ECON31031 Advanced Econometrics

THE UNIVERSITY OF MANCHESTER · LEVEL 3 · SEMESTER 1, 2026/27

Full guide + ask Sia about ECON31031

SIDE 1/2 HANDWRITTEN A4 PLANNER · priority formulas · decision rules · exam answer scaffolds **COPY THIS SIDE BY HAND** **Final 75%** · January 2027

0 · Exam map

COURSE.JSON · BRIEF

Coursework | Final Exam 75%

"Students will be allowed to bring one hand written A4 sheet of notes into the exam room." (ECON31031 course page, Canvas, 2026/27 1st semester)

Printed sheets are not allowed. Use side 1 as the plan for the permitted handwritten page; use both sides of your handwritten A4.

A · microdata: answer 1 of 2

B · macrodata: answer 2 of 3

Paper choice: one Section A microdata answer, two Section B macrodata answers.

PUT ON THE HANDWRITTEN PAGE	WHY
1. Time series	two of three answers
2. MLE + tests	shared machinery
3. Binary calcs	reliable Section A work
4. Matrix OLS	derivations + testing

SIA → *For every answer: set up → calculate/derive → interpret. State the distribution, degrees of freedom and conclusion for every test.*

0b · Scientific-calculator workflow

STUDY STRATEGY

1. Write the model and assumptions before arithmetic.
2. Keep full precision; round only the reported result.
3. Mark known values, forecasts and unobserved shocks distinctly.
4. End with one sentence answering the economic/statistical question.

THREE ANSWER SKELETONS

Derive: define the model/density and dimensions → state the assumptions → show the algebra using a named identity → box the estimator/result → say why it holds. **Calculate:** write the formula → substitute visible numbers → keep the denominator and units explicit → report a rounded value → interpret. **Test:** H_0/H_1 → statistic → reference distribution and df → critical value/decision → conclusion in the language of the question.

WHAT EARNS SPACE

KEEP	COMPRESS
conditions changing the method	introductory prose
denominators + df	software commands
sign/scale interpretation	repeated definitions
one numerical check per family	derivations you can rebuild

1 · Matrix OLS core

CH.1

MODEL + ESTIMATOR

$y = X\beta + \varepsilon$

$\hat\beta = (X'X)^{-1}X'y$

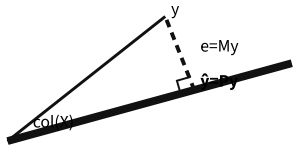
With n observations and k regressors plus intercept: y,ε are $n\times1$; X is $n\times(k+1)$; β is $(k+1)\times1$. Gauss–Markov gives BLUE under correct linear specification, $E(\varepsilon|X)=0$, no perfect collinearity and homoscedastic uncorrelated errors.

MEAN, VARIANCE, PRECISION

$E(\hat\beta|X)=\beta$

$Var(\hat\beta|X)=\sigma^2(X'X)^{-1}$

$s^2=e'e/(n-k-1)$



OLS projects y onto col(X); the residual My is orthogonal to every X column.

PROJECTION IDENTITIES

$P=X(X'X)^{-1}X';\ M=I-P$

$\hat y=Py;\ e=My=M\varepsilon;\ X'e=0$

$P'=P,\ P^2=P;\ M'=M,\ M^2=M;\ MX=0$

$tr(M)=n-k-1$

Where it bites: derive unbiasedness/variance, prove s^2 unbiased, or use $X'e=0$ to catch hand-arithmetic errors.

1b · Fit and restrictions

CH.1 FAQ

$R^2=1-RSS/TSS$

$F=[(RSS_r-RSS_u)/q] \div [RSS_u/(n-k-1)]$

QUESTION	STATISTIC	REFERENCE
one β restriction	$t=(\hat\beta_j-\beta_j^0)/se$	t_{n-k-1} if normal
q joint restrictions	F above	$F_{q,n-k-1}$ if normal
large sample	qF	$\chi^2 q$

Write restrictions as **Rβ=r**. Ordinary R^2 never falls when regressors enter; adjusted R^2 can fall because it accounts for degrees of freedom.

