

CÂU HỎI LUYỆN TẬP

QUANTITATIVES METHODS CFA LEVEL 2

Question 1: During the course of a multiple regression analysis, an analyst has observed several items that she believes may render incorrect conclusions. For example, the coefficient standard errors are too small, although the estimated coefficients are accurate. She believes that these small standard error terms will result in the computed t-statistics being too big, resulting in too many Type I errors. The analyst has most likely observed which of the following assumption violations in her regression analysis?

- A. Positive serial correlation.
- B. Homoscedasticity.
- C. Multicollinearity.

Question 2: When two or more of the independent variables in a multiple regression are correlated with each other, the condition is called:

- A. multicollinearity.
- B. conditional heteroskedasticity.
- C. serial correlation.

Question 3: An analyst is trying to determine whether fund return performance is persistent. The analyst divides funds into three groups based on whether their return performance was in the top third (group 1), middle third (group 2), or bottom third (group 3) during the previous year. The manager then creates the following equation: $R = a + b_1D_1 + b_2D_2 + b_3D_3 + \varepsilon$, where R is return premium on the fund (the return minus the return on the S&P 500 benchmark) and D_i is equal to 1 if the fund is in group i. Assuming no other information, this equation will suffer from:

- A. multicollinearity.
- B. serial correlation.
- C. heteroskedasticity.

Question 4: Which of the following is least likely to result in misspecification of a regression model?

- A. Inappropriate variable form.
- B. Transforming a variable.
- C. Omission of an important independent variable.

Passage (Question 5-8):

Ben Sasse is a quantitative analyst at Gurnop Asset Managers. Sasse is interviewing Victor Sophie for a junior analyst position. Sasse mentions that the firm currently uses several proprietary multiple regression models and wants Sophie's opinion about regression models.

Sophie makes the following statements:

Statement 1: Multiple regression models can be used to forecast independent variables.

Statement 2: Multiple regression models can be used to test existing theories of relationships among variables.

Sasse then discusses a model that the firm uses to forecast credit spread on investment-grade corporate bonds. Sasse states that while the current model parameters are a secret, the following is an older version of the model:

$$\text{CSP} = 0.22 + 1.04 \times \text{DSC} - 0.32 \times \text{index} + 1.33 \times \text{D/E}$$

where:

CSP = credit spread (%)

DSC = EBITDA / unsecured debt

index = 1 if the issuer is part of CDX index; 0 otherwise

D/E = long-term debt / equity

Question 5-8: Regarding Sophie's statement on multiple regression:

- A. only Statement 1 is correct.
- B. only Statement 2 is correct.
- C. both statements are correct.

Question 6-8: everything else the same, which of the following statements most accurately describes the model's forecast?

- A. The credit spread on the rm's issue would decrease by 10 bps.
- B. The credit spread on the rm's issue will increase by 32 bps.
- C. The credit spread on the rm's issue will decrease by 32 bps.

Question 7-8: hich of the following is least likely an assumption of multiple linear regression?

- A. The dependent variable is not serially correlated.
- B. There is no linear relationship between the independent variables.
- C. The error term is normally distributed.

Question 8-8: Which assumption of multiple regression is most likely evaluated using a QQ plot?

- A. Serial correlation of residuals.
- B. Conditional heteroskedasticity.
- C. Error term is normally distributed.

Question 9: Which of the following statements regarding the results of a regression analysis is least accurate? The:

- A. slope coefficient in a multiple regression is the change in the dependent variable for a one-unit change in the independent variable, holding all other variables constant.
- B. slope coefficients in the multiple regression are referred to as partial betas.
- C. slope coefficient in a multiple regression is the value of the dependent variable for a given value of the independent variable.

Question 10: One of the underlying assumptions of a multiple regression is that the variance of the residuals is constant for various levels of the independent variables. This quality is referred to as:

- A. homoscedasticity.
- B. a normal distribution.
- C. a linear relationship.

ĐÁP ÁN

1. A	2. B	3. C	4. C	5. A	6. B	7. C	8. A	9. A	10. C
------	------	------	------	------	------	------	------	------	-------

Question 1: A

Positive serial correlation is the condition where a positive regression error in one time period increases the likelihood of having a positive regression error in the next time period. The residual terms are correlated with one another, leading to coefficient error terms that are too small.

Question 2: A

Multicollinearity refers to the condition when two or more of the independent variables, or linear combinations of the independent variables, in a multiple regression are highly correlated with each other. This condition distorts the standard error of estimate and the coefficient standard errors, leading to problems when conducting t-tests for statistical significance of parameters.

Question 3: A

When we use dummy variables, we have to use one less than the states of the world. In this case, there are three states (groups) possible. We should have used only two dummy variables.

Multicollinearity is a problem in this case. Specifically, a linear combination of independent variables is perfectly correlated. $X_1 + X_2 + X_3 = 1$.

There are too many dummy variables specified, so the equation will suffer from multicollinearity.

Question 4: B

The four types of model specification errors are: omission of an important independent variable, inappropriate variable form, inappropriate variable scaling and data improperly pooled.

Transforming an independent variable is usually done to rectify inappropriate variable scaling.

Question 5: B

Multiple regression models can be used to identify relations between variables, forecast the dependent variable, and test existing theories. Statement 1 is inaccurate in because it mentions forecast independent (and not dependent) variables.

Question 6: C

The coefficient on the index dummy variable is -0.32 , and if the variable takes a value of 1 (inclusion in the index), the credit spread would decrease by 0.32% , or 32 bps.

Question 7: A

The assumption calls for the residual (or errors) to be not serially correlated. The dependent variable can have serial correlation. Other assumptions are accurate.

Question 8: C

A normal QQ plot of the residuals can visually indicate violation of the assumption that the residuals are normally distributed.

Question 9: A

The slope coefficient is the change in the dependent variable for a one-unit change in the independent variable.

Question 10: C

Homoskedasticity refers to the basic assumption of a multiple regression model that the variance of the error terms is constant.