Question 1 of 16

Multiple linear regression describes the variation of the:

A. dependent variable using one or more independent variables.
B. independent variable using two or more dependent variables.
C. dependent variable using two or more independent variables.

Correct Answer: C

Explanation:

C is correct. Multiple linear regression is a method used to model the linear relationship between a dependent variable and ***more than one*** independent variables. Multiple linear regression describes the variation of the dependent variable using two or more independent variables.

A is incorrect. Because simple linear regression describes the variation of the dependent variable using one independent variable. In multiple linear regression, two or more independent variables are used.

B is incorrect. Because the linear regression equation explains the variation of the dependent variables rather than the variations of the independent variables.

Quantitative Methods , Learning Module 1

Question 2 of 16

Multiple linear regression is used to model the:

A. linear relationship between two or more dependent and independent variables.
B. nonlinear relationship between multiple dependent and independent variables.
C. linear relationship between one dependent variable and two or more independent variables.

Correct Answer: C

Explanation:

C is correct. Multiple linear regression is used to model the linear relationship between one dependent variable and two or more independent variables. Multiple linear regression lets analysts estimate using more complex models with multiple explanatory variables.

A is incorrect. See above.

B is incorrect. Because linear regression equation is not used to model the non-linear relationship between dependent and independent variables.

Quantitative Methods , Learning Module 1

Question 3 of 16

A financial analyst wants to determine if financial leverage, profitability, revenue growth, and market share are indicators of impending financial crisis. The most suitable regression model for the analyst to employ is:

A. simple linear regression model.
B. multiple linear regression model.
C. either simple or multiple linear regression model.

Correct Answer: B

Explanation:

B is correct. Financial and economic relations are complex and require models with multiple explanatory variables that must pass rigorous statistical and theoretical scrutiny. Multiple regression can be used to identify relationships between variables to test current or future theories. For example, examining multiple factors such as political stability, economic conditions, and ESG considerations on stock returns.

A is incorrect. Since the financial analyst is working with multiple variables, simple linear regression is inappropriate. Complexity of economic and financial relationships necessitates models with multiple explanatory variables, such as multiple linear regression.

C is incorrect. See above.

Quantitative Methods , Learning Module 1

Question 4 of 16

In a linear regression process, if the dependent variable is discrete, use:

A. logistic regression.
B. nonlinear regression model.
C. traditional regression model.

Correct Answer: A

Explanation:

A is correct. If the dependent variable is discrete, the model can be evaluated as a logistic regression. For example, whether a company is a takeover target or not. The dependent variable can take the value of either 1 or 0.

B is incorrect. Because the question states about linear regression model.

C is incorrect. Because if the dependent variable is continuous, such as returns, the traditional linear regression model is appropriate.

Quantitative Methods , Learning Module 1

Question 5 of 16

Which of the following is the best next step in the regression process if the assumptions of regression are not satisfied?

A. Adjust the model.
B. Analyze the residuals.
C. Examine the goodness of fit for the model.

Correct Answer: A

Explanation:

A is correct. In the regression process, if the assumptions of regression are not satisfied, adjusting the model is the logical next step. All the steps of a regression process are provided below.

The Regression Process

The objective is to use the variation of two or more independent variables to explain the variation in the dependent variable.

Step 1: Find out whether the dependent ***variable is continuous***? If the answer is:

- No, use logistic regression
- Yes, use traditional regression model

Step 2: Estimate the regression model. Examine the model to see ***if it meets the key assumptions***? If the answer is:

- No, adjust the model
- Yes, examine the model's goodness-of-fit

Step 3: Determine whether the ***overall fit is significant***? If the answer is:

- No, adjust the model.
- Yes, determine if this model is the best of all possible models?

Step 4: Check ***if the model is the best of all models***? If the answer is:

- No, adjust the model
- Yes, use the model for analysis and prediction.

Question 6 of 16

A junior analyst is working on a multiple regression model for the automotive industry sector. The model's overall fit is insignificant. Which of the following is the best course of action for him to pursue to solve the issue?

