# FHS Microeconometrics Notes

Harry Folkard, Keble College

2022

**Abstract**

These are my Macroeconometrics notes made for my finals in 2022. They cover all of the topics. Feel free to use these notes and pass them on to others. Please note, however, that these have just been made by a student and not checked over. They likely contain errors, so it will be worth checking things for yourself. Thanks to Kevin Sheppard, Vanessa Berenguer Rico and Bent Nielsen - these notes are just my interpretation of their lectures and tutorials.

## Contents

# Need to know Theorems

## Chebyshev iid LLN

**Theorem** *(Law of Large Numbers by Chebyshev)*

For $i = 1, ..., n$ let $x_i$ be independent and identically distributed with finite mean, $\mu$, and variance $\sigma^2$. Then, as $n \longrightarrow \infty$,

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^{n} x_i \xrightarrow{P} \mu.$$

## Lindeberg-Levy iid CLT

**Theorem** *(Central Limit Theorem by Lindeberg-Levy)*

For $i = 1, ..., n$ let $x_i$ be independent and identically distributed with finite mean, $\mu$, and variance $\sigma^2$. Then, as $n \longrightarrow \infty$,

$$\frac{\sqrt{n}(\bar{x}_n - \mu)}{\sigma} \xrightarrow{D} N(0, 1)$$

## Multivariate iid CLT

**Theorem** *(Multivariate Lindeberg-Levy CLT)*

Let $Z_i$ for $i = 1, ..., n$ be independent and identically distributed m-dimensional random vectors with finite mean vector $\mu_Z = E[Z_i]$, and finite positive definite covariance matrix $\Sigma_Z = E[(Z_i - \mu_Z)(Z_i - \mu_Z)']$. Then

$$\sqrt{n}(\bar{Z}_i - \mu_Z) \xrightarrow{D} N(0_m, \Sigma_Z)$$

where $\bar{Z}_n = n^{-1} \sum_{i=1}^{n} Z_i$ and $N(0_m, \Sigma_Z)$ is multivariate normal.

## Covariance Stationary LLN

**Theorem** *(Law of Large Numbers)*

Let $y_t$ be a covariance-stationary process with $E(y_t) = \mu$ and $\gamma_h = \operatorname{Cov}(y_t, y_{t-h})$ and absolutely summable autocovariances so that $\sum_{h=0}^{\infty} |\gamma_h| < \infty$. Then as $T \longrightarrow \infty$,

$$\frac{1}{T} \sum_{t=1}^{T} y_t \xrightarrow{\mathrm{P}} \mu,$$

**Why require absolutely summable autocovariences?**

Because for a ***weakly stationary process*** a sufficient condition for ***mean square convergence*** (which implies convergence in probability) is,

$$\sum_{h=0}^{\infty} |\gamma_h| < \infty \text{ where } \gamma_h = \operatorname{Cov}(y_t, y_{t-h})$$

What this means is that a sufficient condition for a process to mean square converge is that the covariances, although can initially be non-zero, must at some point tend to zero - hence they are absolutely summable.

The intuition of this is that eventually covariances like $\gamma_{1000} = \operatorname{cov}(y_t, y_{t+1000})$ should be zero if the process is really stationary and $E[y_t]$ really converges to $\mu$. This implies that the sum of all the covariances should be less than infinity.

## Wold Decomposition CLT

**Theorem** *(Central Limit Theorem)*

If $y_t = \mu + \Psi(L)u_t$ where $u_t \sim iid\left(0,\ \sigma^2\right)$ and $\sum_{j=0}^{\infty} |\psi_j| < \infty$, then as $T \longrightarrow \infty$,

$$\sqrt{T}\left(\bar{y}_T - \mu\right) \xrightarrow{\text{D}} \text{N}\left(0, \sum_{h=-\infty}^{\infty} \gamma_h\right)$$

where $\sum_{h=-\infty}^{\infty} \gamma_j = \sigma^2 \Psi^2(1)$ is the long run variance.

## Stationary AR CLT

**Theorem** *(Central Limit Theorem for AR processes)*

Let $y_t$ be a stationary $AR(p)$ process with $E\left(y_t\right) = \mu$ and $\gamma_h = \text{Cov}\left(y_t, y_{t-h}\right)$. Then,

$$\sqrt{T}\left(\bar{y}_T - \mu\right) \xrightarrow{\text{D}} \text{N}\left(0, \sum_{h=-\infty}^{\infty} \gamma_h\right).$$

## Mds LLN

**Theorem** *(Law of Large Numbers)*

Let $(m_t, \mathcal{I}_t)$ for $t \in \mathbb{N}$ be a martingale difference sequence. If one of the following conditions holds

(a) $\sum_{t=1}^{\infty} E\left|m_t\right|^{1+p} / t^{1+p} < \infty$ for some $p \in [0, 1]$,

