Empirical Finance: Week 2

Final Consolidated Exam Study Guide

Time-Series Models: MA, AR, ARMA

Compiled from Lecture Captions & Handout 2

CRITICAL EXAM INSTRUCTIONS FROM PROFESSOR

Exam Format:

- **Pen and paper exam** (NOT computational) — *"Maybe you use a calculator"*
- Multiple choice questions included
- 50% final exam + 50% coursework
- *"Tutorials will be helpful for the coursework, slides for the final exam"*

What You MUST Memorise:

- MA(1) moments: Mean, Variance, First-order autocorrelation
- AR(1) moments: Mean (conditional), First-order autocorrelation
- *"For the exam you have to remember how to calculate the moments of an MA(1) only and AR(1) process only"*

What Will Be Given:

- *"Unless I tell you to remember the formula, you are not supposed to remember it"*
- *"I might give you the formula in the test and ask you to apply it — plug in numbers"*
- AR(1) variance formula will be given

Professor's Warnings

- *"I want to see 100% correct on MA(1) and AR(1) questions"*
- *"If I ask for volatility, I want volatility NOT variance — 50% got this wrong last year"*
- *"The devil is in the details. Read the question carefully!"*
- *"I can always default on my promise [about not testing formulas]... it's a possibility"*

1 Stochastic Processes & Time Series (Pages 6–7)

Professor's Key Teaching Points

- **Stochastic process:** A collection/bundle of random variables $\{Y_t\}$ where $t = 1, 2, \ldots, T$
- **Time series:** A single realisation of a stochastic process — *"like a black box producing data; we only observe one realisation"*
- *"In finance, we don't have the luxury of running experiments multiple times — we only have one single realisation"*
- Observations are dependent over time — *"close in calendar space"*
- *"Standard cross-sectional regression theory does not apply"*
- *"Time ordering is not a detail — it's an important aspect"*

Exam Learning Objectives — You Must Be Able To:

1. Distinguish between a stochastic process (the "black box") and a time series (the observed data)
2. Understand that time series observations are dependent: $y_t \not\perp \mathcal{F}_{t-1}$
3. Explain why time ordering changes inference compared to cross-sectional data

Formulas to MEMORISE:

None required.

Formulas Given on Exam:

- Random walk: $Y_t = Y_{t-1} + \varepsilon_t$ with $\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$
- i.i.d. process: $Y_t \overset{iid}{\sim} D$ where D is any distribution

2 Stationarity (Pages 15–25)

Professor's Key Teaching Points

Why Stationarity Matters:

- *"Stationarity is a measure of regularity in the data"*
- *"Without stationarity, you can't make inference — forget about it, go for a walk"*
- *"If the underlying data generating process changes very often, then you can't make inference"*

Asteroid Analogy:

- *"When the asteroid hit the Earth and all dinosaurs disappeared — this is a permanent shock"*
- *"You can't use past data to say anything about the future after such a game changer"*
- Examples of "game changers": regulation changes, central bank policy shifts, new technology

Why We Use Returns Not Prices:

- *"The S&P 500 is non-stationary — the mean changes with time"*
- *"This is one of the reasons why we don't run regressions on prices"*
- *"If it's not stationary, you can't calculate the mean — you can't make inference"*

Critical Exam Point:

- *"Strong stationarity is NOT nesting weak stationarity — they are not nested"*
- *"In the exam, in a multiple choice question, there could be some relating to this"*
- *"I don't care about the maths behind. But I care mostly about the economic intuition"*

Exam Learning Objectives — You Must Be Able To:

1. Define weak (covariance) stationarity and state its three conditions
2. Define strong (strict) stationarity
3. **Explain that weak and strong stationarity are NOT nested** (likely MCQ topic!)
4. Give examples: Cauchy (strictly but not weakly stationary), time-varying kurtosis (weakly but not strictly)
5. Explain why we work with returns instead of prices
6. Identify whether a process is stationary from its properties

Formulas to MEMORISE:

None required.

Formulas Given on Exam:

Weak Stationarity Conditions:

1. $E(Y_t) = \mu$ for all t (constant mean)
2. $\text{Var}(Y_t) = \sigma^2 < \infty$ for all t (constant, finite variance)
3. $\text{Cov}(Y_t, Y_{t-j}) = \gamma_j$ (covariance depends only on lag j, not on t)

Autocorrelation: $\rho_j = \frac{\gamma_j}{\gamma_0}$

3 Ergodicity (Page 26)

Professor's Key Teaching Points

- *"Ergodicity is a generalisation of the Law of Large Numbers for dependent data"*
- **Key intuition:** Shocks are **temporary** and dissipate over time
- *"Think of the Great Depression — it was a disaster, but its impact on today's data has completely disappeared"*
- *"If you have a shock, the shock should try to vanish at some point in time"*
- Sample moments converge in probability to population moments as errors vanish

Exam Learning Objectives — You Must Be Able To:

1. Understand ergodicity as the property that shocks dissipate over time
2. Distinguish between temporary shocks (ergodic) and permanent shocks (non-ergodic, like the asteroid)
3. Contrast with random walk where shocks have permanent effects

Formulas:

None required.

