Correction and Regression

Test ID: 7440246

Question #1 of 120

Question ID: 461521

Wanda Brunner, CFA, is working on a regression analysis based on publicly available macroeconomic time-series data. The most important limitation of regression analysis in this instance is:

- ✗ **A)** **the error term of one observation is not correlated with that of another observation.**
- ✓ **B)** limited usefulness in identifying profitable investment strategies.
- ✗ **C)** low confidence intervals.

Explanation

Regression analysis based on publicly available data is of limited usefulness if other market participants are also aware of and make use of this evidence.

Question #2 of 120

Question ID: 461464

The standard error of estimate is *closest* to the:

- ✓ **A)** **standard deviation of the residuals.**
- ✗ **B)** standard deviation of the independent variable.
- ✗ **C)** standard deviation of the dependent variable.

Explanation

The standard error of the estimate measures the uncertainty in the relationship between the actual and predicted values of the dependent variable. The differences between these values are called the residuals, and the standard error of the estimate helps gauge the fit of the regression line (the smaller the standard error of the estimate, the better the fit).

Question #3 of 120

Question ID: 461411

A simple linear regression equation had a coefficient of determination (R^2) of 0.8. What is the correlation coefficient between the dependent and independent variables and what is the covariance between the two variables if the variance of the independent variable is 4 and the variance of the dependent variable is 9?

	Correlation coefficient	Covariance
✓ **A)**	**0.89**	**5.34**
✗ **B)**	0.91	4.80
✗ **C)**	0.89	4.80

Explanation

The correlation coefficient is the square root of the R^2, r = 0.89.

To calculate the covariance multiply the correlation coefficient by the product of the standard deviations of the two variables:

COV = 0.89 × √4 × √9 = 5.34

Questions #4-9 of 120

A study of a sample of incomes (in thousands of dollars) of 35 individuals shows that income is related to age and years of education. The following table shows the regression results:

	Coefficient	*Standard Error*	*t-statistic*	*P-value*
Intercept	5.65	1.27	4.44	0.01
Age	0.53	?	1.33	0.21
Years of Education	2.32	0.41	?	0.01
Anova	*df*	*SS*	*MS*	*F*
Regression	?	215.10	?	?
Error	?	115.10	?	
Total	?	?		

Question #4 of 120

Question ID: 461508

The standard error for the coefficient of age and *t*-statistic for years of education are:

✗ **A)** **0.32; 1.65.**

✗ **B)** 0.53; 2.96.

✓ **C)** 0.40; 5.66.

Explanation

standard error for the coefficient of age = coefficient / *t*-value = 0.53 / 1.33 = 0.40

t-statistic for the coefficient of education = coefficient / standard error = 2.32 / 0.41 = 5.66

Question #5 of 120

Question ID: 461509

The mean square regression (MSR) is:

✗ **A)** **6.72.**

✗ **B)** 102.10.

✓ **C)** 107.55.

Explanation

df for Regression = k = 2

MSR = RSS / df = 215.10 / 2 = 107.55

Question #6 of 120

Question ID: 461510

The mean square error (MSE) is:

✗ **A)** **3.58.**

✗ **B)** 7.11.

✓ **C)** 3.60.

Explanation

df for Error = n - k - 1 = 35 - 2 - 1 = 32

MSE = SSE / df = 115.10 / 32 = 3.60

Question #7 of 120

Question ID: 461511

What is the R^2 for the regression?

✓ **A)** **65%.**

✗ **B)** 76%.

✗ **C)** 62%.

Explanation

SST = RSS + SSE

= 215.10 + 115.10

= 330.20

R^2= RSS / SST = 215.10 / 330.20 = 0.65

Question #8 of 120

Question ID: 461512

What is the predicted income of a 40-year-old person with 16 years of education?

✗ **A)** **$62,120.**

✓ **B)** $63,970.

✗ **C)** $74,890.

Explanation

income = 5.65 + 0.53 (age) + 2.32 (education)

= 5.65 + 0.53 (40) + 2.32 (16)

= 63.97 or $63,970

Question #9 of 120

Question ID: 461513

What is the F-value?

✓ **A) 29.88.**

✗ **B)** 1.88.

✗ **C)** 14.36.

Explanation

F = MSR / MSE = 107.55 / 3.60 = 29.88

Question #10 of 120

Question ID: 461458

Assume an analyst performs two simple regressions. The first regression analysis has an R-squared of 0.90 and a slope coefficient of 0.10. The second regression analysis has an R-squared of 0.70 and a slope coefficient of 0.25. Which one of the following statements is *most* accurate?

✓ **A) The first regression has more explanatory power than the second regression.**

✗ **B)** The influence on the dependent variable of a one unit increase in the independent variable is 0.9 in the first analysis and 0.7 in the second analysis.

✗ **C)** Results of the second analysis are more reliable than the first analysis.

Explanation

The coefficient of determination (R-squared) is the percentage of variation in the dependent variable explained by the variation in the independent variable. The larger R-squared (0.90) of the first regression means that 90% of the variability in the dependent variable is explained by variability in the independent variable, while 70% of that is explained in the second regression. This means that the first regression has more explanatory power than the second regression. Note that the Beta is the slope of the regression line and doesn't measure explanatory power.

Question #11 of 120

Question ID: 461481

Paul Frank is an analyst for the retail industry. He is examining the role of television viewing by teenagers on the sales of accessory stores. He gathered data and estimated the following regression of sales (in millions of dollars) on the number of hours watched by teenagers (TV, in hours per week):

> $\text{Sales}_t = 1.05 + 1.6\ \text{TV}_t$

The predicted sales if television watching is 5 hours per week is:

✓ **A) $9.05 million.**

✗ **B)** $8.00 million.

✗ **C)** $2.65 million.

Explanation

The predicted sales are: Sales = $1.05 + [$1.6 (5)] = $1.05 + $8.00 = $9.05 million.

Question #12 of 120

Question ID: 461433

The independent variable in a regression equation is called all of the following EXCEPT:

✗ **A)** **predicting variable.**

✗ **B)** explanatory variable.

✓ **C)** predicted variable.

Explanation

The dependent variable is the predicted variable.

Question #13 of 120

Question ID: 461429

Consider a sample of 60 observations on variables X and Y in which the correlation is 0.42. If the level of significance is 5%, we:

✗ **A)** **cannot test the significance of the correlation with this information.**

✗ **B)** conclude that there is no significant correlation between X and Y.

✓ **C)** conclude that there is statistically significant correlation between X and Y.

Explanation

The calculated t is $t = (0.42 \times \sqrt{58}) / \sqrt{(1-0.42^2)} = 3.5246$ and the critical t is approximately 2.000. Therefore, we reject the null hypothesis of no correlation.

Question #14 of 120

Question ID: 461453

Consider the following estimated regression equation:

> $ROE_t = 0.23 - 1.50\ CE_t$

The standard error of the coefficient is 0.40 and the number of observations is 32. The 95% confidence interval for the slope coefficient, b_1, is:

✗ **A)** **$\{-2.300 < b_1 < -0.700\}$.**

✓ **B)** $\{-2.317 < b_1 < -0.683\}$.

✗ **C)** $\{0.683 < b_1 < 2.317\}$.

Explanation

The confidence interval is $-1.50 \pm 2.042\ (0.40)$, or $\{-2.317 < b_1 < -0.683\}$.

Question #15 of 120

Question ID: 461405

In order to have a negative correlation between two variables, which of the following is *most* accurate?

✗ **A)** **Either the covariance or one of the standard deviations must be negative.**

✗ **B)** The covariance can never be negative.

✓ **C)** The covariance must be negative.

Explanation

In order for the correlation between two variables to be negative, the covariance must be negative. (Standard deviations are always positive.)

Question #16 of 120

Question ID: 461454

Assume you perform two simple regressions. The first regression analysis has an R-squared of 0.80 and a beta coefficient of 0.10. The second regression analysis has an R-squared of 0.80 and a beta coefficient of 0.25. Which one of the following statements is *most* accurate?

✗ **A) The influence on the dependent variable of a one-unit increase in the independent variable is the same in both analyses.**

✗ **B)** Results from the first analysis are more reliable than the second analysis.

✓ **C)** Explained variability from both analyses is equal.

Explanation

The coefficient of determination (R-squared) is the percentage of variation in the dependent variable explained by the variation in the independent variable. The R-squared (0.80) being identical between the first and second regressions means that 80% of the variability in the dependent variable is explained by variability in the independent variable for both regressions. This means that the first regression has the same explaining power as the second regression.

Question #17 of 120

Question ID: 461398

A sample covariance for the common stock of the Earth Company and the S&P 500 is −9.50. Which of the following statements regarding the estimated covariance of the two variables is *most* accurate?

✗ **A) The two variables will have a slight tendency to move together.**

✓ **B)** The relationship between the two variables is not easily predicted by the calculated covariance.

✗ **C)** The two variables will have a strong tendency to move in opposite directions.

Explanation

The actual value of the covariance for two variables is not very meaningful because its measurement is extremely sensitive to the scale of the two variables, ranging from negative to positive infinity. Covariance can, however be converted into the correlation coefficient, which is more straightforward to interpret.

Question #18 of 120

Question ID: 461479

An analyst has been assigned the task of evaluating revenue growth for an online education provider company that specializes in training adult students. She has gathered information about student ages, number of courses offered to all students each year, years of experience, annual income and type of college degrees, if any. A regression of annual dollar revenue on the number of courses offered each year yields the results shown below.

Coefficient Estimates		
Predictor	Coefficient	Standard Error of the Coefficient
Intercept	0.10	0.50
Slope (Number of Courses)	2.20	0.60

Which statement about the slope coefficient is *most correct*, assuming a 5% level of significance and 50 observations?

✗ **A) t-Statistic: 0.20. Slope: Not significantly different from zero.**

✓ **B)** t-Statistic: 3.67. Slope: Significantly different from zero.

✗ **C)** t-Statistic: 3.67. Slope: Not significantly different from zero.

Explanation

t = 2.20/0.60 = 3.67. Since the t-statistic is larger than an assumed critical value of about 2.0, the slope coefficient is statistically significant.

