Question #1 of 18

An analyst has run several regressions hoping to predict stock returns, and wants to translate this into an economic interpretation for his clients.

$$\text{Return} = 0.03 + 0.020\text{Beta} - 0.0001\text{MarketCap (in billions)} + \varepsilon$$

A correct interpretation of the regression *most likely* includes:

A) a billion dollar increase in market capitalization will drive returns down by 0.01%.

B) prediction errors are always on the positive side.

C) a stock with zero beta and zero market capitalization will return precisely 3.0%.

Explanation

The coefficient of MarketCap is 0.01%, indicating that larger companies have slightly smaller returns. Note that a company with no market capitalization would not be expected to have a return at all. Error terms are typically assumed to be normally distributed with a mean of zero.

(Study Session 3, Module 8.9, LOS 8.o)

Related Material

SchweserNotes - Book 1

Question #2 of 18

Which of the following is NOT a model that has a qualitative dependent variable?

A) Logit.

B) Event study.

C) Discriminant analysis.

Explanation

An event study is the estimation of the abnormal returns--generally associated with an informational event—that take on quantitative values.

(Study Session 3, Module 8.9, LOS 8.n)

Related Material

SchweserNotes - Book 1

Question #3 of 18

When utilizing a proxy for one or more independent variables in a multiple regression model, which of the following errors is *most likely* to occur?

A) Model misspecification. ✓

B) Multicollinearity. ✗

C) Heteroskedasticity. ✗

Explanation

By using a proxy for an independent variable in a multiple regression analysis, there is some degree of error in the measurement of the variable.

(Study Session 3, Module 8.9, LOS 8.m)

Related Material

SchweserNotes - Book 1

Question #4 of 18

Which of the following is *least likely* to result in misspecification of a regression model?

A) Measuring independent variables with errors. ✗

B) Using a lagged dependent variable as an independent variable. ✗

C) Transforming a variable. ✓

Explanation

A basic assumption of regression is that the dependent variable is linearly related to each of the independent variables. Frequently, they are not linearly related and the independent variable must be transformed or the model is misspecified. Therefore, transforming an independent variable is a potential solution to a misspecification. Methods used to transform independent variables include squaring the variable or taking the square root.

(Study Session 3, Module 8.9, LOS 8.m)

Related Material

SchweserNotes - Book 1

Question #5 of 18

Mary Steen estimated that if she purchased shares of companies who announced restructuring plans at the announcement and held them for five days, she would earn returns in excess of those expected from the market model of 0.9%. These returns are statistically significantly different from zero. The model was estimated without transactions costs, and in reality these would approximate 1% if the strategy were effected. This is an example of:

A) statistical significance, but not economic significance. ✓

B) a market inefficiency. ✗

C) statistical and economic significance. ✗

Explanation

The abnormal returns are not sufficient to cover transactions costs, so there is no economic significance to this trading strategy. This is not an example of market inefficiency because excess returns are not available after covering transactions costs.

(Study Session 3, Module 8.9, LOS 8.o)

Related Material

SchweserNotes - Book 1

Question #6 of 18

When constructing a regression model to predict portfolio returns, an analyst runs a regression for the past five year period. After examining the results, she determines that an increase in interest rates two years ago had a significant impact on portfolio results for the time of the increase until the present. By performing a regression over two separate time periods, the analyst would be attempting to prevent which type of misspecification?

A) Incorrectly pooling data. ✓

B) Forecasting the past. ✗

C) Using a lagged dependent variable as an independent variable. ✗

Explanation

The relationship between returns and the dependent variables can change over time, so it is critical that the data be pooled correctly. Running the regression for multiple sub-periods (in this case two) rather than one time period can produce more accurate results.

(Study Session 3, Module 8.9, LOS 8.m)

Related Material

SchweserNotes - Book 1

Question #7 of 18

Which of the following is *least likely* to be step in model training?

A) Specify the hyperparameters. ✗

B) Data collection. ✓

C) Evaluation using a performance parameter in the validation sample. ✗

Explanation

Model training steps include specifying the algorithm, specifying the hyperparameters, dividing the data into training and validation samples, evaluating the model using a performance parameter in the validation sample, repeating the training until adequate performance level is reached.

(Study Session 3, Module 8.11, LOS 8.r)

Related Material

SchweserNotes - Book 1

Question #8 of 18

A random forest is *least likely* to:

A) be a classification tree. ✗

B) reduce signal-to-noise ratio. ✓

C) provide a solution to overfitting problem. ✗

Explanation

Random forest is a collection of randomly generated classification trees from the same data set. A randomly selected subset of features is used in creating each tree and hence each tree is slightly different from the others. Since each tree only uses a subset of features, random forests can mitigate the problem of overfitting. Because errors across different trees tend to cancel each other out, using random forests can increase the signal-to-noise ratio.

(Study Session 3, Module 8.11, LOS 8.q)

Related Material

SchweserNotes - Book 1

Question #9 of 18

What is the main difference between probit models and typical dummy variable models?

A) Dummy variable regressions attempt to create an equation to classify items into one of two categories, while probit models estimate a probability.

B) There is no difference--a probit model is simply a special case of a dummy variable regression.

C) A dummy variable represents a qualitative independent variable, while a probit model is used for estimating the probability of a qualitative dependent variable.

