FinTree

JuiceNotes

Quantitative Methods

CFA Level 2

2026

Your CFA Journey with FinTree

A Guided Roadmap from Enrollment to Exam Day

At FinTree, we believe in smart preparation driven by structure, support, and consistency. Our unique LPR Methodology – Learn, Practice, Revise – is designed to guide you through each stage of your CFA prep journey with clarity and confidence.

Roadmap from Enrollment to Exam Day

1. Enrollment & Onboarding:

- Receive a welcome call from your dedicated Success Manager (RM)
- Added to a WhatsApp broadcast list for timely updates
- Join a peer group to engage in discussions, share queries, and stay motivated
- Get a personalized study plan and guidance on how to start your prep journey

2. Learning Phase (Initial Months)

- Watch Main Concept Videos and read the official CFA curriculum
- Focus on understanding foundational concepts topic-wise
- Attend live weekend classes to deepen understanding and clarify doubts

3. Practice Phase (Mid Journey)

- Watch EOC and Blue Box videos to apply concepts to CFA-style questions
- Attempt Learning Evaluation Sessions (LES) for topic-wise testing
- Give Weekly Tests based on a structured test calendar
- Weekly Test results are discussed in class on Sunday for deeper insight into common errors and personal improvement

4. Revision Phase (Final 60 Days)

- Revise thoroughly using Juice Notes and Crash Course videos
- Solve the LES twice –
- Vertically (topic by topic)
- Horizontally (across topics like in actual exams)
- Attempt mock exams as per our 60-day revision schedule
- Review mocks in detail, focusing on time management, accuracy, and weak areas

5. Final Phase - IPASS

Index

Quantitative Methods

Placement Assistance Overview

What is Placement Assistance?

- Placement Assistance at FinTree is a structured support system that helps students transition smoothly from learning to employment.
- Our goal is to connect students with the right opportunities in the finance and analytics sectors through a personalized and practical approach.
- We find relevant job openings according to the skills and experience relevant with the candidates' expectations.

Process of Placement Assistance

All relevant job openings are shared via **WhatsApp Broadcasts** to ensure students receive timely updates and can apply promptly.

How can candidates get added to Placement Broadcasts –

- Placement team will share a form, which needs to be filled.
- Once the candidate has filled the form, he/she will be added to broadcast network within next 48hrs & the candidate will be notified.
- You will start receiving the latest opportunities on your WhatsApp Number.
- Candidates can apply in multiple openings based on their career aspiration.

We will also assist you with resume building,
career counselling & interview preparation.

If you need any assistance, request you to contact on +917447443045

Placement Highlights

- **Top Recruiters:** Goldman Sachs, PWC, Barclays, Crisil, BNP Paribas, HSBC
- **Key Roles:** Financial Analyst, Equity Research, IB Analyst, Risk & Valuation
- **Packages:**
  - Freshers: ₹4.5–8 LPA
  - Experienced: Up to ₹20 LPA and more...
- **Alumni Placed At:** Morningstar, Deutsche Bank, EY, Deloitte, Bank of America, Centrum, Purnartha

Basics of Multiple Regression and Underlying Assumptions

FinTree Fruit 1: INTRODUCTION

Multiple Linear Regression :

Multiple Linear Regression (MLR) is a statistical technique that models the relationship between a dependent variable (Y) and two or more independent variables (X_1, X_2, ..., X_k) using a linear equation.

$$Y = a + b_1x_1 + b_2x_2 + \ldots + b_nx_n + \varepsilon$$

Where,
Y = Dependent Variable
X= Independent Variable
ε = Error Term

FinTree Fruit 2 : USES OF MULTIPLE LINEAR REGRESSION

- **MLR is a flexible tool for modeling complex investment realities — from return attribution to risk forecasting.**
- **In real-world investing, single-factor models fall short. Financial variables are influenced by multiple, interacting drivers — MLR helps capture complexity.**

