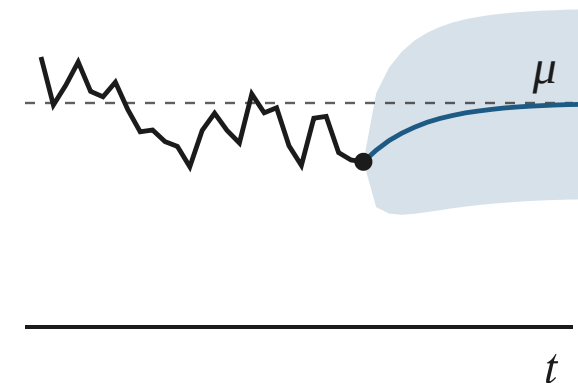


Time series

When today depends on yesterday: measure change, model persistence, and know when a pattern misleads

1 Describe → 2 Model → 3 Interpret



Lecture 1

Probability theory

What can happen, and how does information change it?

Lecture 2

Data, variables & moments

How do measurements vary?

Lecture 3

Statistics

What can one sample tell us about a population?

Lecture 4

Time series

What changes when observations depend on their past?

Lecture 5

Causal inference

Which comparison identifies an effect?

OPENING QUESTION

Two simulated series correlate at 0.71. Did the generator link them?

 X_t, Y_t

two series, generated by a computer

0.71

sample correlation of their levels across the paired dates

?

how each series was generated

YOUR ANSWER

Vote now. Part 3 answers it.

A

Yes: 0.71 is too strong to be chance

B

Yes, if the sample is long

C

The coefficient alone cannot tell

WHY START HERE

Prices, wages, money and output all drift over time. Before reading a link into a strong correlation between two such series, ask how each one moves on its own.

LECTURE 4 · THE ROUTE

Describe the data, model their memory, then interpret their patterns

PART 1

Order, units and change

What does a series record, and how do we measure its change without losing the dates and units?

time series

autocorrelation

difference

growth factor

log change

Notes 4.1–4.2

PART 2

Models with memory

What must stay unchanged for the past to inform the future, and does a shock fade or last?

stationarity

white noise

random walk

AR(1)

forecast band

Notes 4.3–4.5

PART 3

Reversion and spurious correlation

When does movement towards the mean, or a strong correlation between two series, reveal a real mechanism?

mean reversion

regression to the mean

spurious regression

cointegration

Notes 4.6–4.7

BY THE END OF THIS LECTURE YOU CAN

- ✓ Compute differences, growth factors, average growth and log changes, keeping the units and dates.
- ✓ Derive the mean $c/(1 - \phi)$, variance and autocorrelation ϕ^h of a stable AR(1).
- ✓ Forecast an AR(1) with a 95% interval, and explain why its band levels off but a random walk's does not.
- ✓ Explain why regression to the mean, and strong correlations between random walks, need no mechanism.

PART 1 OF 3

Order, units and change

What does a series record, and how do we measure its change without losing the dates and units?

KEY TERMS

time series

autocorrelation

difference

growth factor

log change

[Notes 4.1–4.2 ↗](#)

WHY IT MATTERS

Economic statistics arrive as changes: quarter on quarter, year on year, annualised or in log points. Equal absolute and equal percentage changes answer different questions, and miscompounding, such as multiplying a quarterly rate by four, misstates growth.

IN THIS PART

4.1 Time series, processes and the layers of a plot

4.1 Autocorrelation and the correlogram

▶ Animation: why a 20% rise and a 20% fall do not cancel

4.2 Differences, growth factors and log changes

4.1 | ORDER IS INFORMATION

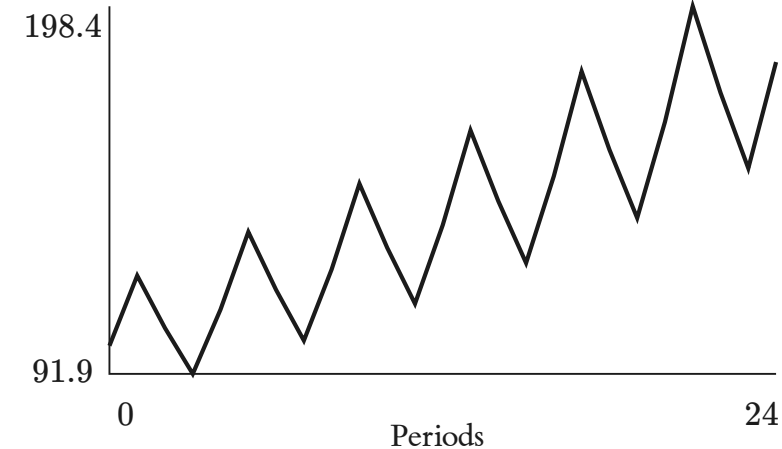
The date belongs to the observation: read a series in layers

DEFINITION

Time series and process

A time series records a variable in chronological order. Its probability model is a *stochastic process* $\{X_t\}$; we observe one path $x_1, \dots, x_T$. At date t , the value at $t + 1$ is not yet known.