(b) $\underset{T \longrightarrow \infty}{lim} \max_{t \leq T} E\left|m_t\right|^{1+p} < \infty$ for some $p \in [0, 1]$,

then, as $T \longrightarrow \infty$,

$$T^{-1} \sum_{t=1}^{T} m_t \xrightarrow{P} 0.$$

## Mds CLT

**Theorem** *(Central Limit Theorem)*

Let $(m_t, \mathcal{I}_t)$ for $t \in \mathbb{N}$ be a martingale difference sequence satisfying $\text{E}m_t^2 < \infty$ for all $t$. Let $S_T^2 = \sum_{t=1}^{T} Em_t^2$. Suppose

$$(i) \quad \sum_{t=1}^{T} m_t^2 / S_T^2 \xrightarrow{\text{P}} 1,$$

$$(ii) \quad \sum_{t=1}^{T} \text{E}\left\{\left(m_t/S_T\right)^2 1_{(|m_t S_T| > \delta)}\right\} \longrightarrow 0 \text{ for all } \delta > 0,$$

then

$$\frac{1}{S_T} \sum_{t=1}^{T} m_t \xrightarrow{\text{D}} \text{N}(0, 1)$$

*Remark:* The Lindeberg condition *(ii)*

$$\sum_{t=1}^{T} \mathrm{E}\left\{ (m_t/S_T)^2 \, 1_{(|m_t/S_T|>\delta)} \right\} \longrightarrow 0,$$

for all $\delta > 0$ follows from the Lyapounov condition

$$(ii') \quad \sum_{t=1}^{T} E \left| m_t/S_T \right|^{2+\delta} \longrightarrow 0$$

for some $\delta > 0$.

## Slutsky's Theorem

**Theorem** *(Slutsky Theorem)*

Let $Y_n \xrightarrow{P} c$ and $X_n \xrightarrow{D} X$, then:

(a) $Y_n + X_n \xrightarrow{D} c + X$
(b) $Y_n X_n \xrightarrow{D} cX$
(c) $Y_n^{-1} X_n \xrightarrow{D} c^{-1} X$ if $c \neq 0$
(d) If $c = 0$ then $Y_n X_n \xrightarrow{P} 0$

## Wold Decomposition

**Theorem** *(The Wold Decomposition)*

If $x_t$ is stationary and non-deterministic, then

$$x_t = \sum_{j=0}^{\infty} \Psi_j u_{t-j} + z_t = \Psi(L) u_t + z_t$$

Where,

- $\Psi_0 = 1$ and $\sum_{j=0}^{\infty} \Psi_j^2 < \infty$
- $u_t$ is $wn(0, \sigma^2)$
- $z_t$ is deterministic
- $\mathrm{Cov}(u_t, z_t) = 0$ for all $s$ and $t$
- $\Psi_j$ and $u_t$ are unique.

# Some Laws, Tricks, and Terms

## Series and Summations

### The 'T' Rules

$$\sum_{t=1}^{T} 1 = T$$
$$\sum_{t=1}^{T} t = \sum_{t=1}^{T}(T+1-t) = \frac{T(T+1)}{2}$$
$$\sum_{t=1}^{T} t^2 = \frac{T(T+1)(2T+1)}{6}$$

### Geometric Series

$$\text{Finite: } \sum_{k=0}^{n-1} ar^k = \frac{a\left(1-r^n\right)}{1-r} = \sum_{k=1}^{n} ar^{k-1} = \frac{a\left(1-r^{(n+1)}\right)}{1-r} \quad [r \neq 1]$$
$$\text{Infinite: } \sum_{k=0}^{\infty} ar^k = \frac{a}{1-r} = \sum_{k=1}^{\infty} ar^{k-1} = \frac{a}{1-r} \quad \left[|r| < 1\right]$$

### Useful Result

$$\frac{1}{T}\sum_{t=1}^{T}\left(\frac{t}{T}\right)^v \longrightarrow \int_0^1 r^v dr = \frac{1}{v+1}$$
$$\frac{1}{T^{v+1}}\sum_{t=1}^{T}(t)^v \longrightarrow \int_0^1 r^v dr = \frac{1}{v+1}$$

## Stationarity

### Strict Stationarity

The time series $\{Y_t, t \in \mathbb{Z}\}$ is strictly stationary if the joint distributions $(Y_t, Y_{t+1}, \ldots, Y_{t+k}) \underset{D}{=} (Y_s, Y_{s+1}, \ldots, Y_{s+k})$ for all $t, s$ and $k$.

### Weak Stationarity

The time series $\{Y_t, t \in \mathbb{Z}\}$ is weakly stationary if:

(i) $E[Y_t] = m$ for all $t$

(ii) $Var(Y_t) = \sigma^2 < \infty$ for all $t$.

(iii) $Cov(Y_t, Y_s) = Cov(Y_{t+h}, Y_{s+h})$ for all $t, s, h \in \mathbb{Z}$.

