Question #1 of 92

Question ID: 1208245

Use the following F-table (5% Level of Significance) for this question:

Denominator Degrees of Freedom	Numerator Degrees of Freedom				
	3	4	5	6	7
92	2.70	2.47	2.31	2.20	2.11
93	2.70	2.47	2.31	2.20	2.11
94	2.70	2.47	2.31	2.20	2.11
95	2.70	2.47	2.31	2.20	2.11
96	2.70	2.47	2.31	2.19	2.11
97	2.70	2.47	2.31	2.19	2.11
98	2.70	2.46	2.31	2.19	2.10

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 2.31. 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.

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #2 of 92

Question ID: 1208191

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) $\{0.683 < b_1 < 2.317\}$.

C) $\{-2.317 < b_1 < -0.683\}$.

Explanation

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

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #3 of 92

Question ID: 1208260

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.

(Study Session 2, Module 4, LOS 4.h)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #4 of 92

Question ID: 1208180

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)	0.91	4.80
B)	0.89	5.34
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:

$$COV = 0.89 \times \sqrt{4} \times \sqrt{9} = 5.34$$

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Question #5 of 92

Question ID: 1208167

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

A) explanatory variable.

B) predicting variable.

C) predicted variable.

Explanation

The dependent variable is the predicted variable.

(Study Session 2, Module 4.1, LOS 4.a)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

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,

Mahakali Book Center
9920411158 / 9820665601

and if the error terms are correlated with each other over time the test statistics may not be reliable.

Question #6 - 11 of 92

Question ID: 1208207

Use the following t-table for this question:

	Area in Right Tail		
	5.0%	2.5%	1.0%
18	1.734	2.101	2.552
19	1.729	2.093	2.539
20	1.725	2.086	2.528
21	1.721	2.080	2.518
22	1.717	2.074	2.508
23	1.714	2.069	2.500
24	1.711	2.064	2.492

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 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Mahakali Book Center 9820411158 / 9820665601

Question #7 - 11 of 92

Question ID: 1208208

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

A) spurious correlation but not the issue of outliers.

B) Both spurious correlation and outliers.

C) issue of outliers but not the issue of spurious correlation.

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 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #8 - 11 of 92

Question ID: 1208209

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

A) 12.2.

B) 13.5.

C) 16.1.

Explanation

The forecast is 12.209 = 0.149 + 1.206 × 10, so the answer is 12.2.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #9 - 11 of 92

Question ID: 1208210

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

Mahakali Book Center
9920411156 / 9820665601

A) 12.84.

B) 2.68.

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 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #10 - 11 of 92

Question ID: 1208211

A 95% confidence interval for the slope coefficient is:

A) 0.734 to 1.677.

B) 0.760 to 1.650.

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 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #11 - 11 of 92

Question ID: 1208212

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

A) multicollinearity leads to inconsistent estimates of the regression coefficients.

B) the relationships of variables change over time.

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

Mahakali Book Center
9920411158 / 9820665601

Explanation

In the presence of multicollinearlity, the regression coefficients would still be consistent but unreliable. The other possible shortfalls listed are valid.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #12 of 92

Question ID: 1208195

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 first regression equation has more explaining power than the second regression equation.

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

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #13 of 92

Question ID: 1208170

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

Mahakali Book Center
9920441158 / 9820665601

A) The dependent variable declines by -1.5 units if X increases by 1 unit.

B) If the value of X is zero, the value of Y will be -1.5.

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.

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Question #14 of 92

Question ID: 1208175

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) +0.65%.

B) +1.00%.

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.

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #15 of 92

Question ID: 1208216

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

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

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

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

Explanation

df = n − k − 1 = 45 − 2 − 1

(Study Session 2, Module 4, LOS 4.d)

Related Material

SchweserNotes - Book 1

Question #16 of 92

Question ID: 1208200

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #17 of 92

Question ID: 1208233

Mahakali Book Center
9920411158 / 9820665601

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) 0.55.

C) 0.30.

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.

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #18 of 92

Question ID: 1208202

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

A) standard deviation of the residuals.

B) standard deviation of the dependent variable.

C) standard deviation of the independent 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).

