**Chapter 9 Linear correlation and regression analysis – MCQ Student**

1. If the Pearson Product Moment Correlation Coefficient shows zero value, this definitely means that there is no relationship between the two variables
   1. True
   2. False

The correct answer is b.

Support comment: Pearson’s Correlation Coefficient was designed just for linear relationships. Technically, if we get a low value, or a zero value, it does not mean that there is no relationship. It just means that there is no linear relationship.

1. To measure ranked variables the following correlation coefficient is used
   1. Pearson’s
   2. Spearman’s
   3. Fisher’s

The correct answer is b.

Support comment: A correlation coefficient that is applied to ranked variables was introduced by Spearman, whilst Pearson and Fisher introduced independently correlation coefficients that measure the strength of the relationship between non-ranked variables.

1. The most commonly used formula to describe linear relationship is
   1. ŷ = b0 + b1x + b2x2
   2. ŷ = b0 + b1x2
   3. ŷ = b0 + b1x

The correct answer is c.

Support comment: The straight line equation y=b0 + b1x is an example of a linear relationship. This means that the changes in one variable are accompanied by the proportional linear changes in another variable. If you increase one variable by 1 and the other by 0.5, this is a linear change.

1. Another expression for constant variance is
   1. Leptokracy
   2. Homoscedacity
   3. Heteroscedacity
   4. Covariance

The correct answer is b.

Support comment: Homoscedacity comes from the Greek words homos (the same) and skedastikos (dispersed). Homoscedacity is an expression that describes a requirement to have a constant variable in some cases, i.e. the variance must not show the tendency to grow, for example.

1. Residual plot in regression analysis is used to mean
   1. Plot of residuals
   2. Plot of marginally residual value
   3. Plot of regressors
   4. Plot of regressing values

The correct answer is a.

Support comment: Residuals in regression analysis are the values that represent the difference between the actual observation values and the regression line values. These residuals need to be plotted and it is expected that they will show no trend or pattern. They have to be random. Only then we know that the regression represents well the actual observations

1. The coefficient of determination and the R-squared (R2) are the same
   1. True
   2. False

The correct answer is b.

Support comment: The coefficient of determination and the R-Squared are just two interchangeable expressions for the same statistic that measures how much of one variable is explained by another variable.

1. Which of the following is true
   1. Correlation measures the strength of relationships between variables, as well as causation
   2. Correlation measures the strength of relationships between variables, but not causation

The correct answer is b.

Support comment:

1. If linear regression is represented by equation y=a+bx, then
   1. a is the slope and b is the intercept
   2. a is the intercept and b is the slope

The correct answer is b.

Support comment:

1. Which one of the following relationship correct
   1. SST=SSR+SSE
   2. SSR=SST+SSE

The correct answer is a.

Support comment:

1. One of the following is not a necessary assumption for regression analysis
   1. Linearity
   2. Independence of errors
   3. Common variances
   4. Normality of errors
   5. Constant variance

The correct answer is c.

Support comment:

1. Prediction interval defines the probability that the actual value will fall within the confidence interval of the estimated value
   1. True
   2. False

The correct answer is a.

Support comment: Prediction interval is essentially a confidence interval for linear regression models. Just like the confidence interval for the population mean states that you are confident that the true mean is somewhere in the interval of values μ=𝑥̅±SE (SE=standard error of the mean), prediction interval states that the true possible data value is somewhere in the interval of values y=𝑦 ̂±SEE (SEE= standard error of the estimate).