(iii') $Cov(Y_t, Y_{t-h}) = \gamma_h$ for all h (this is equivalent to (iii)).

That is the mean, variance, and covariance do no depend on $t$. They are *time invariant.*

If a process is Gaussian Normal then strict and weak stationarity coincide

## Martingale Difference Sequence

Let $m_t$ be a sequence of random scalars with $E\left(m_t\right)=0$ and let $\mathcal{I}_t$ be the information available at date $t$, so $\mathcal{I}_t$ will include current and past values of $\{m_t\}$, as well as current and past values of any other random sequences, such as perhaps $\{x_t\}$.

$$\mathcal{I}_t=\{m_t,m_{t-1},m_{t-2},\ldots,x_t,x_{t-1},x_{t-2},\ldots\}$$

If $E\left(m_t \mid \mathcal{I}_{t-1}\right)=0$ then $\{m_t\}$ is said to be a martingale difference sequence with respect to $\{\mathcal{I}_t\}$.

$E\left(m_t \mid \mathcal{I}_{t-1}\right)=0$ implies that $\{m_t\}$ is serially uncorrelated (stronger assumption than uncorrelatedness but weaker than independence.)

# Lag Operator

## Properties

$$
\begin{aligned}
&Ly_t = y_{t-1} \\
&L\left(Ly_t\right) = L\left(y_{t-1}\right) = y_{t-2} \text{ hence } L^j\left(y_t\right) = y_{t-j} \\
&L\mu = \mu \\
&L^j \mu y_t = \mu y_{t-j} \\
&L\left(y_t + x_t\right) = y_{t-1} + x_{t-1}
\end{aligned}
$$

## Using Lag Polynomials

Let's consider the AR(1) case,

$$y_t = \phi y_{t-1} + u_t$$

We can rewrite the model using the lag operator after doing some basic algebra,

$$
\begin{aligned}
y_t - \phi y_{t-1} &= u_t \\
(1 - \phi L) y_t &= u_t \\
\Phi(L) y_t &= u_t \text{ where } \Phi(L) = 1 - \phi L
\end{aligned}
$$

This doesn't immediately seem useful, but it allows us to simplify more complicated models, for example the AR(p) model,

$$
\begin{aligned}
y_t &= \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + u_t \\
y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - ... - \phi_p y_{t-p} &= u_t \\
(1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p) y_t &= u_t \\
\Phi_p(L) y_t &= u_t \text{ where } \Phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p
\end{aligned}
$$

And the ARMA(p,q) model,

$$
\begin{aligned}
y_t &= \underbrace{\phi_1 y_{t-1} \ldots + \phi_p y_{t-p}}_{AR(p)} + \underbrace{u_t + \theta_1 u_{t-1} + \ldots + \theta_q u_{t-q}}_{MA(q)} \\
y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \ldots - \phi_p y_{t-p} &= u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \ldots + \theta_q u_{t-q} \\
\left(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p\right) y_t &= \left(1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q\right) u_t \\
\Phi_p(L) y_t &= \Theta_q(L) u_t \\
\text{Where } \Phi_p(L) &= \left(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p\right), \\
\text{and } \Theta_q &= \left(1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q\right).
\end{aligned}
$$

## Inverse Lag Polynomial

Using lag polynomials also allows us to 'switch' between representations, such as AR (auto-regressive) and MA (moving-average). We can do this by *inverting* the lag polynomial. Recall in the AR(1) case

$$\Phi(L)y_t = u_t \text{ where } \Phi(L) = 1 - \phi L$$

Well if we want to get the MA representation ($y_t$ in terms of the lags of $u_t$), we simply invert $\Phi(L)$,

$$y_t = \{\Phi(L)\}^{-1} u_t \text{ where } \Phi(L) = 1 - \phi L$$

### Conditions for Invertibility

In order to be able to invert a lag polynomial we *must have roots of the polynomial* ***outside of the unit circle***.

*If the roots of a polynomial are inside the unit circle then the coefficients do not decay sufficiently fast for the error term to be a well-defined random variable with finite variance. More technically when the roots are inside the unit circle the infinite sum does not converge in mean square.*

Note that this only really matters for $AR$ model, since as long as an $MA$ model is finite then you can **always find an equivalent and invertible MA polynomial**. Of course you still need to check roots for both models though; it is just that if an MA polynomial is not invertible you can always find an equivalent model that is invertible.

### Finding Roots of Lag Polynomial

The roots of the lag polynomial are very simply just the values of $L$ when $\Phi(L) = 0$. We will consider two very simple cases for finding these.

(1) First Order Difference Equation

$$\begin{aligned} y_t &= \phi y_{t-1} + u_t \\ (1 - \phi L)y_t &= u_t \\ \Phi(L)y_t &= u_t \end{aligned}$$

So we have the lag polynomial,

$$\Phi(L) = 1 - \phi L$$

Which has roots when,

$$\Phi(L) = 0 \Rightarrow 1 - \phi L = 0 \Rightarrow L = \frac{1}{\phi}$$

Hence the roots of this polynomial are at $L = \frac{1}{\phi}$.