Question #19 of 120

Question ID: 461492

A simple linear regression is run to quantify the relationship between the return on the common stocks of medium sized companies (Mid Caps) and the return on the S&P 500 Index, using the monthly return on Mid Cap stocks as the dependent variable and the monthly return on the S&P 500 as the independent variable. The results of the regression are shown below:

	Coefficient	*Standard Error of Coefficient*	*t-Value*
Intercept	1.71	2.950	0.58
S&P 500	1.52	0.130	11.69
R^2 = 0.599			

Use the regression statistics presented above and assume this historical relationship still holds in the future period. If the expected return on the S&P 500 over the next period were 11%, the expected return on Mid Cap stocks over the next period would be:

✓ **A) 18.4%.**

✗ **B)** 20.3%.

✗ **C)** 33.8%.

Explanation

Y = intercept + slope(X)

Mid Cap Stock returns = 1.71 + 1.52(11) =18.4%

Question #20 of 120

Question ID: 461400

Unlike the coefficient of determination, the coefficient of correlation:

✗ **A) measures the strength of association between the two variables more exactly.**

✗ **B)** indicates the percentage of variation explained by a regression model.

✓ **C)** indicates whether the slope of the regression line is positive or negative.

Explanation

In a simple linear regression the coefficient of determination (R^2) is the squared correlation coefficient, so it is positive even when the correlation is negative.

Question #21 of 120

Question ID: 461477

Consider the regression results from the regression of Y against X for 50 observations:

> Y = 0.78 - 1.5 X
>
> The standard error of the estimate is 0.40 and the standard error of the coefficient is 0.45.

Which of the following reports the correct value of the *t*-statistic for the slope and correctly evaluates H_0: $b_1 \geq 0$ versus H_a: $b_1 < 0$ with 95% confidence?

✗ **A) *t* = 3.750; slope is significantly different from zero.**

✓ **B)** *t* = -3.333; slope is significantly negative.

✗ **C)** *t* = -3.750; slope is significantly different from zero.

Explanation

The test statistic is *t* = (-1.5 - 0) / 0.45 = -3.333. The critical 5%, one-tail *t*-value for 48 degrees of freedom is +/- 1.667. However, in the Schweser Notes you should use the closest degrees of freedom number of 40 df. which is +/-1.684. Therefore, the slope is less than zero. We reject the null in favor of the alternative.

Question #22 of 120

Question ID: 461461

Bea Carroll, CFA, has performed a regression analysis of the relationship between 6-month LIBOR and the U.S. Consumer Price Index (CPI). Her analysis indicates a standard error of estimate (SEE) that is high relative to total variability. Which of the following conclusions regarding the relationship between 6-month LIBOR and CPI can Carroll *most* accurately draw from her SEE analysis? The relationship between the two variables is:

✓ **A) very weak.**

✗ **B)** very strong.

✗ **C)** positively correlated.

Explanation

The SEE is the standard deviation of the error terms in the regression, and is an indicator of the strength of the relationship between the dependent and independent variables. The SEE will be low if the relationship is strong and conversely will be high if the relationship is weak.

Question #23 of 120

Question ID: 461451

The standard error of the estimate measures the variability of the:

- ✓ **A)** **actual dependent variable values about the estimated regression line.**
- ✗ **B)** predicted y-values around the mean of the observed y-values.
- ✗ **C)** values of the sample regression coefficient.

Explanation

The standard error of the estimate (SEE) measures the uncertainty in the relationship between the independent and dependent variables and helps gauge the fit of the regression line (the smaller the standard error of the estimate, the better the fit).

Remember that the SEE is different from the sum of squared errors (SSE). SSE = the sum of (actual value - predicted value)^2. SEE is the the square root of the SSE "standardized" by the degrees of freedom, or SEE = [SSE / (n - 2)]$^{1/2}$

Question #24 of 120

Question ID: 461460

The R^2 of a simple regression of two factors, A and B, measures the:

- ✗ **A)** **impact on B of a one-unit change in A.**
- ✗ **B)** statistical significance of the coefficient in the regression equation.
- ✓ **C)** percent of variability of one factor explained by the variability of the second factor.

Explanation

The coefficient of determination measures the percentage of variation in the dependent variable explained by the variation in the independent variable.

Question #25 of 120

Question ID: 461475

Consider the regression results from the regression of Y against X for 50 observations:

> Y = 0.78 + 1.2 X
>
> The standard error of the estimate is 0.40 and the standard error of the coefficient is 0.45.

Which of the following reports the correct value of the *t*-statistic for the slope and correctly evaluates its statistical significance with 95% confidence?

- ✗ **A)** ***t* = 1.789; slope is not significantly different from zero.**
- ✗ **B)** *t* = 3.000; slope is significantly different from zero.
- ✓ **C)** *t* = 2.667; slope is significantly different from zero.

Explanation

Perform a t-test to determine whether the slope coefficient if different from zero. The test statistic is *t* = (1.2 - 0) / 0.45 = 2.667. The critical *t*-values for 48 degrees of freedom are ± 2.011. Therefore, the slope is different from zero.

Question #26 of 120

Question ID: 461466

Which of the following statements about the standard error of the estimate (SEE) is *least* accurate?

✗ **A)** **The SEE will be high if the relationship between the independent and dependent variables is weak.**

✗ **B)** The SEE may be calculated from the sum of the squared errors and the number of observations.

✓ **C)** The larger the SEE the larger the R^2.

Explanation

The R^2, or coefficient of determination, is the percentage of variation in the dependent variable explained by the variation in the independent variable. A higher R^2 means a better fit. The SEE is smaller when the fit is better.

Question #27 of 120

Question ID: 461457

An analyst performs two simple regressions. The first regression analysis has an R-squared of 0.40 and a beta coefficient of 1.2. The second regression analysis has an R-squared of 0.77 and a beta coefficient of 1.75. Which one of the following statements is *most* accurate?

✗ **A)** **The R-squared of the first regression indicates that there is a 0.40 correlation between the independent and the dependent variables.**

✗ **B)** The first regression equation has more explaining power than the second regression equation.

✓ **C)** The second regression equation has more explaining power than the first regression equation.

Explanation

The coefficient of determination (R-squared) is the percentage of variation in the dependent variable explained by the variation in the independent variable. The larger R-squared (0.77) of the second regression means that 77% of the variability in the dependent variable is explained by variability in the independent variable, while only 40% of that is explained in the first regression. This means that the second regression has more explaining power than the first regression. Note that the Beta is the slope of the regression line and doesn't measure explaining power.

Question #28 of 120

Question ID: 461465

Jason Brock, CFA, is performing a regression analysis to identify and evaluate any relationship between the common stock of ABT Corp and the S&P 100 index. He utilizes monthly data from the past five years, and assumes that the sum of the squared errors is .0039. The calculated standard error of the estimate (SEE) is *closest* to:

✓ **A)** **0.0082.**

✗ **B)** 0.0360.

✗ **C)** 0.0080.

Explanation

The standard error of estimate of a regression equation measures the degree of variability between the actual and estimated Y-values. The SEE may also be referred to as the standard error of the residual or the standard error of the regression. The SEE is equal to the square root of the mean squared error. Expressed in a formula,

SEE = √(SSE / (n-2)) = √(.0039 / (60-2)) = .0082

Question #29 of 120

Question ID: 461403

Determine and interpret the correlation coefficient for the two variables X and Y. The standard deviation of X is 0.05, the standard deviation of Y is 0.08, and their covariance is −0.003.

✓ **A) −0.75 and the two variables are negatively associated.**

✗ **B)** −1.33 and the two variables are negatively associated.

✗ **C)** +0.75 and the two variables are positively associated.

Explanation

The correlation coefficient is the covariance divided by the product of the two standard deviations, i.e. −0.003 / (0.08 × 0.05).

Questions #30-35 of 120

Erica Basenj, CFA, has been given an assignment by her boss. She has been requested to review the following regression output to answer questions about the relationship between the monthly returns of the Toffee Investment Management (TIM) High Yield Bond Fund and the returns of the index (independent variable).

Regression Statistics					
R^2	??				
Standard Error	??				
Observations	20				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	23,516	23,516	?	?
Residual	18	?	7		
Total	19	23,644			
Regression Equation					
		Coefficients	*Std. Error*	*t-statistic*	*P-value*
Intercept		5.2900	1.6150	?	?
Slope		0.8700	0.0152	?	?

Question #30 of 120

Question ID: 485543

What is the value of the correlation coefficient?

✗ **A)** **0.8700.**

✗ **B)** −0.9973.

✓ **C)** 0.9973.

Explanation

R^2 is the correlation coefficient squared, taking into account whether the relationship is positive or negative. Since the value of the slope is positive, the TIM fund and the index are positively related. R^2 is calculated by taking the (RSS / SST) = 0.99459. $(0.99459)^{1/2}$ = 0.9973. (Study Session 3, LOS 9.i)

Question #31 of 120

Question ID: 485544

What is the sum of squared errors (SSE)?

✗ **A)** **23,644.**

✗ **B)** 23,515.

✓ **C)** 128.

Explanation

SSE = SST − RSS = 23,644 − 23,516 = 128. (Study Session 3, LOS 9.k)

Question #32 of 120

Question ID: 485545

What is the value of R^2?

✗ **A)** **0.9471.**

✓ **B)** 0.9946.

✗ **C)** 0.0055.

Explanation

R^2 = RSS / SST = 23,516 / 23,644 = 0.9946. (Study Session 3, LOS 9.k)

Question #33 of 120

Question ID: 485546

Is the intercept term statistically significant at the 5% level of significance and the 1% level of significance, respectively?

	1%	5%
✗ **A)**	**No**	**No**
✗ **B)**	Yes	No
✓ **C)**	Yes	Yes

Explanation

The test statistic is t = b / std error of b = 5.29 / 1.615 = 3.2755.

Critical *t*-values are ± 2.101 for the degrees of freedom = n − k − 1 = 18 for alpha = 0.05. For alpha = 0.01, critical *t*-values are ± 2.878. At both levels (two-tailed tests) we can reject H_0 that b = 0. (Study Session 3, LOS 9.g)

Question #34 of 120

Question ID: 485547

What is the value of the F-statistic?

✓ **A) 3,359.**

✗ **B)** 0.9945.

✗ **C)** 0.0003.

Explanation

F = mean square regression / mean square error = 23,516 / 7 = 3,359. (Study Session 3, LOS 9.k)

Question #35 of 120

Question ID: 485548

Heteroskedasticity can be defined as:

✓ **A) nonconstant variance of the error terms.**

✗ **B)** error terms that are dependent.