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #19 of 92

Question ID: 1208201

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) 15.10.

B) 63.20.

C) 72.10.

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #20 of 92

Question ID: 1208218

Mahakali Book Center
9920411158 / 9820665601

Use the following t-table for this question:

	Area in Right Tail		
	5.0%	2.5%	1.0%
196	1.653	1.972	2.346
197	1.653	1.972	2.345
198	1.653	1.972	2.345
199	1.653	1.972	2.345
200	1.653	1.972	2.345
201	1.652	1.972	2.345
202	1.652	1.972	2.345

A sample of 200 monthly observations is used to run a simple linear regression: Returns = b_0 + b_1Leverage + 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 does not significantly explain returns.

B) do not reject the null hypothesis and conclude that leverage significantly explains returns.

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

Explanation

Do not reject the null since $|-1.09| < 1.972$(critical t-value).

(Study Session 2, Module 4, LOS 4.d)

Related Material

SchweserNotes - Book 1

Question #21 of 92

Question ID: 1208196

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?

Mahakali Book Center
9920411158/9820665601

A) 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.

B) Results of the second analysis are more reliable than the first analysis.

C) The first regression has more explanatory power than the second regression.

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

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-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

Mahakali Book Center
9920411158 / 9820665601

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 #22 - 27 of 92

Question ID: 1220818

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

A) 0.9062.

B) 0.2192.

C) 0.9757.

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 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Mahakali Book Center 9920411158 / 9820665601

Question #23 - 27 of 92

Question ID: 1220819

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

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

B) 0.952, indicating that the variability of industry sales explains about 95.2% 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 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Question #24 - 27 of 92

Question ID: 1220820

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

A) $994.88.

B) $883.72.

C) $900.00.

Explanation

The regression equation is Y = (−94.88) + 0.2796 × X = −94.88 + 0.2796 × (3,500) = 883.72.

(Study Session 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #25 - 27 of 92

Question ID: 1220821

What is the upper limit of a 95% confidence interval for the predicted value of company sales (Y) given industry sales of $3,300? The critical value of tc at 95% confidence and 3 degrees of freedom is 3.182.

A) 827.87.

B) 877.13.

C) 318.42.

Explanation

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

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

(Study Session 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Question #26 - 27 of 92

Question ID: 1220822

What is the lower limit of a 95% confidence interval for the predicted value of company sales (Y) given industry sales of $3,300? The critical value of tc at 95% confidence and 3 degrees of freedom is 3.182.

A) 827.80.

B) 778.47.

C) 1,337.06.

Explanation

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

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

(Study Session 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #27 - 27 of 92

Question ID: 1220823

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

A) 2.9600.

B) 7.7025.

C) 3.1820.

Explanation

$Tb = (b_1\text{hat} - b_1) / s_{b1} = (0.2796 - 0) / 0.0363 = 7.7025.$

(Study Session 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Question #28 of 92

Question ID: 1208242

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 = 20.0.

B) R^2 = 50% and F = 2.0.

C) R^2 = 33% and F = 2.0.

Explanation

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

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #29 of 92

Question ID: 1208179

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

A) The covariance and correlation are always the same sign, positive or negative.

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

C) A zero covariance implies a zero correlation.

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.

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Question #30 of 92

Question ID: 1208204

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 larger the SEE the larger the R^2.

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

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #31 of 92

Question ID: 1208203

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) 0.0360.

B) 0.0082.

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,

$$SEE = \sqrt{(SSE / (n-2))} = \sqrt{(.0039 / (60-2))} = .0082$$

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #32 of 92

Question ID: 1208244

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) 0.82.

B) 0.45.

C) 0.55.

Mahakali Book Center
9920411158 / 9920665601

Explanation

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

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #33 of 92

Question ID: 1220824

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) 33.8%.

B) 18.4%.

C) 20.3%.

Explanation

Y = intercept + slope(X)

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

(Study Session 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #34 of 92

Question ID: 1208173

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

A) independent variable is linearly related to the residuals (or disturbance term).

B) disturbance term is normally distributed with an expected value of 0.

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!*

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Question #35 of 92

Question ID: 1208198

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

A) percent of variability of one factor explained by the variability of the second factor.

