Time Series Analysis and Forecasting

Seminar 1: Stochastic Processes and Stationarity

Daniel Traian PELE

Bucharest University of Economic Studies
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Romanian Academy, Institute for Economic Forecasting
MSCA Digital Finance

Seminar Outline

Today's Activities:

1. **Quick Review** — Stationarity and key processes
2. **Multiple Choice Quiz** — Test your understanding
3. **True/False Questions** — Conceptual checks
4. **Calculation Exercises** — Hands-on practice
5. **Key Formulas** — Reference sheet

Key Concepts to Remember

Weak Stationarity:

- $\mathbb{E}[X_t] = \mu$ (constant mean)
- $\text{Var}(X_t) = \sigma^2$ (constant variance)
- $\text{Cov}(X_t, X_{t+h}) = \gamma(h)$ (depends only on lag)

White Noise:

- $\mathbb{E}[\varepsilon_t] = 0$
- $\text{Var}(\varepsilon_t) = \sigma^2$
- $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for $t \neq s$

Random Walk:

- $X_t = X_{t-1} + \varepsilon_t$
- $\text{Var}(X_t) = t\sigma^2$ (grows!)
- Non-stationary

AR(1) Process:

- $X_t = \phi X_{t-1} + \varepsilon_t$
- Stationary if $|\phi| < 1$
- $\rho(h) = \phi^h$ (ACF)

Quiz 1: Time Series Basics

Question

Which of the following is NOT a characteristic of time series data?

A. Observations are ordered in time
B. Consecutive observations are usually correlated
C. Observations are independent and identically distributed
D. Data has a natural temporal ordering

Quiz 1: Answer

Answer: C

Observations are independent and identically distributed

Explanation:
Time series observations are typically **dependent** (autocorrelated), not independent.

This temporal dependence makes time series analysis unique.

ch1_ts_patterns.pdf

Examples of time series with different patterns

Quiz 2: Stationarity (Random Walk)

Question

A random walk process $X_t = X_{t-1} + \varepsilon_t$ is:

A. Strictly stationary

B. Weakly stationary

C. Non-stationary because variance grows with time

D. Stationary after adding a constant

Quiz 2: Answer

Answer: C

Non-stationary because variance grows with time

Explanation:
For random walk:

$$\mathrm{Var}(X_t) = t\sigma^2$$

Variance depends on $t \Rightarrow$ non-stationary.

Solution: Differencing gives $\Delta X_t = \varepsilon_t$ which IS stationary.

ch1_wn_rw.pdf

Visual: White Noise vs Random Walk

ch1_wn_rw.pdf

Left: White noise (stationary) **Right**: Random walk (non-stationary, variance grows)

ql-logo.png TSA_ch1_wn_rw

Quiz 3: Unit Root Tests

Question

You run ADF and KPSS tests. ADF fails to reject H_0, and KPSS rejects H_0. What do you conclude?

A. The series is stationary

B. The series has a unit root (non-stationary)

C. Results are inconclusive

D. You need to run more tests

Quiz 3: Answer

Answer: B

The series has a unit root (non-stationary)

Explanation:

- ADF: H_0 = unit root
  (fail to reject $\Rightarrow$ evidence FOR)
- KPSS: H_0 = stationary
  (reject $\Rightarrow$ evidence AGAINST)

Both tests: series is **non-stationary**.

ADF	KPSS	Conclusion
Reject	No rej.	Stationary
No rej.	Reject	Non-stat.
Reject	Reject	Uncertain
No rej.	No rej.	Uncertain

ql logo.png TSA_ch1_unit_root_tests

Quiz 4: ACF Interpretation

Question

If the ACF of a time series decays very slowly (remains significant for many lags), this suggests:

A. The series is white noise
B. The series is likely non-stationary
C. The series has no autocorrelation
D. The series is perfectly predictable

Quiz 4: Answer

Answer: B

The series is likely non-stationary

Explanation:

- **Stationary**: ACF decays rapidly ($\rho_k = \phi^k \to 0$)
- **Non-stationary**: ACF stays close to 1

$\Rightarrow$ Differencing needed.

ch1_acf_examples.pdf

Visual: ACF Stationary vs Non-Stationary

ch1_acf_examples.pdf

Left: ACF decays rapidly (stationary) **Right**: ACF decays slowly (non-stationary)

Quiz 5: White Noise Properties

Question

Which property is NOT required for a process to be white noise?

