Quantitative Methods

LEARNING OUTCOMES

Basics of Multiple Regression and Underlying Assumptions

The candidate should be able to:

- describe the types of investment problems addressed by multiple linear regression and the regression process
- formulate a multiple linear regression model, describe the relation between the dependent variable and several independent variables, and interpret estimated regression coefficients
- explain the assumptions underlying a multiple linear regression model and interpret residual plots indicating potential violations of these assumptions

Evaluating Regression Model Fit and Interpreting Model Results

The candidate should be able to:

- evaluate how well a multiple regression model explains the dependent variable by analyzing ANOVA table results and measures of goodness of fit
- formulate hypotheses on the significance of two or more coefficients in a multiple regression model and interpret the results of the joint hypothesis tests
- calculate and interpret a predicted value for the dependent variable, given the estimated regression model and assumed values for the independent variable

Model Misspecification

The candidate should be able to:

- describe how model misspecification affects the results of a regression analysis and how to avoid common forms of misspecification

- □ explain the types of heteroskedasticity and how it affects statistical inference
- □ explain serial correlation and how it affects statistical inference
- □ explain multicollinearity and how it affects regression analysis

Extensions of Multiple Regression

The candidate should be able to:

- □ describe influence analysis and methods of detecting influential data points
- □ formulate and interpret a multiple regression model that includes qualitative independent variables
- □ formulate and interpret a logistic regression model

Time-Series Analysis

The candidate should be able to:

- □ calculate and evaluate the predicted trend value for a time series, modeled as either a linear trend or a log-linear trend, given the estimated trend coefficients
- □ describe factors that determine whether a linear or a log-linear trend should be used with a particular time series and evaluate limitations of trend models
- □ explain the requirement for a time series to be covariance stationary and describe the significance of a series that is not stationary
- □ describe the structure of an autoregressive (AR) model of order p and calculate one- and two-period-ahead forecasts given the estimated coefficients
- □ explain how autocorrelations of the residuals can be used to test whether the autoregressive model fits the time series
- □ explain mean reversion and calculate a mean-reverting level
- □ contrast in-sample and out-of-sample forecasts and compare the forecasting accuracy of different time-series models based on the root mean squared error criterion
- □ explain the instability of coefficients of time-series models
- □ describe characteristics of random walk processes and contrast them to covariance stationary processes
- □ describe implications of unit roots for time-series analysis, explain when unit roots are likely to occur and how to test for them, and demonstrate how a time series with a unit root can be transformed so it can be analyzed with an AR model
- □ describe the steps of the unit root test for nonstationarity and explain the relation of the test to autoregressive time-series models
- □ explain how to test and correct for seasonality in a time-series model and calculate and interpret a forecasted value using an AR model with a seasonal lag
- □ explain autoregressive conditional heteroskedasticity (ARCH) and describe how ARCH models can be applied to predict the variance of a time series
- □ explain how time-series variables should be analyzed for nonstationarity and/or cointegration before use in a linear regression; and
- □ determine an appropriate time-series model to analyze a given investment problem and justify that choice

Machine Learning

The candidate should be able to:

- □ describe supervised machine learning, unsupervised machine learning, and deep learning
- □ describe overfitting and identify methods of addressing it

- describe supervised machine learning algorithms—including penalized regression, support vector machine, k-nearest neighbor, classification and regression tree, ensemble learning, and random forest—and determine the problems for which they are best suited
- describe unsupervised machine learning algorithms—including principal components analysis, k-means clustering, and hierarchical clustering—and determine the problems for which they are best suited
- describe neural networks, deep learning nets, and reinforcement learning

Big Data Projects

The candidate should be able to:

- identify and explain steps in a data analysis project
- describe objectives, steps, and examples of preparing and wrangling data
- evaluate the fit of a machine learning algorithm
- describe objectives, methods, and examples of data exploration
- describe methods for extracting, selecting and engineering features from textual data
- describe objectives, steps, and techniques in model training
- describe preparing, wrangling, and exploring text-based data for financial forecasting