B) impact on B of a one-unit change in A.

C) statistical significance of the coefficient in the regression equation.

Explanation

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

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #36 of 92

Question ID: 1208164

Mahakali Book Center
9920411158 / 9820665601

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

A) capital expenditure is a dependent variable.

B) financial leverage is an independent variable.

C) return on equity is a control variable.

Explanation

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

(Study Session 2, Module 4.1, LOS 4.a)

Related Material

SchweserNotes - Book 1

Question #37 of 92

Question ID: 1208176

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

A) When there is a strong relationship between two variables we can conclude that a change in one will cause a change in the other.

B) The coefficient of determination is defined as the strength of the linear relationship between two variables.

C) An assumption of linear regression is that the residuals are independently distributed.

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.

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #38 of 92

Question ID: 1208172

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

A) The variance of the residuals is constant.

B) The residuals are normally distributed.

C) The independent variable is correlated with the residuals.

Explanation

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

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Question #39 of 92

Question ID: 1208165

The purpose of regression is to:

A) explain the variation in the independent variable.

B) explain the variation in the dependent variable.

C) get the largest R^2 possible.

Explanation

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

(Study Session 2, Module 4.1, LOS 4.a)

Related Material

SchweserNotes - Book 1

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	??				

Mahakali Book Center
9920411158 / 9820665601

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 #40 - 45 of 92

Question ID: 1208235

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 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #41 - 45 of 92

Question ID: 1208236

What is the sum of squared errors (SSE)?

A) 23,515.

Mahakali Book Center 9920411158 / 9820665601

B) 23,644.

C) 128.

Explanation

SSE = SST − RSS = 23,644 − 23,516 = 128.

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #42 - 45 of 92

Question ID: 1208237

What is the value of R^2?

A) 0.9946.

B) 0.0055.

C) 0.9471.

Explanation

R^2 = RSS / SST = 23,516 / 23,644 = 0.9946.

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #43 - 45 of 92

Question ID: 1208238

Mahakali Book Center
9920411158 / 9820665601

Use the following t-table for this question:

	Area in Right Tail		
	5.0%	2.5%	1.0%
14	1.761	2.145	2.624
15	1.753	2.131	2.602
16	1.746	2.120	2.583
17	1.740	2.110	2.567
18	1.734	2.101	2.552
19	1.729	2.093	2.539
20	1.725	2.086	2.528

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

	1%	5%
A)	Yes	Yes
B)	Yes	No
C)	No	No

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 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #44 - 45 of 92

Question ID: 1208239

What is the value of the F-statistic?

A) 0.0003.

B) 3,359.

Mahakali Book Center
9920411158 / 9820665601

C) 0.9945.

Explanation

F = mean square regression / mean square error = 23,516 / 7 = 3,359.

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #45 - 45 of 92

Question ID: 1208240

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 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #46 of 92

Question ID: 1208193

Mahakali Book Center
9920411158 / 9820665601

Use the following t-table for this question:

	Area in Right Tail		
df	5%	2.5%	1%
1	6.314	12.706	31.821
5	2.015	2.571	3.365
10	1.812	2.228	2.764
20	1.725	2.086	2.528
40	1.684	2.021	2.423
80	1.664	1.990	2.374
∞	1.645	1.960	2.326

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) $\{0.444 < b_1 < 2.196\}$.

B) $\{0.480 < b_1 < 2.160\}$.

C) $\{-0.766 < b_1 < 3.406\}$.

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 $\{0.444 < b_1 < 2.196\}$.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #47 of 92

Question ID: 1220816

Mahakali Book Center
9920411158 / 9820665601

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) -0.55

B) 5.83

C) 2.83

Explanation

$Y = 2.83 + (1.5)(2)$

$= 2.83 + 3$

$= 5.83$

(Study Session 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Question #48 of 92

Question ID: 1220815

Use the following t-table for this question:

	Area in Right Tail		
	5.0%	2.5%	1.0%
22	1.717	2.074	2.508
23	1.714	2.069	2.500
24	1.711	2.064	2.492
25	1.708	2.060	2.485
26	1.706	2.056	2.479
27	1.703	2.052	2.473
28	1.701	2.048	2.467

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) 0.59 to 4.30.