GAUSS–MARKOV LINE-BY-LINE

Use a random cross-section and a model linear in β; require full column rank, **E(ε|X)=0** and $Var(\varepsilon|X)=\sigma^2 I$. Then $\hat\beta$ is unbiased and BLUE. Normal errors are an extra condition for exact finite-sample t/F inference; without normality, invoke large-sample normal/χ^2 results.

RESTRICTION WORKFLOW

1. Translate words into each row of Rβ=r.
2. Estimate unrestricted model and record RSS_u.
3. Impose/substitute restrictions; estimate and record RSS_r.
4. Use q rows of R in F; RSS_r cannot be below RSS_u.

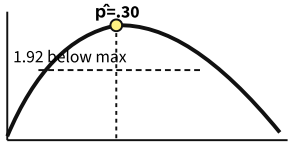
2 · Maximum likelihood recipe

CH.2

1. State one-observation density and support.
2. Use independence: $L(\theta)=\prod f_i(\theta)$.
3. Log: $\ell(\theta)=\Sigma \ln f_i(\theta)$.
4. Score: $s(\theta)=\partial\ell/\partial\theta$; solve $s(\hat\theta)=0$.
5. Curvature: $H(\theta)=\partial^2\ell/\partial\theta\partial\theta'$.
6. Test restrictions and interpret.

LARGE-SAMPLE COVARIANCE

$\widehat{Var}(\hat\theta)=[-H(\hat\theta)]^{-1}$



Six successes in twenty: ℓ peaks at .30; the LR acceptance band is .13–.52.

LIKELIHOOD-RATIO TEST

$LR=2[\ell(\hat\theta_u)-\ell(\hat\theta_r)] \sim \chi^2 q$

q = number of restrictions

Count restrictions, not parameters. A closed form exists only when the score equations solve algebraically; nonlinear index models normally require numerical optimisation.

2b · Two benchmark likelihoods

CH.2

BERNOULLI

$\ell(p)=\Sigma[y_i \ln p+(1-y_i)\ln(1-p)]$

$\hat p=\bar y;\ \widehat{Var}(\hat p)=\hat p(1-\hat p)/n$

NORMAL LINEAR MODEL

$\hat\beta ML=\hat\beta OLS$

$\hat\sigma^2 ML=RSS/n;\ s^2 OLS=RSS/(n-k-1)$

Both variance estimators are consistent; only the OLS version is unbiased in finite samples.

WHAT EACH DERIVATIVE SUPPLIES

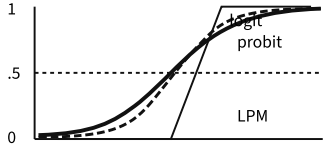
OBJECT	USE	EXAM LINE
$L(\theta)$	sample probability/density	product by IID
$\ell(\theta)$	workable objective	sum of log densities
score	first-order conditions	set every component to zero
Hessian	curvature/precision	invert its negative
ℓ_u,ℓ_r	joint test	$LR=2(\ell_u-\ell_r)$

For binary data, the intercept-only restricted model has fitted probability equal to the sample proportion. This gives ℓ_r directly and reappears in likelihood-based pseudo-R^2 measures. A sharp peak means larger negative curvature and a smaller covariance; a flat likelihood means imprecision.

3 · Binary model chooser

CH.3

MODEL	$P_i=PR(Y_i=1\|X_i)$	KEY EXAM POINT
LPM	$x_i'\beta$	easy OLS; invalid p possible; constant effects; robust SE
Logit	$\Lambda(z)=e^z/(1+e^z)$	log[p/(1−p)]=z; OR=e^β
Probit	$\Phi(z)$	normal table; latent threshold



LPM is linear and can leave [0,1]; logit/probit approach the bounds smoothly.

LPM: $Var(\varepsilon_i|x_i)=p_i(1-p_i)$, so heteroscedasticity is built in. White robust covariance repairs inference, *not* out-of-range predictions or constant effects.

LATENT PROBIT

$y_i^*=x_i'\beta+\varepsilon_i;\ y_i=1[y_i^*>0]$

$\varepsilon_i \sim N(0,1) \Rightarrow Pr(y_i=1)=\Phi(x_i'\beta)$

Only the threshold side is observed, so error variance must be fixed. This scale normalisation explains why logit and probit coefficients differ even when fitted probabilities are close.