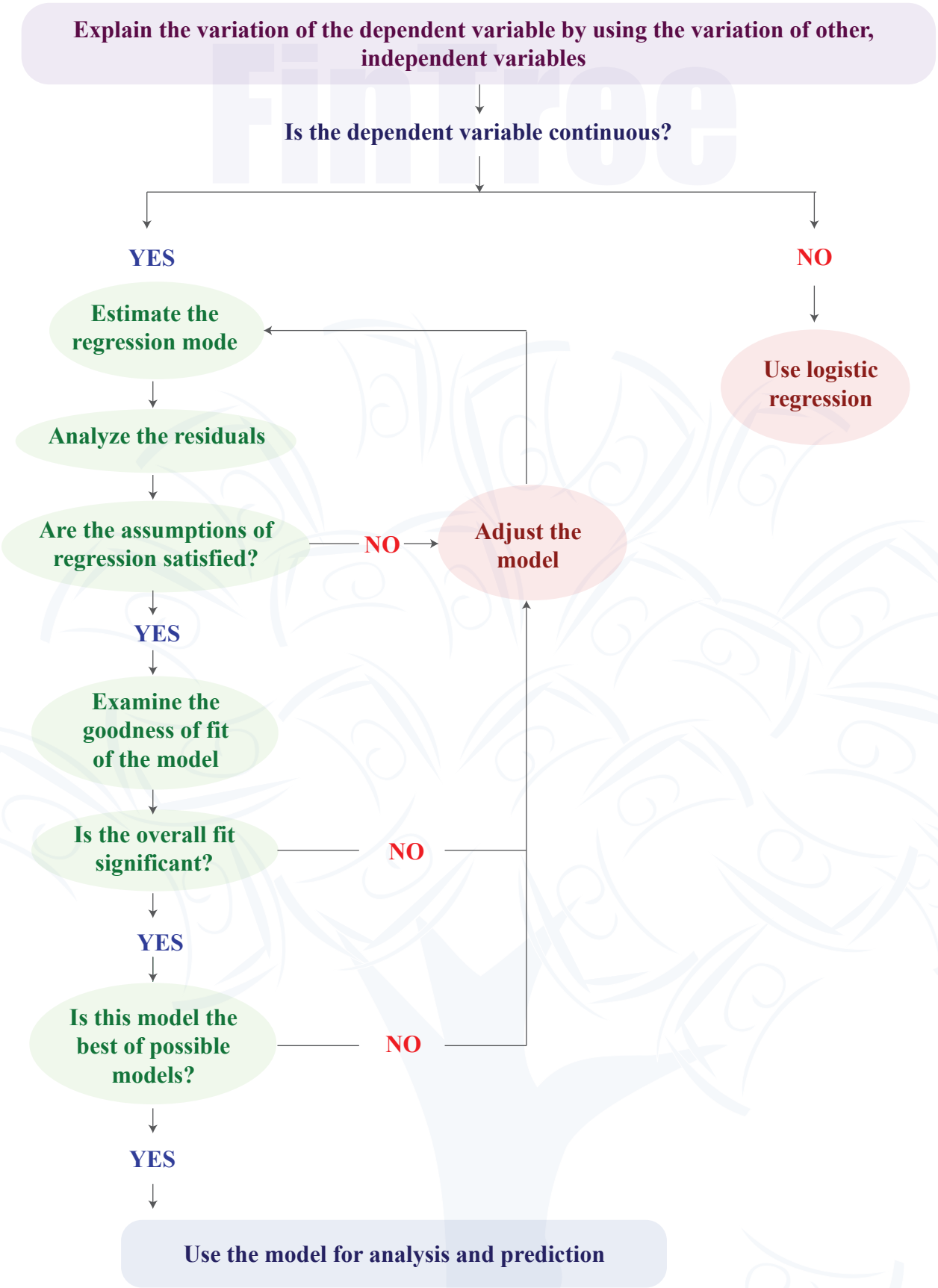
Uses of Multiple linear regression

Factor Models (e.g., Fama–French 5-Factor)	**Understand drivers of stock returns using factors like size, value, profitability, etc.**
Financial Distress Prediction	**Test if metrics like leverage, growth, and profitability predict firm distress.**
Country Risk Analysis	**Assess how political, economic, and ESG risks affect equity returns.**

- **Multiple linear regression is used to Identify relationships, Test financial theories, Forecast future outcomes.**
- **If used incorrectly, MLR can produce:**

1. **Spurious correlations – Detecting false relationships**
2. **Overfitting – Too many variables = poor out-of-sample prediction**
3. **Misleading interpretations – Coefficients may look significant but aren't meaningful**

Regression process :



Example1 :

You are a junior analyst assisting in the development of various multiple regression models for your industry sector. Identify the action you should take to resolve each of the following issues:

Issues :

- **The dependent variable takes on a value of 1 if the company is a merger target and 0 otherwise.**
- **The analyst estimates a model with five independent variables, and none of these variables are significant explanatory variables.**
- **The residuals do not appear to be homoskedastic, thus violating a regression assumption.**
- **The regression assumptions are satisfied, the overall fit is significant, and the model is the best model of the possible models.**

Solution :

Issue	Action
The dependent variable takes on a value of 1 if the company is a merger target and 0 otherwise.	Use logistic regression.
The analyst estimates a model with five independent variables, and none of these variables are significant explanatory variables.	Adjust the model and re-estimate.
The residuals do not appear to be homoskedastic, thus violating a regression assumption.	Adjust the model and re-estimate.
The regression assumptions are satisfied, the overall fit is significant, and the model is the best model of the possible models.	Use the model for analysis and prediction.