Trend	a sustained change in level
Seasonality	repeats at a known calendar frequency
Cycle	rises and falls with no fixed period
Break	a change in mean, slope or volatility at a date



Hypothetical: $X_t = 100 \exp(0.025t + 0.16 \sin(\pi t/2))$

CHECK

Every December is high, and the December peak rises from year to year. Which two features are present?

ANSWER

Seasonality (the timing) and a trend (the change across years). They can coexist.

4.1 | DEPENDENCE ACROSS DATES

Autocorrelation compares a series with its own past

KEY FORMULA

DEFINITION

(1)

Autocorrelation at lag h

$$\rho_t(h) = \frac{\text{Cov}(X_t, X_{t-h})}{\sqrt{\text{Var}(X_t) \text{Var}(X_{t-h})}}$$

IN WORDS $\text{Corr}(X_t, X_{t-h})$: how today moves, linearly, with the value h periods earlier. Write $\rho(h)$ when it depends on the lag alone.

NEEDS finite, positive variances.

RESULT *reading a sample correlogram*

- Lag h leaves only $T - h$ pairs.
- Slow decay can mean persistence, or a trending level.
- Peaks at lags 4, 8, 12 suggest quarterly seasonality.

COMMON MISTAKE

Small sample autocorrelations do not prove independence, nor rule out nonlinear predictability.

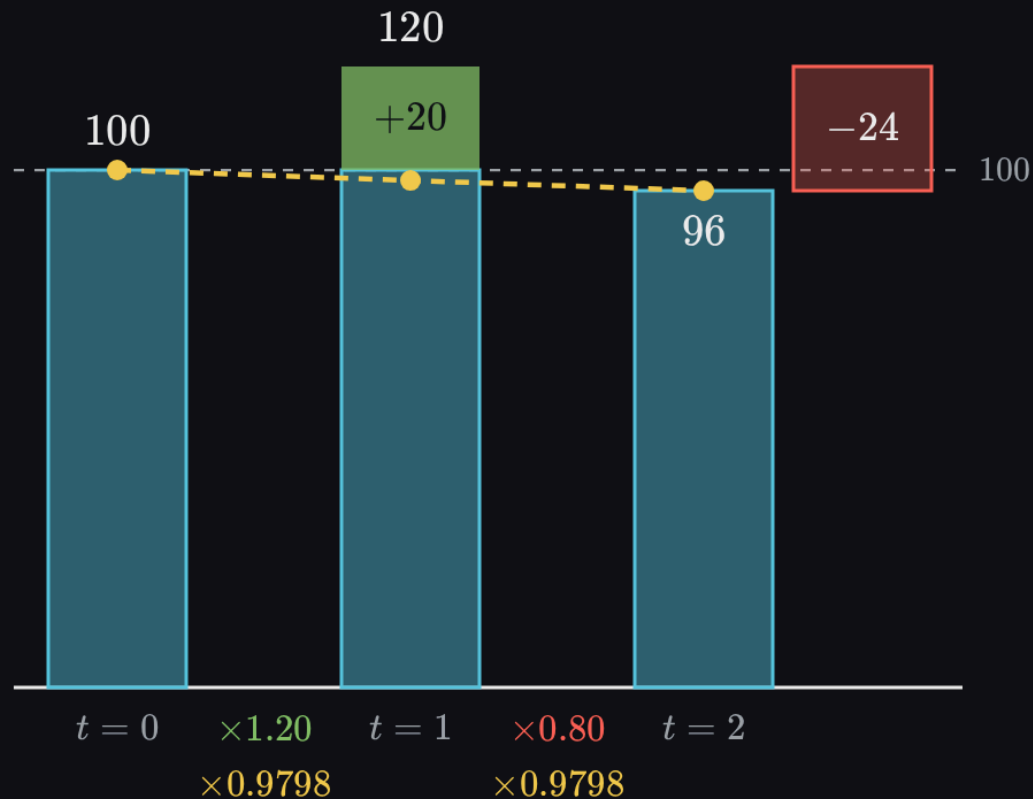
CHECK

Monthly sales are sorted from low to high, the dates deleted and one month is missing. Does a lag-one correlation still measure monthly persistence?

ANSWER

No: sorted neighbours are close by construction, not in time, and a missing month is not a zero. Restore the calendar, then form lag pairs.

Up 20%, then down 20%: back to 100?



start: $X_0 = 100$

$$100 \times 1.20 = 120$$

the rise is 20% of 100: +20

$$120 \times 0.80 = 96$$

the fall is 20% of 120: -24

overall: $96/100 - 1 = -4\%$

$$(1 + \bar{g})^2 = 0.96: \bar{g} = \sqrt{0.96} - 1 \approx -2.02\%$$

not the simple average $(20\% - 20\%)/2 = 0$

4.2 | MEASURING CHANGE

Differences subtract, growth factors multiply, logs add

DEFINITION

Differences

$\Delta X_t = X_t - X_{t-1}$: an absolute change, in units of X per period. $\Delta_4 X_t = X_t - X_{t-4}$ compares matching quarters; it is not $\Delta^4 X_t$.

KEY FORMULA DEFINITION

(2)

Growth factors compound

$$(1 + \bar{g})^h = \frac{X_t}{X_{t-h}}$$

IN WORDS with rate $g_t = X_t/X_{t-1} - 1$, the **growth factor** $1 + g_t$ multiplies the level; $\bar{g}$ compounds to the observed ratio.