A. Adjust the model and re-estimate
B. Analyze the goodness of fit for the model
C. Analyze the residuals and estimate the regression model

Correct Answer: A

Explanation:

A is correct. The best course of action for the junior analyst is to adjust the model and re-estimate the models' overall fit.

B is incorrect. Because the junior analyst was already aware that the model's overall fit is insignificant. This knowledge comes from the previous stage in which he examined the goodness of fit for the model.

C is incorrect. Before determining the overall fit of the model, it is necessary to estimate the regression model and analyze the residuals. So, the analyst has appropriately completed these steps.

The Regression Process

The objective is to use the variation of two or more independent variables to explain the variation in the dependent variable.

Step 1: Find out whether the dependent ***variable is continuous***? If the answer is:

- No, use logistic regression
- Yes, use traditional regression model

Step 2: Estimate the regression model. Examine the model to see ***if it meets the key assumptions***? If the answer is:

- No, adjust the model
- Yes, examine the model's goodness-of-fit

Step 3: Determine whether the ***overall fit is significant***? If the answer is:

- No, adjust the model.
- Yes, determine if this model is the best of all possible models?

Step 4: Check ***if the model is the best of all models***? If the answer is:

- No, adjust the model
- Yes, use the model for analysis and prediction.

Question 7 of 16

The general form of a multiple linear regression model is as follows:

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + \ldots + b_kX_{ki} + \varepsilon_i, \quad i = 1, 2, \ldots n$$

The components involving the k independent variables in this equation represent the:

A. stochastic part of the model.
B. partial regression coefficients.
C. deterministic part of the model.

Correct Answer: C

Explanation:

C is incorrect. The components involving the k independent variables in this equation represent the deterministic element of the model, whereas the error term, ε_i, represents the stochastic or random part of the model.

A is incorrect. See above.

B is incorrect. A slope coefficient, βj is known as ***partial regression coefficients*** or partial slope coefficients. It measures how much the dependent variable, Y, changes when the independent variable, Xj, changes by one unit, holding all other independent variables constant.

Question 8 of 16

The general form of a multiple linear regression model is as follows:

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + \ldots + b_kX_{ki} + \varepsilon_i, \quad i = 1, 2, \ldots n$$

Term b_0 is referred to as the model's:

A. intercept.
B. slope coefficient.
C. independent variable.

Correct Answer: A

Explanation:

A is correct. Term b_0 is referred to as the model's intercept and it represents the expected value of Y if all independent variables are zero.

THE BASICS OF MULTIPLE REGRESSION

A multiple linear regression model has the following general form:

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + \ldots + b_kX_{ki} + \varepsilon_i, i = 1, 2, \ldots n$$

where,

Y_i = i^{th} observation of dependent variable Y
X_{ki} = i^{th} observation of k^{th} independent variable X
β_0 = intercept term
β_k = slope coefficient of k^{th} independent variable
ε_i = error term of ith observation
n = number of observations
k = total number of independent variables

A slope coefficient, β_j is known as partial regression coefficients or partial slope coefficients. It measures how much the dependent variable, Y, changes when the independent variable, X_j, changes by one unit, holding all other independent variables constant.

The intercept term (β_0) is the value of the dependent variable when the independent variables are all equal to zero.

A regression equation has k slope coefficients and one intercept i.e., k + 1 regression coefficients.

Question 9 of 16

The general form of a multiple linear regression model is as follows:

$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + b_3X_{3i} + \varepsilon_i, \qquad i = 1, 2, 3$

The coefficient b_1 measures the change in:

A. Y for one-unit change in X_1 assuming X_2 and X_3 are zero.
B. X_1 for one-unit change in Y assuming X_2 and X_3 are held constant.
C. Y for one-unit change in X_1 assuming X_2 and X_3 are held constant.

Correct Answer: C

Explanation:

C is correct. The coefficient b_1 measures the change in Y for one-unit change in X_1 assuming X_2 and X_3 are held constant.