A. $\mathbb{E}[\varepsilon_t] = 0$

B. $\text{Var}(\varepsilon_t) = \sigma^2$ (constant)

C. $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for $t \neq s$

D. $\varepsilon_t \sim N(0, \sigma^2)$

Quiz 5: Answer

Answer: D

$\varepsilon_t \sim N(0, \sigma^2)$ (normality)

White noise requires only:

- $\mathbb{E}[\varepsilon_t] = 0$ (zero mean)
- $\text{Var}(\varepsilon_t) = \sigma^2$ (constant variance)
- $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for $t \neq s$

Normality is NOT required – Gaussian white noise is a special case.

ch1_wn_rw.pdf

White noise: random fluctuations around zero

Quiz 6: Trend Types

Question

A deterministic trend can be removed by:

A. Differencing

B. Regression on time

C. Seasonal adjustment

D. Moving average smoothing

Quiz 6: Answer

Answer: B

Regression on time

Deterministic trend:

$$Y_t = \alpha + \beta t + \varepsilon_t$$

Removal: regress on t, analyze residuals.

Stochastic trend (unit root):

$$Y_t = Y_{t-1} + \varepsilon_t$$

Requires **differencing**.

Trend type	**Method**
Deterministic	Regression
Stochastic	Differencing

Warning: Using wrong method can introduce problems!

ql-logo.png TSA_ch1_trend_types

Quiz 7: MA(1) Process

Question

For the MA(1) process $X_t = \varepsilon_t + \theta\varepsilon_{t-1}$, the autocorrelation function is:

A. $\rho(h) = \theta^h$ for all h

B. $\rho(1) = \frac{\theta}{1+\theta^2}$, $\rho(h) = 0$ for $h > 1$

C. $\rho(h) = 0$ for all $h \geq 1$

D. $\rho(h) = \theta$ for all h

Quiz 7: Answer

Answer: B

$\rho(1) = \frac{\theta}{1+\theta^2}$, $\rho(h) = 0$ for $h > 1$

Explanation:
MA(1) has **finite memory**: autocorrelation cuts off after lag 1.

Key difference from AR: ACF cuts off sharply (doesn't decay gradually).

ch2_ma1_acf_pacf.pdf

ACF and PACF for MA(1): ACF cuts off at lag 1

Quiz 8: Differencing

Question

If X_t follows a random walk with drift: $X_t = \mu + X_{t-1} + \varepsilon_t$, what is $\Delta X_t = X_t - X_{t-1}$?

A. A random walk

B. A non-stationary process

C. White noise plus a constant: $\Delta X_t = \mu + \varepsilon_t$

D. An AR(1) process

Quiz 8: Answer

Answer: C

White noise plus a constant: $\Delta X_t = \mu + \varepsilon_t$

Explanation:
Differencing removes the unit root.

The result is **stationary**:

- Constant mean μ
- Constant variance σ^2
- No autocorrelation

ch1_stationarity.pdf

Top: prices (non-stat.)

Bottom: returns (stat.)

Quiz 9: Strict vs Weak Stationarity

Question

Which statement about stationarity is TRUE?

A. Weak stationarity implies strict stationarity
B. Strict stationarity implies weak stationarity (if moments exist)
C. They are equivalent concepts
D. Neither implies the other

Quiz 9: Answer

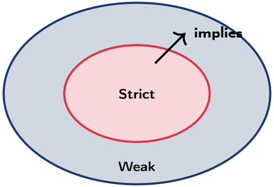
Answer: B

Strict stationarity implies weak stationarity (if moments exist)

Explanation:

- **Strict**: All joint distributions invariant
- **Weak**: Only 1st and 2nd moments

Strict $\Rightarrow$ Weak, but NOT vice versa.



Strict $\subset$ Weak (when moments exist)

ql-logo.png TSA_ch1_Stationarity

True or False? — Questions

Statement	T/F?
1. A time series with constant mean is always stationary.	?
2. Random walk variance increases linearly with time.	?
3. ADF and KPSS have the same null hypothesis.	?
4. Autocorrelation at lag 0 is always equal to 1.	?
5. White noise is always normally distributed.	?
6. The ACF of AR(1) decays exponentially.	?
7. Differencing can make a non-stationary series stationary.	?
8. $\lvert\phi\rvert = 1$ in AR(1) gives a random walk.	?

