ECON4916 Topics in Applied Financial Econometrics

UNIVERSITY OF SYDNEY · ECONOMICS · LEVEL 4

Full guide + ask Sia about ECON4916

SIDE 1/2 HANDWRITTEN A4 PLANNER · conditions → formula → direction check **COPY THIS SIDE BY HAND** S2 2026 · final 50% · 2 hours

1 · Exam map

COURSE DATA · S2 2026

FINAL 50 | Q5 | A1 10 | TEST 25 | A2 10

ASSESSMENT	FORMAT	WT
Final exam	Multiple-choice + short-answer	50%
In-class quizzes	Multiple-choice or fill	5%
Assignment 1	Short/long-answer	10%
In-semester test	Paper-based test	25%
Assignment 2	Short/long-answer	10%

"In-person, closed-book, supervised 2hr exam (two-sided A4 handwritten notes allowed)"

ECMT6006/ECON4916 unit overview, Semester 2 2026

Final: written in the formal examination period, in person and supervised. **No published question count or section split:** prepare portable methods, not a guessed paper.

2 · What earns space

ALL CHAPTERS

Name the task first: moments need support; forecasts need $\mathcal{F}_t$; inference needs H_0; estimation needs an objective; selection needs a holdout rule.

PRIORITY	WRITE BESIDE THE RULE
1	Support / information set / null
2	Existence, stationarity, rank or orthogonality
3	Formula + sum, integral, max or min
4	Direction/unit check + scoped conclusion

COPY-SIDE INVENTORY IN PRIORITY ORDER

TASK	MINIMUM HANDWRITTEN ENTRY
Probability	support; $\Sigma p=1$ or $\int f=1$; E[X], E[X^2], Var
Dependence	Cov definition; independence ⇒ zero covariance only
Stationarity	finite second moment; constant μ; γ depends only on lag
Forecast	$\mathcal{F}_t$; E[future$\|\mathcal{F}_t$]; future shocks only set to zero
OLS	$\hat{\beta}-\beta$ decomposition; LLN, rank, E[xε]=0
Restriction	H_0:Rβ=r; W; q; scoped rejection sentence
AR(1)	$\|\phi\|<1$; μ, variance, ρ_k, h-step forecast
MA(1)	shared-shock γ_0, γ_1; zero after lag 1
ARMA	candidate-specific $\hat{\varepsilon}$`recursion + initial condition
Selection	criterion direction; aligned sample/horizon/loss; no leakage

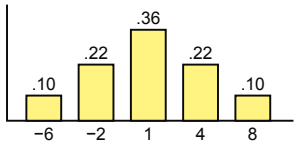
VISIBLE-WORKING SEQUENCE

1. **Name** the stochastic object, model or restriction.
2. **Authorise** the calculation with the required condition.
3. **Substitute** without hiding intermediate moments or residuals.
4. **Audit** sign, units, horizon and direction.
5. **Conclude** only under the maintained assumptions.

For multiple-choice, use the same sequence as elimination logic: reject options with the wrong operator, missing condition, impossible sign, incorrect unit or a conclusion broader than the calculation.

3 · Probability toolkit

CH 1



Discrete masses are nonnegative and sum to one across the complete support.

OBJECT	VALIDITY	USE
pmf p(x)	p≥0; Σp=1	sum support points
density f(x)	f≥0; ∫f=1	integrate interval
CDF F(a)	P(X≤a)	threshold probability

MOMENTS

```
μ = E[X] = Σx p(x) or ∫x f(x)dx
E[g(X)] = Σg(x)p(x) or ∫g(x)f(x)dx
Var(X) = E[X²] - μ²
Cov(X,Y) = E[(X-μₓ)(Y-μᵧ)]
```

Units: E[X] has X-units; Var(X) has squared units; covariance records *linear* comovement. A density height can exceed one because probability is area; for continuous X, P(X=x)=0.

