CHULALONGKORN BUSINESS SCHOOL
FLAGSHIP FOR LIFE

AACSB ACCREDITED
EFMD EQUIS ACCREDITED
EFMD EPAS ACCREDITED

MSF Chula

Financial Econometrics

Lecture VIII

Introduction to Time Series Analysis

Narapong Srivisal, Ph.D.

3 main structures of data

1) Cross sectional data → many individuals , 1 periods ; happen in the same time

2) Time series → 1 individual , many periods
↳ has to be consecutive periods

3) Panal data → many individuals and each individual has many periods

4) Repeated cross section → " ———————————— "
but no connection / link across time

time-series for r.v. y

there should be linle btw period → so we can forecast this

consecutive periods without gap

$\{ y_t \} \Rightarrow y_1, y_2, y_3, y_4, \dots, y_{t-1}, y_t, y_{t+1}, \dots$

correlated

not iid → ~~WLLN~~ ← assumption iid

but there's WLLN for time-series → more assumption

check this "stationary" → pattern that we observe in the past, carries on to the future

---

$\{ y_t \} \Rightarrow y_1, y_2, y_3, y_4, y_5, \dots$

we can forecast / estimate y from its historical data

Outline

Time Series Data

Stationarity

there's only y

Univariate Time Series Models

- AR and MA processes
- ARMA process

combination

Non-stationary Processes

- Random Walk process
- Unit Root process
- Dickey-Fuller Test
- Integration

Relationship across Variables

- Static and Finite Distributed Lag Models
- Cointegration

Time Series Data

- Time series is a sequence of random variables indexed by time. E.g. hourly, daily, monthly, quarterly, yearly, …
- Each time period is treated as an observation.
- A random time-series process is also called a **stochastic process**.
  means random
- Generally for time series, the past can have an impact on the future. Therefore, the observations are unlikely to be independent of one another.
- This means that the iid assumption we made for cross sectional data cannot be used when analyzing time series.
- Some other data structures than iid will be required for conducting time-series analysis.

Stationarity

- Without independence among the observations, the WLLN and CLT used to analyze distribution of the OLS estimator for cross-sectional data are invalid.
- To be applicable for time series data, we need to utilize some other structures of the joint distribution among the observations across periods.
- A commonly employed framework is stationary of time series or stochastic process.
- <u>Definition</u>: a stochastic process $\{Y_t : t = 1, 2, \dots\}$ is **strictly stationary** if for all t, h, k, the joint distribution of $\{Y_t, Y_{t+1}, \dots, Y_{t+k}\}$ is the same as the joint distribution of $\{Y_{\boldsymbol{t+h}}, Y_{\boldsymbol{t+h}+1}, \dots, Y_{\boldsymbol{t+h}+k}\}$
- Intuitively, the strict stationary requires that the joint distribution of observations in adjacent periods does not change overtime. Then, we can use data of any time interval to make analysis.

- Definition: a stochastic process $\{Y_t : t = 1, 2, \dots\}$ is **weak stationary** (or **covariance stationary**, or simply **stationary**) if for all t, k, $E[Y_t], E[Y_t^2]$, and $cov(Y_t, Y_{t+k})$ does not depends on t.
- The weak stationary only requires up to the second moment to be the same. This is often sufficient to study relationship of data across periods.
- Definition: **White-noise Process** is a sequence of random variables $\{e_t\}$ such that $\boldsymbol{E[e_t] = 0}$, $\boldsymbol{E[e_t e_s] = 0}$, $\boldsymbol{E[e_t^2]}$ **is constant.**
- Intuitively, we can think of the white noise as a random shock happening in each period.

Weak / Covariance stationary → required only three things to be the same across history

$y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8$

1) $E[y_t] = \mu$
2) $E[y_t^2]$ or $Var[y_t] \Rightarrow$ constant
3) $cov(y_t, y_{t+k}) \rightarrow$ can depend on k but not t

$k \rightarrow$ gap btw period

e.g. $cov(y_1, y_3) = cov(y_6, y_8)$

1 gap → same gap should get same covariance

but $cov(y_1, y_2)$ may not equal to $cov(y_1, y_3)$

Strict, strong stationary → every moments have to be exactly the same

also include skewness & kurtosis

↓ very hard to prove

⇓ * need to check 3 things from model

Only mean-variance, covariance can prove that model stationary

Population

- cannot be observed, but we want this
- $y = \beta x + u$ model

Sample

OLS
GLS
2SLS
estimators

model construction

- What to explain : y
- What use to explain : x
- % → ln
- diff across groups : dummies

!! check multicollinearity

---

* Model construction for time-series is totally different :

Population

we have just 1 variable the explained by its past

$\{ y_t \}$

↓ only 1 variable y

$y_t = f(y_{t-1}, y_{t-2}, \dots)$

function of itself in the past

model 1 → f( )
model 2 → g( )
model 3 → h( )

which model fit with the data ?

Sample

$\{ y_1, y_2, y_3, \dots, y_t \}$ even we can observe from day 1

AR : auto-regressive
MA : moving average
ARMA Autoregressive moving average

E.g. $y_t = \alpha y_{t-1} + e_t$

$y_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2}$

2 lags
2 orders

→ which one fit our data the most

- how many lags we use ?

how many periods ago ?

Univariate Time Series Models

Autoregressive (AR) Process

- Suppose there is dependence among observations in a stochastic process $\{Y_t\}$. Then, we may want a model that predicts future values of Y based on its past outcomes.
- Since $\{Y_{t-1}, Y_{t-2}, \dots\}$ contains information about Y_t, a simple model is to regress the current value on its lags

what happens today what happens yesterday random shock that happen just that period

$$Y_t = \rho Y_{t-1} + e_t$$

describe
white Noise

where e_t is the white noise process.

order 1 or AR(1)
1 lag

- This model is called **Autoregressive Process of Order One** or **AR(1)**.
- Note that $E[Y_t] = \rho E[Y_{t-1}]$ and $Var[Y_t] = \rho^2 Var[Y_{t-1}] + \sigma_e^2$
- So, to be a stationary process, we need $E[Y_t] = 0$ and $Var[Y_t] = \sigma_Y^2$ for all t. Therefore,

$$\sigma_Y^2 = \rho^2 \sigma_Y^2 + \sigma_e^2 \Rightarrow \sigma_Y^2 = \frac{\sigma_e^2}{1 - \rho^2}$$

- Since variance has to be greater than zero, then we need $|\rho| < 1$.

$$AR(1): \quad y_t = \rho y_{t-1} + e_t$$

$$AR(2): \quad y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + e_t$$

$$\vdots$$

$$AR(p): \quad y_t = \rho_1 y_{t-1} + \dots + \rho_p y_{t-p} + e_t$$

AR(1)

$$y_t = \rho y_{t-1} + e_t$$

1]

3 things to check :

$$E[y_t] = E[\rho y_{t-1} + e_t]$$

$$= \rho E[y_{t-1}] + \cancel{E[e_t]} \to 0 \text{ white-noise}$$

$$\mu = \rho \mu$$

$$\boxed{\mu = 0}$$

2]

$$Var[y_t] = Var[\rho y_{t-1} + e_t]$$

should not correlated to e_{t-1} according to white-noise

$$\sigma_y^2 = Var[\rho y_{t-1}] + \sigma_e^2$$

$$\sigma_y^2 = \rho^2 \sigma_y^2 + \sigma_e^2$$

$$\sigma_y^2 - \rho^2 \sigma_y^2 = \sigma_e^2$$

$$\sigma_y^2 (1-\rho^2) = \sigma_e^2$$

$$\sigma_y^2 = \frac{\sigma_e^2}{(1-\rho^2)}$$

$$+ \sigma_y^2 = \frac{\sigma_e^2}{1-\rho^2} +$$

$\boxed{|\rho| < 1}$ If not, it is not stationary

3]

$$Cov(y_t, y_{t-1}) = Cov(\rho y_{t-1}, y_{t-1}) + \cancel{Cov(e_t, y_{t-1})} \to 0$$

$$\rho y_{t-1} + e_t$$

$$= \rho^1 \sigma_y^2$$

try again

$$Cov(y_t, y_{t-2}) = Cov(\rho^2 y_{t-2} + \rho e_{t-1} + e_t, y_{t-2})$$

$$\rho y_{t-1} + e_t$$

$$= \rho^2 \sigma_y^2$$

$$\rho(\rho y_{t-2} + e_{t-1}) + e_t$$

to have y_{t-2} in common

$$Cov(y_t, y_{t-k})$$

$$\Downarrow$$

$$\rho^k \sigma_y^2$$

$$\text{correlation}(y_t, y_{t-k}) = \frac{\rho^k \cancel{\sigma_y^2}}{\cancel{\sigma_{y_t}} \cancel{\sigma_{y_{t-k}}}} = \rho^k$$

if stationary

AR(1) process :

