DegreeLive

B.Com Honours | Semester VI | Calicut University

Basics of Financial Econometrics

Course Code: COM6FS113 (4) • Module 3 Notes

Basics of Financial Econometrics
Course Code: COM6FS113 (4) • Module 3 Notes

Module III: Time Series Analysis in Finance

Financial asset prices, exchange rates, macroeconomic indicators, and sovereign bond yields unfold sequentially through time. Unlike cross-sectional data where observations are independently sampled, time series data exhibits profound temporal dependence, historical inertia, and dynamic feedback loops. Modeling financial time series requires specialized econometric apparatus capable of handling non-stationarity, stochastic trends, cointegrating equilibria, multivariate dynamic systems, and temporal memory. This module examines the theoretical and empirical mechanics of linear time series modeling: (1) stationarity concepts, random walks, and unit root testing paradigms; (2) structural breaks and testing in the presence of regime shifts; (3) cointegration, Error Correction Models (ECM), and statistical arbitrage pairs trading; (4) Autoregressive (AR) models; (5) Moving Average (MA) models; (6) Autoregressive Moving Average (ARMA) and Seasonal ARIMA (SARIMA) specifications; (7) Vector Autoregression (VAR), Structural VAR, and Granger Causality; and (8) the four-stage Box-Jenkins methodology for Autoregressive Integrated Moving Average (ARIMA) modeling, diagnostic checking, and out-of-sample forecast evaluation.

1. Stationarity, Random Walks, and Unit Roots

The concept of **stationarity** is the cornerstone of time series econometrics. A time series is stationary if its statistical properties do not change over time.

Foundations of Time Series Stationarity Mathematical Definition
1. Strict Stationarity: The joint probability distribution of any collection of observations (Y_{t1}, Y_{t2}, ..., Y_{tk}) is completely invariant under an arbitrary time shift tau. That is, the joint distribution depends only on the lag intervals between observations, not on calendar time. Highly restrictive and difficult to verify empirically.

2. Weak / Covariance Stationarity (Second-Order Stationarity): The operational benchmark in financial econometrics. A stochastic process Y_t is covariance-stationary if it satisfies three conditions:
• **Constant Mean:** E(Y_t) = mu for all t.
• **Constant Finite Variance:** Var(Y_t) = E[(Y_t - mu)^2] = sigma^2 < infinity for all t.
• **Lag-Dependent Autocovariance:** Cov(Y_t, Y_{t-k}) = E[(Y_t - mu)*(Y_{t-k} - mu)] = gamma_k depends strictly on the lag displacement *k*, and not on time *t*.

1.1 Autocovariance and Autocorrelation Function (ACF)

For a covariance-stationary process, the Autocorrelation Function (ACF) measures the linear correlation between observations separated by *k* time periods:

Autocorrelation Coefficient at Lag k (rho_k):
rho_k = gamma_k / gamma_0 = Cov(Y_t, Y_{t-k}) / Var(Y_t)

Core Properties:
• rho_0 = 1.00 (Correlation of an observation with itself).
• Symmetry: rho_k = rho_{-k}.
• Bound: -1.00 ≤ rho_k ≤ +1.00.
• **Bartlett's Asymptotic Standard Error:** Under the null hypothesis that all autocorrelations beyond lag

0 are zero, sample autocorrelation r_k has SE ≈ 1 / sqrt(T). The 95% confidence bands are plotted as **± 1.96 / sqrt(T)**. Any spike extending outside these bands is statistically significant at the 5% level.

1.2 The Pure White Noise Process

A white noise process {e_t} is the fundamental stochastic building block of time series models. A sequence is white noise if:
• Mean is zero: E(e_t) = 0 for all t.
• Variance is constant: Var(e_t) = sigma^2_e for all t.
• No autocorrelation across any lag: Cov(e_t, e_s) = 0 for all t ≠ s.
If the errors are additionally normally distributed, the series is *Gaussian White Noise*: e_t ~ i.i.d. N(0, sigma^2_e).