3b · Dependence traps

CH 1 FAQ

CLAIM	VERDICT / REASON
independent ⇒ Cov=0	Yes, with finite moments
Cov=0 ⇒ independent	No; nonlinear dependence may remain
variance formula always valid	No; require finite second moment
observation = random variable	No; one realisation versus full mapping

If $Y=X^2$ and X is symmetric, covariance may be zero although Y is completely determined by X. Independence is a full-distribution claim; check joint/conditional versus marginal behaviour.

CONTINUOUS-VARIABLE DECISION LINES

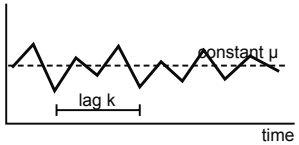
QUESTION WORDING	REQUIRED SETUP	CHECK
"equals x"	P(X=x)=0 for continuous X	point has zero width
"between a and b"	∫ from a to b of f(x)dx	shade only support overlap
"below a"	F(a)=P(X≤a)	CDF is nondecreasing
"expected g(X)"	integrate g(x)f(x), not g(E[X])	existence of the integral

For the course's uniform illustration on [−1,3], the support width is 4 and the density is 1/4 on that interval and zero elsewhere. The height is a density; only an interval's area is probability. Always intersect a requested interval with the stated support before integrating.

MOMENT INTERPRETATION BOUNDARIES

4 · Stationarity & forecasts

CH 2



One realised path is not proof: weak stationarity is a population moment claim.

WEAK STATIONARITY CHECKLIST

```
E[Xₜ]=μ constant; E[Xₜ²]<∞;
Cov(Xₜ,Xₜ₋ₖ)=γₖ depends on lag k, not t.
ρₖ = γₖ/γ₀; ρ₀=1; ρₖ=ρ₋ₖ; |ρₖ|≤1.
```

CONCEPT	LICENCES	DOES NOT MEAN
Strict	joint distributions shift-invariant	flat plot
Weak	stable first/second moments	sample convergence
Ergodicity	time averages learn moments	same as stationarity
Sample ACF	noisy pattern clue	model proof

SQUARED-LOSS FORECAST

```
X̂ₜ₊₁|ₜ = E[Xₜ₊₁ | Fₜ]. MSE = irreducible
conditional variance + squared distance
from conditional mean.
```

List $\mathcal{F}_t$. Keep recovered/current shocks known at t; replace only future zero-mean innovations by their conditional means. State: "optimal under squared loss, given $\mathcal{F}_t$".

4b · MA covariance method

CH 2 WORKED

STEP	ACTION
1	Write X_t and X_{t-k} side by side
2	Circle shared innovations
3	Multiply their coefficients × σ^2
4	No shared shock ⇒ covariance 0

POPULATION VERSUS SAMPLE

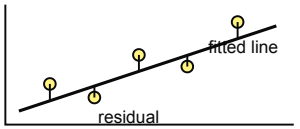
OBJECT	STATUS	SAFE STATEMENT
γ_k, ρ_k	population quantities	model implication at lag k
sample mean	one-path estimator	learns μ only with appropriate convergence
sample ACF	finite-sample estimator	pattern disturbed by sampling noise
residual ACF	post-fit diagnostic	remaining linear dependence check

Long-lag sample correlations can be especially noisy because fewer overlapping pairs contribute. Use the expected population pattern to generate candidates, then inspect residuals rather than declaring a model from one spike.

WHY THE CONDITIONAL MEAN WINS

5 · OLS & consistency

CH 3



OLS chooses $\hat{\beta}$ to minimise total squared vertical residual lengths.

