socrative
A showbie PRODUCT

Name _____
Date _____

Regression

Score _____

1. In simple linear regression, the slope m represents:

(A) The predicted value of Y when $X = 0$

(B) How much Y changes for a one-unit change in X

(C) The average of all Y values

(D) The error between predicted and actual values

2. Which of the following is not an assumption of simple linear regression?

(A) Linearity

(B) Independence of errors

(C) Homoscedasticity

(D) Multicollinearity between X and Y

3. In the regression equation $Y = mX + c$, c represents:

(A) The slope

(B) The intercept

(C) The coefficient of determination

(D) The standard error

4. The purpose of regularization in regression is to:

(A) Increase the number of features in the model

(B) Increase training error

(C) Reduce overfitting by penalizing large coefficients

(D) Eliminate the intercept term

5. Which metric measures how well a regression model fits the data?

(A) RMSE

(B) MAE

(C) R^2 score

(D) p-value

6. In multiple linear regression, multicollinearity occurs when:

(A) The dependent variable is categorical

(B) Independent variables are highly correlated

(C) The residuals are normally distributed

(D) The slope is negative

7. Which of the following is the main objective of linear regression?

(A) To classify data points into categories

(B) To estimate the relationship between variables

(C) To maximize the variance of features

(D) To reduce the dimensionality of the dataset

8. What does a negative slope in a regression equation indicate?

(A) As X increases, Y decreases

(B) As X increases, Y increases

(C) The model is invalid

(D) The intercept is negative

9. In the equation
$Y = \beta_0 + \beta_1 X + \epsilon$
,
ϵ
represents:

(A) Predicted value

(B) Coefficient

(C) Error term (residual)

(D) Standard deviation of X

10. What does an R^2 value of 1 mean?

(A) The model perfectly predicts all data points

(B) The model predicts nothing better than the mean

(C) The residuals are random

(D) The slope is 1

11. Which of the following can help deal with overfitting in regression models?

(A) Increasing the number of predictors

(B) Using regularization methods like Ridge/Lasso

(C) Reducing the size of the training set

(D) Removing the intercept term

12. In multiple linear regression, the interpretation of β_2 is:

(A) Change in Y per unit change in X_2 holding all other predictors constant

(B) Change in Y per unit change in X_2 without considering other predictors

(C) The total change in Y due to all predictors

(D) The residual error for X_2

13. A company is predicting monthly sales based on advertising spend. The regression output shows:
$Sales = 500 + 8 \cdot (Ad\ Spend)$
If Ad Spend = $200, what are the predicted sales?

(A) 2100

(B) 2500

(C) 1600

(D) 21000

14. A student fits a regression model to predict exam scores from hours studied and finds a slope of 5. This means:

(A) Each additional study hour increases score by 5 points on average

(B) The maximum score possible is 5

(C) The minimum score possible is 5

(D) 5% of variation in scores is explained

15. You run a regression model for crop yield and find the coefficient for rainfall is negative. This likely means:

(A) Rainfall reduces yield

(B) Rainfall has no relationship with yield

(C) Rainfall is perfectly correlated with yield

(D) The intercept is negative

16. A regression output shows: Intercept = 100 Slope = -4 If $X = 20$, the predicted Y is:

(A) 180

(B) 20

(C) 20 × -4

(D) 100 - 4

17. In a salary prediction model, Age and Experience are highly correlated. This will likely:

(A) Cause multicollinearity issues

(B) Improve model accuracy

(C) Increase R^2 drastically

(D) Eliminate overfitting

18. Intercept = 50, slope = -2, X = 15

(A) 20

(B) 50

(C) -20

(D) 80

19. A regression model for crop yield finds rainfall coefficient = -0.3. This means

(A) More rainfall reduces yield by 0.3 units per mm

(B) More rainfall increases yield by 0.3 units

(C) Rainfall has no effect

(D) Model is wrong

20. Model has high train R^2 but low test R^2. Cause?

(A) Overfitting

(B) Underfitting

(C) Homoscedasticity

(D) Multicollinearity

21. In multiple regression, coefficient for Fertilizer = 1.2 means

(A) Fertilizer has no effect

(B) For each 1 unit increase in fertilizer, yield increases by 1.2 units (holding others constant)

(C) Fertilizer reduces yield

(D) Fertilizer effect depends on rainfall

22. A student says "Cost function and loss function are same". You say

(A) Correct — they mean exactly the same thing

(B) Incorrect — loss is for one observation, cost is average over dataset

(C) Incorrect — cost is for classification, loss for regression

(D) Both terms mean training error only

23. Homoscedasticity means

(A) Residuals have constant variance

(B) Residuals increase with X

(C) Features are uncorrelated

(D) Errors are normally distributed

24. In a salary model, Age and Experience have correlation = 0.95. This likely causes

(A) Multicollinearity

(B) Underfitting

(C) Overfitting

(D) Homoscedasticity

25. Polynomial regression is chosen when

(A) Relationship between X and Y is clearly non-linear

(B) Relationship between X and Y is perfectly linear

(C) Target variable is categorical

(D) Dataset is too small

26. The main difference between Ridge and Lasso regression is

(A) Ridge shrinks coefficients, Lasso can set some to zero

(B) Lasso shrinks coefficients, Ridge sets some to zero

(C) Ridge uses L1 penalty, Lasso uses L2 penalty

(D) Ridge works only for classification

27. Lasso regression helps by

(A) Eliminating irrelevant features

(B) Increasing coefficients

(C) Removing intercept

(D) Making residuals constant

28. Ridge regression is preferred when

(A) Data has high multicollinearity

(B) Data has no correlation between features

(C) Target is categorical

(D) Features are perfectly uncorrelated

29. In simple regression, if the slope is 0, it means

(A) Perfect positive correlation

(B) Perfect negative correlation

(C) No relationship between X and Y

(D) The intercept is also zero

30. A dataset is modeled with linear regression and residuals show increasing spread with X. This indicates:

(A) Homoscedasticity

(B) Heteroscedasticity

(C) Multicollinearity

(D) Overfitting