4 White Noise (Page 28)

Professor's Key Teaching Points

- *"For white noise, we only require that the errors are uncorrelated — we don't require that they are independent"*
- *"Be careful when you distinguish the concept of independent versus uncorrelated"*
- **Key insight:** Returns can be uncorrelated but NOT independent (squared returns are correlated)
- *"This happens with most financial data — volatility clustering"*
- *"If you plot the ACF and it's basically all zero, this should ring a bell — maybe it's white noise"*
- *"If you say white noise, can you do predictability? No — don't even try"*

Exam Learning Objectives — You Must Be Able To:

1. Distinguish between WN (uncorrelated), IWN (independent), and GWN (Gaussian)
2. **Understand that uncorrelated $\neq$ independent** (critical concept!)
3. Explain the ARCH example: $y_t = \sqrt{h_t}\varepsilon_t$ where returns are uncorrelated but squared returns are correlated
4. Recognize that white noise is unpredictable (no autocorrelation)

Formulas Given on Exam:

White noise conditions: $E(\varepsilon_t) = 0$, $\text{Var}(\varepsilon_t) = \sigma^2$, $\text{Cov}(\varepsilon_t, \varepsilon_\tau) = 0$ for $t \neq \tau$

5 Moving Average (MA) Processes (Pages 35–45)

Professor's Key Teaching Points

Intuition for MA(1):

- *"The current Y depends on the current shock but also depends on the past shock"*
- *"Imagine stock returns a week after COVID — there's a shock today, but the COVID shock is still affecting returns"*
- *"During crises, prices continue to fall for several days because market makers can't absorb all the supply instantly"*
- This creates predictability: observe a large shock today ⇒ predict tomorrow's returns

ACF Pattern:

- *"If you have an MA(q), only the first q bars in the ACF would be non-zero — everything else is zero"*
- *"When you look at the ACF, you can immediately figure out whether you have MA(1), MA(2), MA(10)"*

CRITICAL: Professor's Exam Instructions for MA(1)

"You need to remember this. This is one of these things you must remember."

"In the multiple choice, I might give you an MA(1) process and ask: What is the conditional variance?"

"Only for MA(1). I'm going to ask the same for AR(1). That's all."

"I want to see 100% correct on these questions — lots of students failed these simple questions last year."

Exam Learning Objectives — You Must Be Able To:

1. **Calculate** unconditional mean, variance, and first-order autocorrelation for MA(1)
2. Recognize MA processes from ACF pattern (only first q autocorrelations non-zero)
3. Apply MA(1) formulas to numerical examples (multiple choice format)
4. Know that $\rho_j = 0$ for $j > 1$ in MA(1)

MA(1) FORMULAS — MUST MEMORISE

Model: $Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$, where $\varepsilon_t \sim \text{WN}(0, \sigma^2)$

$$\text{Mean:} \quad E(Y_t) = \mu \tag{1}$$

$$\text{Variance:} \quad \text{Var}(Y_t) = \sigma^2(1+\theta^2) \tag{2}$$

$$\text{Autocovariance:} \quad \gamma_1 = \theta\sigma^2 \tag{3}$$

$$\text{Autocorrelation:} \quad \rho_1 = \frac{\theta}{1+\theta^2} \tag{4}$$

$$\text{Higher lags:} \quad \rho_j = 0 \text{ for } j > 1 \tag{5}$$

Formulas Given on Exam (MA(q)):

- Mean: $E(Y_t) = \mu$
- Variance: $\text{Var}(Y_t) = \sigma^2\left(1 + \sum_{i=1}^{q} \theta_i^2\right)$

6 Autoregressive (AR) Processes (Pages 47–63)

Professor's Key Teaching Points

Stationarity Condition:

- *"As long as $|\phi| < 1$, shocks at some point will vanish, will die out"*
- *"If $\phi = 1$, each shock has a permanent effect — it's going to stay there forever"*
- *"That's exactly what happens in a random walk — shocks never vanish"*

Why You Must Remember the Conditional Mean:

- *"You have to remember this one. Why? Because whenever you work with simulation — and I'm sure if you get a job you're going to do lots of simulation, scenario analysis"*
- *"Often you assume your predictor follows an AR(1) process. You have to initialise the process. What is Y_0? Most of the time you set $Y_0 = \mu$ because you know the AR(1) process will converge to that mean"*

Interest Rates Example:

- *"If you fit an AR(1) on interest rates, ϕ will be 0.9999 — very close to one"*
- *"This is why people don't work with interest rates directly — they work with spreads or first differences"*

CRITICAL: Volatility vs Variance Warning

"For the exam, you don't need to remember the variance formula [for AR(1)]"

*"But I might ask: What is the conditional **volatility**? If I ask for volatility, I want volatility NOT variance!"*

"Last year, 50% of students got this wrong. The devil is in the details. Read the question carefully."