THE BASICS OF MULTIPLE REGRESSION

A multiple linear regression model has the following general form:

$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + \ldots + b_kX_{ki} + \varepsilon_i, i = 1, 2, \ldots n$

where,

$Y_i = i^{th}$ observation of dependent variable Y
$X_{ki} = i^{th}$ observation of k^{th} independent variable X
β_0 = intercept term
β_k = slope coefficient of k^{th} independent variable
ε_i = error term of ith observation
n = number of observations
k = total number of independent variables

A slope coefficient, β_j is known as partial regression coefficients or partial slope coefficients. It measures how much the dependent variable, Y, changes when the independent variable, X_j, changes by one unit, holding all other independent variables constant.

The intercept term (β_0) is the value of the dependent variable when the independent variables are all equal to zero.

A regression equation has k slope coefficients and one intercept i.e., k + 1 regression coefficients.

Question 10 of 16

Rebecca Winter is estimating a regression equation that regresses monthly returns of a bond portfolio BPORT against inflation INF and credit ratings CR. The estimated regression based on 36 monthly observations is shown below.

$$BPORT = 0.5766 - 3.0147INF + 3.4198CR$$

Part 1 of 4

If there are no changes in inflation or credit ratings, what is the average annual return of the bond portfolio BPORT?

A. 0.5766%
B. 6.9192%
C. 11.780%

Correct Answer: B

Explanation:

B is correct. If there are no changes in inflation INF and credit ratings CR, the bond portfolio earns an average of 0.5766 % per month, or about 6.919% (0.5766% x 12) per year.

A is incorrect because 0.5766% is the monthly return on average if the changes in inflation and credit ratings are zero.

C is incorrect. 11.780% comes from the following incorrect formula which does not make any sense. $0.5766 - 3.0147(1) + 3.4198(1) = 0.9817 = 0.9817 \times 12 = 11.78\%$

Quantitative Methods , Learning Module 1

Part 2 of 4

Holding credit rating CR constant, the change in the bond portfolio BPORT return for a given one-unit change in the inflation INF is:

A. 0.5766%.
B. 3.4198%.
C. -3.0147%.

Correct Answer: C

Explanation:

C is correct. Keeping the credit rating CR constant, the change in bond portfolio BPORT return for a one-unit change in inflation INF is -3.0147%. In other words, holding the other variable constant, if the inflation increases by 1%, the bond portfolio's return is expected to decrease by 3.014% as represented by minus 3.104%.

A is incorrect. If there are no changes in inflation INF and credit ratings CR, the bond portfolio earns an average of 0.5766 % per month.

B is incorrect. Because holding inflation INF constant, the change in the bond portfolio BPORT return for a given one-unit change in the credit rating CR is 3.4198%.

Part 3 of 4

For a month in which inflation increases by 0.002% and credit ratings move up by 0.03%, the projected return on the bond portfolio is closest to:

A. 0.5808%.
B. 0.6731%.
C. 0.6852%.

Correct Answer: B

Explanation:

B is correct. For a month in which inflation increases by 0.002% and credit ratings move up by 0.03, the projected return on the bond portfolio is 0.6731%. The calculation is provided below:

$BPORT = 0.5766 - 3.0147INF + 3.4198CR$

$BPORT = 0.5766 - 3.0147(0.002) + 3.4198(0.03) = 0.6731\%$

A is incorrect. The calculation wrongly applies using 0.003 change in the credit rating instead of 0.03.

$BPORT = 0.5766 - 3.0147(0.002) + 3.4198(0.003) = 0.5808\%$

C is incorrect. The calculation fails to consider the minus sign of inflation as provided in the given formula.

$BPORT = 0.5766 + 3.0147(0.002) + 3.4198(0.03) = 0.6852\%.$

Part 4 of 4

Calculate the predicted value of the bond portfolio's return if?

INF = - 1.0% CR = 2.0%

A. 0.61%
B. 4.40%
C. 10.43%

Correct Answer: C

Explanation:

C is correct. Given the expected values of the independent variables, the bond portfolio's expected return is

BPORT = 0.5766 – 3.0147(-1) + 3.4198(2) = 10.43%

A is incorrect. The answer is 0.61 percent if we use the values 0.01 and 0.02 for the independent variables inflation INF and Credit ratings CR, respectively.