B) −1.83 to 6.72.

C) −1.68 to 6.56.

Mahakali Book Center
9920411158 / 9820665601

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:

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

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

= 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.

(Study Session 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Question #49 of 92

Question ID: 1220812

Use the following t-table for this question:

	Area in Right Tail		
	5.0%	2.5%	1.0%
44	1.680	2.015	2.414
45	1.679	2.014	2.412
46	1.679	2.013	2.410
47	1.678	2.012	2.408
48	1.677	2.011	2.407
49	1.677	2.010	2.405
50	1.676	2.009	2.403

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.

Mahakali Book Center
9920411158 / 9820665601

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. Critical t-value for 2-tailed, 5% level of significance and 48 dof = 2.011. The confidence interval is 20 ± 2.011 (0.52) or {18.954 < Y < 21.046}.

(Study Session 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Question #50 of 92

Question ID: 1208214

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

A) Goodness of fit.

B) Coefficient of determination.

C) Coefficient of variation.

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

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

Maharaj Book Center
9920411156 / 9820665601

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 #51 - 56 of 92

Question ID: 1208254

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) the upper tail critical value.

B) slope coefficient.

C) zero.

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 2, Module 4, LOS 4.g)

Mahakali Book Center
9920411158 / 9820665601

Related Material

SchweserNotes - Book 1

Question #52 - 56 of 92

Question ID: 1208255

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

A) 17.321.

B) 3.271.

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 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #53 - 56 of 92

Question ID: 1208256

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 = ±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 2 and 3.

B) Points 1 and 2.

C) Points 1 and 3.

Mahakali Book Center
9920411158 / 9820665601

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 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #54 - 56 of 92

Question ID: 1208257

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) SST = RSS + SSE + MSE. ✗

B) MSE = RSS + SSE. ✗

C) SST = RSS + SSE. ✓

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 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #55 - 56 of 92

Question ID: 1208258

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 correlation between the actual and predicted values of estimated sales is 0.79.

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

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 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #56 - 56 of 92

Question ID: 1208259

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

B) Unconditional heteroskedasticity does create significant problems for statistical inference.

C) Unconditional heteroskedasticity usually causes no major problems with the regression.

Explanation

Mahakali Book Center
9920411158 / 9820665601

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 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #57 of 92

Question ID: 1208177

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) Residuals that are not normally distributed. ✓

C) A zero mean of the residuals. ✗

Explanation

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

(Study Session 2, Module 4, LOS 4.h)

Related Material

SchweserNotes - Book 1

Question #58 of 92

Question ID: 1208168

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:

Mahakali Book Center
9920411158 / 9820665601

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

B) changes, no change in sales is expected.

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

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.

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Question #59 of 92

Question ID: 1220813

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) $8.00 million.

B) $2.65 million.

C) $9.05 million.

Explanation

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

(Study Session 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Question #60 of 92

Question ID: 1208189

Mahakali Book Center
9920411158 / 9820665601

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

A) actual dependent variable values about the estimated regression line.

B) values of the sample regression coefficient.

C) predicted y-values around the mean of the observed y-values.

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 = $[SSE / (n - 2)]^{1/2}$

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

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

Mahakali Book Center
9920411158 / 9820665601

	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 #61 - 66 of 92

Question ID: 1208182

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

A) 62.84.

B) 57.61.

C) 111.72.

Explanation

Mahakali Book Center
9920411158 / 9820665601

Upper Limit

= coefficient + (2.306 × standard error of the coefficient)

= 34.74 + (2.306 x 9.917) = 57.61

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #62 - 66 of 92

Question ID: 1208183

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

A) 12.24. ✗

B) 11.87. ✓

C) 72.84. ✗

Explanation

Lower Limit

= Coefficient - (2.306 × Standard Error of the coefficient)

= 34.74 - (2.306 × 9.917)

= 34.74 - 22.87 = 11.87

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #63 - 66 of 92

Question ID: 1208184

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

A) 0.7780. ✓

B) 0.9500. ✗

C) 0.6053. ✗

Mahakali Book Center
9920411158 / 9820665601

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 R2 in a simple linear regression which is the square root of 0.6053 = 0.7780.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #64 - 66 of 92

Question ID: 1208185

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

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

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 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #65 - 66 of 92

Question ID: 1208186

Mahakali Book Center
9920411158 / 9820665601

Is Nair's statement about conditional heteroskedasticity CORRECT?