✗ **C)** independent variables that are correlated with each other.

Explanation

Heteroskedasticity occurs when the variance of the residuals is not the same across all observations in the sample. Autocorrelation refers to dependent error terms. (Study Session 3, LOS 10.m)

Question #36 of 120

Question ID: 461504

Consider the following analysis of variance (ANOVA) table:

Source	*Sum of squares*	*Degrees of freedom*	*Mean square*
Regression	500	1	500
Error	750	50	15
Total	1,250	51	

The R^2 and the F-statistic are, respectively:

✓ **A) R^2 = 0.40 and F = 33.333.**

✗ **B)** R^2 = 0.40 and F = 0.971.

✗ **C)** R^2 = 0.67 and F = 0.971.

Explanation

R^2 = 500 / 1,250 = 0.40. The F-statistic is 500 / 15 = 33.33.

Question #37 of 120

Question ID: 461408

Consider the case when the Y variable is in U.S. dollars and the X variable is in U.S. dollars. The 'units' of the covariance between Y and

X are:

✓ **A) squared U.S. dollars.**

✗ **B)** U.S. dollars.

✗ **C)** a range of values from −1 to +1.

Explanation

The covariance is in terms of the product of the units of Y and X. It is defined as the average value of the product of the deviations of observations of two variables from their means. The correlation coefficient is a standardized version of the covariance, ranges from −1 to +1, and is much easier to interpret than the covariance.

Question #38 of 120

Question ID: 461503

Consider the following analysis of variance (ANOVA) table:

Source	*Sum of squares*	*Degrees of freedom*	*Mean square*
Regression	200	1	200
Error	400	40	10
Total	600	41	

The R^2 and the F-statistic are, respectively:

✗ **A) R^2 = 33% and F = 2.0.**

✗ **B)** R^2 = 50% and F = 2.0.

✓ **C)** R^2 = 33% and F = 20.0.

Explanation

R^2 = 200 / 600 = 0.333. The F-statistic is 200 / 10 = 20.

Question #39 of 120

Question ID: 461418

Which of the following statements about linear regression is *least* accurate?

✗ **A) The independent variable is uncorrelated with the residuals (or disturbance term).**

✓ **B)** The correlation coefficient, ρ, of two assets x and y = ($\text{covariance}_{x,y}$) × standard deviation_x × standard deviation_y.

✗ **C)** R2 = RSS / SST.

Explanation

The correlation coefficient, ρ, of two assets x and y = ($\text{covariance}_{x,y}$) divided by (standard deviation_x × standard deviation_y). The other statements are true. *For the examination, memorize the assumptions underlying linear regression!*

Question #40 of 120

Question ID: 461478

A sample of 200 monthly observations is used to run a simple linear regression: Returns = $b_0 + b_1$Leverage + u. The t-value for the regression coefficient of leverage is calculated as t = - 1.09. A 5% level of significance is used to test whether leverage has a significant influence on returns. The correct decision is to:

✗ **A) do not reject the null hypothesis and conclude that leverage significantly explains returns.**

✗ **B)** reject the null hypothesis and conclude that leverage does not significantly explain returns.

✓ **C)** do not reject the null hypothesis and conclude that leverage does not significantly explain returns.

Explanation

Do not reject the null since |-1.09|<1.96(critical t-value).

Question #41 of 120

Question ID: 461459

A simple linear regression is run to quantify the relationship between the return on the common stocks of medium sized companies (Mid Caps) and the return on the S&P 500 Index, using the monthly return on Mid Cap stocks as the dependent variable and the monthly return on the S&P 500 as the independent variable. The results of the regression are shown below:

	Coefficient	*Standard Error of coefficient*	*t-Value*
Intercept	1.71	2.950	0.58
S&P 500	1.52	0.130	11.69

R^2= 0.599

The strength of the relationship, as measured by the correlation coefficient, between the return on Mid Cap stocks and the return on the S&P 500 for the period under study was:

✗ **A) 0.599.**

✓ **B)** 0.774.

✗ **C)** 0.130.

Explanation

You are given R^2 or the coefficient of determination of 0.599 and are asked to find R or the coefficient of correlation. The square root of 0.599 = 0.774.

Questions #42-47 of 120

Craig Standish, CFA, is investigating the validity of claims associated with a fund that his company offers. The company advertises the fund as having low turnover and, hence, low management fees. The fund was created two years ago with only a few uncorrelated assets. Standish randomly draws two stocks from the fund, Grey Corporation and Jars Inc., and measures the variances and covariance of their monthly returns over the past two years. The resulting variance covariance matrix is shown below. Standish will test whether it is reasonable to believe that the returns of Grey and Jars are uncorrelated. In doing the analysis, he plans to address the issue of spurious correlation and outliers.

	Grey	Jars
Grey	42.2	20.8
Jars	20.8	36.5

Standish wants to learn more about the performance of the fund. He performs a linear regression of the fund's monthly returns over the past two years on a large capitalization index. The results are below:

ANOVA

	df	*SS*	*MS*	*F*
Regression	1	92.53009	92.53009	28.09117
Residual	22	72.46625	3.293921	
Total	23	164.9963		

	Coefficients	*Standard Error*	*t-statistic*	*P-value*
Intercept	0.148923	0.391669	0.380225	0.707424
Large Cap Index	1.205602	0.227467	5.30011	2.56E-05

Standish forecasts the fund's return, based upon the prediction that the return to the large capitalization index used in the regression will be 10%. He also wants to quantify the degree of the prediction error, as well as the minimum and maximum sensitivity that the fund actually has with respect to the index.

He plans to summarize his results in a report. In the report, he will also include caveats concerning the limitations of regression analysis. He lists four limitations of regression analysis that he feels are important: relationships between variables can change over time, multicollinearity leads to inconsistent estimates of regression coefficients, if the error terms are heteroskedastic the standard errors for the regression coefficient may not be reliable, and if the error terms are correlated with each other over time the test statistics may not be reliable.

Question #42 of 120

Question ID: 485540

Given the variance/covariance matrix for Grey and Jars, in a one-sided hypothesis test that the returns are positively correlated H_0: $\rho \leq 0$ vs. H_1: $\rho > 0$, Standish would:

✗ **A) reject the null at the 5% but not the 1% level of significance.**

✓ **B)** reject the null at the 1% level of significance.

✗ **C)** need to gather more information before being able to reach a conclusion concerning significance.

Explanation

First, we must compute the correlation coefficient, which is $0.53 = 20.8 / (42.2 \times 36.5)^{0.5}$.

The *t*-statistic is: $2.93 = 0.53 \times [(24 - 2) / (1 - 0.53 \times 0.53)]^{0.5}$, and for df = 22 = 24 − 2, the *t*-statistics for the 5% and 1% level are 1.717 and 2.508 respectively. (Study Session 3, LOS 9.g)

Question #43 of 120

Question ID: 485541

In using the correlation coefficient between returns on Grey and Jars, Standish would *most* appropriately question the issue of:

✗ **A)** **issue of outliers but not the issue of spurious correlation.**

✗ **B)** spurious correlation but not the issue of outliers.

✓ **C)** Both spurious correlation and outliers.

Explanation

Both these issues are important in performing correlation analysis. A single outlier observation can change the correlation coefficient from significant to not significant and even from negative (positive) to positive (negative). Even if the correlation coefficient is significant, the researcher would want to make sure there is a reason for a relationship and that the correlation is not spurious (i.e., caused by chance). (Study Session 3, LOS 9.b)

Question #44 of 120

Question ID: 484162

If the large capitalization index has a 10% return, then the forecast of the fund's return will be:

✓ **A)** **12.2.**

✗ **B)** 16.1.

✗ **C)** 13.5.

Explanation

The forecast is 12.209 = 0.149 + 1.206 × 10, so the answer is 12.2. (Study Session 3, LOS 9.h)

Question #45 of 120

Question ID: 484163

The standard deviation of monthly fund returns is *closest to*:

✓ **A)** **2.68.**

✗ **B)** 12.84.

✗ **C)** 7.17.

Explanation

Variance of fund returns = SST/(n-1) = 164.9963/23 = 7.17. Standard deviation = $(7.17)^{0.5}$ = 2.68 (Study Session 3, LOS 9.j)

Question #46 of 120

Question ID: 484164

A 95% confidence interval for the slope coefficient is:

✗ **A)** **0.760 to 1.650.**

✓ **B)** 0.734 to 1.677.

✗ **C)** 0.905 to 1.506.

Explanation

The 95% confidence interval is 1.2056 ± (2.074 × 0.2275). Remember, to use 2-tailed t-statistic for confidence intervals. (Study Session 3, LOS 9.f)

Question #47 of 120

Question ID: 484165

Of the four caveats of regression analysis listed by Standish, the *least* accurate is:

✗ **A)** **the relationships of variables change over time.**

✓ **B)** multicollinearity leads to inconsistent estimates of the regression coefficients.

✗ **C)** if the error terms are heteroskedastic the standard errors for the regression coefficients may not be reliable.

Explanation

In the presence of multicollinearlity, the regression coefficients would still be consistent but unreliable. The other possible shortfalls listed are valid. (Study Session 3, LOS 9.k)

Question #48 of 120

Question ID: 461427

We are examining the relationship between the number of cold calls a broker makes and the number of accounts the firm as a whole opens. We have determined that the correlation coefficient is equal to 0.70, based on a sample of 16 observations. Is the relationship statistically significant at a 10% level of significance, why or why not? The relationship is:

✗ **A)** **not significant; the critical value exceeds the t-statistic by 1.91.**

✗ **B)** significant; the t-statistic exceeds the critical value by 3.67.

✓ **C)** significant; the t-statistic exceeds the critical value by 1.91.

Explanation

The calculated test statistic is t-distributed with n - 2 degrees of freedom:

$t = r\sqrt{(n - 2)} / \sqrt{(1 - r^2)} = 2.6192 / 0.7141 = 3.6678$

From a table, the critical value = 1.76

Question #49 of 120

Question ID: 461462

Which of the following statements about the standard error of estimate is *least* accurate? The standard error of estimate:

✗ **A)** **measures the Y variable's variability that is not explained by the regression equation.**

✓ **B)** is the square of the coefficient of determination.