1) Stationary if $|\rho| < 1$ and $\mu = 0$

2) $\text{corr}(y_t, y_{t-k}) = \rho^k$; how much info in k periods still reflect in y_t

↳ as $|\rho| < 1 \Rightarrow \rho^{k\uparrow} \downarrow\downarrow\downarrow$ value drops as $k\uparrow$

" fading memory "

or fading correlation

- For covariance, note that we can use iterated substitution:

$$
\begin{aligned}
Y_{t+h} &= \rho Y_{t+h-1} + e_{t+h} = \rho(\rho \boldsymbol{Y_{t+h-2}} + e_{t+h-1}) + e_{t+h} \\
&= \rho(\rho(\boldsymbol{\rho Y_{t+h-3}} + \boldsymbol{e_{t+h-2}}) + e_{t+h-1}) + e_{t+h} \\
&\vdots \\
&= \rho^h Y_t + \rho^{h-1} e_{t+1} + \cdots + e_{t+h}
\end{aligned}
$$

- Then, $cov(Y_t, Y_{t+h}) = \rho^h cov(Y_t, Y_t) = \rho^h \sigma_Y^2$

already prove ☺

- In fact, **AR(1) is strictly stationary if $|\rho| < 1$**, but it's more complicated to prove.
- In general, we can model the process with more than one lag: Autoregressive Model of order p or AR(p):

$$Y_t = \rho_1 Y_{t-1} + \cdots + \rho_p Y_{t-p} + U_t$$

Moving Average (MA) Process

- A moving average process is like a weighted average of the past realizations of white noises.
- Moving Average of Order One, or **MA(1):** $\boldsymbol{Y_t = e_t + \alpha_1 e_{t-1}}$
- Note that $E[Y_t] = E[e_t] + \alpha_1 E[e_{t-1}] = 0$ and $Var[Y_t] = (1 + \alpha_1^2)\sigma_e^2$
- For two or more periods apart, i.e. if $h > 1, cov(Y_t, Y_{t+h}) = 0$
- For one period apart,

$$cov(Y_t, Y_{t+1}) = cov(e_t + \alpha_1 e_{t-1}, e_{t+1} + \alpha_1 e_t)$$
$$= cov(e_t, \alpha_1 e_t) = \alpha_1 \sigma_e^2$$

- In general, the **Moving Average of Order q or MA(q)** process is

$$\boldsymbol{Y_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \cdots + \alpha_q e_{t-q}}$$

- MA process is always a stationary process.

MA (1) $y_t = e_t + \alpha_1 e_{t-1}$

MA (2) $y_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2}$

MA (3) $y_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \alpha_3 e_{t-3}$

MA (q) $y_t = e_t + \alpha_1 e_{t-1} + \dots + \alpha_q e_{t-q}$ ⇒ It is like average of shock

$\boxed{1 \quad 2}$ 3 4 5 $\boxed{6 \quad 7}$

If we in this period we use up to 2 lags

moving to this also more lags we're using

Properties of MA(q)

1) $E[y_t] = E[\underset{\substack{\uparrow \\ 0}}{e_t} + \underset{\substack{\uparrow \\ 0}}{\alpha_1 e_{t-1}} + \underset{\substack{\uparrow \\ 0}}{\alpha_2 e_{t-2}} + \dots + \underset{\substack{\uparrow \\ 0}}{\alpha_q e_{t-q}}] = 0$ always no matter how many lags

2) $Var[y_t] = Var[e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \dots + \alpha_q e_{t-q}]$

$= \underset{\sigma_e^2}{\cancel{Var(e_t)}} + \alpha_1^2 \underset{\sigma_e^2}{\cancel{Var(e_{t-1})}} + \alpha_2^2 \underset{\sigma_e^2}{\cancel{Var(e_{t-2})}} + \dots$

no 2cov( ) terms → 0 white noise

white-noise constant

$= (1 + \alpha_1^2 + \alpha_2^2 + \dots)\, \sigma_e^2$

3) MA(1) : $cov(y_t, y_{t-1}) = \alpha_1 \sigma_e^2$

$y_{t-1} = e_{t-1} + \alpha_1 e_{t-2}$

$y_t = e_t + \alpha e_{t-1}$

0 0 0 0

white noise diff period = 0

$cov(y_t, y_{t-2}) = 0$

$y_{t-2} = e_{t-2} + \alpha_1 e_{t-3}$

$y_t = e_t + \alpha_1 e_{t-1}$

0 0 0 0

no overlap

For MA (q) ; $q = 1, 2, 3, 4, 5$

MA (2) ; $y_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2}$

$$\mathrm{cov}(y_t, y_{t-1}) = (\alpha_1 + \alpha_1 \alpha_2)\sigma_e^2$$

$$y_{t-1} = e_{t-1} + \alpha_1 e_{t-2} + \alpha_2 e_{t-3}$$

$$y_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2}$$

$$\mathrm{cov}(y_t, y_{t-2}) = \alpha_2 \sigma_e^2$$

$$\mathrm{cov}(y_t, y_{t-3}) = 0$$

$$\mathrm{cov}(y_t, y_{t-4}) = 0$$

$$\mathrm{cov}(y_t, y_{t-5}) = 0$$

$\boxed{\text{up to } q \text{ period}}$

MA (3) ; $y_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \alpha_3 e_{t-3}$

$y_{t-1} = e_{t-1} + \alpha_1 e_{t-2} + \alpha_2 e_{t-3} + \alpha_3 e_{t-4}$

$y_{t-2} = e_{t-2} + \alpha_1 e_{t-3} + \alpha_2 e_{t-4} + \alpha_3 e_{t-5}$

$y_{t-3} = \underbrace{e_{t-3}}_{\text{that period}} + \underbrace{\alpha_1 e_{t-4}}_{\text{lag 1}} + \underbrace{\alpha_2 e_{t-5}}_{\text{lag 2}} + \underbrace{\alpha_3 e_{t-6}}_{\text{lag 3}}$

* keep parameter, just change time

$$\mathrm{cov}(y_t, y_{t-1}) = (\alpha_1 + \alpha_1\alpha_2 + \alpha_2\alpha_3)\sigma_e^2$$

$$\mathrm{cov}(y_t, y_{t-2}) = (\alpha_2 + \alpha_1\alpha_3)\sigma_e^2$$

$$\mathrm{cov}(y_t, y_{t-3}) = (\alpha_3)\sigma_e^2$$

$$\mathrm{cov}(y_t, y_{t-4}) = 0$$

$$\mathrm{cov}(y_t, y_{t-5}) = 0$$

* $\mathrm{cov}(y_t, y_{t-1}) = \cancel{\mathrm{cov}(e_t, e_{t-1})}^{0} + \cancel{\mathrm{cov}(e_t, \alpha_1 e_{t-2})}^{0} + \cancel{\mathrm{cov}(e_t, \alpha_2 e_{t-3})}^{0}$ white-noise

$+ \mathrm{cov}(\alpha_1 e_{t-1}, e_{t-1}) + \cancel{\mathrm{cov}(\alpha_1 e_{t-1}, \alpha_1 e_{t-2})}^{0} + \cancel{\mathrm{cov}(\alpha_1 e_{t-1}, \alpha_2 e_{t-3})}^{0}$

$+ \cancel{\mathrm{cov}(\alpha_2 e_{t-2}, e_{t-1})}^{0} + \mathrm{cov}(\alpha_2 e_{t-2}, \alpha_1 e_{t-2}) + \cancel{\mathrm{cov}(\alpha_2 e_{t-2}, \alpha_2 e_{t-3})}^{0}$

