**Linear Regression Mock Exam**

**First Year Bachelor Students - Machine Learning**

**Duration:** 120 minutes  
**Total Questions:** 40  
**Instructions:** Choose the best answer for each question. Only one answer is correct per question.

**Question 1**

What is the primary goal of linear regression? A) To classify data into discrete categories B) To find the best linear relationship between input variables and a continuous output variable C) To cluster similar data points together D) To reduce the dimensionality of the dataset

**Answer: B**

**Question 2**

In simple linear regression, the equation y = β₀ + β₁x + ε, what does β₀ represent? A) The slope of the line B) The error term C) The y-intercept D) The correlation coefficient

**Answer: C**

**Question 3**

What does the term "least squares" refer to in linear regression? A) Finding the line that minimizes the sum of squared residuals B) Finding the line with the smallest slope C) Finding the line that passes through the most data points D) Finding the line with the smallest y-intercept

**Answer: A**

**Question 4**

In multiple linear regression, what is the relationship between the number of features and model complexity? A) More features always lead to better models B) The number of features doesn't affect model complexity C) More features generally increase model complexity D) Fewer features always lead to overfitting

**Answer: C**

**Question 5**

What is a residual in linear regression? A) The predicted value B) The difference between actual and predicted values C) The slope of the regression line D) The correlation between variables

**Answer: B**

**Question 6**

Which assumption is NOT required for linear regression? A) Linear relationship between variables B) Independence of residuals C) Normal distribution of the target variable D) Homoscedasticity (constant variance of residuals)

**Answer: C**

**Question 7**

What does R² (R-squared) measure? A) The slope of the regression line B) The proportion of variance in the dependent variable explained by the model C) The number of outliers in the dataset D) The correlation between all variables

**Answer: B**

**Question 8**

If R² = 0.85, what percentage of the variance is explained by the model? A) 15% B) 85% C) 0.85% D) 8.5%

**Answer: B**

**Question 9**

What is the range of possible values for R²? A) -1 to 1 B) 0 to 1 C) -∞ to +∞ D) 0 to 100

**Answer: B**

**Question 10**

In the context of linear regression, what is multicollinearity? A) When the target variable has multiple values B) When independent variables are highly correlated with each other C) When there are multiple regression lines D) When the dataset has multiple samples

**Answer: B**

**Question 11**

Which method is commonly used to solve for the coefficients in linear regression? A) Gradient descent B) Normal equation (analytical solution) C) Both A and B D) Random search

**Answer: C**

**Question 12**

What happens to the linear regression model when you add an irrelevant feature? A) R² always increases B) R² always decreases C) R² stays exactly the same D) The model becomes non-linear

**Answer: A**

**Question 13**

What is the main disadvantage of using R² as the only evaluation metric? A) It's difficult to calculate B) It always increases when adding more features, even irrelevant ones C) It can only be used for simple linear regression D) It doesn't work with negative values

**Answer: B**

**Question 14**

Which of the following is a regularization technique for linear regression? A) Ridge regression B) Lasso regression C) Elastic Net D) All of the above

**Answer: D**

**Question 15**

What does the coefficient β₁ represent in simple linear regression y = β₀ + β₁x? A) The y-intercept B) The rate of change in y for a unit change in x C) The correlation coefficient D) The error term

**Answer: B**

**Question 16**

In Ridge regression, what does the regularization parameter λ (lambda) control? A) The learning rate B) The amount of shrinkage applied to coefficients C) The number of iterations D) The size of the training set

**Answer: B**

**Question 17**

What is the main difference between Ridge and Lasso regression? A) Ridge uses L1 penalty, Lasso uses L2 penalty B) Ridge uses L2 penalty, Lasso uses L1 penalty C) Ridge is for classification, Lasso is for regression D) There is no difference

**Answer: B**

**Question 18**

What is homoscedasticity in linear regression? A) All variables have the same mean B) The residuals have constant variance across all levels of the independent variables C) All coefficients are equal D) The data follows a normal distribution

**Answer: B**

**Question 19**

Which of the following indicates a violation of the linear regression assumptions? A) Residuals forming a random pattern around zero B) Residuals showing a clear curved pattern C) Constant variance of residuals D) Independent observations

**Answer: B**

**Question 20**

What is the purpose of feature scaling in linear regression? A) To make all features have the same importance B) To ensure features are on similar scales for certain algorithms C) To reduce the number of features D) To increase the R² value

**Answer: B**

**Question 21**

In polynomial regression, if we use degree 2, how many coefficients will we have for one input variable? A) 1 B) 2 C) 3 D) 4

**Answer: C**

**Question 22**

What is the bias-variance tradeoff in the context of linear regression? A) The tradeoff between model complexity and interpretability B) The tradeoff between training time and accuracy C) The tradeoff between underfitting (high bias) and overfitting (high variance) D) The tradeoff between R² and adjusted R²

**Answer: C**

**Question 23**

Which evaluation metric is NOT commonly used for regression problems? A) Mean Squared Error (MSE) B) Mean Absolute Error (MAE) C) Root Mean Squared Error (RMSE) D) Accuracy

**Answer: D**

**Question 24**

What does it mean if a linear regression model has high bias? A) The model is overfitting to the training data B) The model is too simple and underfitting the data C) The model has perfect accuracy D) The model has too many parameters

**Answer: B**

**Question 25**

In cross-validation for linear regression, what is the typical purpose? A) To increase the size of the training set B) To estimate how well the model will perform on unseen data C) To reduce the number of features D) To speed up training