1.3 The Random Walk Model and Non-Stationarity

Most raw financial asset prices (stock prices, currency exchange rates, commodity prices) are **non-stationary**. They do not fluctuate around a fixed mean; instead, they wander without mean-reversion over time. The primary econometric model for non-stationary financial prices is the **Random Walk**:

1. Pure Random Walk (Without Drift):
P_t = P_{t-1} + e_t
By recursive substitution: P_t = P_0 + Sum(e_i) from i=1 to t.
• Mean: E(P_t) = P_0 (Constant).
• Variance: Var(P_t) = t * sigma^2_e → **Variance grows linearly with time t and approaches infinity!**
• Autocorrelation: Corr(P_t, P_{t-k}) approaches 1.0 as t becomes large.
The memory of a random walk is infinite: A shock e_t persists permanently in price levels forever.

2. Random Walk with Drift (mu):
P_t = mu + P_{t-1} + e_t = P_0 + mu*t + Sum(e_i)
• Exhibits a deterministic upward or downward trend slope mu, combined with stochastic variance expansion.

1.4 The Spurious Regression Hazard (Granger & Newbold, 1974)

When two independent, completely unrelated non-stationary time series (e.g., Apple stock price and agricultural wheat yields in India) are regressed against each other using standard OLS:

- The regression produces an artificially inflated, high R-squared value (often > 0.80).
- The t-statistics on the slope coefficients are massive (statistically significant at p < 0.001), falsely confirming a causal relationship where none exists.
- **The Granger-Newbold Warning Rule:** Whenever the regression **R-squared exceeds the Durbin-Watson d-statistic (R^2 > d)**, the regression is almost certainly spurious, driven by shared non-stationary stochastic trends rather than economic co-movement.

1.5 Trend-Stationary vs. Difference-Stationary Processes

Trend-Stationary Process (TSP)

Y_t = beta_0 + beta_1 * t + e_t. The series fluctuates around a deterministic linear time trend. Shocks are purely transitory; the series reverts back to its deterministic trend line. Made stationary by **detrending** (regressing Y_t on time t and saving residuals).

DegreeLive

Difference-Stationary Process (DSP)

Y_t = Y_{t-1} + e_t. Contains a **unit root (integrated of order 1, I(1))**. Shocks have permanent effects on future levels. Detrending fails to induce stationarity; it MUST be made stationary by **differencing**: Delta Y_t = Y_t - Y_{t-1}.

1.6 Testing for Unit Roots: Augmented Dickey-Fuller (ADF) & Phillips-Perron

To determine whether a series contains a unit root (phi = 1 in Y_t = phi*Y_{t-1} + e_t), David Dickey and Wayne Fuller (1979) formulated the unit root test by subtracting Y_{t-1} from both sides: Delta Y_t = gamma*Y_{t-1} + e_t, where gamma = phi - 1:

Augmented Dickey-Fuller (ADF) Test Regression:
Delta Y_t = alpha_0 + alpha_1 * t + gamma * Y_{t-1} + Sum[beta_i * Delta Y_{t-i}] + e_t (summed from i=1 to p)

Hypothesis Framework:
• **Null Hypothesis (H0): gamma = 0 (phi = 1) → Series has a Unit Root (Non-Stationary).**
• **Alternative Hypothesis (H1): gamma < 0 (phi < 1) → Series is Stationary.**

Decision Rule:
Under H0, the test statistic (tau) does not follow a standard Student's t-distribution; it follows the non-standard **Dickey-Fuller distribution** (Mackinnon critical values).
• If the calculated ADF test statistic is **more negative** than the critical value (e.g., calculated tau = -3.85 vs critical value of -2.88 at 5%), reject H0 → Series is stationary.
• If calculated ADF is less negative (e.g., -1.24), fail to reject H0 → Series is non-stationary and must be differenced.
Complementary Tests:
• *Phillips-Perron (PP) Test:* Uses non-parametric corrections for serial correlation and heteroskedasticity in residuals rather than adding lagged difference terms.
• *KPSS Test (Kwiatkowski et al., 1992):* Reverses the hypotheses: Null hypothesis is that the series is **stationary**. Rejecting H0 confirms non-stationarity. Used in confirmatory analysis alongside ADF.
• *Zivot-Andrews Test:* Tests for a unit root allowing for a single endogenous structural break in intercept, trend, or both. Protects against falsely identifying stationary series with structural shifts as unit root processes.

