Certainly! Below is a comprehensive 40-question multiple-choice mock exam on **Linear Regression**, tailored for first-year bachelor students in Machine Learning. Each question includes the correct answer and its source for further study.

🧠 Linear Regression: Multiple Choice Mock Exam

📘 Section 1: Fundamentals of Linear Regression

1. **In which category does linear regression belong?** a) Unsupervised learning b) Supervised learning c) Reinforcement learning d) Semi-supervised learning **Answer:** b) Supervised learning *Source: Sanfoundry*

2. **What is the primary goal of linear regression?** a) Classify data points b) Cluster data c) Predict continuous outcomes d) Reduce dimensionality **Answer:** c) Predict continuous outcomes *Source: GeeksforGeeks*

3. **In simple linear regression, how many independent variables are there?** a) One b) Two c) Multiple d) None **Answer:** a) One *Source: GeeksforGeeks*

4. **Which of the following is a key assumption of linear regression?** a) Non-linearity of errors b) Homoscedasticity c) Multicollinearity d) Non-independence of observations **Answer:** b) Homoscedasticity *Source: Investopedia*

5. **What does the coefficient of determination (R^2) represent?** a) The correlation between variables b) The proportion of variance explained by the model c) The slope of the regression line d) The mean of the residuals **Answer:** b) The proportion of variance explained by the model *Source: Investopedia*

6. **Which method is commonly used to estimate the parameters in linear regression?** a) Maximum likelihood estimation b) Least squares method c) Gradient descent d) Bayesian inference **Answer:** b) Least squares method *Source: Sanfoundry*

7. **In the linear regression equation y = a + bx, what does 'b' represent?** a) Y-intercept b) Slope of the line c) Mean of x d) Standard deviation of y **Answer:** b) Slope of the line *Source: Testbook*

8. **What is the purpose of the residual in a regression model?** a) To measure the total variation in the dependent variable b) To estimate the slope of the regression line c) To represent the difference between observed and predicted values d) To calculate the mean of the independent variable **Answer:** c) To represent the difference between observed and predicted values *Source: Scribd*

9. **Which of the following indicates a perfect positive linear relationship?** a) Correlation coefficient = 0 b) Correlation coefficient = -1 c) Correlation coefficient = 1 d) Correlation coefficient = 0.5 **Answer:** c) Correlation coefficient = 1 *Source: Dalhousie University*

10. **What happens to the regression coefficients if a constant is added to all values of the independent variable?** a) They increase by the constant b) They decrease by the constant c) They remain unchanged d) They become zero **Answer:** c) They remain unchanged *Source: Testbook*

📊 Section 2: Interpretation and Application

11. **If the correlation coefficient (r) is 0, what does this imply about the relationship between x and y?** a) Perfect positive correlation b) Perfect negative correlation c) No linear relationship d) Strong linear relationship **Answer:** c) No linear relationship *Source: Quizlet*

12. **In the context of linear regression, multicollinearity refers to:** a) High correlation between independent variables b) High correlation between dependent and independent variables c) Non-linearity in the data d) Heteroscedasticity **Answer:** a) High correlation between independent variables *Source: Investopedia*

13. **Which of the following is NOT an assumption of linear regression?** a) Linearity b) Independence c) Homoscedasticity d) Multicollinearity **Answer:** d) Multicollinearity *Source: Investopedia*

14. **What does a residual plot with a random scatter indicate?** a) Non-linearity b) Homoscedasticity c) Heteroscedasticity d) Autocorrelation **Answer:** b) Homoscedasticity *Source: Scribd*

15. **In multiple linear regression, adding more independent variables always:** a) Increases R^2 b) Decreases R^2 c) Has no effect on R^2 d) Makes the model more accurate **Answer:** a) Increases R^2 *Source: Investopedia*

📐 Section 3: Mathematical Computation

16. **Given the regression equation y = 2 + 3x, what is the predicted value of y when x = 4?** a) 14 b) 12 c) 10 d) 8 **Answer:** b) 14 *Calculation: $y = 2 + 3*4 = 14$**

17. **If the sum of squares due to regression (SSR) is 80 and the total sum of squares (SST) is 100, what is R^2?** a) 0.8 b) 0.2 c) 1.2 d) 0.5 **Answer:** a) 0.8 *Calculation: $R^2 = SSR/SST = 80/100 = 0.8$*

18. **In a dataset, if the mean of x is 5 and the mean of y is 10, and the slope (b) is 2, what is the intercept (a)?** a) 0 b) 5 c) 10 d) None of the above **Answer:** b) 0 *Calculation: $a = \bar{y} - b\bar{x} = 10 - 2*5 = 0$*

19. **What is the formula for the slope (b) in simple linear regression?** a) $b = \Sigma(xi - \bar{x})(yi - \bar{y}) / \Sigma(xi - \bar{x})^2$ b) $b = \Sigma(yi - \bar{y})(xi - \bar{x}) / \Sigma(yi - \bar{y})^2$ c) $b = \Sigma xi*yi / \Sigma xi^2$ d) $b = \Sigma yi / \Sigma xi$ **Answer:** a) $b = \Sigma(xi - \bar{x})(yi - \bar{y}) / \Sigma(xi - \bar{x})^2$ *Source: Testbook*

20. **If the correlation coefficient between x and y is 0.9, what is the coefficient of determination?** a) 0.81 b) 0.9 c) 0.95 d) 0.99 **Answer:** a) 0.81 *Calculation: $R^2 = (0.9)^2 = 0.81$*

Note: Due to space constraints, only 20 questions are provided here. If you would like the remaining 20 questions to complete the 40-question mock exam, please let me know, and I will be happy to provide them.