$= \mathrm{cov}(\alpha_1 e_{t-1}, e_{t-1}) + \mathrm{cov}(\alpha_2 e_{t-2}, \alpha_1 e_{t-2})$

$= \alpha_1 \sigma_e^2 + \alpha_1\alpha_2 \sigma_e^2$

$= (\alpha_1 + \alpha_1\alpha_2)\sigma_e^2$

MA(q) : $E[y_t] = 0$

$Var(y_t) = (1 + \alpha_1^2 + \alpha_2^2 + \dots + \alpha_q)^2 \sigma_e^2$

$cov(y_t, y_{t-k}) \Rightarrow$ doesn't depend on t
only depend on k

MA(q) Always a stationary process

$cov(y_t, y_{t-k}) = 0$ when $k > q$ $\Rightarrow$ the MA is fit model

MA: Short-memory process

3 concepts of not related

- Independent
- Mean independent
- Uncorrelated $cov(\ ) = 0$ $\rightarrow$ there still be relation but it is not correlated

AR(1) : $y_t = \rho y_{t-1} + e_t$

Stationary if $|\rho| < 1$

$\rho = 0.1$

$\rho^{10} = 0.0000000001$

$cov(y_t, y_{t-k}) = \rho^k \sigma_y^2$

$cov(y_t, y_{t-k}) = \rho^k$ as k increases, it drops and approaches to 0 not equal to zero!

unlike MA

$\downarrow$

long-memory process

ออกสิ่งนี้

AR(1) $\Rightarrow y_t = \rho y_{t-1} + e_t \longleftrightarrow |\rho| < 1$

AR(p) $\Rightarrow y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \dots + \rho_p y_{t-p} + e_t$

$\downarrow$ $\rho_1 L y$ (L lag) $\quad \downarrow$ $\rho_2 L^2 y \quad \downarrow$ $\rho_p L^p y$

$$y_t \underbrace{(1 - \rho_1 L - \rho_2 L^2 - \rho_3 L^3 - \dots \rho_p L^p)}_{\text{solve this to know whether we can use the AR(p) model}} = e_t$$

Autocorrelation Function

- Autocorrelation is the correlation between a random variable and its past value.
- Let γ_k denote the **autocorrelation of order k**. Then, for a stationary stochastic process $\{Y_t\}$,

corr(y_t, y_{t-k})

gap = k

$$\gamma_k = \frac{cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)Var(Y_{t-k})}} = \frac{cov(Y_t, Y_{t-k})}{Var(Y_t)} = \frac{cov(y_t, y_{t-k})}{var(y_{t-k})}$$

เหมือน β

which is a function of k but not t.

- Thus, we can construct an **autocorrelation function (ACF)** as the autocorrelation as a function of the distance k.
- ACF can tell us the extent to which one value of a stationary process is correlated with the previous values, or in other words, the length and strength of memory of the process.

Example:

- Recall an AR(1) process: $Y_t = \rho Y_{t-1} + U_t$
- For an AR(1) to be a stationary process, we need $|\rho| < 1$, and the variance of the process is then

$$\sigma_Y^2 = \rho^2 \sigma_Y^2 + \sigma_U^2 \Rightarrow \sigma_Y^2 = \frac{\sigma_U^2}{1-\rho^2}$$

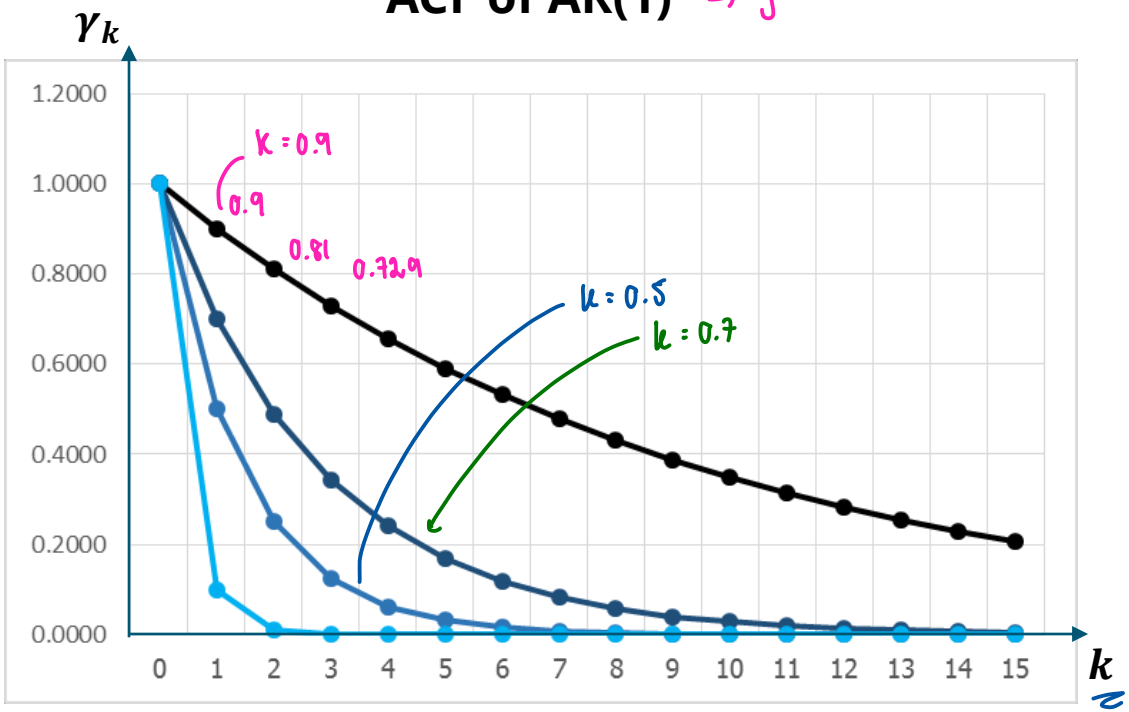
- Since the autocovariance is $cov(Y_t, Y_{t-k}) = \rho^k \sigma_Y^2$, then the autocorrelation function is

$$ACF(k) = \rho^k$$

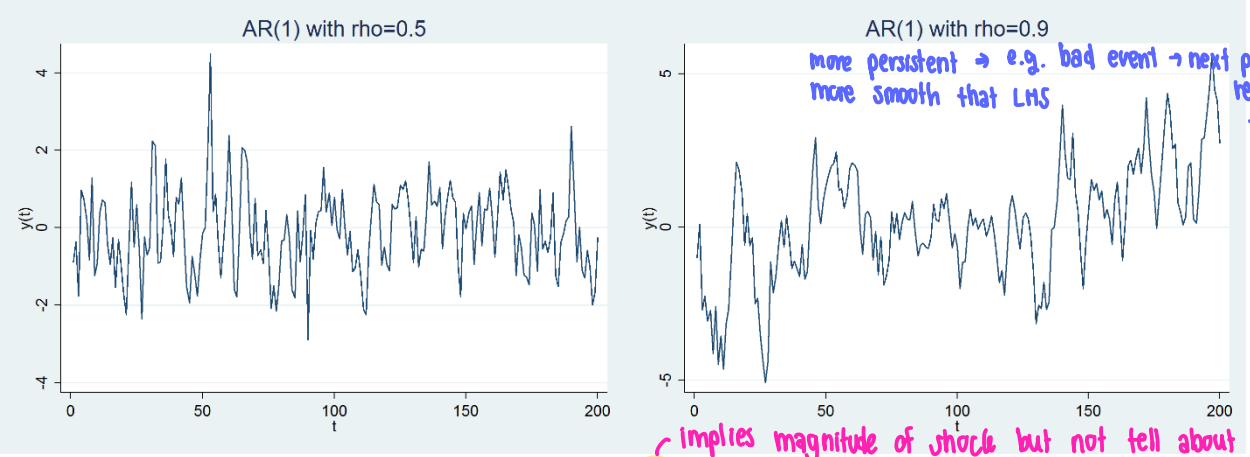
- This implies that the autocorrelation declines overtime toward zero. If $|\rho|$ is low, then the ACF diminishes quickly, but if $|\rho|$ is close to one then the ACF decreases very slowly.
- So, the AR(1) process has a strong memory or is highly persistent if $|\rho|$ approaches to one.