BPORT = 0.5766 – 3.0147(-0.01) + 3.4198(0.02) = 10.43%

B is incorrect. If we fail to put the minus sign to the inflation value as shown below, the answer is 4.4%.

BPORT = 0.5766 – 3.0147(1) + 3.4198(2) = 10.43%

Question 11 of 16

What can result from inappropriate scaling of variables in a regression model?

A. Reduced variance of error terms
B. Heteroskedasticity or multicollinearity
C. Increased predictive power of the model

Correct Answer: B

Explanation:

B is correct. Heteroskedasticity or multicollinearity. Inappropriate scaling can lead to issues such as heteroskedasticity, where error variance is not constant, or multicollinearity, where independent variables are highly correlated, affecting the model's reliability.

A is incorrect because inappropriate scaling typically increases the variance of error terms rather than reducing it, leading to distorted results.

C is incorrect as scaling issues do not inherently increase predictive power; they may lead to misinterpretation if done improperly.

Question 12 of 16

Which of the following *best* presents one of the assumptions underlying linear regression? The independent variable is:

A. uncorrelated with the residuals.
B. positively correlated with the residuals.
C. negatively correlated with the residuals.

Correct Answer: A

Explanation:

The assumptions underlying linear regression are as follows:

- A linear relationship exists between the independent and dependent variables.
- **The independent variable is uncorrelated with the residuals.**
- The expected value of residual term is zero.
- The variance of the residual term is constant for all observations.
- Residual term is independently distributed.
- The residual term is normally distributed.

Question 13 of 16

Linearity assumption underlying the multiple linear regression model states that:

A. the variance of regression residuals should be linear for all observations.
B. the relationship between the dependent and independent variables should be linear.
C. there should be linear relationship between two or more of the independent variables.

Correct Answer: B

Explanation:

B is correct. Linearity assumption underlying the multiple linear regression model states that the relationship between the dependent and independent variables should be linear.

The five key assumptions of the multiple linear regression model are:

1) Linearity
2) Homoskedasticity
3) Independence of errors
4) Normality
5) Independence of independent variables

Assumption 1: Linearity

'Relation between the dependent variable Y and the independent variables (X_1, X_2, . . . ,X_k)is linear.'

Assumption 2: Homoskedasticity

'The variance of residuals is the same for all observations. It is known as Homoskedasticity (same scatter) assumption.'

Assumption 3: Independence of errors

'The observations (pairs of Xs and Ys) are independent of each other, which implies the residuals are uncorrelated across observations. (i.e., **no serial correlation**).'

Assumption 4: Normality

'The regression residuals (error term) must be normally distributed.'

Assumption 5:
Independence of independent variables

a) Independent variables (X_1, X_2, . . . ,X_k) are not random.
b) No exact linear relation exists between two or more of the independent variables.

Note:

- When an exact linear relationship exists between two or more independent variables, linear regression ***cannot*** be estimated.
- Furthermore, when two or more independent variables are highly correlated, the model can be estimated but its interpretation is ***problematic***.

Question 14 of 16

Which of the following assumptions is most likely violated if the variance of the regression residuals is not the same for all observations?

A. Normality
B. Homoskedasticity
C. Independence of errors

Correct Answer: B

Explanation:

B is correct. If the variance of the regression residuals (error term) is not the same for all observations, the **homoscedasticity** assumption is violated.

A is incorrect. If the regression residuals are not normally distributed, **normality** assumption is violated

C is incorrect. When regression residuals are correlated across observations. This occurs when observations are not independent of one another, **independence of errors** assumption is violated.

Question 15 of 16

The purpose of normal Q-Q plot is to explore whether:

A. residuals are normally distributed.
B. there is any relationship between the independent variables.
C. there is a linear relation between dependent and independent variable.

Correct Answer: A

Explanation:

A is correct. The normal Q-Q plot is a tool utilized to examine whether the residuals correspond to a normal distribution, which is an essential assumption of linear regression.

B is incorrect. Pairwise scatterplots are used to explore the relationship between the independent variables.

