Assignment 11 Solution

The correct answer is in bold font

Question 1. The violation of the assumption of constant variance of residuals is known as:

a. **Heteroscedasticity [Hint: Heteroscedasticity is a situation where the variance of the residuals is not constant for different values of the independent variables (X's).]**
b. Multicollinearity [Hint: Multicollinearity occurs when two or more independent variables are highly correlated with one another in a regression model.]
c. Homoscedasticity [Hint: Homoscedasticity is the assumption of equal or similar variances for different values of the independent variable (X's).]
d. Autocorrelation [Hint: Autocorrelation refers to the degree of correlation across error terms for different values of X's.]

Question 2. The property of an estimator $E(\widehat{\beta_1}) = \beta_1$ is termed as....

a. Linearity [Hint: Estimator is linear, if the regression model is linear in parameters.]
b. **Unbiasedness [Hint: Estimator is unbiased, if its average or expected value, $E(\widehat{\beta_1})$, is equal to the true value, β_1.]**
c. Efficiency [Hint: An estimator with a relatively lower variance (or the standard error of estimate) is more efficient compared to the one with higher variance.]
d. Consistency [Hint: Consistency of an estimator means that as the sample size gets larger the estimate converges to the true value of the parameter.]

Question 3. The following statement is correct about the regression analysis

a. Regression analysis also determines the direction of causality. [Hint: the direction of causation is established by a priori theoretical considerations]
b. Regression analysis is employed to examine the correlation between the variables [Hint: Correlation analysis shows the degree of the linear relationship between the two variables. In correlation analysis, both the variables are treated similarly and considered to be random]
c. **Regression analysis measures the statistical significance of the relationship. Hint: Regression only establishes the statistical strength of the relation with the help of F-stat. and t-stat.**
d. Regression analysis also provides the theoretical underpinning of the causality. [Hint: The theoretical underpinings are established a priori.]

Question 4. Which of the following is incorrect in the context of classical normal linear regression model (CLNRM) with OLS procedure.

a. OLS estimates are best linear unbiased and efficient (BLUE) estimates. Hint: OLS estimates are best in linear class of estimators.
b. Normal distribution can be easily defined with just mean and variance. Hint: Normal distribution is bell shaped curve, and it is well defined using the mean and variance of the distribution.

c. **The variance of the error term is not constant. Hint: Homoscedasticity: Error term has constant variance. $Variance = E(u_i)^2 = \sigma^2$**

d. Expected value or the mean of error term is zero. Hint: $Mean = E(u_i) = 0$ under CNLRM.

Question 5. Which of the following is correct in the context of Ordinary Least Square (OLS) regression

a. There are errors in OLS estimation because we work with large populations. Hint: There are errors because we work with samples; since population is not observed.

b. OLS procedure involves minimizing the sum of errors. Hint: Minimizing sum of errors would lead cancellation of positive and negative errors, it is incorrect

c. **OLS procedure involves minimization of residual sum of squares. Hint: Minimization of sum of squares penalizes large errors and is not adversely affected by the positive negative sign of the error term.**

d. The constant term in the regression estimate reflects the slope of regression fit line. Hint: The constant term is the intercept; the coefficient of variable is the slope term.

Question 6. Which of the following is correct in the context of multiple linear regression

a. The coefficients measure the effect of independent variable on the dependent variable irrespective of other variables. Hint: The coefficients are partial regression coefficients and reflect the impact of the corresponding variable, keeping other variables constant.

b. OLS regression procedure assumes that errors are correlated. Hint: Errors are assumed to be not serially correlated.

c. **Constant variance of error terms. Hint: This pertains to the assumption of Homoscedasticity, i.e., constant variance of error terms.**

d. OLS regression assumes a certain correlation between the error term and the independent variables. Hint: OLS regression procedure assumes that there is no correlation between the error term and the independent variable.

Question 7. In the ordinary least squares (OLS) regression [Model:Y=b0+b1X+u] , which of the following is not a required assumption about the error term u?

a. **The expected value of the error (u) term is one. Hint: The expected value of the error term is considered to be zero.]**

b. The variance of the error term (u) is the same for all values of Xi's (i.e., it remains constant). Hint: This is part of homescedasticiy assumption (i.e., constant variance).

c. The values of the error term (u) are independent (i.e., not correlated). Hint: OLS procedure assumes that errors are not correlated for different values of Xi's.

d. The error term (u) is normally distributed. Hint: For hypothesis testing and estimating the significance of the coefficient, OLS procedure assumes normality of errors as a part of classical normal linear regression model.

Question 8. How can the statistical significance of the relationship between two numerical variables be measured? [Y=b0+b1X+u].

a. By visually examining the scatter diagram [Hint: A scatter diagram only visualizes the relationship between two variables]
b. By examining the R-square from the regression. Hint: Regression R-square indicates the proportion of variation explained by the model.
c. **By examining the t-stat. of the coefficient (b1) of the variable Hint: t-stat. indicates the statistical significance of a regression variable's relationship with the dependent variable.**
d. By determining the intercept (or constant) term. Hint: The intercept term in a regression model represents the mean value of the response variable when all of the predictor variables in the model are equal to zero.

Question 9. What econometric issue afflicts the OLS estimation process when the predictors in a multiple regression model are correlated:

a. Heteroscedasticity. Hint: Heteroscedasticity pertains to non-constant variance of error terms.
b. Homoscedasticity. Hint: Homoscedasticity pertains to constant variance of error terms.
c. **Multicollinearity. Hint: Multicollinearity afflicts the OLS estimation when independent variables (or predictors) are highly correlated.**
d. Autocorrelation. Hint: Autocorrelation afflicts the OLS estimation when the error terms are correlated with each other.

Question 10. Which of the following is correct in the context of multiple linear regression

a. A high multicollinearity increases the chances of incorrectly increasing the significance of the variable. Hint: A high multicollinearity leads to high standard errors, which decreases the significance (t-values) of the coefficient.
b. R^2 measure accounts for the fact that the explanatory power of the model can be incorrectly enhanced by increasing the number of variables. Hint: No. R^2 can be artificially inflated just by adding more variables in the model.
c. **$Adjusted - R^2$ measure accounts for the fact that the explanatory power of the model can be incorrectly enhanced by increasing the number of variables. Hint: $Adjusted - R^2$ measure penalizes addition of variables.**
d. Perfect multicollinearity is difficult to identify. Hint: No. In case of perfect multicollinearity, the model is indeterminable and hence easily identified.