MODEL, FIT, ERROR DECOMPOSITION

```
y = Xβ + ε; β̂ = (X'X)⁻¹X'y
β̂-β = (X'X)⁻¹X'ε
= [(X'X)/T]⁻¹[(X'ε)/T]
```

CONDITION	ROLE / FAILURE LOCATION
Time-series LLN	sample moments reach population moments
E[$x_t\varepsilon_t$]=0	orthogonality makes second factor vanish
E[x_tx_t'] nonsingular	rank identifies β; inverse exists

Then $\hat{\beta}-\beta \to_p [E(x_tx_t')]^{-1}E(x_t\varepsilon_t)=0$. Robust standard errors do not repair an inconsistent coefficient. Consistency gives convergence, not a rate or testing law; asymptotic normality supplies the $\sqrt{T}$ rate and approximate normal reference.

5b · Linear restrictions

CH 3

NULL AND WALD STATISTIC

```
H₀: Rβ = r; q = independent rows of R
W=(Rβ̂-r)'[R V̂(β̂)R']⁻¹(Rβ̂-r)
Under H₀, W has asymptotic χ²(q)
reference.
```

CLAIM	R	r
$\beta_1=\beta_2$	[0,1,−1]	0
$\beta_1=0$; $\beta_2=0$	[0,1,0]; [0,0,1]	[0;0]

Read rows back into words. Rejection says the restriction discrepancy is large relative to estimated uncertainty under the maintained null/model; it does not establish causality, importance, forecast value or the whole model.

KEEP FOUR CLAIMS SEPARATE

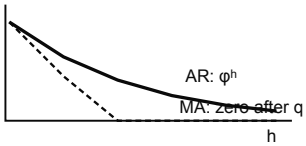
QUESTION	EVIDENCE NEEDED	DO NOT INFER
Is $\hat{\beta}$ consistent?	LLN, rank, E[xε]=0	finite-sample unbiasedness
How uncertain?	valid covariance estimate + asymptotic law	coefficient consistency
Restriction rejected?	Rβ−r scaled by its covariance	alternative theory proven
Economically important?	magnitude and units/context	from p-value alone
Useful forecast?	aligned later-data loss	from in-sample significance

RESTRICTION CONSTRUCTION ROUTINE

1. Write β in its declared coefficient order.

6 · AR/MA dynamics

CH 4



Stationary AR memory decays geometrically; finite MA memory ends exactly.

AR(1): $Y_t=C+\Phi Y_{t-1}+E_t$

```
|ψ|<1; μ=c/(1-ψ)
Var(Yₜ)=σ²ε/(1-ψ²); ρₖ=ψᵏ
Eₜ[Yₜ₊₁]=c+ψYₜ
Eₜ[Yₜ₊ₕ]=μ+ψʰ(Yₜ-μ)
```

Φ	SHOCK/FORECAST CHECK
0<ϕ<1	same sign; monotone reversion
−1<ϕ<0	alternating; oscillating reversion
|ϕ|≥1	no stationary-moment formulae

MA(1): $X_t=M+E_t+\Theta E_{t-1}$

```
γ₀=(1+θ²)σ²; γ₁=θσ²;
γₖ=0 for |k|>1; ρ₁=θ/(1+θ²).
```

FEATURE	AR	MA
regressor	lagged outcome	lagged innovation
shock memory	infinite decay	finite cutoff
population ACF	tails off	cuts off
restriction	stationarity	invertibility for recovery

Stationarity ≠ invertibility. A finite MA with finite-variance white noise is stationary; invertibility concerns unique stable innovation recovery.

6b · Diagnostic order

CH 4 FAQ

1. Use theory/ACF to propose candidates.
2. Estimate under explicit conditions.
3. Inspect residual dependence.
4. Validate on later data.

SHOCK-PATH COMPARISON

HORIZON H	AR(1), Φ=−0.5	MA(2), Θ_1=.6, Θ_2=−.2
0	1	1
1	−.5	.6
2	.25	−.2
3	−.125	0
later	alternates and shrinks	exactly 0

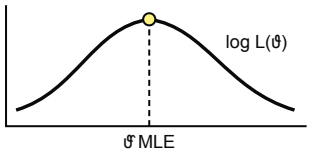
The AR path comes from repeated substitution: one innovation appears h periods later with coefficient ϕ^h. The MA path is read directly from the finite printed innovation window. "Small" is not "cut off".