3b · Probability, odds, interpretation

CH.3 FAQ

odds=p/(1-p)=e^z; p=odds/(1+odds)

one-unit x_j rise ⇒ odds × e^{β_j}

For a dummy, compare odds at $x_j=1$ and 0 holding other regressors fixed. Never multiply the probability by e^β.

GLM AND RESPONSE TYPE

The generalised-linear view changes the **link**: identity for LPM, logit for logistic probabilities, inverse-normal for probit. Binary outcomes have two unordered values; ordered responses rank categories without equal spacing; multinomial responses choose among several exclusive alternatives.

LIMITED DEPENDENT VARIABLES

OBSERVED RULE	MODEL IDEA
only which side of threshold	binary response
threshold recorded below limit	censoring
observations below limit absent	truncation

Probit by hand: calculate $z=x'\beta$, then read $\Phi(z)$. For $z<0$ use symmetry, $\Phi(-a)=1-\Phi(a)$. **Scale warning:** logistic error variance is $\pi^2/3$ while standard normal variance is 1; coefficients are not directly comparable. In the course example, logit coefficients are roughly 1.6–1.8 times probit coefficients, while probability effects are close.

4 · Marginal effects

CH.4

CONTINUOUS REGRESSOR

LPM: $ME_j=\hat\beta_j$

Logit: $ME_j=\hat p(1-\hat p)\hat\beta_j$

Probit: $ME_j=\psi(x'\hat\beta)\hat\beta_j$

DUMMY REGRESSOR

$\Delta p=F(x'\hat\beta;\ x_j=1)-F(x'\hat\beta;\ x_j=0)$

The density is positive, so effect sign matches coefficient sign. For logit, $p(1-p)$ peaks at **.25 when p=.5**; effects shrink near 0 or 1. Report probability changes in percentage points and scale them to the change asked.

SUMMARY	CALCULATION	SAY
effect at average	ME at $\bar x$	effect for mean profile
average effect	$n^{-1}\Sigma ME_i$	mean individual effect
dummy	exact $p(1)-p(0)$	not derivative approximation

Use the **delta method** to transform $Var(\hat\beta)$ into an effect's standard error. Odds ratios are logit-specific and constant across individuals; probability changes depend on the starting index.

5 · Binary-model assessment

CH.5

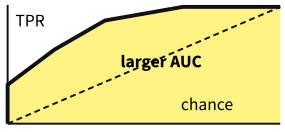
At cut-off c: predict 1 when $\hat p \geq c$. In the hit/miss table let $TP=n_{11}$, $FP=n_{10}$, $FN=n_{01}$, $TN=n_{00}$.

MEASURE	FORMULA
hit count	(TP+TN)/n
recall / TPR	TP/(TP+FN)
FPR	FP/(FP+TN)
Kuipers	TPR−FPR
precision	TP/(TP+FP)
F1	2·precision·recall/(precision+recall)

Brier=$n^{-1}\Sigma(y_i-\hat p_i)^2$ [lower better]

McFadden $R^2=1-\ell_u/\ell_r$

A high hit count misleads with imbalance. Prefer precision/recall/F1 and the **ROC**: plot TPR against FPR across cut-offs; AUC uses trapezia. Guide bands: .5–.7 poor; .7–.8 fair; .8–.9 good; >.9 excellent.



ROC varies the cut-off: move towards top-left for high TPR and low FPR.

PHI COEFFICIENT

$\varphi=(TP\cdot TN-FP\cdot FN)/\sqrt{(TP+FP)(FN+TN)(TP+FN)(FP+TN)}$

In-sample fitted values describe the estimation sample; **out-of-sample** predictions evaluate observations not used in estimation. Do not call in-sample fit evidence of forecast performance.

6 · Stationarity + operators

CH.6

$Ly_t=y_{t-1};\ L^h y_t=y_{t-h}$

$\Delta y_t=(1-L)y_t=y_t-y_{t-1}$

seasonal s: $\Delta_s y_t=(1-L^s)y_t$

Weak stationarity: $E(y_t)=\mu$ finite and constant; $Var(y_t)=\gamma_0$ finite and constant; $Cov(y_t,y_{t-k})=\gamma_k$ depends only on lag k. White noise has mean 0, constant variance σ^2 and zero autocorrelation for $k\neq0$; it need not be normal.