(2) Second Order Difference Equation

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \omega_t$$

Using lags,

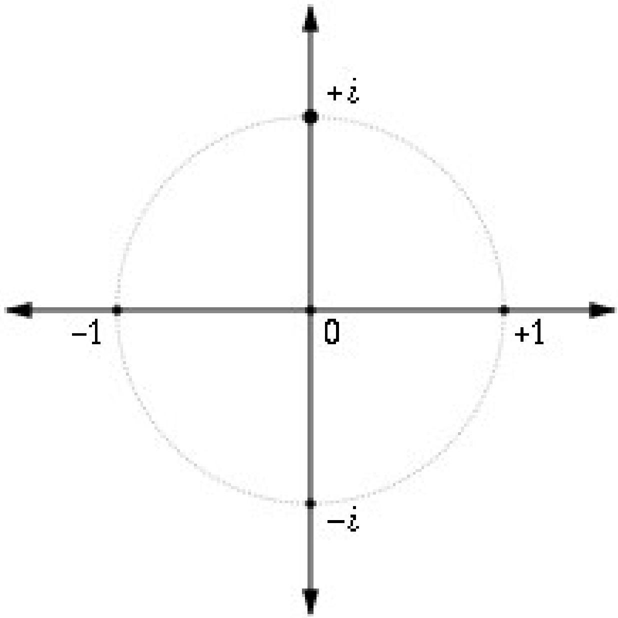
$$\left(1 - \phi_1 L - \phi_2 L^2\right) y_t = \omega_t$$

Find roots of lag polynomial, $1 - \phi_1 L - \phi_2 L^2 = 0$, using,

$$L = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

**The Unit Circle**

To be able to invert a polynomial it must be the case that **all** the polynomials roots are outside of the unit circle. That is we need all the values of $L$ at $\Phi(L) = 0$ to be outside the unit circle.




An imaginary number $bi$ is outside the unit-circle if $|b| > 1$.

A real number $a$ is outside the unit-circle if $|a| > 1$.

A number with a real and imaginary part $a + bi$ is outside the unit circle if $\sqrt{b^2 + a^2} > 1$

**Inverting the Lag Polynomial in the AR(1) case**

$$
\begin{aligned}
y_t &= \phi y_{t-1} + u_t \\
(1 - \phi L) y_t &= u_t \\
\Phi(L) y_t &= u_t
\end{aligned}
$$

We can only invert the polynomial $\Phi(L)$ if its roots are outside of the unit circle. The roots of the lag polynomial, that is the value of $L$ when $\Phi(L) = 0$, are given by,

$$0 = \Phi(L) \Rightarrow 0 = 1 - \phi L \Rightarrow L = \frac{1}{\phi}$$

The roots being outside unit circle requires that $|L| > 1$, hence the lag polynomial $\Phi(L)$ is only invertible in the case in which $|\phi| < 1$.

Supposing that we are in the case in which $|\phi| < 1$, therefore,

$$y_t = [\Phi(L)]^{-1} u_t$$

Where,

$$[\Phi(L)]^{-1} = \frac{1}{1 - \phi L}$$

## Inverse Lag Polynomial: Infinite Sum

We can always turn an inverted lag polynomial into an infinite sum lag polynomial. This is useful when switching between forms, such as from AR to MA, since leaving an inverted lag polynomial has little meaning to us.

What this means is, if we consider a pure AR(p) model,

$$\Phi_p(L)x_t = \varepsilon_t$$

We can also write this as,

$$x_t = \{\Phi_p(L)\}^{-1}\varepsilon_t$$

And we can write this inverted lag polynomial as an infinite lag polynomial, such that we get an MA$\infty$ model,

$$x_t = \Psi_\infty(L)\varepsilon_t$$

In other words it is always the case that

$$\{\Phi_p(L)\}^{-1} = \frac{1}{1+\phi_1 L+\phi_2 L^2 ...+\phi_p L^p} = 1+\psi_1 L+\psi_2 L^2+... = \Psi_\infty(L)$$

### Proof

The proof of this really just comes from the law for infinite geometric sums stated above,

$$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r} \text{ for } |r|<1$$

This law shows that an inverse polynomial of degree one can be written as an infinite sum. This implies that in our AR(1) case,

$$\frac{1}{1-\phi} = \sum_{k=0}^{\infty}\phi^k = 1+\phi+\phi^2+\phi^3+... \text{ for } |\phi|<1$$

In the case of a finite polynomial of degree $p$ we just need to write it in its factorised form,

$$\Phi_p(L) = 1+\phi_1 L+\phi_2 L^2+...\phi_p L^p = \prod_{i=1}^{p}\left(1-\frac{L}{r_i}\right)$$

Where $r_1, ..., r_p$ are the roots of the polynomial.