NEEDS positive levels.

KEY FORMULA IDENTITY

(3)

Log change

$$r_t = \Delta \log X_t = \log(1 + g_t)$$

IN WORDS the log of the growth factor; log changes add over time. Close to g_t only for small rates: 0.20 is growth of $e^{0.20} - 1 \approx 22.14\%$.

NEEDS $X_t > 0$; natural logarithms.

WORKED EXAMPLE *compounding*

Up 20%, down 20%: $(1 + \bar{g})^2 = 0.96$, so $\bar{g} \approx -2.02\%$ per period, not 0. And 2% a quarter is $(1.02)^4 - 1 \approx 8.243\%$ a year, not 8%.

PART 2 OF 3

Models with memory

What must stay unchanged for the past to inform the future, and does a shock fade or last?

KEY TERMS

stationarity

white noise

random walk

AR(1)

forecast band

[Notes 4.3–4.5 ↗](#)

SO FAR a series is one path of a process; autocorrelation measures how it depends on its own past.

WHY IT MATTERS

Any forecast of unemployment, earnings or an exchange rate rests on a model of persistence. Whether a shock is temporary or permanent decides where the forecast goes, and whether its uncertainty stays bounded.

IN THIS PART

- 4.3 Weak stationarity: what must stay unchanged
- 4.4 White noise and the random walk
- 4.3 Choosing a transformation from the model
- 4.4 The AR(1) and its special cases
- 4.5 Mean, variance and autocorrelation of a stable AR(1)
- 4.5 AR(1) forecasts and their uncertainty
- Animation: the forecast returns towards the mean inside a band that levels off

4.3 | WHAT MUST STAY UNCHANGED

Stationarity: the same probability structure at every date

KEY FORMULA

DEFINITION

(4)

Weak stationarity

$$E[X_t] = \mu$$

$$\text{Var}(X_t) = \gamma(0) < \infty$$

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

IN WORDS the same mean and finite variance at every date, and a covariance that depends only on the gap h . Then $\rho(h) = \gamma(h)/\gamma(0)$.

NEEDS finite second moments; it restricts moments, not the whole distribution.

COMMON MISTAKE

Stationary is not flat-looking: long runs above the mean are allowed. Nor is zero autocorrelation enough: $X_t \sim N(0, t)$ is uncorrelated but not stationary.

CHECK

With Z_t iid $N(0, 1)$, which of $3 + Z_t$, $t + Z_t$ and $\sqrt{t} Z_t$ is weakly stationary?

ANSWER

Only $3 + Z_t$: mean 3, variance 1. $t + Z_t$ has mean t ; $\sqrt{t} Z_t$ has variance t .

Stationarity lets one stretch of history inform another: the role that iid sampling played in Lecture 3.

4.4 | TWO BUILDING BLOCKS

White noise is the building block; a random walk adds the shocks up

DEFINITION

White noise

The innovation ε_t is the new component at date t : iid, mean 0, variance $\sigma^2 > 0$, independent of the past. Unforecastable in mean, not in spread.

KEY FORMULA

RESULT

(5)

The random walk accumulates shocks

$$X_t = X_{t-1} + \varepsilon_t = x_0 + \sum_{j=1}^t \varepsilon_j$$

IN WORDS the start plus every shock so far: $E[X_t] = x_0$, but $\text{Var}(X_t) = t\sigma^2$ grows, so it is not stationary.

NEEDS a fixed start x_0 and iid mean-zero shocks.

RESULT *what a walk implies*

- Every shock carries into every later level: a **random walk** has no restoring force.
- $\text{Corr}(X_t, X_{t-h}) = \sqrt{(t-h)/t}$ depends on the date: there is no stationary ACF.
- The forecast of X_{t+h} is X_t , with error variance $h\sigma^2$.
- With drift d , the mean is $x_0 + td$ and $\Delta X_t = d + \varepsilon_t$.

COMMON MISTAKE

“It will come back, because its mean is constant.” The constant mean is an average across possible paths at one date; it does not pull a path back. The best forecast is today’s level.

4.3 | DETREND OR DIFFERENCE

Detrend or difference? The model decides

Detrend a deterministic trend; difference a unit root

Model	Transformation	What is left
$X_t = a + bt + U_t$	subtract the trend	U_t , stationary
$X_t = X_{t-1} + \varepsilon_t$	first difference	$\Delta X_t = \varepsilon_t$
U_t , already white noise	a difference too many	lag-one correlation $-\frac{1}{2}$

RESULT *over-differencing*

$V_t = U_t - U_{t-1}$ has variance $2\sigma^2$ and lag-one covariance $-\sigma^2$: the transformation created the correlation.

COMMON MISTAKE

Fitting a straight line to a random walk does not make what is left stationary.

CHECK

Can a log transformation remove the growing variance of a Gaussian random walk that takes negative values?

ANSWER

No: the log is not even defined at negative values. The model justifies the first difference.

4.4 | AUTOREGRESSION

One recursion: shocks that vanish, last for ever, or fade

KEY FORMULA DEFINITION

(6)

The AR(1) model

$$X_t = c + \phi X_{t-1} + \varepsilon_t$$

IN WORDS an intercept, plus ϕ times yesterday, plus a new shock. An AR(p) adds lags up to X_{t-p} .