Remember: $\text{Volatility} = \sqrt{\text{Variance}}$

Exam Learning Objectives — You Must Be Able To:

1. **Calculate** unconditional mean for AR(1): $\mu = \frac{c}{1-\phi}$
2. **State** the stationarity condition: $|\phi| < 1$
3. Know autocorrelation pattern: $\rho_j = \phi^j$ (geometric decay)
4. Distinguish between stationary AR(1) ($|\phi| < 1$) and random walk ($\phi = 1$)
5. **Calculate volatility if given variance** (take square root!)
6. Know how to initialise simulations using the conditional mean

AR(1) FORMULAS — MUST MEMORISE

Model: $Y_t = c + \phi Y_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim \text{WN}(0, \sigma^2)$

$$\text{Stationarity:} \quad |\phi| < 1 \tag{6}$$

$$\text{Mean:} \quad E(Y_t) = \frac{c}{1-\phi} \quad \textbf{[CRITICAL for simulations!]} \tag{7}$$

$$\text{Autocorrelations:} \quad \rho_j = \phi^j \text{ for } j \geq 1 \tag{8}$$

Formulas Given on Exam:

- AR(1) Variance: $\text{Var}(Y_t) = \frac{\sigma^2}{1-\phi^2}$
- AR(1) Autocovariance: $\gamma_1 = \phi \cdot \frac{\sigma^2}{1-\phi^2}$
- Random Walk: $E(Y_t) = Y_0 + tc$, $\text{Var}(Y_t) = t\sigma^2$ (non-stationary!)

7 AR(p) Stationarity via Eigenvalues (Pages 59–65)

Professor's Key Teaching Points

- *"If you have an AR with multiple lags, you construct the Φ matrix, perform eigenvalue decomposition, and check the eigenvalues"*
- *"All eigenvalues in absolute terms must be less than one — otherwise you have an explosive process"*
- *"If you need to optimise a likelihood, you have to impose the stationarity condition"*
- *"You don't have to remember the details, but at least you have an insight"*

Exam Learning Objectives — You Must Be Able To:

1. Understand that AR(p) can be rewritten as VAR(1) in companion form
2. State the stationarity condition: all eigenvalues $|\lambda_i| < 1$
3. Apply quick rules: necessary condition $\sum \phi_i < 1$; sufficient condition $\sum |\phi_i| < 1$
4. Interpret what $|\lambda| < 1$, $|\lambda| = 1$, $|\lambda| > 1$ mean for shock propagation

Formulas Given on Exam:

- AR(p) Mean: $E(Y_t) = \frac{c}{1-\sum_{i=1}^{p} \phi_i}$
- Stationarity: All eigenvalues of companion matrix satisfy $|\lambda_i| < 1$

8 ACF/PACF Identification (Pages 71–74)

Professor's Key Teaching Points

- *"The ACF tells you if it's MA (cuts off) or AR (decays)"*
- *"The PACF tells you the number of AR lags"*
- *"For random walk, the ACF is always near one — very slow decay"*
- *"You make your data stationary, then plot ACF and PACF, then you can make a guess"*

Parsimony (Occam's Razor):

- *"If you have two models and they're equally good, go with the simplest one"*
- *"Pick the one that has the least number of lags — the simpler one is always better"*
- *"Complex models track data well historically but often perform poorly out-of-sample"*

Exam Learning Objectives — You Must Be Able To:

1. Identify process type (WN, AR, MA, ARMA) from ACF and PACF patterns
2. Know the Box-Jenkins procedure conceptually
3. Understand the principle of parsimony in model selection

Process	**ACF**	**PACF**
White Noise	All zero	All zero
MA(q)	Cuts off after lag q	Decays toward zero
AR(p)	Decays toward zero	Cuts off after lag p
Random Walk	Near 1, very slow decay	Large spike at lag 1
ARMA(p,q)	Both decay	Both decay

Formulas Given on Exam:

- AIC: $-2\ln(L) + 2K$ (choose lowest)
- BIC: $-2\ln(L) + K\ln(T)$ (choose lowest)
- Ljung-Box: $Q = T(T+2)\sum_{i=1}^{s} \frac{\rho_i^2}{T-i} \sim \chi_s^2$; rule of thumb $s = \ln(T)$