FORECAST AND IDENTIFICATION DISTINCTIONS

DIRECTION CHECKS

7 · Likelihood & ARMA

CH 5



MLE is at the peak because the observed sample has greatest log-likelihood there.

ESTIMATOR	USES	MAIN CHECK
OLS	observable regression + squared residuals	rank, orthogonality, errors
Moments	sample analogue	identification + convergence
ML	joint/conditional density	model, maximum, objective

LIKELIHOOD LOGIC

```
L(ϑ)=product of model densities; ℓ(ϑ)=log
L(ϑ)=sum of log densities. Log is
increasing, so argmax is unchanged. Score
= ∂ℓ/∂ϑ; inspect curvature after score=0.
```

ARMA(1,1) GENERATED SHOCKS

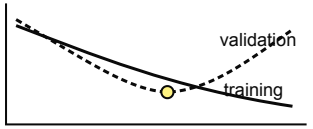
```
Yₜ=c+ψYₜ₋₁+εₜ+θεₜ₋₁
ε̂ₜ(ϑ)=Yₜ-c-ψYₜ₋₁-θε̂ₜ₋₁(ϑ)
```

For **every candidate 9**, select the stated initial condition, recover a new full residual path, then evaluate a comparable objective. Never reuse shocks from another candidate.

CONVERGENCE IS NOT PROOF	AUDIT
local optimum / flat region	alternative starts + nearby values
boundary solution	constraints + bounds
misspecification	moments + residuals + forecasts

8 · Selection & validation

CH 6



Training fit rewards flexibility; validation loss can rise once noise is fitted.

FORECAST-ERROR SPLIT

```
Conditional MSFE combines irreducible
future-shock variance with error from
estimating the conditional mean. Added
complexity changes the second component,
not the future innovation itself.
```

CHOICE	RULE
penalised fit	same sample; reward fit, charge complexity
holdout	same later dates, horizon, target and loss
recursive	growing window; old data retained

Forward this to your tutorial group

SHARED BY _____

Your exam allows one HANDWRITTEN A4 (both sides): copy side 1 by hand; do not bring this printout.

v2 · 2026-10-23 · Semester 2 formal examination period, from 16 November 2026 (the Exams Office publishes the exact date).

ECON4916 Topics in Applied Financial Econometrics

UNIVERSITY OF SYDNEY · ECONOMICS · LEVEL 4

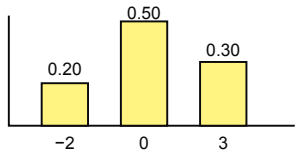
Full guide + ask Sia about ECON4916

SIDE 2/2 DRILLS · carry the numbers to a checked conclusion — **PRACTISE BEFORE THE EXAM** — 5 worked methods · traps · fast recall

9 · Drill 1: discrete moments

CH 1 WORKED

Given: X takes −2, 0, 3 percentage points with probabilities 0.20, 0.50, 0.30. Find validity, mean and variance.



All masses are nonnegative and 0.20+0.50+0.30=1.00.

X	P	XP	X²P
−2	.20	−.40	.80
0	.50	0	0
3	.30	.90	2.70
sum		**.50**	**3.50**

WORKING

```
E[X]=0.50 percentage points
E[X²]=3.50 squared percentage points
Var(X)=3.50-0.50²=3.25.
```

Conclusion: 0.50 is the probability-weighted repeated-draw centre; 3.25 is squared dispersion, not a possible realised return. The standard deviation restores percentage-point units.

