Question #1 of 139

Question ID: 1479918

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) Homoskedasticity.

C) Multicollinearity.

Question #2 of 139

Question ID: 1472066

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 of 139

Question ID: 1479908

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 of 139

Question ID: 1479904

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.

Question #5 of 139

Question ID: 1471891

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 #6 of 139

Question ID: 1471927

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) homoskedasticity.

B) a normal distribution.

C) a linear relationship.

Question #7 of 139

Question ID: 1471867

Consider the following estimated regression equation, with calculated *t*-statistics of the estimates as indicated:

$$AUTO_t = 10.0 + 1.25\ PI_t + 1.0\ TEEN_t - 2.0\ INS_t$$

with a PI calculated *t*-statistic of 0.45, a TEEN calculated *t*-statistic of 2.2, and an INS calculated *t*-statistic of 0.63.

The equation was estimated over 40 companies. Using a 5% level of significance, which of the independent variables significantly different from zero?

A) TEEN only.

B) PI and INS only.

C) PI only.

Question #8 of 139

Question ID: 1471873

Henry Hilton, CFA, is undertaking an analysis of the bicycle industry. He hypothesizes that bicycle sales (SALES) are a function of three factors: the population under 20 (POP), the level of disposable income (INCOME), and the number of dollars spent on advertising (ADV). All data are measured in millions of units. Hilton gathers data for the last 20 years and estimates the following equation (standard errors in parentheses):

SALES	= 0.000	+ 0.004 POP	+ 1.031 INCOME	+ 2.002 ADV
	(0.113)	(0.005)	(0.337)	(2.312)

For next year, Hilton estimates the following parameters: (1) the population under 20 will be 120 million, (2) disposable income will be $300,000,000, and (3) advertising expenditures will be $100,000,000. Based on these estimates and the regression equation, what are predicted sales for the industry for next year?

A) $656,991,000.

B) $509,980,000.

C) $557,143,000.

Question #9 of 139

Question ID: 1479883

One possible problem that could jeopardize the validity of the employment growth rate model is multicollinearity. Which of the following would *most likely* suggest the existence of multicollinearity?

A) The variance of the observations has increased over time.

B) The Durbin–Watson statistic is significant.

C) The F-statistic suggests that the overall regression is significant, however the regression coefficients are not individually significant.

Question #10 of 139

Question ID: 1586006

Consider the following estimated regression equation:

$$AUTO_t = 10.0 + 1.25\ PI_t + 1.0\ TEEN_t - 2.0\ INS_t$$

The equation was estimated over 40 companies. The predicted value of AUTO if PI is 4, TEEN is 0.30, and INS = 0.6 is *closest* to:

A) 14.90.

B) 17.50.

C) 14.10.

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:

CSP = 0.22 + 1.04 × DSC – 0.32 × index + 1.33 × 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 #11 - 14 of 139

Question ID: 1501587

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 #12 - 14 of 139

Question ID: 1501588

Based on the credit spread model, if an issuer gets included in the CDX index and assuming everything else the same, which of the following statements *most accurately* describes the model's forecast?

A) The credit spread on the firm’s issue would decrease by 10 bps.

B) The credit spread on the firm’s issue will increase by 32 bps.

C) The credit spread on the firm’s issue will decrease by 32 bps.

Question #13 - 14 of 139

Question ID: 1501589

Which 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 #14 - 14 of 139

Question ID: 1501590

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 #15 of 139

Question ID: 1471872

Henry Hilton, CFA, is undertaking an analysis of the bicycle industry. He hypothesizes that bicycle sales (SALES) are a function of three factors: the population under 20 (POP), the level of disposable income (INCOME), and the number of dollars spent on advertising (ADV). All data are measured in millions of units. Hilton gathers data for the last 20 years and estimates the following equation (standard errors in parentheses):

SALES	= α	+ 0.004 POP	+ 1.031 INCOME	+ 2.002 ADV
		(0.005)	(0.337)	(2.312)

The critical t-statistic for a 95% confidence level is 2.120. Which of the independent variables is statistically different from zero at the 95% confidence level?

A) INCOME and ADV.

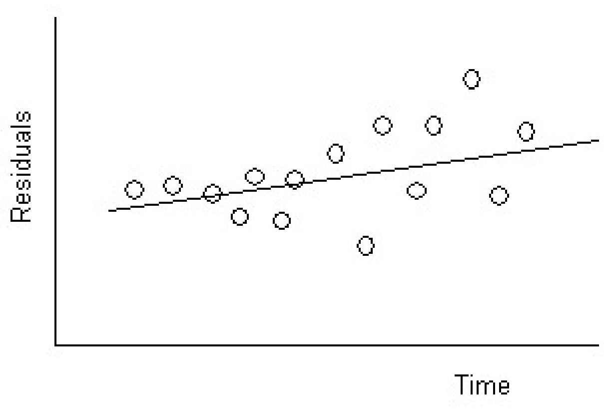
B) ADV only.