True or False? — Answers

Statement	T/F	Explanation
1. A time series with constant mean is always stationary.	F	**Also needs constant variance, covariance**
2. Random walk variance increases linearly with time.	T	$\mathbf{Var}(X_t) = t\sigma^{\mathbf{2}}$
3. ADF and KPSS have the same null hypothesis.	F	**Opposite hypotheses**
4. Autocorrelation at lag 0 is always equal to 1.	T	$\rho(\mathbf{0}) = \mathbf{1}$
5. White noise is always normally distributed.	F	**Gaussian WN = special case**
6. The ACF of AR(1) decays exponentially.	T	$\rho(h) = \phi^{h}$
7. Differencing can make a non-stationary series stationary.	T	**Removes stochastic trends**
8. $\lvert\phi\rvert = 1$ in AR(1) gives a random walk.	T	**Unit root**

Exercise 1: Autocovariance

Problem: For a stationary process with: $\mathbb{E}[X_t] = 5$, $\gamma(0) = 4$, $\gamma(1) = 2$, $\gamma(2) = 1$

Calculate:

a) Autocorrelation function $\rho(0), \rho(1), \rho(2)$

b) $\text{Cov}(X_t, X_{t-1})$

c) $\text{Corr}(X_5, X_7)$

d) If $X_t = 6$, what is $\mathbb{E}[X_{t+1}|X_t = 6]$ assuming AR(1)?

Solution:

- $\rho(0) = 1$, $\rho(1) = 2/4 = 0.5$, $\rho(2) = 1/4 = 0.25$
- $\text{Cov}(X_t, X_{t-1}) = \gamma(1) = 2$
- $\text{Corr}(X_5, X_7) = \rho(2) = 0.25$
- $\mathbb{E}[X_{t+1}|X_t] = 5 + 0.5(6 - 5) = 5.5$

Exercise 2: Random Walk Properties

Problem: $X_t = X_{t-1} + \varepsilon_t$ where $\varepsilon_t \sim WN(0, 4)$ and $X_0 = 100$.

Calculate:

a) $\mathbb{E}[X_{10}]$

b) $\text{Var}(X_{10})$

c) $\text{Cov}(X_5, X_{10})$

d) 95% confidence interval for X_{100}

Solution:

- $\mathbb{E}[X_{10}] = X_0 = 100$
- $\text{Var}(X_{10}) = 10 \times 4 = 40$
- $\text{Cov}(X_5, X_{10}) = \min(5, 10) \times 4 = 20$
- 95% CI: $100 \pm 1.96 \times 20 = [60.8, 139.2]$

Exercise 3: AR(1) Process

Problem: Consider $X_t = 0.8X_{t-1} + \varepsilon_t$ where $\varepsilon_t \sim WN(0, 9)$.

Calculate:

a) Is this process stationary? Why?

b) The unconditional variance $\text{Var}(X_t)$

c) The autocorrelations $\rho(1), \rho(2), \rho(3)$

d) If $X_t = 10$, find $\mathbb{E}[X_{t+1}|X_t]$ and $\mathbb{E}[X_{t+2}|X_t]$

Solution:

- Yes, $|\phi| = 0.8 < 1$ (root outside unit circle)
- $\text{Var}(X_t) = \frac{\sigma^2}{1-\phi^2} = \frac{9}{1-0.64} = 25$
- $\rho(1) = 0.8$, $\rho(2) = 0.64$, $\rho(3) = 0.512$
- $\mathbb{E}[X_{t+1}|X_t = 10] = 8$, $\mathbb{E}[X_{t+2}|X_t = 10] = 6.4$

Exercise 4: MA(1) Process

Problem: Consider $X_t = \varepsilon_t + 0.6\varepsilon_{t-1}$ where $\varepsilon_t \sim WN(0, 4)$.