SAMPLE AUTOCORRELATION

$r_k=\Sigma_{t=k+1}^{T}(y_t-\bar y)(y_{t-k}-\bar y)\ /\ \Sigma_{t=1}^{T}(y_t-\bar y)^2$

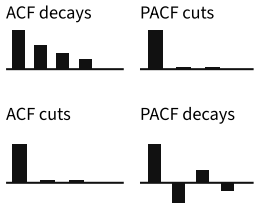
Same denominator at every lag. Slow ACF decay warns of strong persistence or nonstationarity; do not apply stationary formulas until diagnosed.

7 · AR/MA/ARMA identification

CH.7

AR(1) ACF decays PACF cuts

MA(1) ACF cuts PACF decays



Eyeball rule: AR cuts in the PACF; MA cuts in the ACF; mixed decay suggests ARMA.

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

$|\psi|<1;\ \mu=\alpha/(1-\psi)$

$Var(y)=\sigma^2/(1-\psi^2);\ \rho_k=\psi^k$

MA(Q): $Y_t=A+E_t+\Sigma\Theta_j E_{t-j}$

$E(y)=\alpha;\ Var(y)=\sigma^2(1+\Sigma\Theta_j^2)$

MA(1): $\rho_1=\theta/(1+\theta^2);\ \rho_k=0$ for $k>1$

AR(2) stationarity: $\phi_1+\phi_2<1;\ \phi_2-\phi_1<1;\ \phi_2>-1$. Equivalently check all characteristic roots have absolute value below one. MA terms need invertibility.

8 · Forecast, select, difference

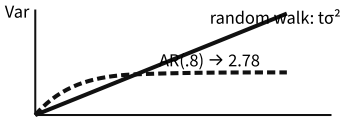
CH.8–9

AR(1): $\hat yT+h|T=\mu+\psi^h(yT-\mu)$

MA(1): h=1 uses $\theta\hat\varepsilon T$; h≥2 → μ

AIC=$-2\ell+2K$; BIC=$-2\ell+K \ln T$

Feed forecasts back when extrapolating; if actual out-of-sample values arrive, use them. Minimise AIC/BIC; BIC penalises dimension more heavily once T exceeds about seven.



With unit shock variance, stationary AR(.8) settles at 2.78; random-walk variance grows.

Two-hour on-campus invigilated examination in January 2027 (check the date in the university exam schedule).
Your exam allows one HANDWRITTEN A4 (both sides): copy side 1 by hand; do not bring this printout.

ECON31031 Advanced Econometrics · Worked Drills

FIVE SOURCE-GROUNDED EXAMPLES · TRAPS · FAST RECALL

Full guide + ask Sia about ECON31031

SIDE 2/2 DRILLS · complete calculations · common traps · method chooser

PRACTISE BEFORE THE EXAM

derive · calculate · interpret

W1 · Matrix OLS by hand

CH.1 WORKED EXAMPLE

Data: (x,y)=(0,1),(1,3),(2,4),(3,6). Estimate intercept/slope, s^2 and test H_0:$\beta_1=0$.

1 · CROSS-PRODUCTS

```
X'X=[[4,6],[6,14]], det=20
(X'X)⁻¹=(1/20)[[14,-6],[-6,4]]
X'y=(14,29)'
```

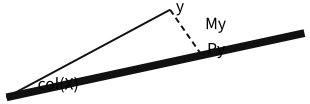
2 · ESTIMATES + RESIDUALS

```
β̂=(22/20,32/20)'=(1.1,1.6)'
ŷ=(1.1,2.7,4.3,5.9)'
e=(-.1,.3,-.3,.1)'
```

3 · UNCERTAINTY + TEST

```
e'e=.20; df=4-1-1=2; s²=.10
Var̂(β̂₁)=.10(4/20)=.02
se=.141; t=1.6/.141=11.3
```

Result: ŷ=1.1+1.6x; s^2=.1; t=11.3 on 2 df rejects a zero slope.



Hand-check geometry: fitted vector lies in col(X); residual is perpendicular.