C) INCOME only.

Question #16 of 139

Question ID: 1472011

Consider the following graph of residuals and the regression line from a time-series regression:



These residuals exhibit the regression problem of:

A) autocorrelation.

B) homoskedasticity.

C) heteroskedasticity.

Using a recent analysis of salaries (in $1,000) of financial analysts, Timbadia runs a regression of salaries on education, experience, and gender. (Gender equals one for men and zero for women.) The regression results from a sample of 230 financial analysts are presented below, with *t*-statistics in parenthesis.

Salary	= 34.98	+ 1.2 Education	+ 0.5 Experience	+ 6.3 Gender
	(29.11)	(8.93)	(2.98)	(1.58)

Timbadia also runs a multiple regression to gain a better understanding of the relationship between lumber sales, housing starts, and commercial construction. The regression uses a large data set of lumber sales as the dependent variable with housing starts and commercial construction as the independent variables. The results of the regression are:

	Coefficient	Standard Error	t-statistics
Intercept	5.337	1.71	3.14
Housing starts	0.76	0.09	8.44
Commercial construction	1.25	0.33	3.78

Finally, Timbadia runs a regression between the returns on a stock and its industry index with the following results:

	Coefficient	Standard Error
Intercept	2.1	2.01
Industry index	1.9	0.31

- Standard error of estimate = 15.1
- Correlation coefficient = 0.849

Question #17 - 19 of 139

Question ID: 1479889

What is the expected salary (in $1,000) of a woman with 16 years of education and 10 years of experience?

A) 65.48.

B) 59.18.

C) 54.98.

Question #18 - 19 of 139

Question ID: 1479890

If the return on the industry index is 4%, the stock's expected return would be:

A) 9.7%.

B) 7.6%.

C) 11.2%.

Question #19 - 19 of 139

Question ID: 1479891

The percentage of the variation in the stock return explained by the variation in the industry index return is *closest* to:

A) 72.1%.

B) 63.2%.

C) 84.9%.

Question #20 of 139

Question ID: 1471980

An analyst runs a regression of monthly value-stock returns on five independent variables over 48 months. The total sum of squares is 430, and the sum of squared errors is 170. Test the null hypothesis at the 2.5% and 5% significance level that all five of the independent variables are equal to zero.

A) Rejected at 5% significance only.

B) Not rejected at 2.5% or 5.0% significance.

C) Rejected at 2.5% significance and 5% significance.

Question #21 of 139

Question ID: 1472075

Jill Wentraub is an analyst with the retail industry. She is modeling a company's sales over time and has noticed a quarterly seasonal pattern. If she includes dummy variables to represent the seasonality component of the sales she must use:

A) three dummy variables.

B) four dummy variables.

C) one dummy variables.

Question #22 of 139

Question ID: 1471947

An analyst regresses the return of a S&P 500 index fund against the S&P 500, and also regresses the return of an active manager against the S&P 500. The analyst uses the last five years of data in both regressions. Without making any other assumptions, which of the following is *most* accurate? The index fund:

A) regression should have higher sum of squares regression as a ratio to the total sum of squares.

B) should have a higher coefficient on the independent variable.

C) should have a lower coefficient of determination.

Question #23 of 139

Question ID: 1479923

Assume that in a particular multiple regression model, it is determined that the error terms are uncorrelated with each other. Which of the following statements is *most* accurate?

A) This model is in accordance with the basic assumptions of multiple regression analysis because the errors are not serially correlated.

B) Unconditional heteroskedasticity present in this model should not pose a problem, but can be corrected by using robust standard errors.

C) Serial correlation may be present in this multiple regression model, and can be confirmed only through a Durbin-Watson test.

Vijay Shapule, CFA, is investigating the application of the Fama-French three-factor model (Model 1) for the Indian stock market for the period 2001–2011 (120 months). Using the dependent variable as annualized return (%), the results of the analysis are shown in **Indian Equities—Fama-French Model**.

Indian Equities—Fama-French Model

Factor	Coefficient	P-Value
Intercept	1.22	<0.001
SMB	0.23	<0.001
HML	0.34	0.003
Rm-Rf	0.88	<0.001
R-squared		0.36
SSE		38.00
AIC		–129.99
BIC		–118.84

Shapule then modifies the model to include a liquidity factor. Results for this four-factor model (Model 2) are shown in **Revised Fama-French Model With Liquidity Factor**

Revised Fama-French Model With Liquidity Factor

Factor	Coefficient	P-Value
Intercept	1.56	<0.001

SMB	0.22	<0.001
HML	0.35	0.012
Rm-Rf	0.87	<0.001
LIQ	–0.12	0.02
R-squared		0.39
SSE		34.00
AIC		–141.34
BIC		–127.40

Question #24 - 27 of 139

Question ID: 1501592

The adjusted R^2 of Model 2 is *closest* to:

A) 0.39.