2. Cointegration and Error Correction Models (ECM)

While differencing non-stationary I(1) series eliminates spurious regression, it completely discards vital long-run equilibrium information contained in price levels. Clive Granger and Robert Engle (1987 Nobel laureates) proved that two or more non-stationary series can move together in long-run harmony through **Cointegration**.

Theoretical Concept of Cointegration Long-Run Equilibrium
Formal Definition:
Two time series Y_t and X_t are cointegrated of order CI(1, 1) if:
1. Both individual series are non-stationary integrated of order 1: Y_t ~ I(1) and X_t ~ I(1).
2. There exists a unique linear combination: **e_t = Y_t - beta_0 - beta_1 * X_t** such that the resulting residual equilibrium error e_t is **stationary integrated of order 0: e_t ~ I(0)**.

Economic Intuition: Cointegrated asset prices may wander wanderlust in the short term, but they are

bound by an invisible economic tether that prevents them from drifting permanently apart (e.g., Spot and Futures prices, dual-listed shares, sovereign yield spreads).

2.1 The Engle-Granger Two-Step Methodology

- **Step 1 (Cointegrating OLS Regression):** Estimate the static long-run relationship in levels: Y_t = beta_0 + beta_1 * X_t + e_t. Save the residuals e_t. Perform an ADF unit root test on residuals e_t (using Engle-Yoo cointegration critical values). If e_t is stationary I(0), cointegration is confirmed!
- **Step 2 (Vector Error Correction Model - VECM):** By the **Granger Representation Theorem**, if two series are cointegrated, they can be expressed in an Error Correction specification:
  Delta Y_t = alpha_0 + alpha_ec * e_{t-1} + Sum[gamma_i * Delta Y_{t-i}] + Sum[delta_i * Delta X_{t-i}] + u_t.
  • *e_{t-1} = (Y_{t-1} - beta_0 - beta_1 * X_{t-1}):* The **Error Correction Term (ECT)** measuring deviation from equilibrium in period t-1.
  • *alpha_ec:* The **Speed of Adjustment Coefficient**. Must be statistically significant and **negative (-1 < alpha_ec < 0)**. It indicates the speed at which any temporary market divergence is corrected back toward long-run equilibrium each period.

2.2 Cointegration in Quantitative Pairs Trading (Statistical Arbitrage)

In quantitative hedge funds, cointegration forms the mathematical foundation of **Pairs Trading**:

- **Constructing the Spread:** Estimate the cointegrating vector between two economically linked assets (e.g., Chevron and ExxonMobil): Spread_t = Price_A,t - beta * Price_B,t.
- **Normalized Z-Score Tracking:** Calculate Z_t = (Spread_t - Mean(Spread)) / SD(Spread).
- **Algorithmic Trading Rules:**
  • *Divergence Entry:* When Z_t > +2.0, Asset A is overpriced relative to Asset B → Short Asset A, Long Asset B.
  • *Convergence Exit:* When Z_t crosses back to 0.0, the temporary pricing anomaly has resolved → Close both positions and lock in arbitrage profit.
  • *Risk Stop-Loss:* If |Z_t| > 3.5, structural cointegration has broken down → Cut positions immediately.

2.3 The Johansen System Cointegration Framework

When modeling systems of three or more non-stationary variables (e.g., stock index, exchange rate, and interest rate), the Engle-Granger bivariate technique is inadequate because multiple cointegrating vectors may exist. Soren Johansen (1988) formulated a Maximum Likelihood system approach based on Vector Autoregressions:

- **Trace Statistic:** Tests the null hypothesis of at most *r* cointegrating vectors against the alternative of more than *r* vectors.
- **Maximum Eigenvalue Statistic:** Tests the null hypothesis of *r* cointegrating vectors against the specific alternative of *r + 1* vectors.

3. Autoregressive (AR) Models

An autoregressive model expresses the current value of a stationary time series as a linear function of its own past historical values plus a random disturbance.