CHECK → *$Xe=0$ and $X'e=0$ before using residuals. These catch a wrong inverse or multiplication.*

T1 · OLS traps

CH.1 FAQ

TRAP	CORRECTION
divide RSS by n	unbiased s^2 uses n−k−1
forget q in joint test	F numerator divides by q; large sample $qF\sim\chi^2 q$
restriction unclear	write Rβ=r before algebra
add regressor ⇒ fit "improves"	R^2 cannot fall; inspect adjusted R^2

RESTRICTION SELF-CHECK

Four coefficients β_1 (intercept) to β_4; impose $\beta_2=2\beta_3$ and $\beta_1+\beta_4=0$:

```
R=[[0,1,-2,0],[1,0,0,1]]; r=(0,0)'
β₂=2β₃; β₄=-β₁
y=β₁(1-x₄)+β₃(2x₂+x₃)+ε
```

Estimate the transformed equation with two constructed regressors. Two independent rows of R mean **q=2**. If restrictions already hold in the sample, imposing them leaves estimates/RSS unchanged and F=0.

RESIDUAL-MAKER DERIVATION DRILL

```
e=My=M(Xβ+ε)=Mε because MX=0
E(e'e|X)=E[tr(Mεε')|X]
=tr[M·σ²I]=σ²tr(M)
```

TWO-BY-TWO INVERSE CARD

TEST ANSWER CHECKLIST

W2 · Bernoulli MLE + LR

CH.2 WORKED EXAMPLE

Data: 20 successes in 80 independent trials. Estimate p; test H_0:p=.5 at 5%.

1 · LOG-LIKELIHOOD + SCORE

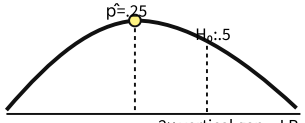
```
ℓ(p)=20 ln p+60 ln(1-p)
s(p)=20/p-60/(1-p)=0
p̂=20/80=.25
```

2 · TWO MAXIMA

```
ℓᵤ=20 ln .25+60 ln .75=-44.99
ℓᵣ=80 ln .5=-55.45
```

3 · LR

```
LR=2(-44.99+55.45)=20.93
χ²₁,.95=3.841 ⇒ reject H₀
```



LR doubles the log-likelihood gap between unrestricted p̂ and restricted p=.5.

Verdict sentence: "LR=20.93 exceeds 3.841, so reject p=.5 at 5%."

T2 · Likelihood traps

CH.2

- Write the **one-observation density** first; support errors poison every later line.
- LR is $2(\ell_u-\ell_r)$, never reverse the signs.
- Use q restrictions for χ^2 df.
- At a maximum H is negative; covariance uses **−H inverse**.
- ML normal variance uses n; OLS unbiased variance uses n−k−1.

BERNOULLI CURVATURE CHECK

```
s(p)=Σ[yᵢ/p-(1-yᵢ)/(1-p)]
H(p)=-Σ[yᵢ/p²+(1-yᵢ)/(1-p)²]<0
se(p̂)=√[p̂(1-p̂)/n]
```

For W2, se=√[.25(.75)/80]=.0484. The negative Hessian confirms a maximum; the inverse curvature supplies the same large-sample variance.

LR ANSWER FRAME

1. Fit/unpack the unrestricted model and record ℓ_u.
2. Impose H_0, re-maximise and record ℓ_r.
3. Compute $2(\ell_u-\ell_r)$; it must be non-negative.
4. Use $\chi^2 q$, where q counts imposed restrictions.
5. Compare, decide and interpret the restrictions jointly.

Closed form vs numerical: Bernoulli $\hat{p}=\bar{y}$ and normal-regression $\hat{\beta}$ have algebraic solutions. Logit, probit and ARMA parameters generally sit inside nonlinear sums, so say "solve numerically"; do not pretend the score isolates β.

W3 · Logit probability + odds

COURSE WORKED EXAMPLE

Model: z=−.8+.06(study hours)+.7(course). At 15 hours, compare course=1 vs 0.

1 · INDICES

```
z₁=-.8+.06(15)+.7=.8
z₀=-.8+.06(15)=.1
```

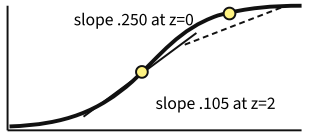
2 · PROBABILITIES

```
p₁=e^.8/(1+e^.8)=2.2255/3.2255=.690
p₀=e^.1/(1+e^.1)=1.1052/2.1052=.525
Δp=.165 = 16.5 percentage points
```

3 · ODDS RATIO

```
odds₁=2.2255; odds₀=1.1052
OR=2.2255/1.1052=2.014=e^.7
```

The course doubles the odds for every profile, but its probability effect depends on the starting index.