TRAP *Validate support and masses first; subtract the squared mean; do not report variance in unsquared units.*

9b · Probability rapid checks

CH 1 FAQ

IF ASKED…	SHOW…
P(a<X<b), continuous	support + interval area, not point height
moment	existence + units
dependence	joint/conditional logic, not covariance alone
interpret mean	repeated-draw centre, not guaranteed outcome

Nonlinear warning: zero covariance removes linear association only. Independence must be assessed through the full joint distribution or conditional versus marginal probabilities.

REBUILD THE ANSWER, DO NOT PATTERN-MATCH

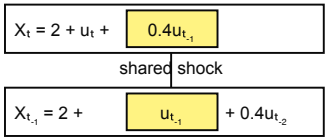
LINE	WHAT EARNS IT	FAILURE CAUGHT
support	{−2,0,3} and corresponding masses	probability on impossible values
validity	each mass ≥0; total exactly 1	incoherent distribution
first moment	show every xp contribution	missing sign or outcome

DEFINITION-TO-METHOD RECALL

10 · Drill 2: MA(1) stationarity

CH 2 WORKED

Given: $X_t=2+u_t+0.4u_{t-1}$, white noise $Var(u_t)=9$. Find μ, γ_0, γ_1 and γ_k for k≥2; establish weak stationarity.



Only u_{t-1} is shared; multiply its coefficients and innovation variance.

WORKING

```
E[X_t]=2.
γ₀=(1²+0.4²)9=10.44.
γ₁=(0.4)(1)(9)=3.6.
γ_k=0 for k≥2.
ρ₁=3.6/10.44≈0.345.
```

Conclusion: finite second moment, constant mean and covariance depending only on separation establish weak stationarity under the stated white-noise assumptions.

WRONG MOVE	REPAIR
"looks stable"	derive mean, variance, lag covariance
add γ_t terms	shared-shock coefficient products only
stationary ⇒ ergodic	separate dependence requirement

10b · Forecast logic

CH 2

AT FORECAST ORIGIN T	TREATMENT
observed/recovered shock in $\mathcal{F}_t$	keep it
future mean-zero shock	replace by its conditional mean, zero

STATIONARITY PROOF CHECKLIST

CONDITION	CALCULATION IN THIS MA(1)	CONCLUSION
finite second moment	white-noise variance 9; finite coefficient squares	γ_0=10.44 exists
constant mean	$E[u_t]=E[u_{t-1}]=0$	$E[X_t]=2$ for every t
lag-only covariance	overlap depends on separation, not date	$\gamma_t=3.6$; $\gamma_k=0$ for k≥2

The conclusion is *weakly stationary under the stated white-noise assumptions*. It is not proved by the visual appearance of one path, and it does not by itself show that a sample mean converges to the population mean.

ACF ARITHMETIC AND DIAGNOSTICS

NORMALISATION CHECK

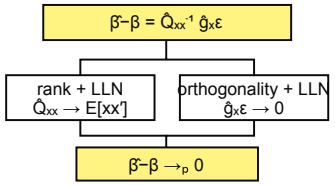
```
ρ₁=γ₁/γ₀=3.6/10.44≈0.345; ρ₀=1; ρ_k=0
for k≥2 in this population MA(1).
```

FORECAST WORDING

11 · Drill 3: OLS consistency

CH 3 WORKED

Given: a time-series LLN, $E[x_t\varepsilon_t]=0$ and nonsingular $E[x_tx_t']$. Explain why OLS is consistent.



Rank authorises the inverse; orthogonality kills the second factor.

1. Substitute $y=X\beta+\varepsilon$ into $\hat\beta-\beta=(X'X)^{-1}X'y$.
2. Obtain $\hat\beta-\beta=(X'X)^{-1}X'\varepsilon$.
3. Scale: $[(X'X)/T]^{-1}[(X'\varepsilon)/T]$.
4. LLN gives $E[x_tx_t']$ and $E[x_t\varepsilon_t]$.
5. Nonsingularity + orthogonality give $[E(xx')]^{-1}0=0$; hence $\hat\beta\to_p\beta$.