Having written the $p^{th}$ degree polynomial as a product of first degree polynomials, invert and apply the law from above,

$$\Big\{\Phi_p(L)\Big\}^{-1} = \prod_{i=1}^{p}\left(1-\frac{L}{r_i}\right)^{-1} = \prod_{i=1}^{p}\left(\sum_{j=0}^{\infty}\left(\frac{L}{r_i}\right)^j\right) = \sum_{j=0}^{\infty}\psi_j L^j = 1+\psi_1 L+\psi_2 L^2+... = \Psi_\infty(L)$$

Which holds if $|r_i|>1$ for all $i=1,...,p$ (if the roots are outside of the unit circle).

Hence,

$$\{\Phi_p(L)\}^{-1} = \frac{1}{1+\phi_1 L+\phi_2 L^2 ...+\phi_p L^p} = 1+\psi_1 L+\psi_2 L^2+... = \Psi_\infty(L)$$

**Example: General Case**

We have already shown the AR(1) case, so lets consider the $AR(2)$ case,

$$\begin{aligned}
y_t &= \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t \\
y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} &= u_t \\
(1 - \phi_1 L - \phi_2 L^2) y_t &= u_t \\
y_t &= \frac{1}{1 - \phi_1 L - \phi_2 L^2} u_t \\
&= \Big\{ \Phi_2(L) \Big\}^{-1} u_t
\end{aligned}$$

We now know that is possible to write this model as an MA$\infty$ model by writing the inverted polynomial of degree two as an infinite geometric sum.

$$\Big\{ \Phi_2(L) \Big\}^{-1} = \frac{1}{1 - \phi_1 L - \phi_2 L^2} = \psi_0 + \psi_1 L + \psi_2 L + \psi_3 L^3 + ... = \Psi_\infty(L)$$

To calculate the values of $\phi_1, \phi_2$, etc we know it will the case that,

$$\left(1 - \phi_1 L - \phi_2 L^2\right) \left(\psi_0 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \ldots\right) = 1$$

Or perhaps more obviously that,

$$\left(1 - \phi_1 L - \phi_2 L^2\right) \left(\psi_0 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \ldots\right) = 1 + 0L + 0L^2 + 0L^3 + \ldots$$

By multiplying out the brackets and matching up coefficients of $L$ we find that,

$$\begin{aligned}
&\Rightarrow \psi_0 = 1 \\
&\Rightarrow \psi_1 - \phi_1 \psi_0 = 0 \\
&\Rightarrow \psi_2 - \psi_1 \phi_1 - \phi_2 \psi_0 = 0 \\
&\Rightarrow \psi_3 - \psi_2 \phi_1 - \psi_1 \phi_2 = 0 \\
&\qquad \ldots \\
&\Rightarrow \psi_j - \psi_{j-1} \phi_1 - \psi_{j-2} \phi_2 = 0
\end{aligned}$$

**Example: Specific Case**

Now we can consider a specific example. Take this ARMA(2,2) model and express it into MA form,

$$x_t = 1.1 x_{t-1} - 0.8 x_{t-2} + u_t - 1.7 u_{t-1} + 0.72 u_{t-2}$$

We can begin by writing it using lag polynomials,

$$\left(1 - 1.1L + 0.8L^2\right) x_t = \left(1 - 1.7L + 0.72L^2\right) u_t$$

Of course before we try and invert the polynomial on $x_t$ we need to check that it can be inverted: that the roots are outside of the unit circle. Given that, in this case, they are, we can write it as,

$$x_t = \frac{\left(1 - 1.7L + 0.72L^2\right)}{(1 - 1.1L + 0.8L^2)} u_t$$

Where we know it will be the case that,

$$\begin{aligned}
\frac{\left(1 - 1.7L + 0.72L^2\right)}{(1 - 1.1L + 0.8L^2)} &= \left(\delta_0 + \delta_1 L + \delta_2 L^2 + \delta_3 L^3 + \ldots\right) \\
\left(1 - 1.7L + 0.72L^2\right) &= \left(1 - 1.1L + 0.8L^2\right) \left(\delta_0 + \delta_1 L + \delta_2 L^2 + \delta_3 L^3 + \ldots\right)
\end{aligned}$$

Multiplying out the brackets on the RHS,

$$\begin{aligned}
&= \delta_0 \quad + \delta_1 L + \quad\ \delta_2 L^2 \quad\ + \delta_3 L^3 + \dots \\
&\qquad - 1.1\delta_0 L - 1.1\delta_1 L^2 - 1.1\delta_2 L^3 + \dots \\
&\qquad\qquad\qquad + 0.8\delta_0 L^2 + 0.8\delta_1 L^3 + \dots \\
&= \delta_0 + (\delta_1 - 1.1\delta_0)\, L + (\delta_2 - 1.1\delta_1 + 0.8\delta_0)\, L^2 + (\delta_3 - 1.1\delta_2 + 0.8\delta_1)\, L^3 + ...
\end{aligned}$$