autocorrelation function for model

ACF of AR(1) ⇒ ρ^k



Autoregressive of Order 1



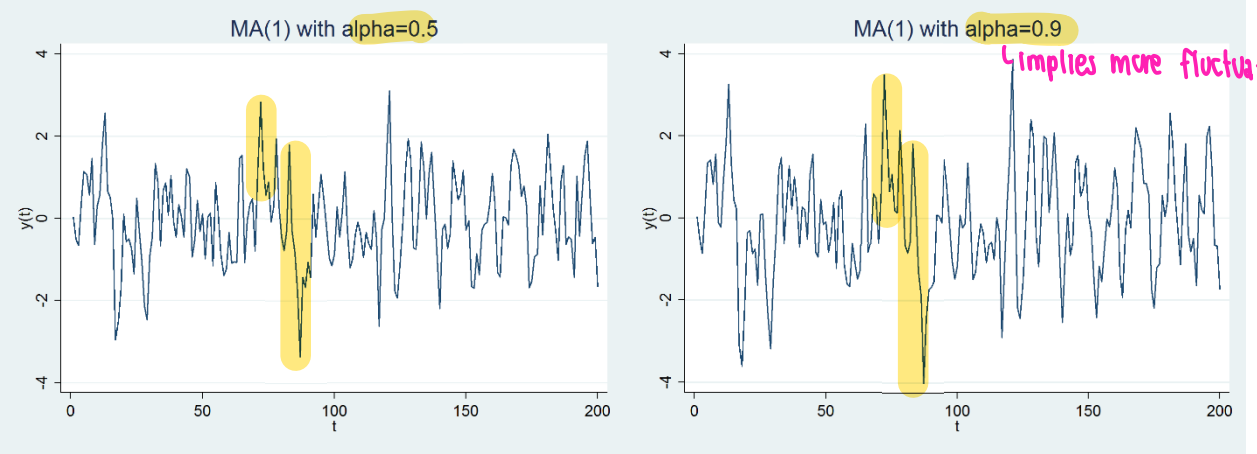
more persistent → e.g. bad event → next period still receives effect from previous one

more smooth that LHS

implies magnitude of shock but not tell about persistent

$y_t = e_t + \alpha e_{t-1} \longrightarrow Var(y_t) = (1+\alpha^2)\sigma_e^2$

Moving Average of Order 1



implies more fluctuate → not use picture to select model

Example:

- Recall an MA(1) process: $Y_t = U_t + \alpha_1 U_{t-1}$
- We know that it is a stationary process with $Var[Y_t] = (1 + \alpha_1^2)\sigma_U^2$
- Recall the autocovariance is $cov(Y_t, Y_{t-1}) = \alpha\sigma_U^2$, and $cov(Y_t, Y_{t-k}) = 0$ if $k > 1$
- Then the autocorrelation function is simply $\alpha_1\sigma_U^2/(1 + \alpha_1^2)$ for one period and zero otherwise.
- Therefore, MA(1) is a one-period-memory process.
- In general, **MA(q)** is like a weighted average of white noises of q periods; so, its memory is shortly limited to just q periods.
- In other words, the ACF of MA(q) is zero after q lags.

ACF and Model Selection

$$\gamma_k = \frac{cov(y_t, y_{t-k})}{Var(y_{t-k})} = \beta$$

- If a stochastic process has ACF becoming roughly zero after q periods, we may use MA(q) model to represent it.
- So, a starting point to choose the number of lags for MA is to examine consistent sample estimates of ACF and test when it becomes zero.
- More precisely, to test whether ACF(k) = 0, we can estimate the OLS model and test:

autocorrelation of order 1

$$Y_t = \beta_0 + \boxed{\gamma_k} Y_{t-k} + U_t; t = (k+1), \dots, T$$

$$H_0: \gamma_k = 0 \; vs \; H_1: \gamma_k \neq 0$$

- Under the null,

spectral variance

$$\sqrt{T}\,(\hat{\gamma}_k - \gamma_k) \xrightarrow{d} N(0, v_k)$$

$$v_k = 1 + 2\gamma_1^2 + 2\gamma_2^2 + \cdots + 2\gamma_{k-1}^2$$

- Hence, the test statistics:

$$\frac{\sqrt{T}\hat{\gamma}_k}{\sqrt{1 + 2\hat{\gamma}_1^2 + 2\hat{\gamma}_2^2 + \cdots + 2\hat{\gamma}_{k-1}^2}} \sim N(0,1)$$

reg	y_t	y_{t-1}	$\Rightarrow$	$\hat{\gamma}_1^{OLS}$	$= 0$?	if $\neq 0$
reg	y_t	y_{t-2}	$\Rightarrow$	$\hat{\gamma}_2^{OLS}$	$= 0$?	$\neq 0$
reg	y_t	y_{t-k} ($k \to q$)	$\Rightarrow$	$\hat{\gamma}_k^{OLS}$	$= 0$?	$\neq 0$

$\hat{\gamma}_{k+1} = 0$

e.g. $\hat{\gamma}_{10} = 0$

select MA(q)

sample not population

If get $\gamma_1 = 0$ may not get $\hat{\gamma}_1^{OLS} = 0.003$ close to 0 $\Rightarrow$ need to do hypothesis test

Partial Autocorrelation Coefficient

- For AR(p), ACF is not helpful as γ_q doesn't cut off at some points.
- Instead, we can investigate **Partial Autocorrelation Coefficient**, which is defined as the coefficient of Y_{t-p} in the AR(p) model.
- **Partial autocorrelation coefficient function (PACF)** is the partial autocorrelation coefficient as a function of number of lag p
  - **PACF(1) is the** ρ_{11} **in** $Y_t = \rho_{11}Y_{t-1} + U_t$ AR(1)
  - **PACF(2) is the** ρ_{22} **in** $Y_t = \rho_{21}Y_{t-1} + \rho_{22}Y_{t-2} + U_t$ AR(2)
    0 if this is actually AR(1)
  - **PACF(3) is the** ρ_{33} **in** $Y_t = \rho_{31}Y_{t-1} + \rho_{32}Y_{t-2} + \rho_{33}Y_{t-3} + U_t$ AR(3)
    0 if this is AR(2)
  - **PACF(k) is the** ρ_{kk} **in** $Y_t = \rho_{k1}Y_{t-1} + \rho_{k2}Y_{t-2} + \cdots + \rho_{kk}Y_{t-k} + U_t$ AR(k)
    called PACF(k)
- If the AR(p) is the right model, then the coefficient of the terms Y_{t-k} for $k > p$ should be zero. So, the PACF should become roughly zero after p lags.

↗ used to select number of lag

PACF

AR(1) : $y_t = \rho_{11} y_{t-1} + e_t$

AR(2) : $y_t = \rho_{21} y_{t-1} + \cancel{\rho_{22}}^{\,0} y_{t-2} + e_t$

AR(3) : $y_t = \rho_{31} y_{t-1} + \rho_{32} y_{t-2} + \cancel{\rho_{33}}^{\,0} y_{t-3} + e_t$

AR(4) : $y_t = \rho_{41} y_{t-1} + \rho_{42} y_{t-2} + \rho_{43} y_{t-3} + \cancel{\rho_{44}}^{\,0} y_{t-4} + e_t$

AR(p) : PACF becomes zero from p+1 on

RULE : we select lag k of term k+1 on become zero

PACF ⇓ AR(p)

ACF ⇓ MA(q)

PACF and Model Selection

- We can estimate PACF and test when it becomes zero to help determine an appropriate number of lag for a finite AR process.
- To test whether PACF(k) is zero, we can use OLS to estimate the AR model and test:

$$Y_t = \rho_{k1} Y_{t-1} + \cdots + \rho_{kk} Y_{t-k} + U_t$$

$$H_0: \rho_{kk} = 0 \; vs \; H_1: \rho_{kk} \neq 0$$

- Under the null, the test statistic is

$$\sqrt{T}\,\hat{\rho}_{kk} \xrightarrow{d} N(0,1)$$

- If we fail to reject the null from k on, then we should use AR($k-1$)

Estimating Correlogram (ACF and PACF)