NEEDS white-noise shocks; a stable solution needs $|\phi| < 1$.

CHECK

Compare $X_t = 2 + X_{t-1} + \varepsilon_t$ with $Y_t = 2 + 0.5Y_{t-1} + \varepsilon_t$. Which has a stationary long-run mean?

ANSWER

Only Y : $\mu = 2 + 0.5\mu$ gives $\mu = 4$. X is a walk with drift 2; its mean $x_0 + 2t$ keeps rising.

Restriction	Model	A shock ...
$c = 0, \phi = 0$	white noise	vanishes at once
$c = 0, \phi = 1$	random walk	lasts for ever
$c = d, \phi = 1$	walk with drift	lasts, around a drifting mean
$ \phi < 1$	stable AR(1)	fades: $1, \phi, \phi^2, \dots$

4.5 | THE STATIONARY AR(1)

A stable AR(1): the intercept is not the mean

KEY FORMULA RESULT

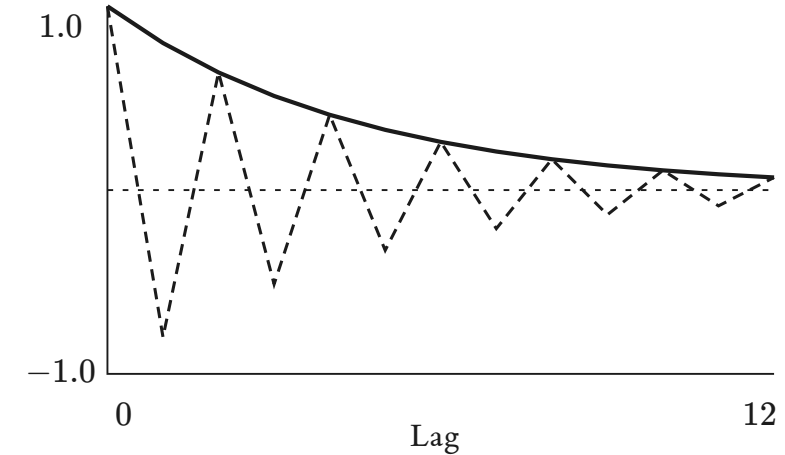
(7)

Stationary mean and variance

$$\mu = \frac{c}{1 - \phi}, \quad v = \frac{\sigma^2}{1 - \phi^2}$$

IN WORDS expectations of the recursion give $\mu = c + \phi\mu$; variances give $v = \phi^2 v + \sigma^2$.

NEEDS $|\phi| < 1$, the stationary solution.



Exact $\rho(h)$: $\phi = 0.8$ solid, $\phi = -0.8$ dashed.

KEY FORMULA RESULT

(8)

Autocorrelation of an AR(1)

$$\rho(h) = \phi^h \quad (h \geq 0)$$

IN WORDS the new shock is independent of the past, so $\gamma(h) = \phi \gamma(h - 1)$; and $\rho(1) = \phi$.

CHECK

$X_t = 3 + 0.6X_{t-1} + \varepsilon_t$ with innovation variance 2: find μ , v and $\rho(2)$.

ANSWER

$\mu = 3/0.4 = 7.5$; $v = 2/0.64 = 3.125$; $\rho(2) = 0.36$.

4.5 | FORECASTING

Forecast the centre, then state the uncertainty

KEY FORMULA RESULT

(9)

AR(1) forecast and error variance

$$\widehat{X}_{t+h|t} = \mu + \phi^h (X_t - \mu)$$

$$V_h = \sigma^2 \frac{1 - \phi^{2h}}{1 - \phi^2}$$

IN WORDS the mean plus ϕ^h times today's deviation. The error is the sum of the h unseen shocks, weighted $1, \phi, \dots, \phi^{h-1}$.

NEEDS information through t and known parameters; normal shocks for the interval $\pm 1.96\sqrt{V_h}$.

RESULT as the horizon grows

The forecast tends to μ and V_h to the stationary variance: the band levels off. For a random walk the forecast stays at X_t and $h\sigma^2$ grows without limit.

WORKED EXAMPLE a worked forecast

$\mu = 10, \phi = 0.8, \sigma^2 = 1$, and today $X_t = 15$.