DegreeLive

3.1 The First-Order Autoregressive Process: AR(1)

AR(1) Model Specification:
Y_t = c + phi * Y_{t-1} + e_t

Mathematical Properties:
• **Stationarity Condition:** The autoregressive parameter must satisfy **|phi| < 1** (-1 < phi < 1). If |phi| = 1, it becomes a random walk; if |phi| > 1, the series explodes exponentially.
• **Unconditional Mean:** mu = c / (1 - phi).
• **Unconditional Variance:** Var(Y_t) = sigma^2_e / (1 - phi^2).
• **Autocovariance:** gamma_k = phi^k * gamma_0.
• **Autocorrelation Function (ACF):** rho_k = phi^k. Decays geometrically toward zero. If phi > 0, ACF exhibits smooth exponential decay; if phi < 0, ACF alternates signs while decaying.
• **Partial Autocorrelation Function (PACF):** Cuts off cleanly to zero after lag 1!

3.2 Second-Order AR(2) Process & Yule-Walker Equations

The AR(2) process is: Y_t = c + phi_1 * Y_{t-1} + phi_2 * Y_{t-2} + e_t.

- **Stationarity Triangle:** An AR(2) process is stationary if and only if three joint conditions hold: (1) phi_1 + phi_2 < 1; (2) phi_2 - phi_1 < 1; and (3) |phi_2| < 1.
- **Pseudo-Cyclical Dynamics:** If (phi_1^2 + 4*phi_2) < 0, the characteristic roots are complex conjugates, generating damped cyclical wave patterns that model business cycles and interest rate fluctuations.
- **Yule-Walker Equations:** Establishes the relationship between autocorrelation coefficients and AR parameters:
  rho_1 = phi_1 / (1 - phi_2)
  rho_2 = phi_1 * rho_1 + phi_2 = (phi_1^2 + phi_2 - phi_2^2) / (1 - phi_2).

4. Moving Average (MA) Models

While AR models express dependence on past values of the series, Moving Average models express the current value as a linear combination of current and past white noise disturbance shocks.

4.1 The First-Order Moving Average Process: MA(1)

MA(1) Model Specification:
Y_t = mu + e_t + theta * e_{t-1}

Mathematical Properties:
• **Stationarity:** Any finite MA process is **always stationary**, regardless of the value of theta, because it is a finite linear combination of stationary white noise shocks.
• **Unconditional Mean:** E(Y_t) = mu.
• **Unconditional Variance:** Var(Y_t) = sigma^2_e * (1 + theta^2).
• **Autocorrelation Function (ACF):**
- At lag 1: rho_1 = theta / (1 + theta^2). (Maximum possible |rho_1| = 0.50 when |theta| = 1).
- At lag k ≥ 2: rho_k = 0. **The ACF strictly cuts off to zero after lag 1!**
• **PACF:** Tails off geometrically toward zero.

Degreelive

4.2 The Invertibility Condition and Wold Decomposition

An MA process is **invertible** if it can be equivalently rewritten as an infinite-order autoregressive AR(infinity) process: Y_t = c - Sum[(-theta)^j * Y_{t-j}] + e_t. Invertibility requires that the moving average parameter satisfies **|theta| < 1**. Invertibility ensures that past disturbance shocks can be recovered from historical observations of Y_t, with distant past shocks exerting diminishing influence on current pricing.

Wold's Decomposition Theorem: Fundamental theorem of time series stating that any covariance-stationary time series can be uniquely represented as the sum of a deterministic component plus an infinite moving average MA(infinity) of uncorrelated white noise innovations.

5. Autoregressive Moving Average (ARMA) & Seasonal ARIMA (SARIMA)

George Box and Gwilym Jenkins demonstrated that combining autoregressive and moving average terms creates parsimonious models that require far fewer parameters than pure AR or MA models.

ARMA(p, q) Specification:
Y_t = c + phi_1*Y_{t-1} + ... + phi_p*Y_{t-p} + e_t + theta_1*e_{t-1} + ... + theta_q*e_{t-q}

Seasonal ARIMA Model Formulation: SARIMA(p, d, q) x (P, D, Q)_s:
Accommodates seasonal cycles of length *s* (e.g., s = 4 for quarterly data, s = 12 for monthly data):
Phi_P(L^s) * phi_p(L) * (1 - L^s)^D * (1 - L)^d * Y_t = Theta_Q(L^s) * theta_q(L) * e_t

• Essential for modeling quarterly corporate earnings announcements, festival-driven seasonal retail revenues, and agricultural commodity price cycles.