- diagram of correlation

correlogram

use to select lag

don't see it becomes zero from any point

```
. corrgram consump
```

estimated ACF

patial autocorrelation

picture of ACF graph

PACF

LAG	AC	PAC	Q	Prob>Q	-1 0 1 [Autocorrelation]	-1 0 1 [Partial Autocor]
1	0.9825	1.0138	206.61	0.0000	|-------	|--------
2	0.9651	-0.1634	406.92	0.0000	|-------	-|
3	0.9480	-0.3003	601.11	0.0000	|-------	--|
4	0.9310	-0.3723	789.32	0.0000	|-------	--|
5	0.9140	0.1024	971.58	0.0000	|-------	|
6	0.8973	-0.0584	1148.1	0.0000	|-------	|
7	0.8804	-0.2066	1318.9	0.0000	|-------	-|
8	0.8638	-0.1296	1484.1	0.0000	|------	-|
9	0.8472	-0.1157	1643.8	0.0000	|------	|
10	0.8306	-0.0134	1798	0.0000	|------	|
11	0.8142	-0.0217	1947	0.0000	|------	|

not significantly diff from zero

MA(8)

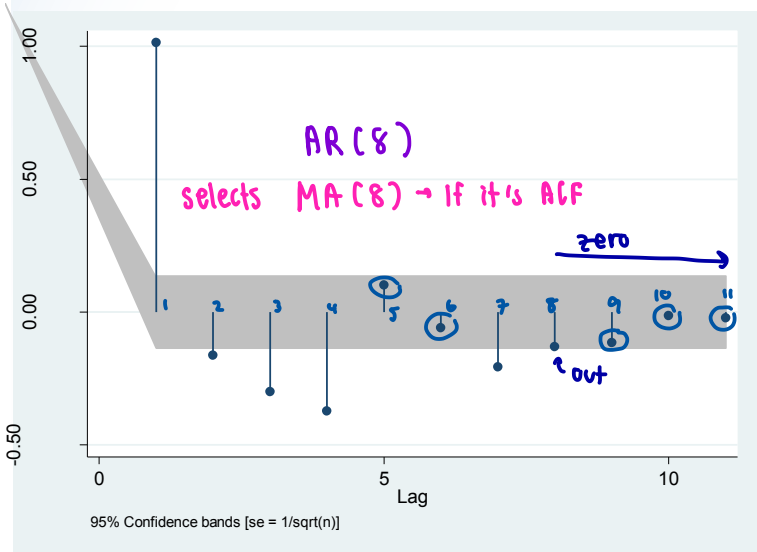
need to do hypothesis test whether it's significantly diff from zero

Graphing Correlogram

plot PACF graph

to select number of lag

```
. pac consump, lag(11)
```



tell data to plot ACF graph

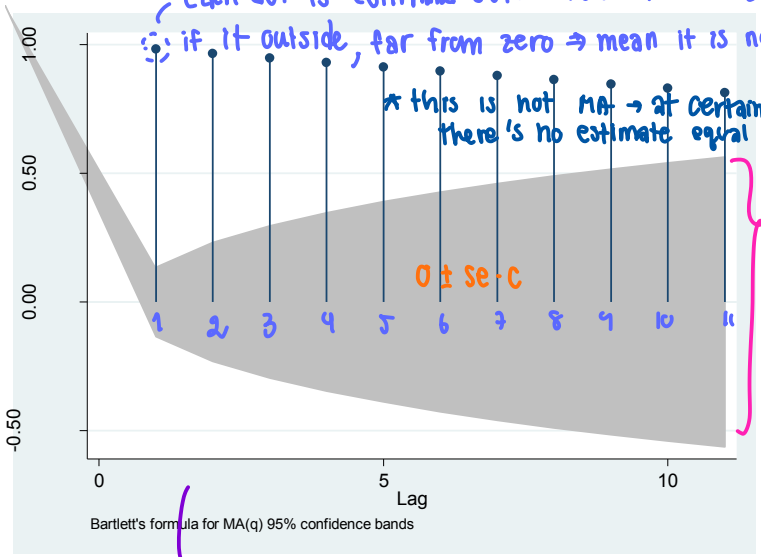
name of random variable y

```
. ac consump, lag(11)
```

each dot is estimated auto correlation of each order

if it outside, far from zero ⇒ mean it is not equal to 0

* this is not MA → at certain point, there's no estimate equal to 0



don't see it becomes zero from certain point on

Review < for 26 Oct 2023 >

Time-series (no iid anymore) ⇒ need stationary < strict : whole joint distribution to be the same
weak

1) $E(y_t)$
2) $E(y_t^2)$ or $Var(y_t)$
3) $cov(y_t, y_{t+k})$

forecast y from its historical data

univariate model

$\{y_t\}$

↓ select model :

① AR(p)

② MA(q) — autocorrelation function γ_k : become zero when $k > q$ (short-term memory, not correlate anymore)

— always stationary regardless parameter in the model

$$y_t = e_t + \alpha_1 e_{t-1} + \dots + \alpha_q e_{t-q}$$

③ ARMA(p,q)

AR(p) ⇒ $y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \dots + \rho_p y_{t-p} + e_t$

stationary or not depends on ρ

AR(1) : stationary when $|\rho| < 1$

ACF : ρ^k as k increases, this approaches to zero but not exactly equal to zero

when ρ is small → it fading faster!!

PURE AR(p) → use PACF to select it
PURE MA(q) → use ACF "———— "

} CORRELOGRAM

ARMA (p,q)

may have more than 1 ARMA representatives

ARMA Process

- An **ARMA(p, q) process is a combination of AR(p) and MA(q),** that is

has to be white-noise

$$Y_t = \underbrace{\rho_1 Y_{t-1} + \cdots + \rho_p Y_{t-p}}_{AR(p)} + \underbrace{\overset{e}{\cancel{U}}_t + \alpha_1 \overset{e}{\cancel{U}}_{t-1} + \cdots + \alpha_q \overset{e}{\cancel{U}}_{t-q}}_{MA(q)}$$

- However, it is important to note that AR and MA can be thought of as the same model but with different representational forms.
- Consider the AR(1) model:

$$\begin{aligned} Y_t &= \rho_1 Y_{t-1} + U_t = \rho_1(\rho_1 Y_{t-2} + U_{t-1}) + U_t \\ &= \rho_1(\rho_1(Y_{t-3} + U_{t-2}) + U_{t-1}) + U_t \\ &= \cdots \\ &= U_t + \rho_1 U_{t-1} + \rho_1^2 U_{t-2} + \rho_1^3 U_{t-3} + \cdots \end{aligned}$$

- Therefore, AR(1) can be written as an MA(∞): $Y_t = \sum_{j=0}^{\infty} \rho_1^j U_{t-j}$
- This implies that **a stochastic process $\{Y_t\}$ does not have a unique ARMA representation.**

$AR(1) = ARMA(1,0)$

$y_t = \rho y_{t-1} + e_t \Rightarrow MA(\infty)$

$\rho(y_{t-2}) + e_{t-1} \Rightarrow ARMA(2,1)$

shock in that period

$\rho(\rho(\rho y_{t-3} + e_{t-2}) + e_{t-1}) + e_t \Rightarrow ARMA(3,2)$

it's not unique!!

what kind of number we choose select for model??

$$y_t = \underbrace{\rho^k y_{t-k}}_{AR(k)} + \underbrace{\rho^{k-1} e_{t-(k-1)} + \rho^{k-2} e_{t-(k-2)} + \dots + e_t}_{MA(k-1)}$$

can write this as infinite lag of MA

* we want less lags!!!

model with more lags → lose more observations
= lower degree of freedom
→ not good for estimation

