. This method performs L2 regularization. When the issue of multicollinearity occurs, least-squares are unbiased, and variances are large, this results in predicted values being far away from the actual values.

13. Describe the concept of lasso regression.

Lasso regression is a type of linear regression that uses shrinkage. Shrinkage is where data values are shrunk towards a central point, like the mean. The lasso procedure encourages simple, sparse models (i.e. models with fewer parameters).

14. What is polynomial regression and how does it work?

In statistics, polynomial regression is a form of regression analysis in which the relationship between the independent variable x and the dependent variable y is modelled as an nth degree polynomial in x.

15. Describe the basis function.

16. Describe how logistic regression works.

Logistic regression is a Machine Learning classification algorithm that is used to predict the probability of certain classes based on some dependent variables. In short, the logistic regression model computes a sum of the input features (in most cases, there is a bias term), and calculates the logistic of the result.