Model Type	Autocorrelation Function (ACF)	Partial Autocorrelation Function (PACF)
AR(p)	Tails off (infinite geometric or damped sinusoidal decay).	**Cuts off sharply after lag p** (zero for all lags > p).
MA(q)	**Cuts off sharply after lag q** (zero for all lags > q).	Tails off (infinite geometric or damped sinusoidal decay).
ARMA(p, q)	Tails off (geometric decay beginning after lag q - p).	Tails off (geometric decay beginning after lag p - q).

6. Vector Autoregressive (VAR) Modeling and Granger Causality

When multiple financial time series interact dynamically (such as stock index returns, foreign exchange rates, and crude oil prices), univariate ARMA models fail to capture inter-variable feedback. Christopher Sims (1980) pioneered **Vector Autoregression (VAR)**, treating all variables symmetrically as endogenous:

Bivariate VAR(1) Model:
Y_t = c_1 + a_{11} * Y_{t-1} + a_{12} * X_{t-1} + e_{1,t}
X_t = c_2 + a_{21} * Y_{t-1} + a_{22} * X_{t-1} + e_{2,t}

Core Diagnostic & Analytical Tools of VAR:
• **Lag Order Selection:** Selected by minimizing multivariate AIC, SIC, or Hannan-Quinn information criteria.
• **Granger Causality Testing:** Tests whether past lagged values of X help predict Y beyond past values of Y

itself (H0: a_{12,1} = a_{12,2} = ... = 0).
• **Impulse Response Function (IRF):** Traces the dynamic temporal response of variable Y to a one-standard-deviation shock in variable X across future time horizons.
• **Forecast Error Variance Decomposition (FEVD):** Quantifies what percentage of the unexpected forecast error variance in Y is driven by shocks to X versus shocks to Y itself.
• **Structural VAR (SVAR):** Imposes economic theory restrictions on contemporaneous interactions, overcoming the arbitrary ordering limitations of Cholesky factorization.

7. The Box-Jenkins Methodology for ARIMA Modeling

Box and Jenkins (1970) formulated a systematic four-stage iterative methodology for constructing **Autoregressive Integrated Moving Average (ARIMA(p, d, q))** models, where *d* represents the order of differencing required to achieve stationarity.

The Four-Stage Box-Jenkins Protocol Iterative Cycle
Stage 1: Identification:
• Plot the raw time series to inspect trends, cycles, and seasonality.
• Perform ADF and PP unit root tests. If the series is non-stationary, difference the series *d* times (typically d = 1 for asset prices) until stationarity is attained.
• Compute sample ACF and PACF correlograms of the stationary differenced series Delta^d Y_t to select candidate lag orders *p* and *q*.

Stage 2: Estimation:
• Estimate candidate ARIMA(p, d, q) specifications using Maximum Likelihood Estimation (MLE).
• Verify parameter significance (t-statistics > 1.96, p < 0.05).
• Compare competing candidate specifications using Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC/BIC). Select the model that minimizes information criteria.

Stage 3: Diagnostic Checking:
• Evaluate whether the residuals of the estimated model behave as pure white noise.
• Inspect residual ACF and PACF (all residual autocorrelations should fall inside the 95% Bartlett confidence bands ± 1.96 / sqrt(T)).
• Calculate the **Ljung-Box Q-Statistic:** Q = T*(T + 2) * Sum[r_k^2 / (T - k)]. Under H0 (residuals are white noise), Q follows Chi-Square(m - p - q). A non-significant p-value (p > 0.05) validates the model.

Stage 4: Forecasting & Out-of-Sample Evaluation:
• Generate dynamic point forecasts and prediction confidence intervals (fan charts).
• Evaluate out-of-sample forecast accuracy using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Theil's U-Statistic.