$$\widehat{X}_{t+1|t} = 10 + 0.8(5) = 14$$

$$\widehat{X}_{t+2|t} = 10 + 0.8^2(5) = 13.2$$

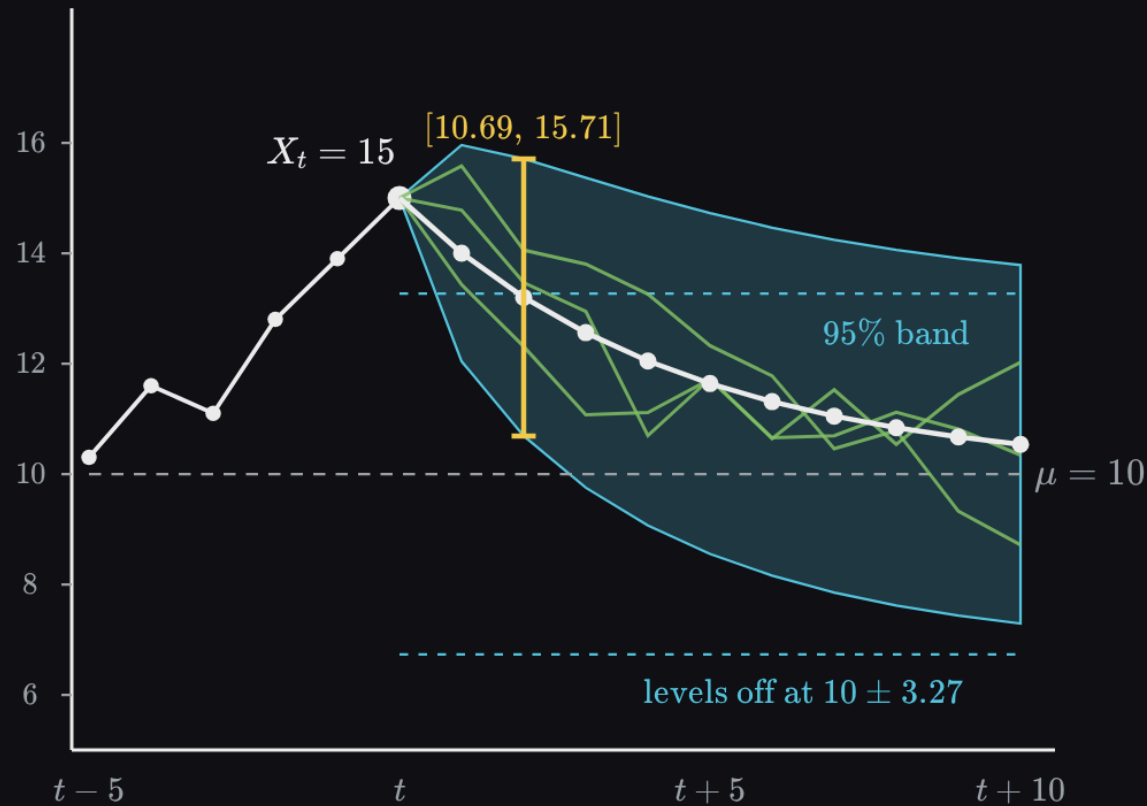
$$V_2 = 1 + 0.8^2 = 1.64$$

95% interval: $13.2 \pm 1.96\sqrt{1.64} = [10.69, 15.71]$.

WHY ECONOMISTS CARE

A **forecast** is a distribution: report its centre and its band. Estimating μ, ϕ and σ^2 from data adds further uncertainty.

Where will an AR(1) series go next?



$$\mu = 10, \phi = 0.8, \sigma^2 = 1, X_t = 15$$

$$\begin{aligned}\hat{X}_{t+h|t} &= 10 + 0.8^h(15 - 10) \\ &= 14, 13.2, 12.56, \dots \rightarrow 10\end{aligned}$$

$$V_2 = 1 + 0.8^2 = 1.64$$

$$13.2 \pm 1.96\sqrt{1.64} = [10.69, 15.71]$$

$$V_h \rightarrow 1/(1 - 0.8^2) \approx 2.78: \text{the band levels off}$$

The expected deviation shrinks;
a realised value can still rise.

PART 3 OF 3

Reversion and spurious correlation

When does movement towards the mean, or a strong correlation between two series, reveal a real mechanism?

KEY TERMS

mean reversion

regression to the mean

spurious regression

cointegration

[Notes 4.6–4.7 ↗](#)

SO FAR a stable AR(1) forecast returns towards μ ; a random walk's stays at today's level while its variance grows.

WHY IT MATTERS

Programmes that select extreme cases, such as tutoring the lowest scorers or treating road sites after unusually many accidents, can show before–after improvements with no effect. Regressing one trending level on another can give a high fit and a small p-value with no relationship at all.

IN THIS PART

- 4.6 Mean reversion is a property of the expected path
- ▶ Animation: why extreme first scores drift back on retest
- 4.6 Regression to the mean, and what it means for programmes
- 4.7 Independent random walks can correlate strongly
- 4.7 Before trusting a correlation between two series

4.6 | MEAN REVERSION

Mean reversion describes the expected path, not the next value

DEFINITION

Mean reversion

In a stable AR(1) the expected deviation after h periods, $\phi^h(X_t - \mu)$, tends to zero. Next period's expected change is

$$E[X_{t+1} - X_t \mid \mathcal{F}_t] = (\phi - 1)(X_t - \mu)$$

which is negative when today is above the mean.

WORKED EXAMPLE *half-life*

The expected deviation halves when $\phi^h = 1/2$, so $h_{1/2} = \log(1/2)/\log \phi$. At $\phi = 0.8$ that is about 3.11 periods.

CHECK

$\mu = 0$, $\phi = 0.8$, and today's value is 5. Must the next value be below 5?