FinTree Fruit 3 : THE BASICS OF MULTIPLE REGRESSION

Simple Linear Regression :

Simple regression explains how changes in one independent variable (X) affect the dependent variable (Y).

$$Y = a + bx + \varepsilon$$

Multiple Linear Regression :

The goal of multiple regression is to explain the variation in the dependent variable, Y, using variations in multiple independent variables.

$$Y = a + b_1x_1 + b_2x_2 + \ldots + b_nx_n + \varepsilon$$

Intercept — Slope — Error Term

Dependent Variable — Independent Variable

Interpretation of Slope :

- Slope coefficient in a multiple regression, known as a partial regression coefficient or a partial slope coefficient.
- Each slope coefficient in a regression model indicates the expected change in the dependent variable (Y) for a one-unit change in the corresponding independent variable (X_n), assuming all other variables are held constant.

Interpretation of Intercept :

- The intercept term in a regression model represents the expected value of the dependent variable (Y) when all independent variables (X) are equal to zero.

Example2 :

You have estimated a regression of monthly excess returns on a portfolio (RETRF) using the following three-factor model:

MOM: Momentum factor (monthly return difference between past winners and losers)
LIQ: Liquidity factor (return difference between illiquid and liquid stocks)
VOL: Volatility factor (return spread between low-volatility and high-volatility stocks)

The regression equation, using returns stated in percentages, is:
RETRF = 2.1053 + 0.7124MOM + 0.3840LIQ − 0.2178VOL

Questions :

1. Interpret the intercept.
2. Interpret each slope coefficient.
3. Calculate the predicted value of the portfolio's return when:
   MOM = 1.5, LIQ = 0.5, VOL = 2

Solution :

1. **Intercept Interpretation:**

If MOM, LIQ, and VOL are all zero, the expected portfolio return is 2.1053%.

2. **Slope Coefficient Interpretations:**

- MOM (0.7124): A 1% increase in the momentum factor increases the portfolio return by 0.7124%, holding other factors constant.
- LIQ (0.3840): A 1% increase in the liquidity factor increases the return by 0.3840%, all else equal.
- VOL (−0.2178): A 1% increase in the volatility factor decreases the return by 0.2178%, assuming other factors remain unchanged.

3. **Predicted Return Calculation:**

Use the regression equation:
$R = 2.1053 + 0.7124(1.5) + 0.3840(0.5) - 0.2178(2)$
$= 2.1053 + 1.0686 + 0.1920 - 0.4356$
$= 2.9303\%$

FinTree Fruit 4 : ASSUMPTIONS UNDERLYING MULTIPLE LINEAR REGRESSION

Assumptions :

Linearity

The relationship between the dependent variable (Y) and each independent variable (X_1, X_2, …, X_k) is linear.

Homoskedasticity (Constant Variance of Residuals)

The variance of the residuals (errors) is constant across all levels of independent variables.

Independence of Errors

The residuals (ε) are independent across observations.

Normality of Residuals

The residuals are normally distributed, which is important for valid hypothesis testing and confidence intervals.

Independence Among Independent Variables

a. The independent variables are not random.
b. There is no perfect multicollinearity:
-No exact linear relationship among X_1, X_2, …, X_k.

Assumption Violation :

1. **Heteroskedasticity :**

- Variance of error term is not constant.
- Detected by Graph or BP (Chi-square) Test.
- Impact : Type I Error.
- Resolve : Robust standard error.