Finally we can set these coefficients equal to the coefficients that we know from the LHS,

$$\begin{aligned}
\delta_0 &= 1 \\
\delta_1 - 1.1\delta_0 &= -1.7 \\
\delta_2 - 1.1\delta_1 + 0.8\delta_0 &= 0.72 \\
\delta_3 - 1.1\delta_2 + 0.8\delta_1 &= 0
\end{aligned}$$

And so this implies,

$$\begin{aligned}
&\delta_0 = 1 \\
&\Rightarrow \delta_1 = -0.6 \\
&\Rightarrow \delta_2 = -0.74 \\
&\Rightarrow \delta_3 = -1.614 \\
&\Rightarrow \delta_j = 1.1\delta_{j-1} - 0.8\delta_{j-2} \;\; (j > 2)
\end{aligned}$$

# AR(p)

AR(p) is an autoregressive model of order $p$. That is we model $y_t$ as a function of its last $p$ lags $y_{t-1}, ..., y_{t-p}$, a constant $\alpha$, and an error $u_t$.

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + u_t$$

## Usefulness of the Model

- Forecasting,
- Stochastic properties of the data,
- Modelling time series.

## Stationary AR(1)

We will studies the properties of the stationary AR(1) process,

$$y_t = \alpha + \phi y_{t-1} + u_t \ , \ u_t \sim iid\left(0, \sigma^2\right)$$

For simplicity we omit the constant $\alpha$, but it could be included. When $\alpha = 0$ the model above is,

$$y_t = \phi y_{t-1} + u_t$$

Using the fact that $y_{t-i} = \phi y_{t-i-1} + u_{t-i}$ and backward substitution, the AR(1) model can be represented as,

$$\begin{aligned}
y_t &= \phi y_{t-1} + u_t \\
&\quad y_{t-1} = \phi y_{t-2} + u_{t-1} \\
y_t &= \phi\left(\phi y_{t-2} + u_{t-1}\right) + u_t \\
&= \phi^2 y_{t-2} + \phi u_{t-1} + u_t \\
&= \phi^2\left(\phi y_{t-3} + u_{t-2}\right) + \phi u_{t-1} + u_t \\
&= \phi^3 y_{t-3} + \phi^2 u_{t-2} + \phi u_{t-1} + u_t \\
&\ldots \\
y_t &= \phi^h y_{t-h} + \sum_{i=0}^{h-1} \phi^i u_{t-i} \\
y_t &= \phi^t y_0 + \sum_{i=0}^{t-1} \phi^i u_{t-i}
\end{aligned}$$

### Requirements for Stationarity

For an AR(1) to be stationary we require,

(1) $|\phi| < 1$,

(2) $E\left[y_0\right] = 0, \ \mathrm{Var}\left(y_0\right) = \frac{\sigma^2}{1-\phi^2}$.

**Expectation**

$$E\left[y_t\right] = \phi^t E\left[y_0\right] + \sum_{j=0}^{t-1} \phi^j E\left[u_{t-j}\right] = 0$$

**Variance**

$$\begin{aligned}
\operatorname{Var}\left(y_t\right) &= \phi^{2t} \operatorname{Var}\left(y_0\right) + \operatorname{Var}\left(\sum_{j=0}^{t-1} \phi^j u_{t-j}\right) \\
&= \phi^{2t} \operatorname{Var}\left(y_0\right) + \sum_{j=0}^{t-1} \phi^2 \operatorname{Var}\left(u_{t-j}\right) + 2 \sum_{j=0}^{t-2} \sum_{i=j+1}^{t-1} \operatorname{Cov}\left(u_{t-j}, u_{t-i}\right) \\
&= \phi^{2t}\left(\frac{\sigma^2}{1-\phi^2}\right) + \sigma^2 \sum_{j=0}^{t-1} \phi^{2j} \\
&= \phi^{2t}\left(\frac{\sigma^2}{1-\phi^2}\right) + \sigma^2\left(\frac{1-\phi^{2t}}{1-\phi^2}\right) \\
&= \frac{\sigma^2}{1-\phi^2}
\end{aligned}$$

**Covariance**

$$\begin{aligned}
\operatorname{Cov}\left(y_t, y_{t-h}\right) &= \operatorname{Cov}\left(\phi^h y_{t-h} + \sum_{j=0}^{h-1} \phi^j u_{t-j},\ y_{t-h}\right) \\
&= \phi^h \operatorname{Cov}\left(y_{t-h}, y_{t-h}\right) + \underbrace{\operatorname{Cov}\left(\sum_{j=0}^{h-1} \phi^j u_{t-j},\ y_{t-h}\right)}_{=\operatorname{Cov}(u_t+\phi u_{t-1}+\ldots+\phi^{h-1}u_{t-h+1})=0} \\
&= \phi^h \operatorname{Var}\left(y_{t-h},\ y_{t-h}\right) \\
&= \phi^h\left(\frac{\sigma^2}{1-\phi^2}\right)
\end{aligned}$$

In case it isn't obvious why $\operatorname{Cov}(u_t+\phi u_{t-1}+\ldots+\phi^{h-1}u_{t-h+1}) = 0$, this is because $y_{t-h}$ is only correlated with error terms that come temporally *before* it, that is before period $(t-h)$. In our case here we are considering error terms from period $t-h+1$ to period $t$, and notice that both of these periods are temporally *after* $t-h$, hence there is no covariances between $y_{t-h}$ and these error terms.