Explanation

Dummy variables are used to represent a qualitative independent variable. Probit models are used to estimate the probability of occurrence for a qualitative dependent variable.

(Study Session 3, Module 8.9, LOS 8.n)

Related Material

SchweserNotes - Book 1

Question #10 of 18

Which of the following statements about supervised learning is *most accurate*?

A) Supervised learning requires human intervention in machine learning process.

B) Typical data analytics tasks for supervised learning include classification and prediction.

C) Supervised learning does not differentiate between tag and features.

Explanation

Supervised learning utilizes labeled training data to guide the ML program but does not need "human intervention." Typical data analytics tasks for supervised learning include classification and prediction.

(Study Session 3, Module 8.10, LOS 8.p)

Related Material

SchweserNotes - Book 1

Question #11 of 18

Which of the following about unsupervised learning is *most accurate*?

A) Unsupervised learning has lower forecasting accuracy as compared to supervised learning. ✗

B) There is no tag variable. ✓

C) Classification is an example of unsupervised learning algorithm. ✗

Explanation

In unsupervised learning, the ML program is not given labeled training data. Instead, inputs are provided without any conclusions about those inputs. In the absence of any tagged data, the program seeks out structure or inter-relationships in the data. Clustering is one example of the output of unsupervised ML program while classification is suited for supervised learning.

(Study Session 3, Module 8.10, LOS 8.p)

Related Material

SchweserNotes - Book 1

Question #12 of 18

What is the appropriate remedy in the presence of excessive number of features in a data set?

A) Dimension reduction. ✓

B) Big data analysis. ✗

C) Unsupervised learning. ✗

Explanation

Big Data refers to very large data sets which may include both structured (e.g. spreadsheet) data and unstructured (e.g. emails, text, or pictures) data and includes a large number of features as well as number of observations. Dimension reduction seeks to remove the noise (i.e., those attributes that do not contain much information) when the number of features in a data set (its dimension) is excessive.

(Study Session 3, Module 8.11, LOS 8.q)

Related Material

SchweserNotes - Book 1

Question #13 of 18

Overfitting is *least likely* to result in:

A) higher number of features included in the data set. ✗

B) higher forecasting accuracy in out-of-sample data. ✓

C) inclusion of noise in the model. ✗

Explanation

Overfitting results when a large number of features (i.e., independent variables) are included in the data sample. The resulting model can use the "noise" in the dependent variables to improve the model fit. Overfitting the model in this way will actually decrease the accuracy of model forecasts on other (out-of-sample) data.

(Study Session 3, Module 8.11, LOS 8.q)

Related Material

SchweserNotes - Book 1

Question #14 of 18

A high-yield bond analyst is trying to develop an equation using financial ratios to estimate the probability of a company defaulting on its bonds. Since the analyst is using data over different economic time periods, there is concern about whether the variance is constant over time. A technique that can be used to develop this equation is:

A) logit modeling. ✓

B) multiple linear regression adjusting for heteroskedasticity. ✗

C) dummy variable regression. ✗

Explanation

The only one of the possible answers that estimates a probability of a discrete outcome is logit modeling.

(Study Session 3, Module 8.9, LOS 8.n)

Related Material

SchweserNotes - Book 1

Question #15 of 18

Which of the following questions is *least likely* answered by using a qualitative dependent variable?

A) Based on the following executive-specific and company-specific variables, how many shares will be acquired through the exercise of executive stock options? ✓

B) Based on the following company-specific financial ratios, will company ABC enter bankruptcy? ✗

C) Based on the following subsidiary and competition variables, will company XYZ divest itself of a subsidiary? ✗

Explanation

The number of shares can be a broad range of values and is, therefore, not considered a qualitative dependent variable.

(Study Session 3, Module 8.9, LOS 8.n)

Related Material

SchweserNotes - Book 1

Question #16 of 18

An analyst is building a regression model which returns a qualitative dependant variable based on a probability distribution. This is *least likely* a:

A) probit model. ✗

B) logit model. ✗

C) discriminant model. ✓

Explanation

A probit model is a qualitative dependant variable which is based on a normal distribution. A logit model is a qualitative dependant variable which is based on the logistic distribution. A discriminant model returns a qualitative dependant variable based on a linear relationship that can be used for ranking or classification into discrete states.

(Study Session 3, Module 8.9, LOS 8.n)

Related Material

SchweserNotes - Book 1

Question #17 of 18

Dimension reduction is an example of:

A) supervised learning. ✗

B) unsupervised learning. ✓

C) clustering. ✗

Explanation

Dimension reduction and clustering are examples of supervised learning algorithms.

(Study Session 3, Module 8.11, LOS 8.q)

Related Material

SchweserNotes - Book 1

Question #18 of 18

Which supervised learning model is *most appropriate* (1) when the Y-variable is continuous and (2) when the Y-variable is categorical

	Continuous Y-variable	Categorical Y-variable	
A)	Decision trees	Regression	✗
B)	Classification	Neural Networks	✗
C)	Regression	Classification	✓

Explanation

When the Y-variable is continuous, the appropriate approach is that of regression (used in a broad, ML context). When the Y-variable is categorical (i.e., belonging to a category or classification) or ordinal (i.e., ordered or ranked), a classification model is used.

(Study Session 3, Module 8.11, LOS 8.q)

Related Material

SchweserNotes - Book 1