2. **Serial Correlation :**

- Errors term are correlated to each other.
- Detected by graph or Durbin Watson Test.
- Impact : Positive serial correlation - Type I Error
  Negative serial correlation - Type II Error
- Resolve : Robust standard error

3. **Multicollinearity :**

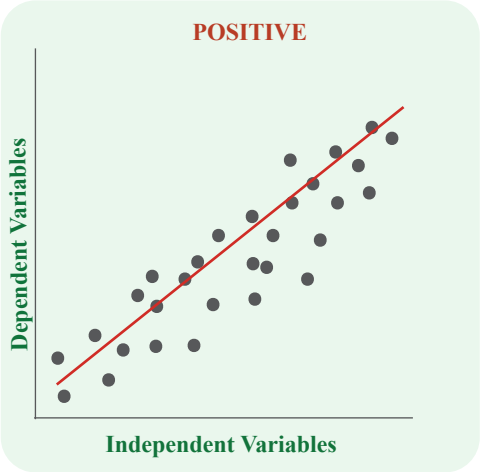
- Two or more independent variable are linearly related to each other.
  Detected by R^2 high, F test significnt, but none of the slope coeffiicients are significant.
- Impact : Type II Error
- Resolve : Drop one of the correlated variables

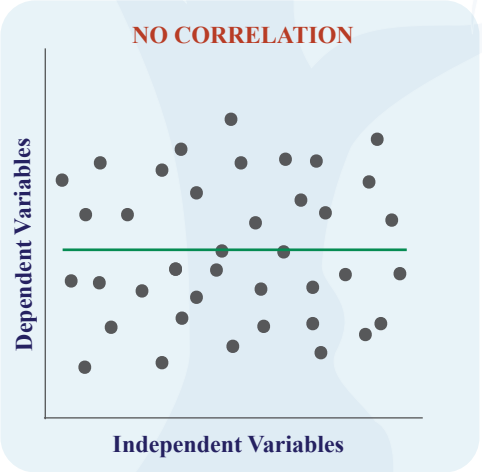
Scatterplot :

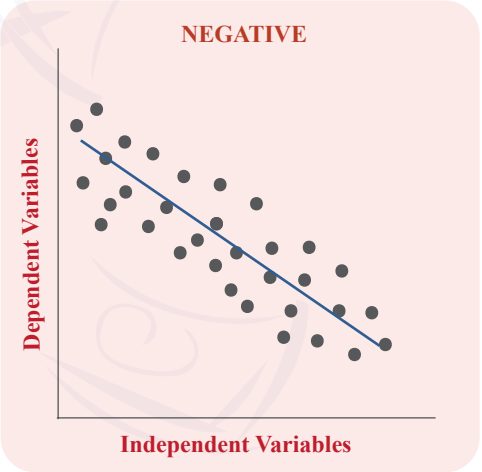
A scatterplot is a graph that displays pairs of values for two variables (X and Y) as dots.

Each point represents an observation:
X-axis: Independent variable
Y-axis: Dependent variable

- Visualize the relationship between variables
- Identify patterns, trends, outliers, and potential model issues
- Check if the relationship is linear (required for linear regression)
- Identify possible nonlinear patterns, clusters, or heteroskedasticity







Normal QQ Plot :

A Normal Quantile-Quantile (Q-Q) plot helps determine whether the residuals (errors) from a regression model follow a normal distribution, which is a critical assumption in linear regression for:

- Validity of t-tests and F-tests
- Accurate confidence intervals and prediction intervals

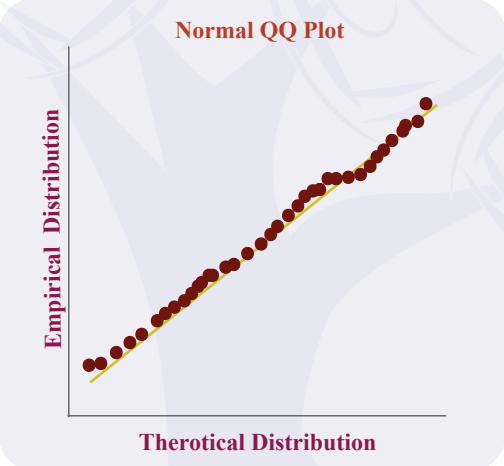
The Q-Q plot plots:

- Theoretical quantiles from a standard normal distribution (x-axis)
- Observed standardized residuals from your regression model (y-axis)

If residuals are normally distributed:
- Points fall approximately along the diagonal line

If not:
- Points deviate from this line, especially in the tails

Observation Pattern	Interpretation
Points closely follow the line	Residuals ≈ Normal → Normality assumption holds
Points curve away at both ends	Heavy tails (e.g., kurtosis or outliers)
S-shaped curve	Skewness in residuals
Few extreme points far off	Outliers in residuals



Evaluating Regression Model Fit and Interpreting Model Results

FinTree Fruit 1: INTRODUCTION

To assess the quality and reliability of a multiple regression model and interpret its output meaningfully:

We use following methods to evaluate how well the model explains the dependent variable and how to interpret the significance of its results -

- Adjusted R^2 – improves on R^2 by accounting for the number of predictors.
- AIC and BIC – used for comparing and selecting the best model among alternatives.
- t-tests and F-tests – assess the statistical significance of individual and groups of coefficients.
- Prediction techniques – guide how to forecast outcomes and construct confidence intervals by incorporating both model and sampling uncertainty.