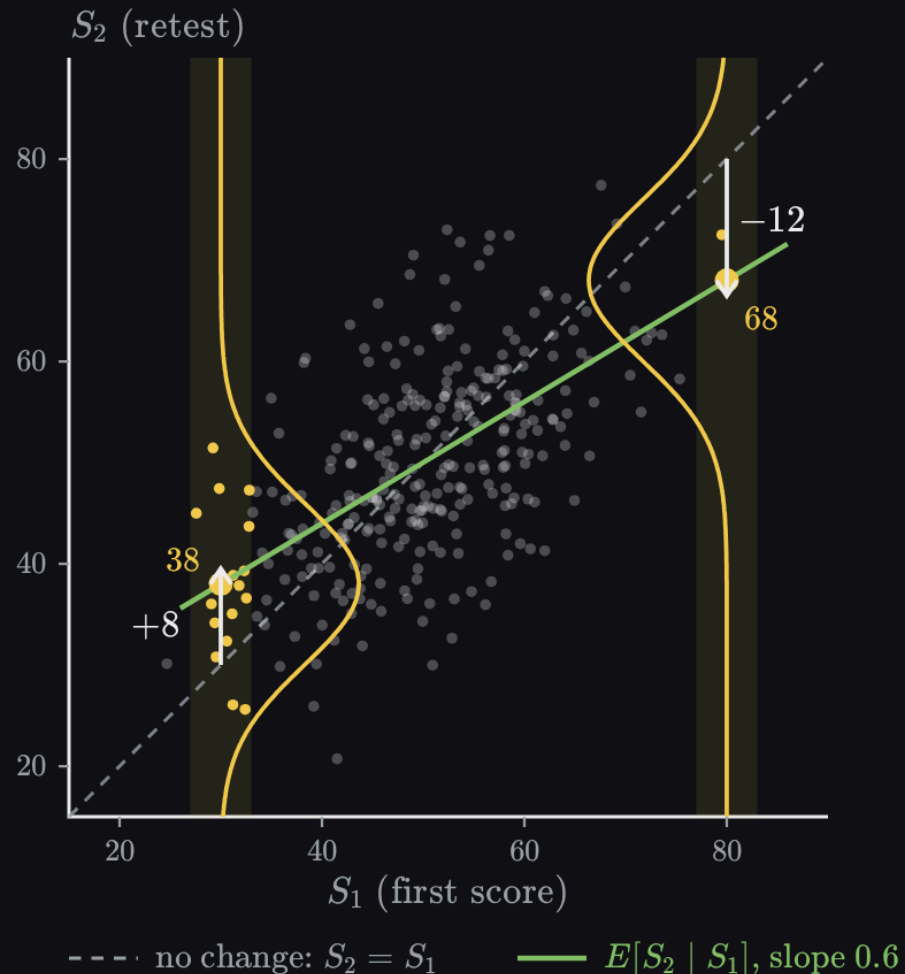
ANSWER

No. The forecast is 4, but $X_{t+1} = 4 + \varepsilon_{t+1}$ exceeds 5 whenever the shock exceeds 1. The expected change is -1 ; the realised change is not restricted.

COMMON MISTAKE

“It is far above its historical average, so it will fall.” That needs an identified restoring model. A random walk has a constant unconditional mean but no restoring forecast.

Do the lowest scorers improve without help?



$$S_1 = 50 + A + U_1, \quad S_2 = 50 + A + U_2$$

$\text{Var}(A) = 60$ persists; $\text{Var}(U_i) = 40$ is fresh

$$E[S_2 | S_1 = 30] = 50 + 0.6(30 - 50) = 38$$

+8 with no programme at all

$$E[S_2 | S_1 = 80] = 50 + 0.6(80 - 50) = 68$$

$r = 60/(60 + 40) = 0.6$: slope 0.6, not 1

Extreme first scores carry unusual noise;
the retest draws fresh noise.

A before–after gain is not a programme effect.

4.6 | REGRESSION TO THE MEAN

Selecting extreme first scores also selects favourable noise

KEY FORMULA RESULT

(10)

Regression to the mean

$$E[S_2 \mid S_1 = s] = \mu + r(s - \mu)$$

$$r = \frac{\tau^2}{\tau^2 + \omega^2}$$

IN WORDS a score is μ plus a lasting part (variance τ^2) plus fresh noise (variance ω^2); the retest keeps the share r of the first deviation.

NEEDS the Gaussian model $S_k = \mu + A + U_k$, all parts independent.

RESULT *two extremes*

Pure guessing: the top scorers' expected retest is the mean.

No new noise: no regression at all.

WHY ECONOMISTS CARE

A school tutors its lowest scorers. With $\mu = 50$ and $r = 0.6$, a first score of 30 retests at 38 on average, with no effect at all. A before–after gain mixes an effect, **regression to the mean**, a common trend and noise. Ask what the same selection gives without treatment (Lecture 5).

The same form as the $AR(1)$ forecast, a different mechanism: a lasting component measured twice, and selection on the first measurement.