⇓

more coefficient to estimate

ARMA(1,1) ⇒ $\rho_1 y_{t-1} + e_t + \alpha_1 e_{t-1}$

ARMA(1,3) ⇒ $\rho_1 y_{t-1} + e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \alpha_3 e_{t-3}$

e.g. suppose data:

lose 1 observation
(bcs don't have data before Jan 1

t	y_t	y_{t-1}	y_{t-2}
Jan 1	0.1	-	-
Jan 2	0.2	0.1	-
Jan 3	0.15	0.2	0.1
Jan 4	0.2	0.15	0.2
Jan 5	0.3	0.2	0.15
⋮	⋮	⋮	

lost

∴ prefer not to have too many lags but more lags ⇒ more regressors

⇓ lower error of estimation

↳ better fit

- In general AR(p) is an MA(∞), and MA(q) can also be written as AR(∞). Thus, how many lags to use is a matter of **parsimony**: use not too many lags but well-enough describe the data.
- We use MA(q) if the ACF is not significantly different from zero for lags larger than q, while the PACF seems to last infinitely.
- We use AR(p) if the ACF seems to last infinitely, while the PACF is not significantly different from zero for lags larger than p.
- If the conditions for MA(q) and the conditions for AR(p) are not satisfied, an ARMA model may be a better parsimonious representation of the interested stochastic process.

penalty of # of lags & error

Information Criteria

- Another widely used method for determining the number of lags is to compare information criterion, based on the idea of trade-off between goodness-of-fit (low $\hat{\sigma}_U^2$) and parsimony (the number of lags).

penalty function → lower the better → choose which one has lowest information criteria

- **Akaike's Information Criterion (AIC)** for ARMA(p, q), proposed by Akaike (1973), is

↳ tends to suggest large # of lags → not informative for us to use

2 on result แปลว่าไม่แม่นไหม?

$$AIC = \log \hat{\sigma}_U^2 + \boxed{\frac{2(p+q)}{T \uparrow}}$$

when T is large → approach 0

if T is a lot → use BIC

- **Schwarz's Bayesian Information Criterion (SBIC or BIC)** for ARMA(p, q), proposed by Schwarz (1978), is given by

$$BIC = \log \hat{\sigma}_U^2 + \frac{(p+q)}{T} \underbrace{\log(T)}_{\text{increasing function}}$$

T → # observation

not approach to zero quickly

where $\hat{\sigma}_U^2$ is the estimated variance for U_t.

- HIC

maximum number of lags is 2

more terms → better

as guideline for # of lags but not absolutely right

AIC → (2, 2)

BIC → (1, 2)

lose 2 observations equally

if not much similar → may choose either

AIC → (10, 9)

✓ BIC → (2, 3)

AR, MA, ARMA, and more

- We can think of the information criteria as the regression costs, which consist of the variance of the error and the number of parameters to be estimated.
- **The model with the lowest value of information criterion is preferred.**
- Usually, AIC and BIC provides similar results; so, using either one is fine. However, there are also cases that AIC and BIC disagree on the number of lags.
- Theoretically, BIC is better as it will almost surely select the true model when T is large, while AIC tends to suggest more lags because the second term that penalizes the number of lags goes to zero when T is large.

Cross section → don't need to tell STATA
time-serie & panel → need to tell

used for time-serie

STATA needs to know time index

* not OLS estimator but "maximum likelihood estimator"
consistent estimators but typically biased
need large data set

Estimating ARMA(1, 1) Model

equal to zero for stationary process

name of r.v.

integration

```
. tsset t
. arima consump, arima(1,0,1)
```

AR i ma

ARMA (1,1)

```
ARIMA regression
Sample:  1946q1 - 1998q3                        Number of obs     =        211
                                                Wald chi2(2)      =  165877.35
Log likelihood = -997.5516                      Prob > chi2       =     0.0000
------------------------------------------------------------------------------
             |                 OPG
     consump |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
consump      |
       _cons |   2857.642   4001.692     0.71   0.475     -4985.53    10700.81
-------------+----------------------------------------------------------------
ARMA         |
          ar |
       L1. |   .9998755   .0024725   404.40   0.000     .9950295    1.004721
             |
          ma |
       L1. |   1.455106    .088814    16.38   0.000     1.281034    1.629178
-------------+----------------------------------------------------------------
      /sigma |   18.35787   1.530803    11.99   0.000     15.35755    21.35819
------------------------------------------------------------------------------
Note: The test of the variance against zero is one sided, and the two-sided
      confidence interval is truncated at zero.
```

Computing AIC and BIC

(This is to run after estimating the model)

```
. estat ic

Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
    Model |        Obs  ll(null)  ll(model)      df         AIC        BIC
----------+------------------------------------------------------------------
        . |        211         .  -997.5516       4    2003.103   2016.511
-----------------------------------------------------------------------------
          Note: N=Obs used in calculating BIC; see [R] BIC note.
```

Ex Model

- VAR
- VARSOC

	AIC	BIC
1		
2		
3		
4		
⋮		

tell the whole thing :)

Non-stationary Processes

Random Walk Process

- It is an AR(1) with $\rho = 1$: $Y_t = \overset{\rho = 1}{Y_{t-1}} + e_t$
- This process says that the best guess for the next period value, $E[Y_t]$, is just the current value.
- By iterating substitution, we can see that

$$
\begin{aligned}
Y_{t+h} &= Y_{t+h-1} + e_{t+h} = (Y_{t+h-2} + e_{t+h-1}) + e_{t+h} \\
&= \big((Y_{t+h-3} + e_{t+h-2}) + e_{t+h-1}\big) + e_{t+h} \\
&= Y_t + \underbrace{e_{t+1} + \cdots + e_{t+h}}_{\text{white-noise to h period}}
\end{aligned}
$$

- So, the expected value $E[Y_{t+h}|Y_t] = Y_t$ best guess is today value bcs random walk goes randomly
  data point of some certain period not estimator → is unbiased เท่า ไม่ค่อยเป็น
- Similar iterating substitution gives $Y_t = e_t + e_{t-1} + \cdots + e_1 + Y_0$. So, the variance of Y_t is $t\sigma_e^2$ which changes with time.
- Therefore, the **random walk process is non-stationary**.

$$y_{t+3} = y_{t+2} + e_{t+3}$$

$$= (y_{t+1} + e_{t+2}) + e_{t+3}$$

$$= (y_t + e_{t+1}) + e_{t+2} + e_{t+3}$$

$$E[y_{t+3} \mid y_t] = E[y_t + \cancel{e_{t+1}} + \cancel{e_{t+2}} + \cancel{e_{t+3}} \mid y_t]$$

$$= y_t$$

↳ best guess is what's the return today

$$y_t = y_{t-1} + e_t$$

t	y_t
0	0
1	$\cancel{y_0}^{\,0} + e_1 = e_1$
2	$y_1 + e_2 = e_1 + e_2$
3	$y_2 + e_3 = (e_1 + e_2) + e_3$
4	$y_3 + e_4 = e_1 + e_2 + e_2 + e_4$
5	
⋮	
t	$y_t = e_1 + e_2 + e_3 + e_4 + \dots + e_t$

shock that accumulated along the time

$$Var(y_t) = Var(\quad)$$

; white-noise no 2cov( )

$$= Var(e_1) + Var(e_2) + \dots + Var(e_t)$$

$$= t \cdot \sigma_e^2$$

Unit Root Process

- A unit root process $\{Y_t\}$ is of the form:

$$\boldsymbol{Y_t = Y_{t-1} + \eta_t}$$

- If $\{\eta_t\}$ is a white noise process, the unit root process is the random walk process or AR(1) with $\rho = 1$.
- This ($\rho = 1$) is why people call it "unit root".
- In general, $\{\eta_t\}$ is allowed to be any stationary process, for example, MA(q).
- Recall that AR(1) process is stationary if $|\rho| < 1$. Hence, **a unit process is non-stationary**
- Note that a non-stationary process like unit root is highly persistent, but not necessarily trending.