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

B) No, because coefficients will not be biased. ✓

C) No, because the t-statistics will not be biased.

Explanation

Conditional heteroskedasticity will result in consistent coefficient estimates but inconsistent standard errors resulting in biased t-statistics.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #66 - 66 of 92

Question ID: 1208187

What is the calculated F-statistic?

A) 92.2840.

B) 0.1250.

C) 12.2700.

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 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #67 of 92

Question ID: 1208188

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

A) the alpha coefficient will be zero.

B) the regression line will be sloped upward.

C) the standard error of estimate will be zero.

Explanation

Mahakali Book Center
9920411158 / 9820665601

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #68 of 92

Question ID: 1208194

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^2measures:	Correlation coefficient

A) percent of variability of the

B) percent of variability of the

C) percent of variability of the

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #69 of 92

Question ID: 1208197

Question #69 of 92 Question ID: 1208197

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) 0.130.

C) 0.774.

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #70 of 92 Question ID: 1208213

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

A) is the percentage of the total variation in the dependent variable that is explained by the independent variable.

B) cannot decrease as independent variables are added to the model.

C) may range from −1 to +1.

Mahakali Book Center
9920411158 / 9820665601

Explanation

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

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #71 of 92

Question ID: 1208171

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) If GSTERN increases by one unit, RCRANTZ should increase by 5.9 units. ✔

B) The covariance of RCRANTZ and GSTERN is negative. ✘

C) The intercept term implies that if GSTERN is zero, RCRANTZ is 61.4. ✘

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.

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Question #72 of 92

Question ID: 1208178

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. ✘

Mahakali Book Center
9920411158 / 9820665601

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.

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Question #73 of 92

Question ID: 1208232

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

Use the following t-table for this question:

	Area in Right Tail		
	5.0%	2.5%	1.0%
45	1.679	2.014	2.412
46	1.679	2.013	2.410
47	1.678	2.012	2.408
48	1.677	2.011	2.407
49	1.677	2.010	2.405
50	1.676	2.009	2.403
51	1.675	2.008	2.402

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

Marakali Book Center
9920411158 / 9820665601

A) $4 million to $40 million.

B) $12 million to $33 million.

C) $6 million to $39 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

(Study Session 2, Module 4, LOS 4.f)

Related Material

SchweserNotes - Book 1

Question #74 of 92

Question ID: 1208215

Mahakali Book Center
9920411158 / 9820665601

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.

Use the following t-table for this question:

	Area in Right Tail		
	5.0%	2.5%	1.0%
48	1.677	2.011	2.407
49	1.677	2.010	2.405
50	1.676	2.009	2.403
51	1.675	2.008	2.402
52	1.675	2.007	2.400
53	1.674	2.006	2.399
54	1.674	2.005	2.397

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) $t = 2.667$; slope is significantly different from zero.

B) $t = 1.789$; slope is not significantly different from zero.

C) $t = 3.000$; 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.

(Study Session 2, Module 4, LOS 4.d)

Related Material

SchweserNotes - Book 1

Question #75 of 92

Question ID: 1208192

Mahakali Book Center 9920411158 / 9820665601

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) Explained variability from both analyses is equal.

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

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

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

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 #76 - 81 of 92

Question ID: 1208247

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

A) 0.40; 5.66.

B) 0.53; 2.96.

C) 0.32; 1.65.

Mahakali Book Center
9920411158 / 9820665601

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

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #77 - 81 of 92

Question ID: 1208248

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

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #78 - 81 of 92

Question ID: 1208249

The mean square error (MSE) is:

A) 3.58.

B) 7.11.

C) 3.60.

Explanation

Mahakali Book Center
9920411158 / 9820665601

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

MSE = SSE / df = 115.10 / 32 = 3.60

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #79 - 81 of 92

Question ID: 1208250

What is the R^2 for the regression?