✗ **C)** is the square root of the sum of the squared deviations from the regression line divided by (n − 2).

Explanation

Note: The coefficient of determination (R^2) is the *square of the correlation coefficient* in simple linear regression.

Question #50 of 120

Question ID: 461480

Consider the regression results from the regression of Y against X for 50 observations:

> Y = 5.0 + 1.5 X

The standard error of the coefficient is 0.50 and the standard error of the forecast is 0.52. The 95% confidence interval for the predicted value of Y if X is 10 is:

- ✗ **A)** **{18.980 < Y < 21.019}.**
- ✓ **B)** {18.954 < Y < 21.046}.
- ✗ **C)** {19.480 < Y < 20.052}.

Explanation

The predicted value of Y is: Y = 5.0 + [1.5 (10)] = 5.0 + 15 = 20. The confidence interval is 20 ± 2.011 (0.52) or {18.954 < Y < 21.046}.

Question #51 of 120

Question ID: 461442

Which of the following statements about linear regression analysis is *most* accurate?

- ✓ **A)** **An assumption of linear regression is that the residuals are independently distributed.**
- ✗ **B)** The coefficient of determination is defined as the strength of the linear relationship between two variables.
- ✗ **C)** When there is a strong relationship between two variables we can conclude that a change in one will cause a change in the other.

Explanation

Even when there is a strong relationship between two variables, we cannot conclude that a causal relationship exists. The coefficient of determination is defined as the percentage of total variation in the dependent variable explained by the independent variable.

Question #52 of 120

Question ID: 461420

A sample covariance of two random variables is most commonly utilized to:

- ✓ **A)** **calculate the correlation coefficient, which is a measure of the strength of their linear relationship.**
- ✗ **B)** identify and measure strong nonlinear relationships between the two variables.
- ✗ **C)** estimate the "pure" measure of the tendency of two variables to move together over a period of time.

Explanation

Since the actual value of a sample covariance can range from negative to positive infinity depending on the scale of the two variables, it is most commonly used to calculate a more useful measure, the correlation coefficient.

Question #53 of 120

Question ID: 461522

Regression analysis has a number of assumptions. Violations of these assumptions include which of the following?

✗ **A) Independent variables that are not normally distributed.**

✗ **B)** A zero mean of the residuals.

✓ **C)** Residuals that are not normally distributed.

Explanation

The assumptions include a normally distributed residual with a constant variance and a mean of zero.

Question #54 of 120

Question ID: 461414

For the case of simple linear regression with one independent variable, which of the following statements about the correlation coefficient is *least* accurate?

✗ **A) If the correlation coefficient is negative, it indicates that the regression line has a negative slope coefficient.**

✗ **B)** The correlation coefficient can vary between −1 and +1.

✓ **C)** If the regression line is flat and the observations are dispersed uniformly about the line, the correlation coefficient will be +1.

Explanation

Correlation analysis is a statistical technique used to measure the strength of the relationship between two variables. *The measure of this relationship is called the coefficient of correlation.*

If the regression line is flat and the observations are dispersed uniformly about the line,there is **no linear relationship** between the two variables and the **correlation coefficient will be zero.**

Both of the other choices are *CORRECT.*

Question #55 of 120

Question ID: 461436

In the estimated regression equation Y = 0.78 - 1.5 X, which of the following is *least* accurate when interpreting the slope coefficient?

✓ **A) If the value of X is zero, the value of Y will be -1.5.**

✗ **B)** The dependent variable declines by -1.5 units if X increases by 1 unit.

✗ **C)** The dependent variable increases by 1.5 units if X decreases by 1 unit.

Explanation

The slope represents the change in the dependent variable for a one-unit change in the independent variable. If the value of X is zero, the value of Y will be equal to the intercept, in this case, 0.78.

Question #56 of 120

Question ID: 461435

Which of the following is *least likely* an assumption of linear regression? The:

✓ **A) residuals are mean reverting; that is, they tend towards zero over time.**

✗ **B)** residuals are independently distributed.

✗ **C)** expected value of the residuals is zero.

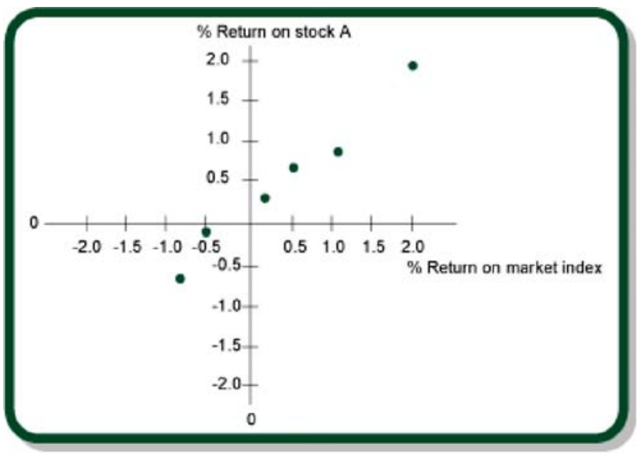
Explanation

The assumptions regarding the residuals are that the residuals have a constant variance, have a mean of zero, and are independently distributed.

Question #57 of 120

Question ID: 461399

In the scatter plot below, the correlation between the return on stock A and the market index is:



✓ **A) positive.**

✗ **B)** not discernable using the scatter plot.

✗ **C)** negative.

Explanation

In the scatter plot, higher values of the return on stock A are associated with higher values of the return on the market, i.e. a positive correlation between the two variables.

Question #58 of 120

Question ID: 461437

An analyst is examining the relationship between two random variables, RCRANTZ and GSTERN. He performs a linear regression that produces an estimate of the relationship:

> RCRANTZ = 61.4 − 5.9GSTERN

Which interpretation of this regression equation is *least* accurate?

✗ **A) The intercept term implies that if GSTERN is zero, RCRANTZ is 61.4.**

✗ **B)** The covariance of RCRANTZ and GSTERN is negative.

✓ **C)** If GSTERN increases by one unit, RCRANTZ should increase by 5.9 units.

Explanation

The slope coefficient in this regression is -5.9. This means a one unit increase of GSTERN suggests a *decrease* of 5.9 units of RCRANTZ. The slope coefficient is the covariance divided by the variance of the independent variable. Since variance (a squared term) must be positive, a negative slope term implies that the covariance is negative.

Question #59 of 120

Question ID: 461421

Ron James, CFA, computed the correlation coefficient for historical oil prices and the occurrence of a leap year and has identified a statistically significant relationship. Specifically, the price of oil declined every fourth calendar year, all other factors held constant. James has *most* likely identified which of the following conditions in correlation analysis?

✗ **A)** **Positive correlation.**

✓ **B)** Spurious correlation.

✗ **C)** Outliers.

Explanation

Spurious correlation occurs when the analysis erroneously indicates a linear relationship between two variables when none exists. There is no economic explanation for this relationship; therefore this would be classified as spurious correlation.

Question #60 of 120

Question ID: 461467

The *most* appropriate measure of the degree of variability of the actual *Y*-values relative to the estimated *Y*-values from a regression equation is the:

✗ **A)** **sum of squared errors (SSE).**

✓ **B)** standard error of the estimate (SEE).

✗ **C)** coefficient of determination (R^2).

Explanation

The SEE is the standard deviation of the error terms in the regression, and is an indicator of the strength of the relationship between the dependent and independent variables. The SEE will be low if the relationship is strong, and conversely will be high if the relationship is weak.

Question #61 of 120

Question ID: 461483

A variable Y is regressed against a single variable X across 24 observations. The value of the slope is 1.14, and the constant is 1.3. The mean value of X is 1.10, and the mean value of Y is 2.67. The standard deviation of the X variable is 1.10, and the standard deviation of the Y variable is 2.46. The sum of squared errors is 89.7. For an X value of 1.0, what is the 95% confidence interval for the Y value?

✗ **A)** **−1.68 to 6.56.**

✓ **B)** −1.83 to 6.72.

✗ **C)** 0.59 to 4.30.

Explanation

First the standard error of the estimate must be calculated-it is equal to the square root of the mean squared error, which is equal to the sum of squared errors divided by the number of observations minus 2 = $(89.7 / 22)^{1/2}$ = 2.02. The variance of the prediction is equal to:

$$S_f^2 = SEE^2\left[1 + \frac{1}{n} + \frac{(X - \bar{X})^2}{(n - 1)\,s_x^2}\right]$$

$$= \left[2.02^2\left(1 + \frac{1}{24} + \frac{(1.0 - 1.1)^2}{(24 - 1)1.1^2}\right)\right]^{½}$$

= 2.06

The prediction value is 1.3 + (1.0 × 1.14) = 2.44. The *t*-value for 22 degrees of freedom is 2.074. The endpoints of the interval are 2.44 ± 2.074 × 2.06 = −1.83 and 6.72.

Question #62 of 120

Question ID: 461426

Suppose the covariance between Y and X is 10, the variance of Y is 25, and the variance of X is 64. The sample size is 30. Using a 5% level of significance, which of the following statements is *most* accurate? The null hypothesis of:

✗ **A)** **no correlation is rejected.**

✓ **B)** no correlation cannot be rejected.

✗ **C)** significant correlation is rejected.

Explanation

The correlation coefficient is r = 10 / (5 × 8) = 0.25. The test statistic is t = (0.25 × √28) / √(1 − 0.0625) = 1.3663. The critical t-values are ± 2.048. Therefore, we cannot reject the null hypothesis of no correlation.

Question #63 of 120

Question ID: 461505

Consider the following analysis of variance (ANOVA) table:

Source	*Sum of squares*	*Degrees of freedom*	*Mean square*
Regression	556	1	556
Error	679	50	13.5
Total	1,235	51	

The R^2 for this regression is:

✗ **A)** **0.55.**

✓ **B)** 0.45.

✗ **C)** 0.82.

Explanation

R^2 = RSS/SST = 556/1,235 = 0.45.

Question #64 of 120

Question ID: 461406

Rafael Garza, CFA, is considering the purchase of ABC stock for a client's portfolio. His analysis includes calculating the covariance between the returns of ABC stock and the equity market index. Which of the following statements regarding Garza's analysis is *most* accurate?

✗ **A)** **A covariance of +1 indicates a perfect positive covariance between the two variables.**

✗ **B)** The covariance measures the strength of the linear relationship between two variables.