SYMPTOM	WHERE TO LOOK
collinearity	first matrix: rank
endogeneity	second moment: E[xε]
uncertain inference	covariance / CLT conditions

TRAP *More data does not remove $E[x_t\varepsilon_t]\neq0$. Standard-error changes do not fix coefficient inconsistency.*

11b · Wald translation

CH 3

WORDS	MATRIX LINE
β_1 equals β_2	[0,1,−1]β=0
β_1 and β_2 jointly zero	[0,1,0]β=0; [0,0,1]β=0

Count independent rows q; calculate W with Cov(Rβ̂); use $\chi^2(q)$ asymptotically; interpret the exact combination. A small p-value is neither $P(H_0$ true) nor an effect-size measure.

WHY EVERY CONDITION IS VISIBLE

REMOVED CONDITION	BREAK IN THE DERIVATION	DIAGNOSIS
time-series LLN	sample averages need not reach fixed limits	dependence/convergence problem
nonsingularity	$E[x_tx_t']$ cannot be inverted	insufficient independent regressor variation
$E[x_t\varepsilon_t]=0$	second factor tends to a nonzero vector	systematic regressor–error relation

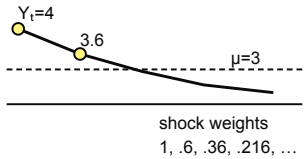
Slutsky combines the converging inverse moment matrix and the vanishing orthogonality moment. Attach "under the stated LLN, rank and orthogonality conditions" to the consistency conclusion.

CONSISTENCY VERSUS INFERENCE

12 · Drill 4: stationary AR(1)

CH 4 WORKED

Given: $Y_t=1.2+0.6Y_{t-1}+\varepsilon_t$, $Var(\varepsilon_t)=4$; $Y_1=4$. Find μ, variance, ρ_2 and the one-step forecast.



Positive $|\phi|<1$ keeps an above-mean forecast above μ while moving toward it.

WORKING

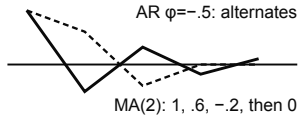
```
|0.6|<1 ⇒ stationary solution.
μ=1.2/(1-0.6)=3.
Var(Y)=4/(1-0.6²)=4/0.64=6.25.
ρ₂=0.6²=0.36.
E_t[Y_{t+1}]=1.2+0.6(4)=3.6.
```

Direction audit: 4>3 and 0<φ<1, so the forecast must lie between current value 4 and mean 3. The result 3.6 passes.

COMMON ERROR	REPAIR
μ=c/φ	μ=c+φμ ⇒ c/(1−φ)
$Var=\sigma^2/(1-\phi)$	variance recursion squares φ
$\rho_2=2\phi$	AR(1) ACF is ϕ^k
replace Y_t by μ	Y_t is observed in $\mathcal{F}_t$

12b · AR versus MA response

CH 4 COMPARE



Negative AR powers alternate forever while shrinking; finite MA reaches an exact cutoff.

GENERAL AR(1) MARKING ROUTE

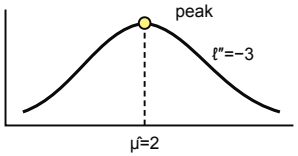
1. Check $|\phi|<1$ before using stationary moments.
2. Take unconditional means and solve μ=c+φμ.
3. Demean the recursion; use uncorrelated innovations to solve $\gamma_0=\phi^2\gamma_0+\sigma^2\varepsilon$.
4. Use $\rho_k=\phi^k$ for the population ACF.
5. For a forecast, condition on $\mathcal{F}_t$ and set only future innovations to zero.
6. Check direction against μ and the sign of φ.