9 Spurious Regressions (Pages 84–87)

Professor's Key Teaching Points

The Problem:

- *"You find that something does not exist — it's a statistical mirage"*
- *"Y and X have no relationship whatsoever, but OLS tells you beta is large and significant, R^2 is very high"*
- *"You can't believe how many spurious regressions I have to go through when marking dissertations"*

Why It Happens:

- *"Random walks generate lots of local trends — by chance they move together"*
- *"You keep adding shocks, and you get sequences of shocks moving in the same direction by chance"*

How to Detect:

- *"Always check the stationarity of the residuals"*
- *"The regression is spurious if the residuals are non-stationary"*
- *"If you add the lagged Y as a regressor, the significance on X will disappear"*

Practical Advice:

- *"Unit root tests are not powerful — they often fail"*
- *"Most powerful approach: plot the residuals and eyeball them"*
- *"If residuals trend up or trend down, that's a random walk behaviour — spurious regression"*

Exam Learning Objectives — You Must Be Able To:

1. Understand why spurious regressions occur with non-stationary data
2. List the red flags: high R^2, significant $\hat{\beta}$, persistent residuals, results change when adding Y_{t-1}
3. Know to always check stationarity of residuals
4. Understand that visual inspection (plotting residuals) is often more powerful than formal tests

10 Cointegration (Pages 92–94)

Professor's Key Teaching Points

The Exception to Spurious Regression:

- *"When two variables are non-stationary but they move together in the long run, you can run regression with data as they are"*
- *"If cointegrated, check the residuals — residuals are stationary"*

Dog and Owner Analogy:

- *"I walk in a straight line through the park; my dog wanders around but always comes back to me"*
- *"The dog moves away from me, then converges back — that's cointegration"*
- *"This is exactly what happens with pairs trading — you wait for Y to move away from X, then bet it will converge back"*

Dynamic OLS:

- *"You can use OLS for the point estimate — it's super consistent"*
- *"But you can't use the standard 1.96 critical value for significance"*
- *"Run Dynamic OLS with leads and lags of Δx — then you can use normal distribution"*

Exam Learning Objectives — You Must Be Able To:

1. Define cointegration (both non-stationary, linear combination is stationary)
2. Distinguish between spurious regression (residuals non-stationary) and cointegration (residuals stationary)
3. Understand when regression with non-stationary variables is valid
4. Know that DOLS allows standard inference with cointegrated variables

Formulas Given on Exam:

Dynamic OLS: $y_t = \alpha + \beta x_t + \sum_{i=-p}^{p} \gamma_i \Delta x_{t-i} + \varepsilon_t$

11 Unit Root Tests (Pages 96–99)

Professor's Key Teaching Points

- *"Unit root tests have low power — you need a lot of samples to identify a stationary process"*
- *"The test is going to tell you more often than not that the variable is not stationary, even when it is"*
- *"They are not very nice — often people just assume cointegration based on theory"*
- *"Critical values change with constants and trends — use non-standard distributions"*

Exam Learning Objectives — You Must Be Able To:

1. State unit root test hypotheses: $H_0 : \phi = 1$ (non-stationary) vs $H_A : |\phi| < 1$ (stationary)
2. Know when to use Case I (constant only) vs Case II (constant + trend)
3. Recognize the limitations of unit root tests (low power)

12 Final Exam Checklist

Formulas You MUST Memorise

MA(1): $Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$

- $E(Y_t) = \mu$
- $\text{Var}(Y_t) = \sigma^2(1 + \theta^2)$
- $\rho_1 = \frac{\theta}{1+\theta^2}$, $\rho_j = 0$ for $j > 1$

AR(1): $Y_t = c + \phi Y_{t-1} + \varepsilon_t$

- Stationarity: $|\phi| < 1$
- $E(Y_t) = \frac{c}{1-\phi}$ [**Use to initialise simulations!**]
- $\rho_j = \phi^j$

Critical: Volatility $= \sqrt{\text{Variance}}$

Conceptual Understanding Required

- Weak vs Strong stationarity **(NOT nested — likely MCQ!)**
- White noise vs i.i.d. **(uncorrelated $\neq$ independent)**
- ACF/PACF patterns for AR vs MA vs ARMA
- Spurious regressions **(always check residual stationarity!)**
- Cointegration **(exception: residuals ARE stationary)**
- Why we work with returns/differences instead of prices
- How to initialise AR(1) simulations using $Y_0 = c/(1 - \phi)$

Professor's Final Warnings

- *"Read the question carefully!"* (especially volatility vs variance)
- *"I want to see 100% correct on MA(1) and AR(1) questions"*
- Practice applying formulas with numbers (multiple choice format)
- Plot residuals to check for spurious regression (more powerful than tests!)
- Remember: tutorials → coursework; slides → final exam