C is incorrect. Because the pairwise scatterplot is used to determine whether a linear relationship exists between the dependent and independent variables.

Question 16 of 16

Courtney Berrios works as an analyst at Mellon Morgan Management (MMM), an asset management firm. She is tasked with analyzing the return drivers for a firm's portfolio by building a regression model of the portfolio's monthly excess returns (PMER) against the following three factors:

I. The market factor (MKT) - the expected market return minus the risk-free rate
II. The small minus big factor (SMB) - the average return on small portfolios minus the average return on large portfolios
III. The robust minus weak factor (RMW) - the average return on portfolios with robust operating profits minus those with weak operating profits

She gathers the necessary data and runs the regression. The resulting model is as follows:

$$Y_{PMER} = 0.075 + 1.671X_{MKT} + 0.851X_{SMB} + 1.987X_{RMW}$$

She then created some charts intended to help determine how accurately the model predicted the data.

Chart 1

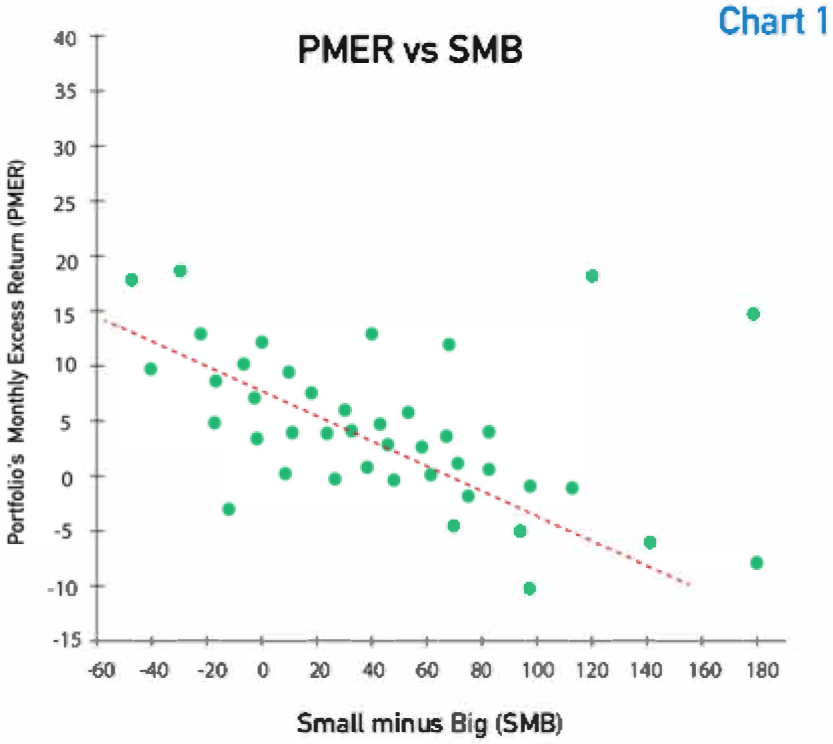


Chart 2

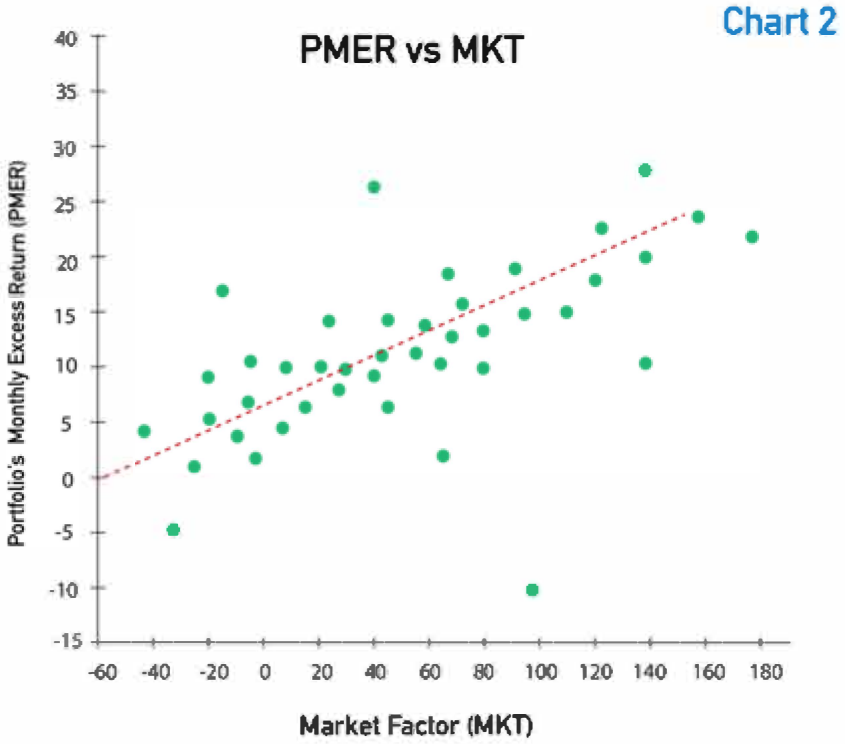


Chart 3

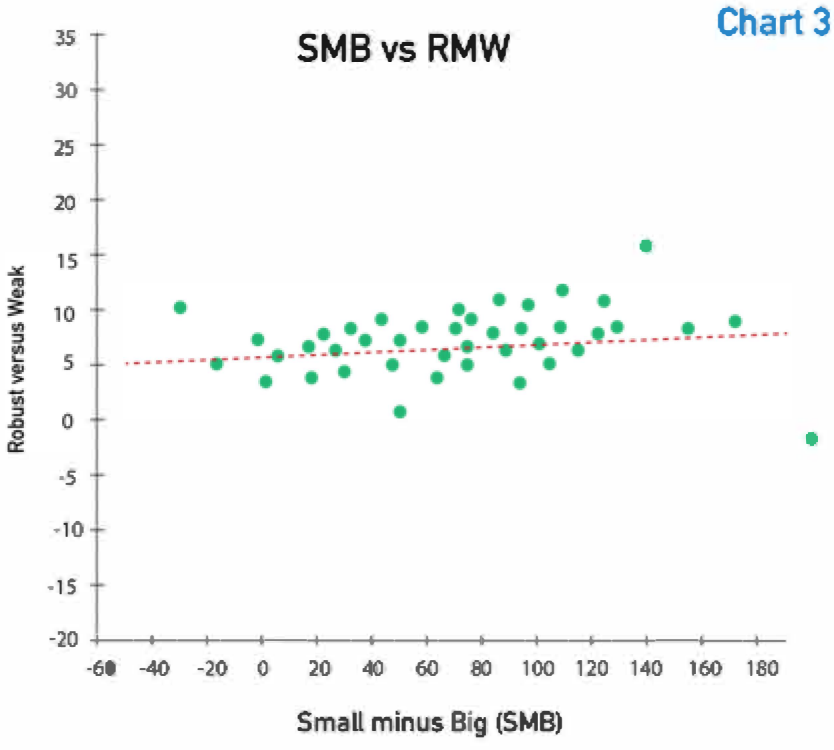
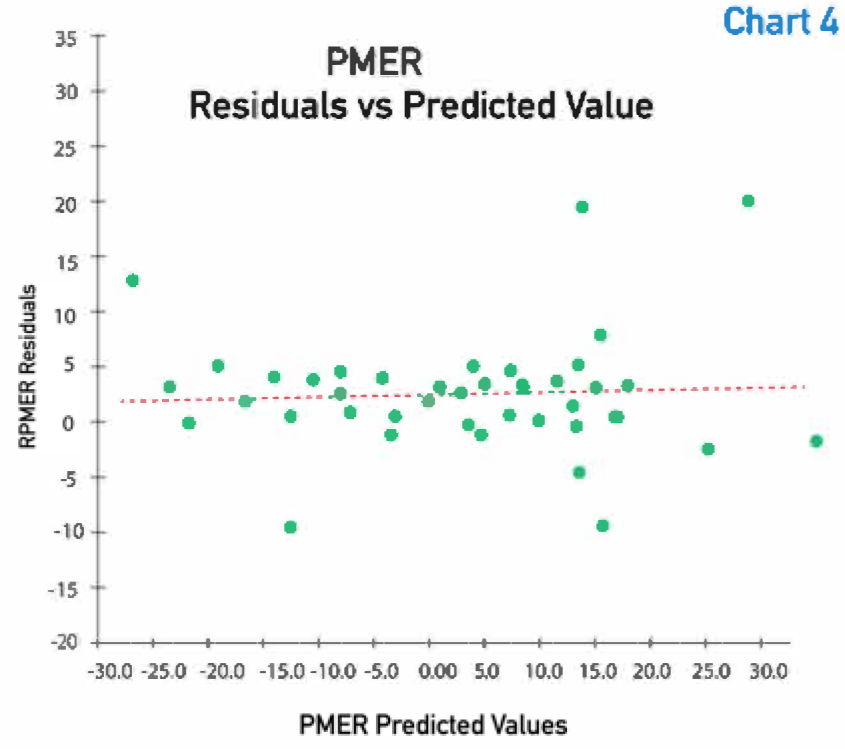