✓ **C)** The actual value of the covariance is not very meaningful because the measurement is very sensitive to the scale of the two variables.

Explanation

Covariance is a statistical measure of the linear relationship of two random variables, but the actual value is not meaningful because the measure is extremely sensitive to the scale of the two variables. Covariance can range from negative to positive infinity.

Question #65 of 120

Question ID: 461401

Which of the following statements regarding scatter plots is *most* accurate? Scatter plots:

✗ **A)** **are used to examine the third moment of a distribution (skewness).**

✗ **B)** illustrate the scatterings of a single variable.

✓ **C)** illustrate the relationship between two variables.

Explanation

A scatter plot is a collection of points on a graph where each point represents the values of two variables. They are used to examine the relationship between two variables.

Question #66 of 120

Question ID: 461463

A regression between the returns on a stock and its industry index returns gives the following results:

	Coefficient	*Standard Error*	*t-value*
Intercept	2.1	2.01	1.04
Industry Index	1.9	0.31	6.13

- The *t*-statistic critical value at the 0.01 level of significance is 2.58
- Standard error of estimate = 15.1
- Correlation coefficient = 0.849

The regression statistics presented indicate that the dispersion of stock returns about the regression line is:

✗ **A)** **72.10.**

✓ **B)** 15.10.

✗ **C)** 63.20.

Explanation

The standard deviation of the differences between the actual observations and the projection of those observations (the regression line) is called the standard error of the estimate. The standard error of the estimate is the unsystematic risk.

Question #67 of 120

Question ID: 461409

Which of the following statements about covariance and correlation is *least* accurate?

✗ **A)** **A zero covariance implies a zero correlation.**

✓ **B)** There is no relation between the sign of the covariance and the correlation.

✗ **C)** The covariance and correlation are always the same sign, positive or negative.

Explanation

The other two choices are accurate statements. The correlation is the ratio of the covariance to the product of the standard deviations of the two variables. Therefore, the covariance and the correlation have the same sign, and a zero covariance implies a zero correlation.

Question #68 of 120

Question ID: 461476

The *most* appropriate test statistic to test statistical significance of a regression slope coefficient with 45 observations and 2 independent variables is a:

✗ **A)** **one-tail *t*-statistic with 42 degrees of freedom.**

✓ **B)** two-tail *t*-statistic with 42 degrees of freedom.

✗ **C)** one-tail *t*-statistic with 43 degrees of freedom.

Explanation

df = n − k − 1 = 45 − 2 − 1

Question #69 of 120

Question ID: 461438

Which of the following is *least likely* an assumption of linear regression?

✗ **A)** **The residuals are normally distributed.**

✓ **B)** The independent variable is correlated with the residuals.

✗ **C)** The variance of the residuals is constant.

Explanation

The assumption is that the independent variable is uncorrelated with the residuals.

Question #70 of 120

Question ID: 461450

If X and Y are perfectly correlated, regressing Y onto X will result in which of the following:

✗ **A) the regression line will be sloped upward.**

✗ **B)** the alpha coefficient will be zero.

✓ **C)** the standard error of estimate will be zero.

Explanation

If X and Y are perfectly correlated, all of the points will plot on the regression line, so the standard error of the estimate will be zero. Note that the sign of the correlation coefficient will indicate which way the regression line is pointing (there can be perfect negative correlation sloping downward as well as perfect positive correlation sloping upward). Alpha is the intercept and is not directly related to the correlation.

Question #71 of 120

Question ID: 461419

The table below shows a sample of returns on two securities:

Period	1	2	3	4	*Mean*
Security P	0.2%	0.5%	1.1%	−0.6%	*0.3%*
Security Q	−0.3%	0.9%	1.5%	−0.5%	*0.4%*

The sample covariance between the two securities' returns is closest to:

✗ **A) 0.78.**

✓ **B)** 0.62.

✗ **C)** 0.47.

Explanation

Period	1	2	3	4	*Sum*
$P_i - \bar{P}$	−0.1	0.2	0.8	−0.9	
$Q_i - \bar{Q}$	−0.7	0.5	1.1	−0.9	
$(P_i - \bar{P})(Q_i - \bar{Q})$	0.07	0.10	0.88	0.81	*1.86*

$$\text{Sample covariance} = \frac{\sum_{i=1}^{n}(P_i - \bar{P})(Q_i - \bar{Q})}{n-1} = \frac{1.86}{3} = 0.62$$

Question #72 of 120

Question ID: 461494

A dependent variable is regressed against a single independent variable across 100 observations. The mean squared error is 2.807, and the mean regression sum of squares is 117.9. What is the correlation coefficient between the two variables?

✗ **A) 0.99.**

✗ **B)** 0.30.

✓ **C)** 0.55.

Explanation

The correlation coefficient is the square root of the R^2, which can be found by dividing the regression sum of squares by the total sum of squares. The regression sum of squares is the mean regression sum of squares multiplied by the number of independent variables, which is 1, so the regression sum of squares is equal to 117.9. The residual sum of squares is the mean squared error multiplied by the denominator degrees of freedom, which is the number of observations minus the number of independent variables, minus 1, which is equal to 100 − 1 − 1 = 98. The residual sum of squares is then 2.807 × 98 = 275.1. The total sum of squares is the sum of the regression sum of squares and the residual sum of squares, which is 117.9 + 275.1 = 393.0. The R^2 = 117.9 / 393.0 = 0.3, so the correlation is the square root of 0.3 = 0.55.

Question #73 of 120

Question ID: 461456

What does the R^2 of a simple regression of two variables measure and what calculation is used to equate the correlation coefficient to the coefficient of determination?

	R^2 measures:	Correlation coefficient
✓ **A)**	**percent of variability of the dependent variable that is explained by the variability of the independent variable**	**$R^2 = r^2$**
✗ **B)**	percent of variability of the independent variable that is explained by the variability of the dependent variable	$R^2 = r \times 2$
✗ **C)**	percent of variability of the independent variable that is explained by the variability of the dependent variable	$R^2 = r^2$

Explanation

R^2, or the Coefficient of Determination, is the square of the coefficient of correlation (r). The coefficient of correlation describes the strength of the relationship between the X and Y variables. The standard error of the residuals is the standard deviation of the dispersion about the regression line. The t-statistic measures the statistical significance of the coefficients of the regression equation. In the response: "percent of variability of the independent variable that is explained by the variability of the dependent variable," the definitions of the variables are reversed.

Question #74 of 120

Question ID: 461432

The capital asset pricing model is given by: $R_i = R_f$ + Beta ($R_m - R_f$) where R_m = expected return on the market, R_f = risk-free market and

R_i = expected return on a specific firm. The dependent variable in this model is:

✗ **A)** **R_f.**

✗ **B)** R_m - R_f.

✓ **C)** R_i.

Explanation

The dependent variable is the variable whose variation is explained by the other variables. Here, the variation in R_i is explained by the variation in the other variables, R_f and R_m.

Question #75 of 120

Question ID: 461422

One major limitation of the correlation analysis of two random variables is when two variables are highly correlated, but no economic relationship exists. This condition *most* likely indicates the presence of:

✓ **A)** **spurious correlation.**

✗ **B)** nonlinear relationships.

✗ **C)** outliers.

Explanation

Spurious correlation occurs when the analysis erroneously indicates a relationship between two variables when none exists.

Question #76 of 120

Question ID: 461502

Consider the following analysis of variance (ANOVA) table:

Source	*Sum of squares*	*Degrees of freedom*	*Mean square*
Regression	550	1	550.000
Error	750	38	19.834
Total	1,300	39	

The F-statistic for the test of the fit of the model is *closest* to:

✗ **A)** **0.965.**

✗ **B)** 0.423.

✓ **C)** 27.730.

Explanation

F = Mean Square of Regression / Mean Square of Error = 550 / 19.834 = 27.730.

Question #77 of 120

Question ID: 461430

Joe Harris is interested in why the returns on equity differ from one company to another. He chose several company-specific variables to explain the return on equity, including financial leverage and capital expenditures. In his model:

✗ **A) return on equity is the independent variable, and financial leverage and capital expenditures are dependent variables**

✓ **B)** return on equity is the dependent variable, and financial leverage and capital expenditures are independent variables.

✗ **C)** return on equity, financial leverage, and capital expenditures are all independent variables.

Explanation

The dependent variable is return on equity. This is what he wants to explain. The variables he uses to do the explaining (i.e., the independent variables) are financial leverage and capital expenditures.

Question #78 of 120

Question ID: 461415

The Y variable is regressed against the X variable resulting in a regression line that is horizontal with the plot of the paired observations widely dispersed about the regression line. Based on this information, which statement is *most likely accurate*?

✗ **A) X is perfectly positively correlated to Y.**

✓ **B)** The correlation between X and Y is close to zero.

✗ **C)** The R^2 of this regression is close to 100%.

Explanation

Perfect correlation means that all the observations fall on the regression line. An R^2 of 100% means perfect correlation. When there is no correlation, the regression line is horizontal.

Question #79 of 120

Question ID: 461410

Which of the following statements regarding the coefficient of determination is *least* accurate? The coefficient of determination:

✗ **A) cannot decrease as independent variables are added to the model.**

✗ **B)** is the percentage of the total variation in the dependent variable that is explained by the independent variable.

✓ **C)** may range from −1 to +1.

Explanation

In a simple regression, the coefficient of determination is calculated as the correlation coefficient squared and ranges from 0 to +1.

Question #80 of 120

Question ID: 461484

Given: Y = 2.83 + 1.5X

What is the predicted value of the dependent variable when the value of an independent variable equals 2?

✓ **A) 5.83**

✗ **B)** -0.55

✗ **C)** 2.83

Explanation

Y = 2.83 + (1.5)(2)

= 2.83 + 3

= 5.83

Question #81 of 120

Question ID: 461434

Paul Frank is an analyst for the retail industry. He is examining the role of television viewing by teenagers on the sales of accessory stores. He gathered data and estimated the following regression of sales (in millions of dollars) on the number of hours watched by teenagers (in hours per week):

> $Sales_t = 1.05 + 1.6\ TV_t$

Which of the following is the *most* accurate interpretation of the estimated results? If TV watching:

✓ **A) goes up by one hour per week, sales of accessories increase by $1.6 million.**

✗ **B)** changes, no change in sales is expected.