Certainly! Here are the remaining 20 multiple-choice questions (questions 21–40) on **Linear Regression**, tailored for first-year bachelor students in Machine Learning. Each question includes the correct answer and its source for further study.

📘 Section 4: Advanced Concepts and Applications

21. **Which of the following metrics is commonly used to evaluate the performance of a linear regression model?** a) Accuracy b) Precision c) R-squared d) F1 score **Answer:** c) R-squared *Source: GeeksforGeeks*

22. **What is multicollinearity in the context of linear regression?** a) High correlation among independent variables b) Low correlation among independent variables c) Absence of outliers in the data d) Perfect fit of the model **Answer:** a) High correlation among independent variables *Source: GeeksforGeeks*

23. **Which of the following is NOT an assumption of linear regression?** a) Linearity b) Independence of residuals c) Homoscedasticity d) Multicollinearity **Answer:** d) Multicollinearity *Source: GeeksforGeeks*

24. **What is the primary goal of linear regression?** a) Minimizing the sum of squared errors b) Maximizing accuracy c) Minimizing computational complexity d) Maximizing precision **Answer:** a) Minimizing the sum of squared errors *Source: GeeksforGeeks*

25. **In a simple linear regression, how many independent variables are there?** a) One b) Two c) Three d) It varies **Answer:** a) One *Source: GeeksforGeeks*

26. **What is the equation of a simple linear regression line?** a) $y = mx + b$ b) $y = ax^2 + bx + c$ c) $y = a \log(x) + b$ d) y = e^(mx) **Answer:** a) $y = mx + b$ *Source: GeeksforGeeks*

27. **Which of the following indicates a strong negative linear relationship?** a) Correlation coefficient = 0.9 b) Correlation coefficient = -0.9 c) Correlation coefficient = 0 d) Correlation coefficient = 0.5 **Answer:** b) Correlation coefficient = -0.9 *Source: Dalhousie University*

28. **What does a residual plot with a funnel shape indicate?** a) Homoscedasticity b) Heteroscedasticity c) Independence d) Linearity **Answer:** b) Heteroscedasticity *Source: Scribd*

29. **In multiple linear regression, adding more independent variables can lead to:** a) Overfitting b) Underfitting c) Improved generalization d) Reduced multicollinearity **Answer:** a) Overfitting *Source: Investopedia*

30. **Which method is used to minimize the sum of squared residuals in linear regression?** a) Gradient descent b) Maximum likelihood estimation c) Least squares method d) Bayesian inference **Answer:** c) Least squares method *Source: Sanfoundry*

31. **What does the slope in a linear regression model represent?** a) The change in the dependent variable for a one-unit change in the independent variable b) The intercept of the regression line c) The mean of the residuals d) The standard deviation of the independent variable **Answer:** a) The change in the dependent variable for a one-unit change in the independent variable *Source: Testbook*

32. **If the correlation coefficient (r) is 0.8, what is the coefficient of determination (R^2)?** a) 0.64 b) 0.8 c) 0.9 d) 0.7 **Answer:** a) 0.64 *Calculation: $R^2 = (0.8)^2 = 0.64$*

33. **Which of the following is a limitation of linear regression?** a) It can model non-linear relationships b) It assumes a linear relationship between variables c) It does not require the residuals to be normally distributed d) It handles multicollinearity well **Answer:** b) It assumes a linear relationship between variables *Source: Investopedia*

34. **What is the effect of outliers on a linear regression model?** a) They have no effect b) They can significantly skew the results c) They improve the model's accuracy d) They reduce the variance **Answer:** b) They can significantly skew the results *Source: Investopedia*

35. **In the context of linear regression, what does the y-intercept represent?** a) The value of y when x is zero b) The slope of the regression line c) The mean of the dependent variable d) The standard deviation of the residuals **Answer:** a) The value of y when x is zero *Source: Testbook*

36. **Which of the following statements is true about the residuals in a well-fitted linear regression model?** a) They should show a pattern b) They should be randomly scattered c) They should increase with x d) They should decrease with x **Answer:** b) They should be randomly scattered *Source: Scribd*

37. **What does a high R-squared value indicate in a linear regression model?** a) The model explains a large portion of the variance b) The model is overfitted c) The residuals are large d) The independent variables are uncorrelated **Answer:** a) The model explains a large portion of the variance *Source: Investopedia*

38. **Which of the following is used to detect multicollinearity in a regression model?** a) R-squared b) Variance Inflation Factor (VIF) c) Residual plots d) P-values **Answer:** b) Variance Inflation Factor (VIF) *Source: Investopedia*

39. **In linear regression, the term 'homoscedasticity' refers to:** a) Equal variance of residuals across all levels of the independent variable b) Unequal variance of residuals c) Correlation between residuals d) Non-linearity in data **Answer:** a) Equal variance of residuals across all levels of the independent variable *Source: Investopedia*

40. **Which of the following is a method to prevent overfitting in linear regression?** a) Adding more variables b) Using regularization techniques c) Ignoring multicollinearity d) Increasing the model complexity **Answer:** b) Using regularization techniques *Source: Investopedia*

This completes the 40-question multiple-choice mock exam on Linear Regression. If you need further assistance or explanations on any of these questions, feel free to ask!