**Autocorrelation Function**

$$\begin{aligned}
\operatorname{Corr}\left(y_t, y_{t-h}\right) &= \frac{\operatorname{Cov}\left(y_t, y_{t-h}\right)}{\operatorname{Var}\left(y_t\right)} \\
&= \frac{\phi^h\left(\frac{\sigma^2}{1-\phi^2}\right)}{\left(\frac{\sigma^2}{1-\phi^2}\right)} \\
&= \phi^h
\end{aligned}$$

**Converting AR(1) to MA**

$$
\begin{aligned}
y_t &= \phi y_{t-1} + u_t \\
y_t - \phi y_{t-1} &= u_t \\
(1-\phi L) y_t &= u_t \\
y_t &= \frac{1}{1-\phi L} u_t \\
&\quad \text{Where } \frac{1}{1-\phi L} = \left(1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + \ldots\right) \\
y_t &= \sum_{j=0}^{\infty} \phi^j L^j u_t \\
y_t &= \sum_{j=0}^{\infty} \phi^j u_{t-j}
\end{aligned}
$$

**Wold Decomposition**

Recall the theorem,

**Theorem** *(The Wold Decomposition)*

If $x_t$ is stationary and non-deterministic, then

$$x_t = \sum_{j=0}^{\infty} \Psi_j u_{t-j} + z_t = \Psi(L) u_t + z_t$$

Where,

- $\Psi_0 = 1$ and $\sum_{j=0}^{\infty} \Psi_j^2 < \infty$
- $u_t$ is $wn(0, \sigma^2)$
- $z_t$ is deterministic
- $\text{Cov}(u_t, z_t) = 0$ for all $s$ and $t$
- $\Psi_j$ and $u_t$ are unique

In the case of the stationary AR(1) process (with no constant) the Wold Decomposition is

$$y_t = \sum_{j=0}^{\infty} \phi^j u_{t-j}$$

Which is also the same as the MA representation.

# ARMA(p,q)

$$y_t = \underbrace{\phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p}}_{AR(p)} + \underbrace{\alpha}_{constant} + \underbrace{u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \ldots + \theta_q u_{t-q}}_{MA(q)}$$

## Lag Polynomials

We can write an ARMA(p,q) using lag polynomials as,

$$\begin{aligned} y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \ldots - \phi_p y_{t-p} &= \alpha + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \ldots + \theta_q u_{t-q} \\ \left(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p\right) y_t &= \alpha + \left(1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q\right) u_t \\ \Phi_p(L) y_t &= \alpha + \Theta_q(L) u_t \end{aligned}$$

Where we define $\Phi_p(L) = \left(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p\right)$ and $\Theta_q(L) = \left(1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q\right)$.

## MA Representation

**Must** check for **stationarity** of the polynomial we are going to invert, in this case $\Phi_p(L)$, before trying to invert it. This is done by checking that the roots of the polynomial $\Phi_p(L)$ are outside of the unit circle.

$$\begin{aligned} y_t &= \{\Phi_p(L)\}^{-1} \alpha + \{\Phi_p(L)\}^{-1} \Theta_q(L) u_t \\ &= \alpha_{MA} + \{\Phi_p(L)\}^{-1} \Theta_q(L) u_t \end{aligned}$$

Where we know that,

$$\{\Phi_p(L)\}^{-1} \Theta_q(L) = \frac{1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q}{1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p} = 1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \ldots$$

The $\alpha_{MA}$ term is just a constant, since the Lag operator passes through constants,

$$\{\Phi_p(L)\}^{-1} \alpha = \frac{1}{1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p} \alpha = \{\Phi_p(1)\}^{-1} \alpha = \frac{\alpha}{1 - \phi_1 - \phi_2 - \ldots - \phi_p} = \alpha_{MA}$$

## Wold Decomposition

Recalling the theorem from above, for a stationary and non-deterministic $y_t$, then

$$y_t = \sum_{j=0}^{\infty} \Psi_j u_{t-j} + z_t = \Psi(L) u_t + z_t$$

In the ARMA(p,q) case, the Wold Decomposition is,

$$y_t = \{\Phi_p(L)\}^{-1} \Theta_q(L) u_t + \{\Phi_p(L)\}^{-1} \alpha$$

Which is exactly the MA representation.