Chart 4



Part 1 of 4

Using the regression model, the predicted value of the portfolio's monthly excess returns if the MKT, SMB and RMW are each 1, is equal to:

A. 0.075.
B. 0.046.
C. 4.584.

Correct Answer: C

Explanation:

C is correct. Using the regression model, the predicted values of the portfolio's monthly excess returns if the MKT, SMB and RMW are each 1, is equal to 4.5165 as calculated below:

$$Y_{PMER} = 0.075 + 1.671X_{MKT} + 0.851X_{SMB} + 1.987X_{RMW}$$

By incorporating the following values to the above equation

$X_{MKT} = 1$
$X_{SMB} = 1$
$X_{RMW} = 1$

$$Y_{PMER} = 0.075 + 1.671 + 0.851 + 1.987 = 4.584$$

Quantitative Methods , Learning Module 1

Part 2 of 4

Based on the data provided in charts, Berrios can conclude that there is a positive relationship between portfolio's monthly excess return and the:

A. MKT factor.
B. SMB factor.
C. RMW factor.

Correct Answer: A

Explanation:

A is correct. Because the slope of the scatter plot between PMER and MKT in chart 2 is positive, we can conclude that there is a positive relationship between the two variables.

B is incorrect. There is a negative relationship between PMER and SMB (see chart 1).

C is incorrect. There is no chart that depicts the link between PMER and RMW.

Quantitative Methods , Learning Module 1

Part 3 of 4

Based on Chart 3, Berrios can *most likely* conclude that the regression assumption:

A. 'linearity' is not violated.
B. 'homoskedasticity' is violated.
C. 'independence of independent variables' is violated.

Correct Answer: C

Explanation:

C is correct. Based on chart 3, Berrios can conclude that the regression assumption independence of independent variables' is violated. Chart 3 shows an approximate linear relationship between SMB and RMW, which would be a violation of the independence of independent variables and should be investigated further.

A is incorrect. Linearity assumption states that the relation between the dependent variable Y and the independent variables (X_1, X_2, . . . ,X_k)is linear.

B is incorrect. In order to assess homoskedasticity, Berrios is required to examine whether the variance of the residuals of the regression model remains consistent across all observations.

Part 4 of 4

Which chart is *most likely* to be used to assess homoskedasticity ?

A. Chart 1
B. Chart 3
C. Chart 4

Correct Answer: C

Explanation:

C is correct. One of the assumptions of the simple linear regression is homoskedasticity which states that the variance of residuals should be the same for all observations. The Chart 4 depicts a scatterplot that showcases the regression residuals of PMER in relation to the predicted values of PMER. This chart serves as a valuable tool for visually evaluating the uniformity of the residuals' variance across all observations.

A is incorrect. Chart 1 can be used to see the linearity assumption of the simple linear regression as it is used to depict the relation between the dependent variable PMER and the independent variable SMB.

B is incorrect. Chart 3 can be used to see whether the independence of the independent variable is violated.