- However, in some cases, a non-stationary process can also have a trend. For example,

$$Y_t = \beta_0 + Y_{t-1} + e_t$$

- This process is called the random walk with drift, with $\boldsymbol{\beta_0}$ **is called the drift term.**
- We can do iterating substitution similar to what we did for the random walk without drift to prove that, with drift,

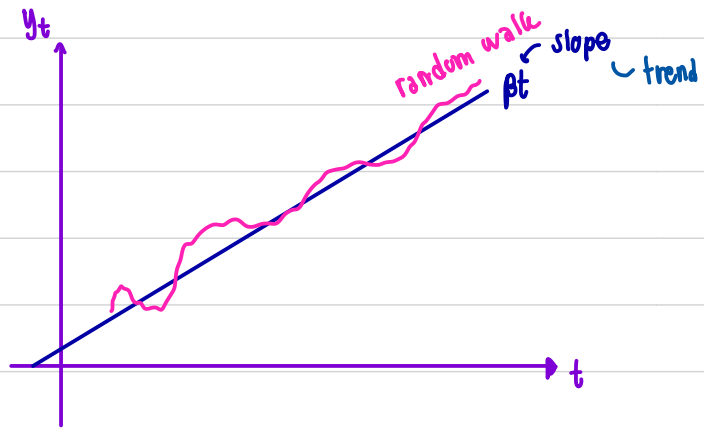
$$Y_t = \beta_0 t + e_t + e_{t-1} + \cdots + e_1 + Y_0$$

- Therefore, if $Y_0 = 0, E[Y_t] = \beta_0 t$, which trivially is a linear trend.

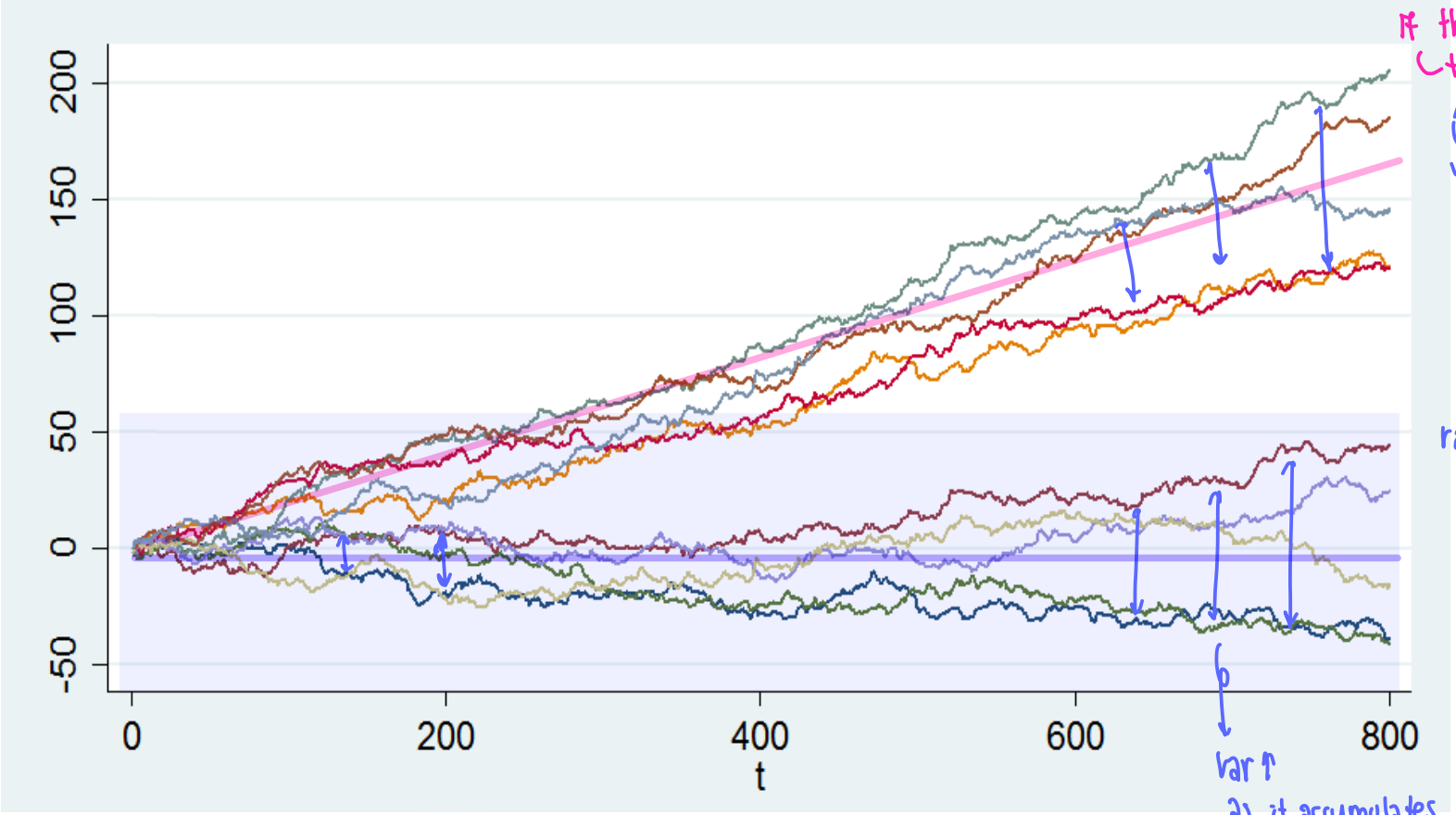
Random walk process with drift : $y_t = \beta + y_{t-1} + e_t$

assume $y_0 = 0$

t	y_t
1	$y_1 = \beta + 0 + e_1$
2	$y_2 = \beta + (\beta + e_1) + e_2 = 2\beta + e_1 + e_2$
3	$y_3 = \beta + (2\beta + e_1 + e_2) + e_3 = 3\beta + e_1 + e_2 + e_3$
4	$y_4 = \beta + (3\beta + e_1 + e_2 + e_3) + e_4 = 4\beta + e_1 + e_2 + e_3 + e_4$
5	
6	
⋮	
t	$y_t = \beta t + \overbrace{e_1 + e_2 + \dots + e_t}^{\text{random walk without drift}}$



Random Walk vs Random Walk with Drift



e.g. price goes up along inflation

add linear tread

if there's trend there's is drift

variance still the same

random walk can go anywhere

var ↑ as it accumulates shock

not the same

Trend vs Persistency

Trend → capture by "drift"

Persistency → capture by $|\rho|$

correlation of itself in the past

when $\rho \uparrow$ → more persistent

pattern not fluctuate, smooth

may not be trend but persistency

▷ Non-stationary processes

AR(1) with $|\rho| \geq 1$

random walk : $\rho = 1$

unit root walk

$$y_t = y_{t-1} + \square$$

can be any stationary process

doesn't have to be white-noise

$$\boxed{y_t = y_{t-1} + \underbrace{e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2}}_{\text{not white-not but MA(2)}}}$$

unit root process → more general than random walk

test whether process is stationary or not

Dickey-Fuller Test for a Unit Root

non-stationary process

- Consider a process: $\boxed{Y_t = \rho Y_{t-1} + \eta_t.}$
- This is a unit root process if $\rho = 1$; so, we want to test the null hypothesis $H_0: \rho = 1$. → hard to do
- The alternative hypothesis would be $H_1: |\rho| < 1$, as $|\rho| > 1$ implies an explosive series which is usually not considered or needed transformation before analysis.
- However, we cannot do simple OLS with t –Test, because under H_0 that the process has unit root, OLS estimator is not asymptotically normal but has some unknown distribution.
- Dickey and Fuller (1979) instead transforms the model to

$$\underbrace{\Delta Y_t}_{y_t - y_{t-1}} = \underbrace{(\rho - 1)}_{=0\,?} Y_{t-1} + e_t \equiv \theta Y_{t-1} + e_t$$

- This model can be estimated by OLS and test $\boldsymbol{H_0: \theta = 0}$ **vs** $\boldsymbol{H_1: \theta < 0}$, by the usual t –statistics:

doesn't follow t-distribution!! cannot use t-table

$$t_{\hat{\theta}} = \frac{\hat{\theta}_T^{OLS}}{SE\left(\hat{\theta}_T^{OLS}\right)}$$

- However, the t –statistic has a special distribution which has now been called **Dickey-Fuller distribution**.
- This means that we can compute the t-stat in the same way, but use critical values from the Dickey-Fuller distribution instead of Student t or standard Normal.
- We **reject** $\boldsymbol{H_0}$ **if** $\boldsymbol{t_{\hat{\theta}} < c}$**, and rejecting** $\boldsymbol{H_0}$ **implies** $\boldsymbol{|\rho| < 1}$, or there's evidence that $\{Y_t\}$ **is not a unit root process.**
- There is also the Augmented Dickey-Fuller test which has the same idea but allowing for more than one lag in the AR regression model.