Same coefficient, different probability effect: the logistic density shrinks in the tails.

D3 · Probit effects drill

CH.4 WORKED EXAMPLE

At z=0, βcontinuous=.25, βdummy=.5, ϕ(0)=.3989, Φ(.5)=.6915:

```
continuous ME=.3989(.25)=.0997
exact dummy=Φ(.5)-Φ(0)=.1915
derivative approx=.3989(.5)=.1995
```

Use the exact probability difference for a dummy. Here the derivative slightly overstates it because the curve flattens from 0 to .5.

ODDS-UPDATE SELF-CHECK

Starting logit probability .8 gives odds .8/.2=**4**. A dummy coefficient .5 multiplies odds by $e^{.5}$=1.6487, giving **6.595**; convert back: 6.595/7.595=**.868**. The probability rises 6.8 points, not by 64.87%.

EFFECT REPORTING GRID

INPUT	ACTION	INTERPRETATION
continuous x	density × β	p change per unit x
dummy x	F(z₁)−F(z₀)	percentage-point switch
logit β	e^{β}	odds multiplier
change Δx	ME × Δx	scale to question

Scale check: a probit effect −.065 per unit becomes −.0065 for a .1 rise—about two-thirds of a percentage point. Keep "probability", "percentage points" and "odds" distinct.

MODEL TRAPS

- Robust LPM SEs repair heteroscedastic inference, not invalid fitted probabilities.

W4 · Classification + AUC

CH.5 WORKED MATERIAL

At cut-off .5, the guide's twelve-person table has TP=5, FP=2, FN=1, TN=4.

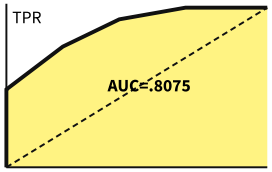
	observed 1	observed 0	
pred 1	TP 5	FP 2	TPR=.833
pred 0	FN 1	TN 4	FPR=.333
			Prec=.714
			F1=.769

Confusion-matrix denominators: TPR uses actual 1s; precision uses predicted 1s.

```
hit=(5+4)/12=.750
Kuipers=5/6-2/6=.500
precision=5/7=.714; recall=5/6=.833
F1=2(5)/(10+2+1)=.769
```

AUC drill points: (0,0),(.1,.5),(.3,.8),(.6,.95),(1,1).

```
.1(.25)+.2(.65)+.3(.875)+.4(.975)
=.025+.130+.2625+.390=.8075 ⇒ good
```



Sort by FPR, then sum width × average endpoint height for each trapezium.

T4 · Assessment traps

CH.5 FAQ

- One cut-off is one ROC point, not the whole curve.
- Lower Brier is better; applying it to 0/1 predictions gives 1−hit count.
- With rare successes, accuracy can look high while separation is poor.
- Link test: add η^2 and η^3; significant powers signal misspecification but do not identify its cause.

PHI AND DENOMINATOR AUDIT

```
φ=(5·4-2·1)/√[(7)(5)(6)(6)]
=18/√1260=.507
```

MEASURE	CONDITION ON	QUESTION ANSWERED
recall	actual success	how many successes found?
precision	predicted success	how many alerts correct?
FPR	actual failure	how many false alarms?
Kuipers	TPR−FPR	skill beyond false alarms?

Threshold direction: raising the cut-off usually predicts fewer successes: recall/TPR falls, FPR falls and precision often rises. The ROC traces this trade-off without committing to one cut-off.