✗ **C)** goes up by one hour per week, sales of accessories increase by $1.60.

Explanation

The interpretation of the slope coefficient is the change in the dependent variable (sales in millions of dollars) for a given one-unit change in the independent variable (TV hours per week). The intercept of 1.05 means that 1.05 million dollars worth of accessories is expected to be sold even if TV watching is zero.

Question #82 of 120

Question ID: 461413

A sample of paired points A and B is shown below. What is the covariance between the values of A and B?

Sample	*A*	*B*
1	1	2
2	4	5
3	9	7
4	11	10
5	14	12

✓ **A) 20.55.**

✗ **B)** 29.76.

✗ **C)** 7.80.

Explanation

Sample	*A*	*A-mean (A)*	*B*	*B-mean (B)*	*Product*
1	1	−6.8	2	−5.2	35.36
2	4	−3.8	5	−2.2	8.36
3	9	1.2	7	−0.2	−0.24
4	11	3.2	10	2.8	8.96
5	14	6.2	12	4.8	29.76
mean	7.8		7.2	Sum	82.2

Cov = 82.20 / (5 - 1) = 20.55

Questions #83-88 of 120

Cynthia Jones is Director of Marketing at Vancouver Industries, a large producer of athletic apparel and accessories. Approximately three years ago, Vancouver experienced increased competition in the marketplace, and consequently sales for that year declined nearly 20%. At that time, Jones proposed a new marketing campaign for the company, aimed at positioning Vancouver's product lines toward a younger target audience. Although the new marketing effort was significantly more costly than previous marketing campaigns, Jones assured her superiors that the resulting increase in sales would more than justify the additional expense. Jones was given approval to proceed with the implementation of her plan.

Three years later, in preparation for an upcoming shareholders meeting, the CEO of Vancouver has asked Jones for an evaluation of the marketing campaign. Sales have increased since the inception of the new marketing campaign nearly three years ago, but the CEO is questioning whether the increase is due to the marketing expenditures or can be attributed to other factors. Jones is examining the following data on the firm's aggregate revenue and marketing expenditure over the last 10 quarters. Jones plans to demonstrate the effectiveness of marketing in boosting sales revenue. She chooses to utilize a simple linear regression model. The output is as follows:

	Aggregate Revenue (Y)	*Advertising Expenditure (X)*	*Y^2*	*XY*	*X^2*
	300	7.5	90,000	2,250	56.25
	320	9.0	102,400	2,880	81.00
	310	8.5	96,100	2,635	72.25
	335	8.2	112,225	2,747	67.24
	350	9.0	122,500	3,150	81.00
	400	8.5	160,000	3,400	72.25
	430	10.0	184,900	4,300	100.00
	390	10.5	152,100	4,095	110.25
	380	9.0	144,400	3,420	81.00
	430	11.0	184,900	4,730	121.00
TOTAL	**3,645**	**91.2**	**1,349,525**	**33,607**	**842.24**

> Slope coefficient = 34.74 Standard error of slope coefficient = 9.916629313 Standard error of intercept = 92.2840128

ANOVA			
	Df	*SS*	*MS*
Regression	1	12,665.125760	12,665.12576
Residual	8	8,257.374238	1,032.17178
Total	9	20,922.5	

Jones discusses her findings with her market research specialist, Mira Nair. Nair tells Jones that she should check her model for heteroskedasticity. She explains that in the presence of conditional heteroskedasticity, the model coefficients and t-statistics will be biased.

For the questions below, assume a t-value of 2.306.

Question #83 of 120

Question ID: 461444

Which of the following is *closest to* the upper limit of the 95% confidence interval for the slope coefficient?

✓ **A) 57.61.**

✗ **B)** 111.72.

✗ **C)** 62.84.

Explanation

Upper Limit = coefficient + (2.306 x standard error of the coefficient)

= 34.74 + (2.306 x 9.917) = 57.61

(Study Session 3, LOS 11.f)

Question #84 of 120

Question ID: 461445

Which of the following is *closest to* the lower limit of the 95% confidence interval for the slope coefficient?

✓ **A) 11.87.**

✗ **B)** 72.84.

✗ **C)** 12.24.

Explanation

Lower Limit = Coefficient - (2.306 x Standard Error of the coefficient)

= 34.74 - (2.306 x 9.917)

= 34.74 - 22.87 = 11.87

(Study Session 3, LOS 11.f)

Question #85 of 120

Question ID: 461446

Which of the following is the CORRECT value of the correlation coefficient between aggregate revenue and advertising expenditure?

✗ **A)** **0.9500.**

✓ **B)** 0.7780.

✗ **C)** 0.6053.

Explanation

The R^2 = (SST - SSE)/SST = RSS/SST = (20,922.5 - 8,257.374) / 20,922.5 = 0.6053.

The correlation coefficient is the square root of the R^2 in a simple linear regression which is the square root of 0.6053 = 0.7780. (Study Session 3, LOS 11.j)

Question #86 of 120

Question ID: 461447

Which of the following reports the CORRECT value and interpretation of the R^2 for this regression? The R^2 is:

✗ **A)** **0.3947 indicating that the variability of advertising expenditure explains about 39.47% of the variability of aggregate revenue.**

✗ **B)** 0.3947 indicating that the variability of aggregate revenue explains about 39.47% of the variability in advertising expenditure.

✓ **C)** 0.6053 indicating that the variability of advertising expenditure explains about 60.53% of the variability in aggregate revenue.

Explanation

The R^2 = (SST - SSE)/SST = (20,922.5 - 8,257.374) / 20,922.5 = 0.6053.

The interpretation of this R^2 is that 60.53% of the variation in aggregate revenue (Y) is explained by the variation in advertising expenditure (X). (Study Session 3, LOS 11.j)

Question #87 of 120

Question ID: 461448

Is Nair's statement about conditional heteroskedasticity CORRECT?

✗ **A)** **No, because the t-statistics will not be biased.**

✗ **B)** Yes, because both the coefficients and t-statistics will be biased.

✓ **C)** No, because coefficients will not be biased.

Explanation

Conditional heteroskedasticity will result in consistent coefficient estimates but inconsistent standard errors resulting in biased t-statistics. (Study Session 3, LOS 12.k)

Question #88 of 120

Question ID: 461449

What is the calculated F-statistic?

✗ **A)** **92.2840.**

✓ **B)** 12.2700.

✗ **C)** 0.1250.

Explanation

The computed value of the F-Statistic = MSR/MSE = 12,665.12576 / 1,032.17178 = 12.27, where MSR and MSE are from the ANOVA table. (Study Session 3, LOS 11.j)

Question #89 of 120

Question ID: 461440

Linear regression is based on a number of assumptions. Which of the following is *least* likely an assumption of linear regression?

✗ **A)** **Values of the independent variable are not correlated with the error term.**

✓ **B)** There is at least some correlation between the error terms from one observation to the next.

✗ **C)** The variance of the error terms each period remains the same.

Explanation

When correlation (between the error terms from one observation to the next) exists, autocorrelation is present. As a result, residual terms are not normally distributed. This is inconsistent with linear regression.

Question #90 of 120

Question ID: 461424

Suppose the covariance between Y and X is 0.03 and that the variance of Y is 0.04 and the variance of X is 0.12. The sample size is 30. Using a 5% level of significance, which of the following is *most* accurate? The null hypothesis of:

✗ **A)** **no correlation is not rejected.**

✗ **B)** significant correlation is rejected.

✓ **C)** no correlation is rejected.

Explanation

The correlation coefficient is $r = 0.03 / (\sqrt{0.04} * \sqrt{0.12}) = 0.03 / (0.2000 * 0.3464) = 0.4330$.

The test statistic is $t = (0.4330 \times \sqrt{28}) / \sqrt{(1 - 0.1875)} = 2.2912 / 0.9014 = 2.54$.

We can find the critical t-values from a t-table, using df = 28 and two-tailed 95% significance (recall that for a t-test the degrees of freedom = n − 2).

The critical *t*-values are ± 2.048. Therefore, we reject the null hypothesis of no correlation.

Question #91 of 120

Question ID: 461402

If the correlation between two variables is −1.0, the scatter plot would appear along a:

✗ **A)** **a curved line running from southwest to northeast.**

✗ **B)** straight line running from southwest to northeast.

✓ **C)** straight line running from northwest to southeast.

Explanation

If the correlation is −1.0, then higher values of the y-variable will be associated with lower values of the x-variable. The points would lie on a straight line running from northwest to southeast.

Questions #92-97 of 120

Rebecca Anderson, CFA, has recently accepted a position as a financial analyst with Eagle Investments. She will be responsible for providing analytical data to Eagle's portfolio manager for several industries. In addition, she will follow each of the major public corporations within each of those industries. As one of her first assignments, Anderson has been asked to provide a detailed report on one of Eagle's current investments. She was given the following data on sales for Company XYZ, the maker of toilet tissue, as well as toilet tissue industry sales ($ millions). She has been asked to develop a model to aid in the prediction of future sales levels for Company XYZ. She proceeds by recalling some of the basic concepts of regression analysis she learned while she was preparing for the CFA exam.

Year	*Industry Sales (X)*	*Company Sales (Y)*	*$(X-\overline{X})^2$*
1	$3,000	$750	84,100
2	$3,200	$800	8,100
3	$3,400	$850	12,100
4	$3,350	$825	3,600
5	$3,500	$900	44,100
Totals	$16,450	$4,125	152,000

Coefficient Estimates			
Predictor	*Coefficient*	*Stand. Error of the Coefficient*	*t-statistic*
Intercept	-94.88	32.97	??
Slope (Industry Sales)	0.2796	0.0363	??

Analysis of Variance Table (ANOVA)				
Source	*df (Degrees of Freedom)*	*SS (Sum of Squares)*	*Mean Square (SS/df)*	*F-statistic*
Regression	1 (# of independent variables)	11,899.50 (SSR)	11,899.50 (MSR)	59.45
Error	3 (n-2)	600.50 (SSE)	200.17 (MSE)	
Total	4 (n-1)	12,500 (SS Total)		

Abbreviated Two-tailed t-table		
df	*10%*	*5%*
2	2.920	4.303

3	2.353	3.182
4	2.132	2.776

Standard error of forecast is 15.5028.

Question #92 of 120

Question ID: 461486

Which of the following is the correct value of the correlation coefficient between industry sales and company sales?