4.7 | BACK TO THE OPENING QUESTION

Back to 0.71: independent random walks can correlate strongly

WORKED EXAMPLE *two independent walks*

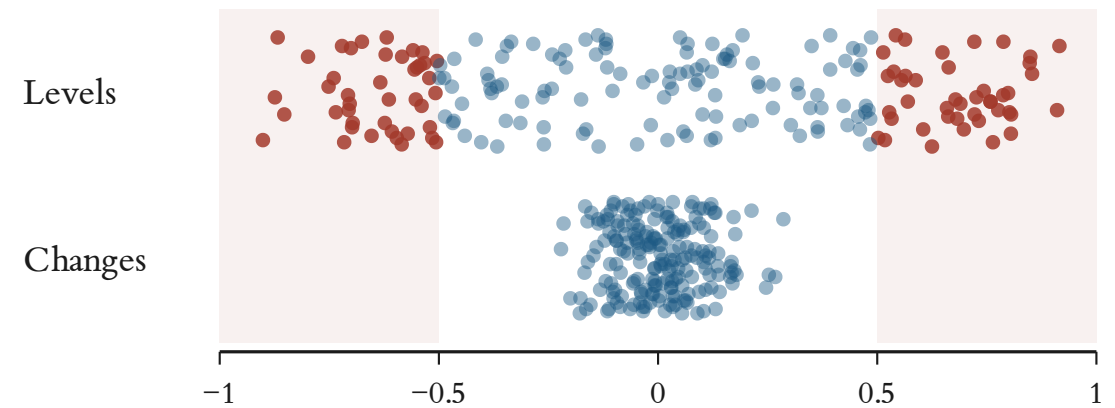
$X_t = X_{t-1} + \varepsilon_t$ and $Y_t = Y_{t-1} + \eta_t$, both from 0, with independent $N(0, 1)$ shocks. An early positive shock lifts X for many later dates; an unrelated early shock may do the same for Y . Across one realised sample, their levels then move together.

Neighbouring dates share shocks, so T date pairs are not T independent pieces of information.

RESULT *the opening question*

(c). The generator did not link the series. Nor does a longer sample help: levels correlations stay widely spread across independent pairs.

The notes' lab: 200 prespecified pairs of independent walks, $T = 100$.



Levels (X_t, Y_t) $|r| > 0.5$ in 85 of 200 pairs

Changes ($\Delta X_t, \Delta Y_t$) $|r| > 0.5$ in 0 of 200: independent iid pairs, correlation zero

4.7 | CARRY THE QUESTION THROUGH

A high coefficient does not choose its own explanation

DEFINITION

Spurious regression

Misleading evidence of a relationship when unrelated persistent, nonstationary levels are regressed on one another, using inference designed for a different error structure.

RESULT *four sources of a high coefficient*

1. A common driver of both series.
2. Two deterministic trends, such as $X_t = t + \varepsilon_t$ and $Y_t = 2t + \eta_t$.
3. A search across many pairs that selects an extreme one.
4. Nonstationary levels, which align even when the processes are independent.

COMMON MISTAKE

“The p-value is tiny, so the relation is real.” iid correlation tests and regression standard errors are wrong for random-walk levels; a small p-value from the wrong reference distribution is not reassuring.

*Transform because the model says so. Differencing changes the question and can remove a real long-run relation: some nonstationary series have a stationary combination, called **cointegration**.*