W5 · Time-series chain

CH.6–9 WORKED EXAMPLES

A · SAMPLE AUTOCORRELATION

y=(2,4,3,5,1), $\bar{y}$=3, deviations=(−1,1,0,2,−2); denominator=10. Neighbour products sum −5:

```
Δy=(2,-1,2,-4); r₁=-5/10=-.5
```

B · AR(1) AND MA(1) MOMENTS

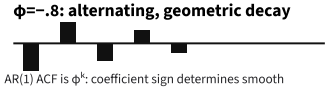
AR: $y_t=3+.5y_{t-1}+\varepsilon_t$, Var(ε)=3. MA coefficient θ=−.4.

```
μAR=3/(1-.5)=6
VarAR=3/(1-.25)=4
ρ₂=.5²=.25
ρ₁,MA=-.4/(1+.16)=-.345; ρk>1=0
```

ϕ=.8: same sign, geometric decay



ϕ=−.8: alternating, geometric decay



AR(1) ACF is ϕ^k: coefficient sign determines smooth persistence or oscillation.

C · AR FORECAST RECURSION

$y_t=1+.8y_{t-1}+\varepsilon_t$, last y=10:

```
ŷ₁=1+.8(10)=9.0
ŷ₂=1+.8(9)=8.2
ŷ₃=1+.8(8.2)=7.56
μ=1/(1-.8)=5
```

When extrapolating, forecast feeds forecast. Never use an unobserved value as though it were data.

D · IDENTIFICATION AND STATIONARITY CHECKS

CANDIDATE	ACF	PACF
AR(p)	decays	cuts after p
MA(q)	cuts after q	decays
ARMA(p,q)	decays	decays
unit-root warning	very slow decay	do not force stationary rule

MA logic: autocovariances vanish beyond q because no shocks are shared. Invertibility rewrites the shock polynomial as an autoregression in observed y.

E · OPERATOR MICRO-DRILLS

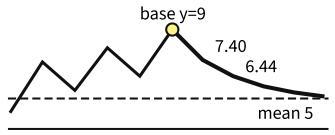
```
trend: yₜ=2t+εₜ ⇒ not stationary
since E(yₜ)=2t
Δyₜ=2+εₜ-εₜ₋₁ ⇒ E=2, Var=2σ²
quarterly annual change:
(1-L⁴)yₜ=yₜ-yₜ₋₄
```

Stationarity answers must name the failing condition. A plot or slow correlogram is evidence to investigate; a time-dependent mean/variance or a unit root is the formal demonstration.

F · MA AND MIXED FORECASTS

W5 continued · selection + unit root

CH.8–9



Source forecast path: α=2, ϕ=.6, yT=9; forecasts converge from 7.40 to 5.

D · INFORMATION CRITERIA

T=100; maximised ℓ: ARMA(1,0) −142.3; (2,0) −140.1; (1,1) −139.0; (2,2) −136.9. Count K=p+q+1.

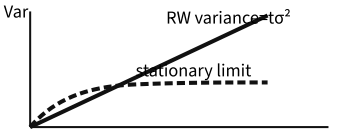
MODEL	K	AIC	BIC
(1,0)	2	288.6	293.81
(2,0)	3	286.2	294.02
(1,1)	3	284.0	**291.82**
(2,2)	5	**283.8**	296.83

AIC chooses (2,2); BIC chooses (1,1). Circle the minimum and explain BIC's heavier penalty.

E · RANDOM WALK WITH DRIFT

$y_t=.5+y_{t-1}+\varepsilon_t$; y_0=10; Var(ε)=2. At t=8:

```
y₈=10+8(.5)+Σε ⇒ E(y₈)=14
Var(y₈)=8(2)=16
Δyₜ=.5+εₜ ⇒ E(Δy)=.5, Var(Δy)=2
```



A random walk's growing variance violates stationarity; first differencing removes the unit root.

M · 20-second method chooser

FAST RECALL

IF ASKED…	WRITE FIRST
derive estimator	density/model + dimensions
test restrictions	H_0, q, restricted/unrestricted
binary effect	index z, then p, then effect
assess classifier	TP/FP/FN/TN with denominators
identify AR/MA	ACF vs PACF cut-off/decay
forecast	base, horizon, data vs forecast lags
nonstationarity	moment depending on t or unit root; then difference

UNIT-ROOT ANSWER PLAN

1. State weak-stationarity conditions.
2. Back-substitute or solve characteristic equation.
3. Show the mean/variance depends on t or a root equals one.

INFORMATION-CRITERION AUDIT

Two-hour on-campus invigilated examination in January 2027 (check the date in the university exam schedule).
Your exam allows one HANDWRITTEN A4 (both sides): copy side 1 by hand; do not bring this printout.