FinTree Fruit 2: GOODNESS OF FIT

Simple linear regression :

Goodness of fit : Coefficient of determination (R^2)

$$R^2 = \frac{\text{Sum of squares regression}}{\text{Sum of squares total}} = \frac{\sum (\hat{Y} - \bar{Y})^2}{\sum (Y - \bar{Y})^2}$$

$\hat{Y}$ → Forecasted Y

$\bar{Y}$ → Mean

Y → Actual Y

Multiple Linear Regression :

Limitations of R^2 in multiple linear regression as a measure of a model's goodness of fit :

- Always Increases or Stays the Same: Adding more independent variables to a model will never decrease R^2, even if the new variables are irrelevant.
- Does Not Indicate Significance: R^2 doesn't provide information about the statistical significance of individual predictors.
- Cannot Detect Bias: It doesn't reveal biases in coefficient estimates or predictions.
- Misleading Fit Assessment: A high R^2 doesn't necessarily mean a good model fit; overfitting can inflate R^2 without improving predictive power.

Overfitting of a regression model is a situation in which the model is too complex, meaning there may be too many independent variables relative to the number of observations in the sample.

In multiple linear regression - Goodness of fit : Adjusted R^2

Adjusted R^2 modifies the standard R^2 to account for the number of predictors in the model, providing a more accurate measure of model fit.

$$R^2 = 1 - (1 - R^2) \times \frac{n - 1}{n - k - 1}$$

If, $k \geq 1$: R^2 is strictly greater than adjusted R^2
Adjusted R^2 may be negative, whereas the R^2 has a minimum of zero.

If the coefficient's t-statistic > |1.0|, then adjusted R^2 increases.
If the coefficient's t-statistic < |1.0|, then adjusted R^2 decreases.

Adjusted R^2 : Higher the better.

Limitations of Adjusted R^2

- **No Clear Interpretation as % Explained: Unlike R^2, Adjusted R^2 can't be directly interpreted as % variance explained.**
- **Doesn't Show Significance or Bias: It doesn't indicate coefficient significance or model bias—use residual plots and hypothesis tests for that.**
- **Not for Testing Model Fit: Adjusted R^2 doesn't assess overall model significance—use ANOVA and F-tests instead.**

Model Selection & Overfitting Risk

R^2 & Adjusted R^2 Can Mislead: Both may increase with more variables, risking overfitting.

Use Parsimony Criteria: AIC (Akaike Info. Criterion) and BIC (Bayesian Info. Criterion) help compare models and favor simpler ones with better explanatory power.

Both AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) help select the best model (same dependent variable) by balancing fit and complexity.

Akaike's information criterion (AIC) :

- **Evaluate collection of models that explain the same dependent variable.**

$$AIC = n \cdot \ln\left(\frac{\text{Sum of squares error}}{n}\right) + 2(k + 1)$$

Lower AIC indicates better fitting model.
The term 2 (k + 1) is the penalty assesed for adding independent variable to model.

Schwarz's Bayesian information criterion (BIC or SBC) :

- This allows comparison of models with the same dependent variable.

$$BIC = n \cdot \ln\left(\frac{\text{Sum of squares error}}{n}\right) + \ln(n)(k+1)$$

Lower BIC : Better it is.
BIC assesses a greater penalty for having more parameters in a model

AIC Model is used for prediction purposes, BIC model is preferred when the best goodness of fit is desired.

Assessing Model Fit Using Multiple Regression Statistics

Statistic	Criterion to use in assessment
Adjusted R^2	The higher the better
Akaike's information criterion (AIC)	The lower the better
Schwarz's Bayesian information criterion (BIC)	The lower the better
t-Statistic on a slope coefficient	Outside bounds of critical t-value(s) for the selected significance level
F-test for joint test of slope coefficients	Exceeds the critical F-value for the selected significance level