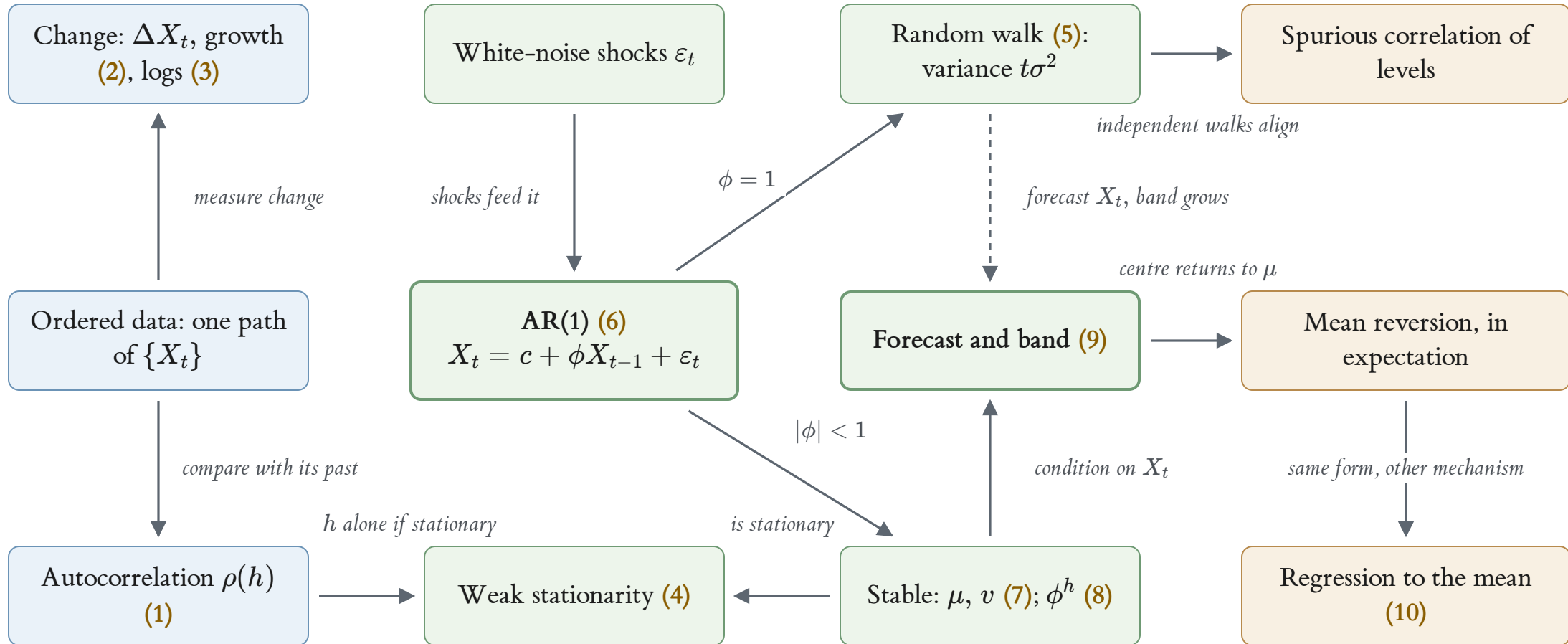
REVIEW · HOW THE IDEAS CONNECT

Describe, model the memory, then interpret through the model

PART 1 · DESCRIBE

PART 2 · MODEL

PART 3 · INTERPRET



REVIEW · SUMMARY SHEET

What to take away: the terms and the ten formulas

KEY TERMS

1 ORDER, UNITS AND CHANGE

Time series and process ordered data; the process $\{X_t\}$ is their probability model, and we observe one path.

Differences $\Delta X_t = X_t - X_{t-1}$; $\Delta_4 X_t$ compares matching quarters.

Growth factor $1 + g_t$ multiplies the level; rates compound, they do not add.

2 MODELS WITH MEMORY

White noise the new shock ε_t : iid, mean 0, variance σ^2 .

Random walk $X_t = X_{t-1} + \varepsilon_t$: every shock is permanent and the variance $t\sigma^2$ grows.

Forecast the expected X_{t+h} given information at t ; V_h is its error variance.

3 REVERSION AND SPURIOUS CORRELATION

Mean reversion the expected deviation $\phi^h(X_t - \mu)$ shrinks to zero when $|\phi| < 1$.

Regression to the mean selecting extreme noisy scores selects favourable noise too, so retests move towards the mean.

Spurious regression misleading evidence of a link between persistent levels, from iid-style inference.

KEY FORMULAS

1 ORDER, UNITS AND CHANGE

(1) Autocorrelation at lag h DEFINITION

$$\rho(h) = \text{Corr}(X_t, X_{t-h})$$

finite variances; a function of h alone if stationary

(2) Growth factors compound DEFINITION

$$(1 + \bar{g})^h = X_t / X_{t-h}$$

positive levels

(3) Log change IDENTITY

$$r_t = \log(1 + g_t), g_t = e^{r_t} - 1$$

$X_t > 0$; natural logs

2 MODELS WITH MEMORY

(4) Weak stationarity DEFINITION

$$E[X_t] = \mu, \text{Cov}(X_t, X_{t-h}) = \gamma(h)$$

finite variance $\gamma(0)$

(5) The random walk accumulates shocks RESULT

$$X_t = x_0 + \sum_{j \leq t} \varepsilon_j, \text{Var}(X_t) = t\sigma^2$$

fixed start; iid mean-zero shocks

(6) The AR(1) model DEFINITION

$$X_t = c + \phi X_{t-1} + \varepsilon_t$$

iid shocks; stable if $|\phi| < 1$

(7) Stationary mean and variance RESULT

$$\mu = c/(1 - \phi), v = \sigma^2/(1 - \phi^2)$$

$|\phi| < 1$; the stationary solution

(8) Autocorrelation of an AR(1) RESULT

$$\rho(h) = \phi^h \quad (h \geq 0)$$

the stationary AR(1)

(9) AR(1) forecast and error variance RESULT

$$\widehat{X}_{t+h|t} = \mu + \phi^h(X_t - \mu), V_h = \sigma^2 \frac{1 - \phi^{2h}}{1 - \phi^2}$$

known parameters; normal shocks for $\pm 1.96\sqrt{V_h}$

3 REVERSION AND SPURIOUS CORRELATION

(10) Regression to the mean RESULT

$$E[S_2 | S_1 = s] = \mu + r(s - \mu)$$

Gaussian model; $r = \tau^2/(\tau^2 + \omega^2)$

REVIEW · REMEMBER

Keep the order, model the memory, question the pattern

RESULT *the opening question*

(c). The series were independent random walks. Their levels share shocks across dates, so a correlation of 0.71 is unremarkable; their changes are independent pairs.

REMEMBER

1. Keep the dates, units and information set: differences subtract, growth factors compound, logs add.
2. In $X_t = c + \phi X_{t-1} + \varepsilon_t$, $|\phi| < 1$ makes shocks fade and forecasts return to μ inside a band that levels off; $\phi = 1$ makes them last.
3. Reversion and strong level correlations can arise with no mechanism: ask what the same model or selection gives without the claimed link.

BEFORE NEXT TIME

- Notes 4.1–4.7 with the worked solutions.
- The four labs: scales, the AR(1), retest scores and independent walks.
- Rework the forecast with $\mu = 20$, $\phi = -0.5$, $\sigma^2 = 4$, $X_t = 24$.

NEXT · LECTURE 5

Causal inference. A before–after gain after selection, or a correlation between two series, is not an effect. Which comparison identifies one?