✗ **A)** **0.9062.**

✓ **B)** 0.9757.

✗ **C)** 0.2192.

Explanation

The R^2 = (SST − SSE) / SST = (12,500 − 600.50) / 12,500 = 0.952

The correlation coefficient is $\sqrt{R^2}$ in a simple linear regression, which is $\sqrt{0.952}$ = 0.9757. (Study Session 3, LOS 11.a)

Question #93 of 120

Question ID: 461487

Which of the following reports the correct value and interpretation of the R^2 for this regression? The R^2 is:

✓ **A)** **0.952, indicating that the variability of industry sales explains about 95.2% of the variability of company sales.**

✗ **B)** 0.048, indicating that the variability of industry sales explains about 4.8% of the variability of company sales.

✗ **C)** 0.952, indicating the variability of company sales explains about 95.2% of the variability of industry sales.

Explanation

The R^2 = (SST − SSE) / SST = (12,500 − 600.50) / 12,500 = 0.952

The interpretation of this R^2 is that 95.2% of the variation in company XYZ's sales is explained by the variation in tissue industry sales. (Study Session 3, LOS 11.a)

Question #94 of 120

Question ID: 461488

What is the predicted value for sales of Company XYZ given industry sales of $3,500?

✗ **A)** **$900.00.**

✗ **B)** $994.88.

✓ **C)** $883.72.

Explanation

The regression equation is Y = (−94.88) + 0.2796 × X = −94.88 + 0.2796 × (3,500) = 883.72. (Study Session 3, LOS 11.h)

Question #95 of 120

Question ID: 461489

What is the upper limit of a 95% confidence interval for the predicted value of company sales (Y) given industry sales of $3,300?

✗ **A)** **318.42.**

✓ **B)** 877.13.

✗ **C)** 827.87.

Explanation

The predicted value is $\hat{Y}$ = −94.88 + 0.2796 × 3,300 = 827.8.

The upper limit for a 95% confidence interval = $\hat{Y} + t_c s_f$ = 827.8 + 3.182 × 15.5028 = 827.8 + 49.33 = 877.13.

The critical value of t_c at 95% confidence and 3 degrees of freedom is 3.182.

(Study Session 3, LOS 11.h)

Question #96 of 120

Question ID: 461490

What is the lower limit of a 95% confidence interval for the predicted value of company sales (Y) given industry sales of $3,300?

✗ **A)** **827.80.**

✗ **B)** 1,337.06.

✓ **C)** 778.47.

Explanation

The predicted value is $\hat{Y}$ = -94.88 + 0.2796 × 3,300 = 827.8.

The lower limit for a 95% confidence interval = $\hat{Y} - t_c s_f$ = 827.8 − 3.182 × 15.5028 = 827.8 − 49.33 = 778.47.

The critical value of t_c at 95% confidence and 3 degrees of freedom is 3.182.

(Study Session 3, LOS 11.h)

Question #97 of 120

Question ID: 461491

What is the *t*-statistic for the slope of the regression line?

✗ **A)** **3.1820.**

✓ **B)** 7.7025.

✗ **C)** 2.9600.

Explanation

Tb = ($b_1$hat − b_1) / s_{b1} = (0.2796 − 0) / 0.0363 = 7.7025. (Study Session 3, LOS 11.g)

Question #98 of 120

Question ID: 461452

The standard error of the estimate in a regression is the standard deviation of the:

✓ **A)** **residuals of the regression.**

✗ **B)** dependent variable.

✗ **C)** differences between the actual values of the dependent variable and the mean of the dependent variable.

Explanation

The standard error is $s_e = \sqrt{[SSE/(n-2)]}$. It is the standard deviation of the residuals.

Question #99 of 120

Question ID: 461423

One of the limitations of correlation analysis of two random variables is the presence of outliers, which can lead to which of the following erroneous assumptions?

✗ **A)** **The presence of a nonlinear relationship between the two variables, when in fact, there is a linear relationship.**

✗ **B)** The presence of a nonlinear relationship between the two variables, when in fact, there is no relationship whatsoever between the two variables.

✓ **C)** The absence of a relationship between the two variables, when in fact, there is a linear relationship.

Explanation

Outliers represent a few extreme values for sample observations in a correlation analysis. They can either provide statistical evidence that a significant relationship exists, when there is none, or provide evidence that no relationship exists when one does.

Question #100 of 120

Question ID: 479060

A study of 40 men finds that their job satisfaction and marital satisfaction scores have a correlation coefficient of 0.52. At 5% level of significance, is the correlation coefficient significantly different from 0?

✗ **A)** **No, t = 2.02.**

✓ **B)** Yes, t = 3.76.

✗ **C)** No, t = 1.68.

Explanation

We want to test whether the correlation between the population of two variables is equal to zero. The appropriate null and alternative hypotheses can be structured as a two-tailed test as follows:

H_0: r = 0 vs. H_a: r ≠ 0

Assuming that the two populations are normally distributed, we can use a t-test to determine whether the null hypothesis should be rejected. The test statistic is computed using the sample correlation, r, with n - 2 degrees of freedom:

$t = [r\sqrt{(n - 2)}] / \sqrt{(1 - r^2)} = [(0.52\sqrt{(40 - 2)}] / \sqrt{(1 - 0.52^2)} = 3.75$

t_c (α = 0.05 and degrees of freedom = 38) = 2.021

To make a decision, the calculated test statistic is compared with the critical t-value.

$t > t_c$ hence we reject H_0.

Questions #101-106 of 120

Milky Way, Inc. is a large manufacturer of children's toys and games based in the United States. Their products have high name brand recognition, and have been sold in retail outlets throughout the United States for nearly fifty years. The founding management team was bought out by a group of investors five years ago. The new management team, led by Russell Stepp, decided that Milky Way should try to expand its sales into the Western European market, which had never been tapped by the former owners. Under Stepp's leadership, additional personnel are hired in the Research and Development department, and a new marketing plan specific to the European market is implemented. Being a new player in the European market, Stepp knows that it will take several years for Milky Way to establish its brand name in the marketplace, and is willing to make the expenditures now in exchange for increased future profitability.

Now, five years after entering the European market, Stepp is reviewing the results of his plan. Sales in Europe have slowly but steadily increased over since Milky Way's entrance into the market, but profitability seems to have leveled out. Stepp decides to hire a consultant, Ann Hays, CFA, to review and evaluate their European strategy. One of Hays' first tasks on the job is to perform a regression analysis on Milky Way's European sales. She is seeking to determine whether the additional expenditures on research and development and marketing for the European market should be continued in the future.

Hays begins by establishing a relationship between the European sales of Milky Way (in millions of dollars) and the two independent variables, the number of dollars (in millions) spent on research and development (R&D) and marketing (MKTG). Based upon five years of monthly data, Hays constructs the following estimated regression equation:

> Estimated Sales = 54.82 + 5.97 (MKTG) + 1.45 (R&D)

Additionally, Hays calculates the following regression estimates:

	Coefficient	*Standard Error*
Intercept	54.82	3.165
MKTG	5.97	1.825
R&D	1.45	0.987

Question #101 of 120

Question ID: 485550

Hays begins the analysis by determining if both of the independent variables are statistically significant. To test whether a coefficient is statistically significant means to test whether it is statistically significantly different from:

✓ **A) zero.**

✗ **B)** the upper tail critical value.

✗ **C)** slope coefficient.

Explanation

The magnitude of the coefficient reveals nothing about the importance of the independent variable in explaining the dependent variable. Therefore, it must be determined if each independent variable is statistically significant. The null hypothesis is that the slope coefficient for each independent variable equals zero. (Study Session 3, LOS 9.a)

Question #102 of 120

Question ID: 485551

The *t*-statistic for the marketing variable is calculated to be:

✓ **A) 3.271.**

✗ **B)** 17.321.

✗ **C)** 1.886.

Explanation

The t-statistic for the marketing coefficient is calculated as follows: (5.97- 0.0) / 1.825 = 3.271. (Study Session 3, LOS 9.g)

Question #103 of 120

Question ID: 485552

Hays formulates a test structure where the decision rule is to reject the null hypothesis if the calculated test statistic is either larger than the upper tail critical value or lower than the lower tail critical value. At a 5% significance level with 57 degrees of freedom, assume that the two-tailed critical t-values are $t_c = \pm 2.004$. Based on this information, Hays makes the following conclusions:

- Point 1: The intercept term is statistically significant.
- Point 2: Both independent variables are statistically significant in the model explaining sales for Milky Way, Inc.
- Point 3: If an F-test were being used, the null hypothesis would be rejected.

Which of Hays' conclusions are CORRECT?

✗ **A) Points 1 and 2.**

✗ **B)** Points 2 and 3.

✓ **C)** Points 1 and 3.

Explanation

Hays' Point 1 is correct. The t-statistic for the intercept term is (54.82 - 0) / 3.165 = 17.32, which is greater than the critical value of 2.004, so we can conclude that the intercept term is statistically significant.

Hays' Point 2 is incorrect. The t-statistic for the R&D term is (1.45 - 0) / 0.987 = 1.469, which is not greater than the critical value of 2.004. This means that only MKTG can be said to statistically significant.

Hays' Point 3 is correct. An F-test tests whether at least one of the independent variables is significantly different from zero, where the null hypothesis is that all none of the independent variables are significant. Since we know that MKTG is a significant variable (t-statistic of 3.271), we can reject the hypothesis that none of the variables are significant. (Study Session 3, LOS 9.j)

Question #104 of 120

Question ID: 485553

Hays is aware that part, but not all, of the total variation in expected sales can be explained by the regression equation. Which of the following statements correctly reflects this relationship?

✗ **A) MSE = RSS + SSE.**

✓ **B)** SST = RSS + SSE.

✗ **C)** SST = RSS + SSE + MSE.

Explanation

RSS (Regression sum of squares) is the portion of the total variation in Y that is explained by the regression equation. The SSE (Sum of squared errors), is the portion of the total variation in Y that is not explained by the regression. The SST is the total variation of Y around its average value. Therefore, SST = RSS + SSE. These sums of squares will always be calculated for you on the exam, so focus on understanding the interpretation of each. (Study Session 3, LOS 9.j)

Question #105 of 120

Question ID: 485554

Hays decides to test the overall effectiveness of the both independent variables in explaining sales for Milky Way. Assuming that the total sum of squares is 389.14, the sum of squared errors is 146.85 and the mean squared error is 2.576, then:

✗ **A) The R^2 equals 0.242, indicating that the two independent variables account for 24.2% of the variation in monthly sales.**

✓ **B)** The correlation between the actual and predicted values of estimated sales is 0.79.