$y_t = \rho y_{t-1} + y_t$

$\Delta y_t = \theta y_{t-1} + y_t$

↳ if $\rho = 1 \Rightarrow$ unit root → not stationary process

↳ cannot predict

Conducting Dickey-Fuller Test in STATA

2SLS?

```
. dfuller consump, nocon

Dickey-Fuller test for unit root                   Number of obs   =       210
                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             36.742  >       -2.585            -1.950            -1.617
```

↳ fail to reject $H_0 \Rightarrow$ not stationary process

e.g. look at return instead of price

growth diff

```
. dfuller d.consump, nocon
```

$d.c \Rightarrow c_t - c_{t-1}$

↳ stationary process

```
Dickey-Fuller test for unit root                   Number of obs   =       209
                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -3.111          -2.585            -1.950            -1.617
```

↳ reject

Integration

- Consider a unit root process: $Y_t = Y_{t-1} + \eta_t$. If we take the first difference:

$$Y_t - Y_{t-1} = \eta_t,$$

stationary

it will become a stationary process.

- A non-stationary stochastic process whose first difference is stationary is said to be integrated of order one or $I(1)$.
- In general, a non-stationary process that becomes stationary after p −differenced is called an **integration of order p** or $\boldsymbol{I(p)}$ process.
- All the stationary processes are said to be $I(0)$ or integrated of order zero.

how many times you have to do diff to get stationary process

$\{y_t\}$ → Dickey-Fuller → reject ✓ → stationary process by itself , integration of order zero · $I(0)$

fail to reject

ปกติแค่ order 1 หรือ 2 ก็ได้ stationary process แล้ว

$\Delta y_t = \{y_t - y_{t-1}\}$ → Dickey fuller test → reject : $I(1)$

fail to reject

$\Delta y_t - \Delta y_{t-1}$ → Dickey fuller test ⟶ reject $I(2)$

Fail to reject

$\{x_t\}$, $\{y_t\}$

divide into several cases :

- Both are stationary → reg y x
- One serie is stationary ; another not → no relationship
  growth, change
- Both non-stationary ; same order of integration
  don't run reg y x !!!
  get wrong conclusion
  may have relationship
  if we can find constant α → $\{y_t + \alpha x_t\}$ is stationary
  linear combination
  they have long run relationship
  cointegration factor

Relationship across Stationary Variables

- **Static time-series model** is used to study contemporaneous relationship between stationary time-series variables:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_k X_{kt} + U_t$$

where $t = 1, 2, \ldots, T$ index periods of the observed data.

- For this model, we can estimate the parameter by using OLS as the cross-sectional case.
- **Finite Distributed Lag (FDL) model** is used to study contemporaneous and lagged relationship between stationary time-series variables, as some changes may not produce immediate effect on the regressand that we are interested in.
- For example, an FDL model of order 2 to study impact of tax on trade volume:

$$Vol_t = \beta_0 + \gamma_0 tax_t + \gamma_1 tax_{t-1} + \gamma_2 tax_{t-2} + U_t$$

Interpretation of the FDL Model

- **Temporary Change**: consider if the government increases tax 1 unit from c to $c+1$ only in period t and revert back. Then, predicted volumes are:

suppose change happen at time t

$$Vol_{t-1} = \beta_0 + \gamma_0 c + \gamma_1 c + \gamma_2 c$$
$$Vol_t = \beta_0 + \gamma_0 (c+1) + \gamma_1 c + \gamma_2 c$$
$$Vol_{t+1} = \beta_0 + \gamma_0 c + \gamma_1 (c+1) + \gamma_2 c$$
$$Vol_{t+2} = \beta_0 + \gamma_0 c + \gamma_1 c + \gamma_2 (c+1)$$
$$Vol_{t+3} = \beta_0 + \gamma_0 c + \gamma_1 c + \gamma_2 c$$

$y_t = \beta_0 + \gamma_0 x_t + \gamma_1 x_{t-1} + \gamma_2 x_{t-2} + u_t$

do it as a table

change happens

t	x_t	x_{t-1}	x_{t-2}	$\rightarrow y_t$
-2	0	0	0	
-1	0	0	0	β_0
0 temporary change	1	0	0	$\beta_0 + \gamma_0$
1	0	1	0	$\beta_0 + \gamma_1$
2	0	0	1	$\beta_0 + \gamma_2$
3	0	0	0	β_0
4	0	0	0	β_0
5	0	0	0	β_0

1-period lag effect

remaining effect

2-period lag effect of x on y

ออกข้อ

- **Impact propensity** or **impact multiplier**, which is the immediate effect, is $Vol_t - Vol_{t-1} = \gamma_0$.
- **One-period lagged effect** is $Vol_{t+1} - Vol_{t-1} = \gamma_1$
- **Two-period lagged effect** is $Vol_{t+2} - Vol_{t-1} = \gamma_2$
- There is no effect after the second period in this model.

$$y_t = \beta_0 + \gamma_0 x_t + \gamma_1 x_{t-1} + \gamma_2 x_{t-2} + u_t$$

t	x_t	x_{t-1}	x_{t-2}	$\Rightarrow y_t$
-2	0	0	0	
-1	0	0	0	β_0
0	1	0	0	$\beta_0 + \gamma_1$
1	1	1	0	$\beta_0 + \gamma_1 + \gamma_2$
2	1	1	1	$\beta_0 + \gamma_1 + \gamma_2 + \gamma_3$
3	1	1	1	$\beta_0 + \gamma_1 + \gamma_2 + \gamma_3$
4	1	1	1	
5	1	1	1	

$$y_t = \beta_0 + \gamma_0 x_t + \gamma_1 x_{t-1} + \gamma_2 x_{t-2} + u_t$$

t	x_t	x_{t-1}	x_{t-2}	$\Rightarrow y_t$
-2				
-1				
0	1			
1	2			
2	3			
3	2			
4	1			
5	0			

test whether there's long run effect <1st part before midterm>

$H_0 : \gamma_1 + \gamma_2 + \gamma_3 = 0$ \ F-test for linear combination

$H_1 : \gamma_1 + \gamma_2 + \gamma_3 \neq 0$

- **Permanent Change**: consider if the government increases tax 1 unit from c to $c+1$ in period t and keeps it permanently. Then, predicted volumes are:

$$Vol_{t-1} = \beta_0 + \gamma_0 c + \gamma_1 c + \gamma_2 c$$

$$Vol_t = \beta_0 + \gamma_0 (c+1) + \gamma_1 c + \gamma_2 c$$

$$Vol_{t+1} = \beta_0 + \gamma_0 (c+1) + \gamma_1 (c+1) + \gamma_2 c$$

$$Vol_{t+2} = \beta_0 + \gamma_0 (c+1) + \gamma_1 (c+1) + \gamma_2 (c+1)$$

$$Vol_{t+3} = \beta_0 + \gamma_0 (c+1) + \gamma_1 (c+1) + \gamma_2 (c+1)$$

- The immediate effect is $Vol_t - Vol_{t-1} = \gamma_0$.
- One-period effect is $Vol_{t+1} - Vol_{t-1} = \gamma_0 + \gamma_1$
- Two-period effect is $Vol_{t+2} - Vol_{t-1} = \gamma_0 + \gamma_1 + \gamma_2$
- The effect also remains $\boldsymbol{\gamma_0 + \gamma_1 + \gamma_2}$ after two period. Thus, this is also called the **long-run propensity** or **long-run multiplier**.

permanent effect of x on y
in respond to permanent shock

R, STATA, EVIEW

Cointegration

- If the a process $\{X_t\}$ is stationary but a process $\{Y_t\}$ is not stationary, they should not move together.
- If two processes $\{X_t\}$ and $\{Y_t\}$ are both non-stationary, typically they do not move together especially in the short-run.
- However, there may be long-run relationship.
- $\{\boldsymbol{X_t}\}$ **and** $\{\boldsymbol{Y_t}\}$ **are cointegrated** if both are $I(1)$ processes, and there is a constant γ such that $Y_t - \gamma X_t$ is $I(0)$.
- This cointegration factor γ describes how the two series $\{X_t\}$ and $\{Y_t\}$ move together in the long-run.
- A well-known example of cointegrated series are aggregate consumption and income.