7.1 Step-by-Step Empirical ARIMA Case Study

Consider an econometrician modeling the monthly 10-year Indian Government Bond Yield over T = 120 months:

- **Identification:** The ADF test on yield levels produces tau = -1.42 (p = 0.57 > 0.05), confirming non-stationarity. First-differencing yields Delta Y_t, for which ADF produces tau = -7.84 (p < 0.001), confirming stationarity (d = 1). The PACF of Delta Y_t displays a significant spike at lag 1, while ACF tails off. Candidate models: ARIMA(1,1,0), ARIMA(0,1,1), and ARIMA(1,1,1).

DegreeLive

Candidate Model	Log-Likelihood	Akaike (AIC)	Schwarz (SIC)	Ljung-Box Q(12) Sig.
ARIMA(1, 1, 0)	142.35	-280.70	-275.14	p = 0.421 (White Noise)
ARIMA(0, 1, 1)	140.80	-277.60	-272.04	p = 0.285 (White Noise)
ARIMA(1, 1, 1)	142.85	-279.70	-271.36	p = 0.512 (White Noise)

Model Selection & Diagnostic Conclusion: ARIMA(1, 1, 0) achieves the minimum AIC (-280.70) and minimum SIC (-275.14). The estimated AR(1) parameter is phi = 0.284 (SE = 0.088, t = 3.23, p = 0.002). The Ljung-Box Q-statistic at lag 12 has p = 0.421 > 0.05, confirming that residuals are pure white noise without remaining autocorrelation. ARIMA(1, 1, 0) is selected as the parsimonious forecasting engine.

7.2 Out-of-Sample Forecast Evaluation Metrics

To evaluate how well an econometric forecasting model performs on future data untouched during estimation, researchers compute formal predictive metrics across an evaluation test window of length *h*:

Forecasting Accuracy Performance Metrics:
• **Root Mean Squared Error (RMSE):** sqrt[(1/h) * Sum((Actual_t - Forecast_t)^2)]. Penalizes large forecast errors heavily.
• **Mean Absolute Error (MAE):** (1/h) * Sum(|Actual_t - Forecast_t|). Robust to isolated outliers.
• **Mean Absolute Percentage Error (MAPE):** (1/h) * Sum(|(Actual_t - Forecast_t) / Actual_t|) * 100%. Expresses error as a percentage.
• **Theil's U-Statistic:** Compares model forecasts against a naive benchmark (e.g., no-change random walk forecast). U < 1 indicates that the econometric model outperforms a naive random walk; U = 1 indicates equal accuracy; U > 1 indicates the model is inferior to a naive forecast.
• **Diebold-Mariano Test:** Evaluates whether the difference in loss functions (e.g., squared error loss) between two competing forecasting models is statistically significant.

8. Fractional Integration and Long Memory Processes (ARFIMA)

Standard time series models enforce a sharp dichotomy: a series is either strictly stationary I(0) where shocks decay exponentially, or non-stationary I(1) where shocks persist permanently. However, many financial series (especially market volatility, trading volume, and real exchange rates) exhibit **Long Memory**—autocorrelations decay at an exceedingly slow, hyperbolic power-law rate rather than an exponential geometric rate.

Autoregressive Fractionally Integrated Moving Average (ARFIMA) Fractional Integration
ARFIMA(p, d, q) Model:
Phi(L) * (1 - L)^d * Y_t = Theta(L) * e_t

Properties of Fractional Differencing Parameter d:
• *d = 0:* Standard short-memory stationary ARMA process.
• *0 < d < 0.50:* Stationary and invertible process exhibiting **Long Memory / Persistence**. Autocorrelations are all positive and decay hyperbolically: rho_k ≈ k^(2*d - 1).
• *-0.50 < d < 0:* Stationary and invertible process exhibiting **Anti-Persistence / Intermediate Memory**.
• *0.50 ≤ d < 1.0:* Non-stationary process, but **mean-reverting** (shocks eventually dissipate, unlike a pure random walk).
• *d = 1.0:* Standard unit root random walk process.

8.1 The Hurst Exponent (H) and Rescaled Range (R/S) Analysis

Originally developed by hydrologist Harold Edwin Hurst and introduced to finance by Benoit Mandelbrot, the **Hurst Exponent (H)** relates directly to fractional integration: **$H = d + 0.50$**.