FEATURE	STATIONARY AR(1)	FINITE MA(Q)
one-shock weight at h	ϕ^h	printed θ_h through q; then 0
positive coefficient	same-sign decay	sign follows each θ
negative coefficient	alternating decay	negative only where printed

13 · Drill 5: Gaussian MLE

CH 5 WORKED

Given: independent y=(1,2,3), normal with unknown μ and known variance 1. Derive $\hat\mu$ and verify a maximum.



Zero score occurs at the sample mean; negative curvature confirms a maximum.

LOG-LIKELIHOOD UP TO A CONSTANT

```
ℓ(μ)=-½Σ_t(y_t-μ)².
Score: ∂ℓ/∂μ=Σ_t(y_t-μ).
Set to 0: 1+2+3-3μ=0 ⇒ μ̂=2.
Curvature: ∂²ℓ/∂μ²=-3<0 ⇒ maximum.
```

STEP	EXAM LINE
objective	write what is maximised; omit only parameter-free constants
stationary point	score=0
maximum	negative second derivative / compare objective
ARMA	candidate ϑ generates candidate $\hat\varepsilon$ path first

TRAP *A zero score is not automatically a maximum. "Converged" is a stopping message, not statistical validation.*

13b · ARMA recovery

CH 5



repeat across sample

then evaluate $\ell(\vartheta)$ or CSS

change ϑ ⇒ rebuild every $\hat\varepsilon_t$

Initialisation and parameter values are part of the generated shock path.

AUDIT	QUESTION
objective	same likelihood/initialisation?
numerics	multiple starts, boundaries, nearby values?
model	moments and residual dependence plausible?
forecast	later aligned observations?

ESTIMATOR CHOICE BY OBSERVABLES

METHOD	SAMPLE TARGET	WHY RESULTS MAY DIFFER
OLS	minimum sum of squared observable regression residuals	uses conditional-mean/orthogonality structure

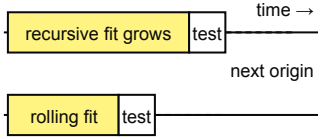
14 · Penalised-fit drill

CH 6 WORKED

Given: T=120. Models A/B/C have RSS 150/138/132 and k 2/5/9. Rank by s^2=RSS/(T−k), minimised.

MODEL	RSS	K	T−K	S^2
A	150	2	118	1.2712
B	138	5	115	1.2000
C	132	9	111	1.1892

Conclusion: C minimises this in-sample penalised residual variance. Its advantage over B is narrow; this does not establish best forecasts. Check residuals and time-ordered holdout loss.

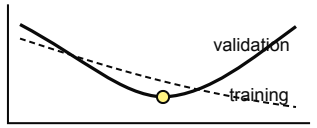


At each historical origin, fit only on available past data and score the later target.

FAIR-COMPARISON GATE	REQUIRED
information	no future scaling, features, tuning or ranking
sample	same effective forecast dates
target	same variable and horizon
criterion	same loss; direction stated
window	same rolling/recursive rule

15 · Fast method chooser

ALL CHAPTERS



choose by later loss

Extra flexibility helps fit but can worsen later performance through estimation error.

PROMPT CUE	METHOD / FIRST LINE
support, probabilities	pmf/density/CDF; normalise
time-invariant moments	weak-stationarity checklist
optimal MSE forecast	E[future | $\mathcal{F}_t$]
OLS consistency	$\hat\beta-\beta$ decomposition
coefficient claim	H_0: Rβ=r; Wald
shock persistence	AR powers / MA window
unobserved shocks	recursive recovery for each ϑ
fit versus forecast	penalty or ordered holdout

THIRTY-SECOND ANSWER AUDIT

1. Object and notation match?
2. Condition before formula?
3. Operator/direction correct?
4. $\mathcal{F}_t$, H_0 or window explicit?

Forward this to your tutorial group

SHARED BY _____

Your exam allows one HANDWRITTEN A4 (both sides): copy side 1 by hand; do not bring this printout.

v2 · 2026-10-23 · Semester 2 formal examination period, from 16 November 2026 (the Exams Office publishes the exact date).