✗ **C)** The R^2 equals 0.623, indicating that the two independent variables together account for 37.7% of the variation in monthly sales.

Explanation

The R^2 is calculated as (SST - SSE) / SST. In this example, R^2 equals (389.14-146.85) / 389.14 = .623 or 62.3%. Multiple R is the square root of multiple R-squared i.e. $(0.623)^{0.5}$ = 0.79. Mutiple R is the correlation between the predicted and actual values of the dependent variable. The value for mean squared error is not used in this calculation. (Study Session 3, LOS 9.j)

Question #106 of 120

Question ID: 485555

Stepp is concerned about the validity of Hays' regression analysis and asks Hays if he can test for the presence of heteroskedasticity. Hays complies with Stepp's request, and detects the presence of unconditional heteroskedasticity. Which of the following statements regarding heteroskedasticity is *most correct*?

✗ **A) Unconditional heteroskedasticity does create significant problems for statistical inference.**

✓ **B)** Unconditional heteroskedasticity usually causes no major problems with the regression.

✗ **C)** Heteroskedasticity can be detected either by examining scatter plots of the residual or by using the Durbin-Watson test.

Explanation

Unconditional heteroskedasticity occurs when the heteroskedasticity is not related to the level of the independent variables. This means that it does not systematically increase or decrease with changes in the independent variable(s). Note that heteroskedasticity occurs when the variance of the residuals is different across all observations in the sample and can be detected either by examining scatter plots or using a Breusch-Pagen test. (Study Session 3, LOS 10.k)

Question #107 of 120

Question ID: 461428

Consider a sample of 32 observations on variables X and Y in which the correlation is 0.30. If the level of significance is 5%, we:

✗ **A) conclude that there is significant correlation between X and Y.**

✓ **B)** conclude that there is no significant correlation between X and Y.

✗ **C)** cannot test the significance of the correlation with this information.

Explanation

The calculated $t = (0.30 \times \sqrt{30}) / \sqrt{(1 - 0.09)} = 1.72251$ and the critical t values are ± 2.042. Therefore, we fail to reject the null hypothesis of no correlation.

Question #108 of 120

Question ID: 461439

The assumptions underlying linear regression include all of the following EXCEPT the:

- ✗ **A)** **disturbance term is normally distributed with an expected value of 0.**
- ✓ **B)** independent variable is linearly related to the residuals (or disturbance term).
- ✗ **C)** disturbance term is homoskedastic and is independently distributed.

Explanation

The independent variable is *uncorrelated with the* residuals (or disturbance term).

The other statements are true. The disturbance term is homoskedastic because it has a constant variance. It is independently distributed because the residual for one observation is not correlated with that of another observation. *Note:* The opposite of homoskedastic is heteroskedastic. *For the examination, memorize the assumptions underlying linear regression!*

Question #109 of 120

Question ID: 461455

Consider the following estimated regression equation:

> $AUTO_t = 0.89 + 1.32\ PI_t$

The standard error of the coefficient is 0.42 and the number of observations is 22. The 95% confidence interval for the slope coefficient, b_1, is:

- ✗ **A)** **$\{-0.766 < b_1 < 3.406\}$.**
- ✗ **B)** $\{0.480 < b_1 < 2.160\}$.
- ✓ **C)** $\{0.444 < b_1 < 2.196\}$.

Explanation

The degrees of freedom are found by n-k-1 with k being the number of independent variables or 1 in this case. DF = 22-1-1 = 20. Looking up 20 degrees of freedom on the student's t distribution for a 95% confidence level and a 2 tailed test gives us a critical value of 2.086. The confidence interval is 1.32 ± 2.086 (0.42), or $\{0.444 < b_1 < 2.196\}$.

Question #110 of 120

Question ID: 461523

Limitations of regression analysis include all of the following EXCEPT:

- ✗ **A)** **outliers may affect the estimated regression line.**
- ✗ **B)** parameter instability.
- ✓ **C)** regression results do not indicate anything about economic significance.

Explanation

The estimated coefficients tell us something about economic significance - they tell us the expected or average change in the dependent variable for a given change in the independent variable.

Question #111 of 120

Question ID: 461412

Which term is *least likely* to apply to a regression model?

- ✗ **A) Goodness of fit.**
- ✓ **B)** Coefficient of variation.
- ✗ **C)** Coefficient of determination.

Explanation

Goodness of fit and coefficient of determination are different names for the same concept. The coefficient of variation is not directly part of a regression model.

Question #112 of 120

Question ID: 461493

Dan Gates, CFA is forecasting price elasticity of demand for GMX Inc's products. Gates used monthly revenues for the past four years as the dependent variable ($ millions) and price per unit as the independent variable. The results are shown below.

Sales = 23.45 − 0.6 Price

Standard error (intercept) = 10.22

Standard error (slope) = 0.03

Standard error of estimate = 8.32

Standard error of forecast = 8.93

The 95% confidence interval for predicted value of monthly sales given price was $2.00 per unit is *closest* to:

- ✗ **A) $6 million to $39 million.**
- ✓ **B)** $4 million to $40 million.
- ✗ **C)** $12 million to $33 million.

Explanation

The predicted value of sales when price = $2.00 is 23.45 − 0.6(2) = $22.25 million

There are 4 years × 12 = 48 monthly observations. D.O.F = n−2 = 46

t_C (95%, 46, 2-tailed) = 2.013

For the confidence interval of predicted value, make sure to use the standard error of forecast

95% confidence interval = 22.25 ± (2.013)(8.93) or $4.27 to $40.23 million

---

Question #113 of 120

Question ID: 461431

The purpose of regression is to:

- ✓ **A) explain the variation in the dependent variable.**

✗ **B)** get the largest R^2 possible.

✗ **C)** explain the variation in the independent variable.

Explanation

The goal of a regression is to explain the variation in the dependent variable.

Question #114 of 120

Question ID: 461441

Sera Smith, a research analyst, had a hunch that there was a relationship between the percentage change in a firm's number of salespeople and the percentage change in the firm's sales during the following period. Smith ran a regression analysis on a sample of 50 firms, which resulted in a slope of 0.72, an intercept of +0.01, and an R^2 value of 0.65. Based on this analysis, if a firm made no changes in the number of sales people, what percentage change in the firm's sales during the following period does the regression model predict?

✓ **A) +1.00%.**

✗ **B)** +0.65%.

✗ **C)** +0.72%.

Explanation

The slope of the regression represents the linear relationship between the independent variable (the percent change in sales people) and the dependent variable, while the intercept represents the predicted value of the dependent variable if the independent variable is equal to zero. In this case, the percentage change in sales is equal to: 0.72(0) + 0.01 = +0.01.

Question #115 of 120

Question ID: 461416

Thomas Manx is attempting to determine the correlation between the number of times a stock quote is requested on his firm's website and the number of trades his firm actually processes. He has examined samples from several days trading and quotes and has determined that the covariance between these two variables is 88.6, the standard deviation of the number of quotes is 18, and the standard deviation of the number of trades processed is 14. Based on Manx's sample, what is the correlation between the number of quotes requested and the number of trades processed?

✗ **A) 0.18.**

✓ **B)** 0.35.

✗ **C)** 0.78.

Explanation

Correlation = Cov (X,Y) / (Std. Dev. X)(Std. Dev. Y)

Correlation = 88.6 / (18)(14) = 0.35

Question #116 of 120

Question ID: 461506

Which statement is *most accurate*? Assume a 5% level of significance. The F-statistic is:

Analysis of Variance Table (ANOVA)				
Source	*Degrees of freedom (df)*	*Sum of Squares*	*Mean Square (SS/df)*	*F-statistic*
Regression	5	18,500	3,700	
Error	94	600.66	6.39	
Total	99	19,100.66		

✗ **A) 579.03 and the regression is said to be statistically insignificant.**

✓ **B)** 579.03 and the regression is said to be statistically significant.

✗ **C)** 0.0017 and the regression is said to be statistically significant.

Explanation

F =3,700/6.39 = 579.03 which is significant since the critical F value is between 2.29 and 2.37. The critical F value is found by using a 5% level of significance F-table and looking up the value that corresponds with 5 = k = the number of independent variables in the numerator and 100 − 5 − 1 = 94 df in the denominator resulting in a critical value between 2.29 and 2.37.

Question #117 of 120

Question ID: 461482

Consider the regression results from the regression of Y against X for 50 observations:

> Y = 5.0 - 1.5 X

The standard error of the estimate is 0.40 and the standard error of the coefficient is 0.45. The predicted value of Y if X is 10 is:

✗ **A) 10.**

✗ **B)** 20.

✓ **C)** -10.

Explanation

The predicted value of Y is: Y = 5.0 - [1.5 (10)] = 5.0 - 15 = -10

Question #118 of 120

Question ID: 461407

Which model does not lend itself to correlation coefficient analysis?

✓ **A) Y = X^3.**

✗ **B)** Y = X + 2.

✗ **C)** X = Y × 2.

Explanation

The correlation coefficient is a measure of linear association. All of the functions except for Y = X^3 are linear functions.

Question #119 of 120

Question ID: 461417

Suppose the covariance between Y and X is 12, the variance of Y is 25, and the variance of X is 36. What is the correlation coefficient (r), between Y and X?

✗ **A)** **0.013.**

✗ **B)** 0.160.

✓ **C)** 0.400.

Explanation

The correlation coefficient is:

$$r = \frac{Cov_{xy}}{\sqrt{x}\sqrt{y}}$$

$$r = \frac{12}{(5)(6)} = 0.40$$

Question #120 of 120

Question ID: 461404

Which of the following statements regarding a correlation coefficient of 0.60 for two variables Y and X is *most* accurate? This correlation:

✓ **A)** **indicates a positive covariance between the two variables.**

✗ **B)** is significantly different from zero.

✗ **C)** indicates a positive causal relation between the two variables.

Explanation

A test of significance requires the sample size, so we cannot conclude anything about significance. There is some positive relation between the two variables, but one may or may not cause the other.