- **$H = 0.50$ ($d = 0$):** The series is an independent random walk (Brownian motion). Past price changes have zero correlation with future price changes. Validates the weak-form Efficient Market Hypothesis.
- **$0.50 < H \leq 1.00$ ($0 < d \leq 0.50$):** The series exhibits **persistent long memory and trend-reinforcing behavior**. If the asset has been rising in the past, it has a statistically higher probability of continuing to rise. Typical of commodity supercycles and implied volatility indexes.
- **$0.00 \leq H < 0.50$ ($-0.50 \leq d < 0$):** The series exhibits **anti-persistence and strong mean-reversion**. High returns are systematically followed by low returns, creating choppy, range-bound oscillator behavior. Typical of high-frequency bid-ask spread reversals and interest rate spreads.

9. Practical Time Series Implementation in R and Python

Empirical econometricians deploy script-driven automated pipelines to execute unit root tests, determine optimal lag lengths, and forecast ARIMA models:

```
# R Implementation: Automated ARIMA Pipeline
library(urca)
library(forecast)

# 1. Augmented Dickey-Fuller Test with Trend and Drift
adf_result <- ur.df(yield_series, type = "trend", selectlags = "AIC")
summary(adf_result)

# 2. Automated Box-Jenkins Model Identification and Estimation
fit_arima <- auto.arima(yield_series, seasonal = FALSE, ic = "aic", stepwise = FALSE)
summary(fit_arima)

# 3. Diagnostic Residual White Noise Checking
checkresiduals(fit_arima)

# 4. Generating 12-Month Out-of-Sample Dynamic Forecast
forecast_values <- forecast(fit_arima, h = 12)
plot(forecast_values)
```

```
# Python Implementation: Statsmodels Time Series Pipeline
import statsmodels.api as sm
from statsmodels.tsa.stattools import adfuller
from statsmodels.tsa.arima.model import ARIMA

# 1. Unit Root Test
adf_stat, p_val, lags, nobs, crit_vals, _ = adfuller(yield_series, autolag='AIC')
print(f"ADF Statistic: {adf_stat:.4f}, p-value: {p_val:.4f}")

# 2. Fit ARIMA(1, 1, 0) Model
model = ARIMA(yield_series, order=(1, 1, 0))
results = model.fit()
print(results.summary())
```

DegreeLive

```
# 3. Ljung-Box Diagnostic Test on Residuals
lb_test = sm.stats.acorr_ljungbox(results.resid, lags=[12], return_df=True)
print(lb_test)

# 4. Dynamic Forecast for 12 Steps Ahead
forecast_df = results.get_forecast(steps=12).summary_frame()
```

Summary Synthesis

Time Series Concept	Mathematical / Operational Rule	Primary Financial Application
Covariance Stationarity	Constant mean, constant variance, lag-dependent autocovariance.	Guarantees that statistical parameters remain stable and predictable over time.
Unit Root / Non-Stationarity	AR coefficient phi = 1; variance expands linearly with time t.	Asset prices follow random walks; requires differencing to avoid spurious regression.
Cointegration & VECM	Linear combination of I(1) variables is stationary I(0); speed of adjustment alpha < 0.	Models long-run asset pricing equilibria, pairs trading arbitrage, and term structures.
Vector Autoregression (VAR)	Symmetric multi-equation system with lagged feedbacks across variables.	Captures macroeconomic-financial linkages, Granger causality, and impulse responses.
Autoregressive Model (AR)	Y_t modeled on own historical lags; stationary if |phi| < 1.	PACF cuts off at lag p; models momentum and mean-reverting interest rates.
Moving Average (MA)	Y_t modeled on past white noise shocks; invertible if |theta| < 1.	ACF cuts off at lag q; captures short-lived transitory market microstructure noise.
Box-Jenkins Methodology	Four-stage cycle: Identification, Estimation, Diagnostics, Forecasting.	Generates optimal linear forecasts for asset returns, exchange rates, and yields.

Acing your exams is just a click away!
Visit www.degreelive.in to download the next module for free.

DegreeLive