A) 76%.

B) 62%.

C) 65%.

Explanation

SST = RSS + SSE

= 215.10 + 115.10

= 330.20

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

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #80 - 81 of 92

Question ID: 1208251

What is the predicted income of a 40-year-old married female with 16 years of education and 18 years of work experience?

A) $106,930.

B) $82,706.

C) $63,970.

Explanation

Mahakali Book Center
9920411158 / 9920665601

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

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

= 63.97 or $63,970

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #81 - 81 of 92

Question ID: 1208252

What is the F-value?

A) 14.36.

B) 1.88.

C) 29.88.

Explanation

F = MSR / MSE = 107.55 / 3.60 = 29.88

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #82 of 92

Question ID: 1208205

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) coefficient of determination (R^2).

B) sum of squared errors (SSE).

C) standard error of the estimate (SEE).

Explanation

Mahakali Book Center
9920411158 / 9820665601

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.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #83 of 92

Question ID: 1208166

The capital asset pricing model is given by: $R_i = R_f + \text{Beta} \ (R_m - R_f)$ 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.

(Study Session 2, Module 4.1, LOS 4.a)

Related Material

SchweserNotes - Book 1

Question #84 of 92

Question ID: 1208174

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

A) The variance of the error terms each period remains the same.

B) Values of the independent variable are not correlated with the error term.

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

Mahakali Book Center
9920411158 / 9820665601

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.

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Question #85 of 92

Question ID: 1208261

Limitations of regression analysis include all of the following EXCEPT:

A) regression results do not indicate anything about economic significance. ✔

B) outliers may affect the estimated regression line. ✘

C) parameter instability. ✘

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.

(Study Session 2, Module 4, LOS 4.h)

Related Material

SchweserNotes - Book 1

Question #86 of 92

Question ID: 1208217

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) $t = -3.333$; slope is significantly negative. ✔

B) $t = 3.750$; slope is significantly different from zero. ✘

C) $t = -3.750$; slope is significantly different from zero. ✘

Mahakali Book Center
9920411158 / 9820665601

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.

(Study Session 2, Module 4, LOS 4.d)

Related Material

SchweserNotes - Book 1

Question #87 of 92

Question ID: 1208243

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 = 0.67 and F = 0.971.

B) R^2 = 0.40 and F = 0.971.

C) R^2 = 0.40 and F = 33.333.

Explanation

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

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Question #88 of 92

Question ID: 1220814

Mahakali Book Center
9920411158 / 9820665601

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

(Study Session 2, Module 4.3, LOS 4.e)

Related Material

SchweserNotes - Book 1

Question #89 of 92

Question ID: 1208199

Bea Carroll, CFA, has performed a regression analysis of the relationship between two economic measures. 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 the two economic measures can Carroll most accurately draw from her SEE analysis? The relationship between the two variables is *most likely*:

A) linear and inverse.

B) logarithmic.

C) very weak.

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. Essentially, the standard error of estimate is a measure of how well the regression model fits the data. If the SEE is small, the model fits well. The SEE will be low if the relationship is strong and conversely will be high if the relationship is weak. We do not have enough evidence to conclude that the relationship between the two variables is inverse or logarithmic.

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601

Question #90 of 92

Question ID: 1208190

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

A) dependent variable.

B) differences between the actual values of the dependent variable and the mean of the dependent variable.

C) residuals of the regression.

Explanation

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

(Study Session 2, Module 4.2, LOS 4.c)

Related Material

SchweserNotes - Book 1

Question #91 of 92

Question ID: 1208169

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

A) residuals are independently distributed.

B) expected value of the residuals is zero.

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

Explanation

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

(Study Session 2, Module 4.1, LOS 4.b)

Related Material

SchweserNotes - Book 1

Question #92 of 92

Question ID: 1208241

Mahakali Book Center
9920411158 / 9820665601

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) 0.423.

B) 0.965.

C) 27.730.

Explanation

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

(Study Session 2, Module 4, LOS 4.g)

Related Material

SchweserNotes - Book 1

Mahakali Book Center
9920411158 / 9820665601