Calculate:

a) $\mathbb{E}[X_t]$

b) $\text{Var}(X_t)$

c) $\gamma(1) = \text{Cov}(X_t, X_{t-1})$

d) $\rho(1)$ and $\rho(2)$

Solution:

- $\mathbb{E}[X_t] = 0$ (white noise has zero mean)
- $\text{Var}(X_t) = (1 + \theta^2)\sigma^2 = (1 + 0.36) \times 4 = 5.44$
- $\gamma(1) = \theta\sigma^2 = 0.6 \times 4 = 2.4$
- $\rho(1) = \frac{\theta}{1+\theta^2} = \frac{0.6}{1.36} = 0.44$, $\rho(2) = 0$

Python Exercise 1: Loading and Visualization

Task: Load S&P 500 data and create a plot.

Starter Code

```
import yfinance as yf
import matplotlib.pyplot as plt

# Download S&P 500 data
sp500 = yf.download('^GSPC', start='2020-01-01', end='2025-01-01')

# TODO: Plot closing prices
# TODO: Calculate and display basic statistics
# TODO: Does the series appear stationary? Why or why not?
```

ql_logo.png TSA_ch1_python_viz

Python Exercise 2: Stationarity Testing

Task: Test stationarity using ADF and KPSS tests.

Starter Code

```
from statsmodels.tsa.stattools import adfuller, kpss

# Test S&P 500 prices
prices = sp500['Close']
returns = prices.pct_change().dropna()

# TODO: Run ADF test on prices and returns
# TODO: Run KPSS test on prices and returns
# TODO: Interpret the results
```

Questions: Are prices stationary? Are returns stationary?

ql_logo.png TSA_ch1_python_adf

Case Study: S&P 500 Index

ch1_stationarity.pdf

Top: Prices – clear upward trend (non-stationary) **Bottom**: Returns – stationary, fluctuate around zero

ql_log

TSA_ch1_stationarity

Stationarity Comparison: Prices vs Returns

ADF Test Results

Series	ADF Statistic	p-value	Conclusion
S&P 500 Prices	-0.82	0.812	Non-stationary
S&P 500 Returns	-45.3	< 0.001	Stationary

Key Observation

Financial prices are typically $I(1)$ – integrated of order 1. Taking first differences (returns) achieves stationarity. That's why we model **returns**, not prices!

Discussion Question 1

Scenario

You are analyzing daily stock prices for a technology company. The ACF of the price series decays very slowly, remaining significant even at lag 50. The ADF test gives a p-value of 0.73.

Discuss:

1. Is this series stationary? What evidence supports your answer?
2. What transformation would you apply before modeling?
3. After transformation, what properties would you expect the ACF to have?

Discussion Question 2

Scenario

A colleague claims: "I ran the ADF test on my stock price data and got a p-value of 0.65, so my data is stationary and I can fit an ARMA model directly."

Discuss:

1. What is wrong with this interpretation?
2. What do the ADF hypotheses actually test?
3. What should the colleague do before fitting an ARMA model?

Key Takeaways

1. **Time series are dependent** – not i.i.d. like cross-sectional data
2. **Stationarity is crucial** – constant mean, variance, and autocovariance
3. **Random walk is non-stationary** – variance grows with time ($t\sigma^2$)
4. **Test for stationarity** – use both ADF and KPSS together
5. **Differencing removes unit roots** – transforms $I(1)$ to $I(0)$
6. **ACF reveals structure** – slow decay suggests non-stationarity

Next Seminar

ARMA/ARIMA model identification, estimation and forecasting

Key Formulas Summary

Concept	Formula
Weak Stationarity	$\mathbb{E}[X_t] = \mu$, $\mathrm{Var}(X_t) = \sigma^2$, $\mathrm{Cov}(X_t, X_{t+h}) = \gamma(h)$
Autocorrelation	$\rho(h) = \gamma(h)/\gamma(0)$
White Noise	$\mathbb{E}[\varepsilon_t] = 0$, $\mathrm{Var}(\varepsilon_t) = \sigma^2$, $\mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0$
Random Walk	$X_t = X_{t-1} + \varepsilon_t$, $\mathrm{Var}(X_t) = t\sigma^2$
AR(1)	$X_t = \phi X_{t-1} + \varepsilon_t$
AR(1) Stationarity	$\|\phi\| < 1$
AR(1) Variance	$\mathrm{Var}(X_t) = \sigma^2/(1-\phi^2)$
AR(1) ACF	$\rho(h) = \phi^h$
ADF Test	H_0: unit root (non-stationary)
KPSS Test	H_0: stationary

Questions?

Good luck with the exercises!