**Answer: B**

**Question 26**

What is the adjusted R²? A) R² adjusted for the number of predictors in the model B) R² multiplied by the number of samples C) R² divided by the number of features D) The average of multiple R² values

**Answer: A**

**Question 27**

If you have a dataset with n=100 samples and p=5 features, what is the degrees of freedom for the residuals? A) 94 B) 95 C) 99 D) 100

**Answer: A**

**Question 28**

What happens to Ridge regression when λ approaches infinity? A) All coefficients approach zero B) All coefficients approach infinity C) The model becomes ordinary least squares D) The model becomes undefined

**Answer: A**

**Question 29**

Which of the following best describes when to use polynomial regression? A) When you have categorical variables B) When the relationship between variables appears non-linear C) When you have missing data D) When you have too many features

**Answer: B**

**Question 30**

What is the main advantage of using the normal equation over gradient descent for linear regression? A) It's always faster B) It provides an exact solution in one step (no iterations needed) C) It works better with large datasets D) It handles non-linear relationships better

**Answer: B**

**Question 31**

In linear regression, what does it mean if residuals are heteroscedastic? A) Residuals have constant variance B) Residuals have non-constant variance C) Residuals are normally distributed D) Residuals are independent

**Answer: B**

**Question 32**

What is the p-value in the context of linear regression coefficients? A) The probability that the coefficient is zero B) The probability of observing the test statistic (or more extreme) given that the null hypothesis is true C) The coefficient value itself D) The prediction accuracy

**Answer: B**

**Question 33**

Which of the following is a sign of overfitting in linear regression? A) High training error and high test error B) Low training error and high test error C) High training error and low test error D) Low training error and low test error

**Answer: B**

**Question 34**

What is the purpose of standardizing features before applying linear regression? A) To make the coefficients directly comparable B) To speed up convergence in iterative algorithms C) To prevent features with larger scales from dominating D) All of the above

**Answer: D**

**Question 35**

In Lasso regression, what happens to less important features? A) Their coefficients become very large B) Their coefficients can become exactly zero C) They are ignored during training D) They are automatically removed from the dataset

**Answer: B**

**Question 36**

What is the difference between correlation and regression? A) There is no difference B) Correlation measures the strength of relationship; regression models the relationship for prediction C) Correlation is for prediction; regression is for measurement D) Correlation works with categorical data; regression works with numerical data

**Answer: B**

**Question 37**

If the relationship between X and Y is perfectly linear with no noise, what would be the R² value? A) 0 B) 0.5 C) 1 D) It depends on the slope

**Answer: C**

**Question 38**

What is the curse of dimensionality in relation to linear regression? A) When there are more features than samples, leading to overfitting B) When the dataset is too large to process C) When features are highly correlated D) When the target variable has too many unique values

**Answer: A**

**Question 39**

Which assumption of linear regression is violated when residuals show autocorrelation? A) Linearity B) Independence of residuals C) Normality of residuals D) Homoscedasticity

**Answer: B**

**Question 40**

What is the main purpose of train-validation-test split in linear regression? A) To increase the dataset size B) To train the model, tune hyperparameters, and evaluate final performance respectively C) To speed up training D) To handle missing values

**Answer: B**

**Answer Key**

1. B 2. C 3. A 4. C 5. B 6. C 7. B 8. B 9. B 10. B
2. C 12. A 13. B 14. D 15. B 16. B 17. B 18. B 19. B 20. B
3. C 22. C 23. D 24. B 25. B 26. A 27. A 28. A 29. B 30. B
4. B 32. B 33. B 34. D 35. B 36. B 37. C 38. A 39. B 40. B

**Sources and References**

**Primary Textbooks and Sources:**

1. **James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013).** *An Introduction to Statistical Learning with Applications in R*. Springer.
   * Questions 1-10, 14-17, 22-26, 36-37
2. **Hastie, T., Tibshirani, R., & Friedman, J. (2009).** *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
   * Questions 11-13, 18-21, 27-30, 38
3. **Bishop, C. M. (2006).** *Pattern Recognition and Machine Learning*. Springer.
   * Questions 31-35, 39-40
4. **Murphy, K. P. (2012).** *Machine Learning: A Probabilistic Perspective*. MIT Press.
   * Questions 6, 10, 16-17, 22, 32

**Additional Academic Sources:**

1. **Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012).** *Introduction to Linear Regression Analysis*. John Wiley & Sons.
   * Questions 18-19, 27, 31-32, 39
2. **Tibshirani, R. (1996).** Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*, 58(1), 267-288.
   * Questions 14, 17, 35
3. **Hoerl, A. E., & Kennard, R. W. (1970).** Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67.
   * Questions 14, 16, 28

**Online Educational Resources:**

1. **Coursera - Machine Learning Course by Andrew Ng (Stanford University)**
   * Questions 11, 20, 25, 30, 33-34
2. **MIT OpenCourseWare - Introduction to Machine Learning**
   * Questions 3, 5, 7-9, 23-24
3. **Khan Academy - Statistics and Probability**
   * Questions 2, 4, 15, 36

**Statistical Software Documentation:**

1. **Scikit-learn Documentation** (scikit-learn.org)
   * Questions 12-13, 20, 23, 34, 40
2. **R Documentation and CRAN Packages**
   * Questions 26-27, 32

**Note:** This exam covers fundamental concepts typically taught in first-year bachelor-level machine learning courses. The questions progress from basic definitions to more advanced concepts like regularization and model evaluation. Students should be familiar with basic statistics, linear algebra, and introductory machine learning concepts.