B) 0.37.

C) 0.36.

Question #25 - 27 of 139

Question ID: 1501593

The model better suited for prediction is:

A) Model 1 because it has a lower Bayesian information criterion.

B) Model 2 because it has a higher Akaike information criterion.

C) Model 2 because it has a lower Akaike information criterion.

Question #26 - 27 of 139

Question ID: 1501594

The *F*-statistic for testing H_0: coefficient of LIQ = 0 versus H_a: coefficient of LIQ ≠ 0 is *closest* to:

A) 13.33.

B) 5.45.

C) 2.11.

Question #27 - 27 of 139

Question ID: 1501595

What is the predicted return for a stock using Model 1 when SMB = 3.30, HML = 1.25 and Rm-Rf = 5?

A) 6.80%.

B) 7.88%.

C) 9.58%.

Question #28 of 139

Question ID: 1479901

A multiple regression model has included independent variables that are not linearly related to the dependent variable. The model is *most likely* misspecified due to:

A) incorrect data pooling.

B) incorrect variable form.

C) incorrect variable scaling.

Question #29 of 139

Question ID: 1472077

Consider the following model of earnings (EPS) regressed against dummy variables for the quarters:

$$EPS_t = \alpha + \beta_1 Q_{1t} + \beta_2 Q_{2t} + \beta_3 Q_{3t}$$

where:

EPS_t is a quarterly observation of earnings per share

Q_{1t} takes on a value of 1 if period t is the second quarter, 0 otherwise

Q_{2t} takes on a value of 1 if period t is the third quarter, 0 otherwise

Q_{3t} takes on a value of 1 if period t is the fourth quarter, 0 otherwise

Which of the following statements regarding this model is *most* accurate? The:

A) significance of the coefficients cannot be interpreted in the case of dummy variables.

B) coefficient on each dummy tells us about the difference in earnings per share between the respective quarter and the one left out (first quarter in this case).

C) EPS for the first quarter is represented by the residual.

Question #30 of 139

Question ID: 1479913

One of the main assumptions of a multiple regression model is that the variance of the residuals is constant across all observations in the sample. A violation of the assumption is *most likely* to be described as:

A) positive serial correlation.

B) heteroskedasticity.

C) unstable remnant deviation.

Question #31 of 139

Question ID: 1472012

Which of the following statements regarding heteroskedasticity is *least* accurate?

A) Heteroskedasticity may occur in cross-sectional or time-series analyses.

B) The assumption of linear regression is that the residuals are heteroskedastic.

C) Heteroskedasticity results in an estimated variance that is too small and, therefore, affects statistical inference.

Werner Baltz, CFA, has regressed 30 years of data to forecast future sales for National Motor Company based on the percent change in gross domestic product (GDP) and the change in retail price of a U.S. gallon of fuel. The results are presented below.

Predictor	Coefficient	Standard Error of the Coefficient
Intercept	78	13.710
Δ GDP	30.22	12.120
Δ $ Fuel	−412.39	183.981

Analysis of Variance Table (ANOVA)		
Source	**Degrees of Freedom**	**Sum of Squares**
Regression		291.30
Error	27	132.12
Total	29	423.42

Question #32 - 34 of 139

Question ID: 1479910

If GDP rises 2.2% and the price of fuels falls $0.15, Baltz's model will predict Company sales to be (in $ millions) *closest* to:

A) $82.00.

B) $128.00.

C) $206.00.

Question #33 - 34 of 139

Question ID: 1479911

Baltz proceeds to test the hypothesis that none of the independent variables has significant explanatory power. He concludes that, at a 5% level of significance:

A) all of the independent variables have explanatory power, because the calculated *F*-statistic exceeds its critical value.

B) none of the independent variables has explanatory power, because the calculated *F*-statistic does not exceed its critical value.

C) at least one of the independent variables has explanatory power, because the calculated *F*-statistic exceeds its critical value.

Question #34 - 34 of 139

Question ID: 1479912

Presence of conditional heteroskedasticity is *least likely* to affect the:

A) computed *F*-statistic.

B) coefficient estimates.

C) computed *t*-statistic.

Question #35 of 139

Question ID: 1471928

Which of the following statements *least* accurately describes one of the fundamental multiple regression assumptions?

A) The error term is normally distributed.

B) The independent variables are not random.

C) The variance of the error terms is not constant (i.e., the errors are heteroskedastic).

Question #36 of 139

Question ID: 1479878