## AR Representation

**Must** check for **stationarity** of the polynomial we are going to invert, in this case $\Theta_p qL)$, before trying to invert it. This is done by checking that the roots of the polynomial $\Theta_q(L)$ are outside of the unit circle. Note, however, that even if this MA polynomial is non-invertible we can find an equivalent and invertible MA polynomial since it is finite.

$$\begin{aligned}\{\Theta_q(L)\}^{-1}\Phi_p(L)y_t - \{\Theta_q(L)\}^{-1}\alpha &= u_t \\ \{\Theta_q(L)\}^{-1}\Phi_p(L)y_t - a_{AR} &= u_t\end{aligned}$$

Where we know that,

$$\{\theta_q(L)\}^{-1}\Phi_p(L) = \frac{1-\phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p}{1-\theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q} = 1 + \Gamma_1 L + \Gamma_2 L^2 + \Gamma_3 L^3 + \ldots$$

The $\alpha_{AR}$ term is just a constant, since the Lag operator passes through constants

$$\{\Theta_q(L)\}^{-1}\alpha = \frac{1}{1-\theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q}\alpha = \{\Theta_q(1)\}^{-1}\alpha = \frac{\alpha}{1-\theta_1-\theta_2-\ldots-\theta_q} = \alpha_{AR}$$

## Autocorrelation Function

Consider a stationary ARMA(1,1), that is where $|\phi| < 1,\ |\theta| < 1,\ \phi \neq 0$,

$$x_t = \phi x_{t-1} + u_t - \theta u_{t-1} \text{ with } u_t \sim iid(0,\sigma^2)$$

Using lag polynomials this can be rewritten as,

$$\begin{aligned}x_t - \phi x_{t-1} &= u_t - \theta u_{t-1} \\ (1-\phi L)x_t &= (1-\theta L)u_t \\ x_t &= \frac{1-\theta L}{1-\phi L}u_t \\ &= (1-\theta L)\frac{1}{1-\phi L}u_t\end{aligned}$$

Recall that,

$$\frac{1}{1-\phi L} = 1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + ...$$

Hence we can write the ARMA(1,1) as,

$$\begin{aligned}x_t &= (1-\theta L)\frac{1}{1-\phi L}u_t \\ &= (1-\theta L)(1+\phi L+\phi^2 L^2+\ldots)u_t \\ &= (1+\phi L+\phi^2 L^2+\ldots \\ &\qquad -\theta L - \theta\phi L^2 - \theta\phi^2 L^3 + \ldots)u_t \\ &= (1+\phi L - \theta L + \phi^2 L^2 - \theta\phi L^2 + \ldots)u_t \\ &= u_t + (\phi-\theta)u_{t-1} + (\phi-\theta)\phi u_{t-2} + \ldots \\ x_t &= u_t + (\phi-\theta)\sum_{j=1}^{\infty}\phi^{j-1}u_{t-j}\end{aligned}$$

This form above will be useful in showing the autocovariance function.

The autocovariance function is found starting from,

$$x_t = \phi x_{t-1} + u_t - \theta u_{t-1}$$

Then multiplying through by $x_{t-h}$ and taking expectations,

$$\begin{aligned} x_t x_{t-h} &= \phi x_{t-1} x_{t-h} + u_t x_{t-h} - \theta u_{t-1} x_{t-h} \\ E\left[x_t x_{t-h}\right] &= \phi E\left[x_{t-1} x_{t-h}\right] + E\left[u_t x_{t-h}\right] - \theta E\left[u_{t-1} x_{t-h}\right] \end{aligned}$$

Given that $u_t \sim iid(0, \sigma^2)$, we know from the MA form of the ARMA(1,1) that $E(x_t) = 0$,

$$E[x_t] = E[u_t] + (\phi - \theta) \sum_{j=1}^{\infty} \phi^{j-1} E[u_{t-j}] = 0$$

This implies that the autocovariance which usually is given by $\gamma_x(h) = E[x_t x_{t-h}] - E[x_t]E[x_{t-h}]$ is given by $\gamma_x(h) = E[x_t x_{t-h}]$ instead. Therefore,

$$\gamma_x(h) = \phi\gamma_x(h-1) + E\left[u_t x_{t-h}\right] - \theta E\left[u_{t-1} x_{t-h}\right]$$

And so,

$$\begin{aligned} h = 0 : \gamma_x(0) &= \phi\gamma_x(1) + E\left[u_t x_t\right] - \theta E\left[u_{t-1} x_t\right] \\ h = 1 : \gamma_x(1) &= \phi\gamma_x(0) + E\left[u_t x_{t-1}\right] - \theta E\left[u_{t-1} x_{t-1}\right] \\ h \geq 2 : \gamma_x(h) &= \phi\gamma_x(h-1) + E\left[u_t x_{t-h}\right] - \theta E\left[u_{t-1} x_{t-h}\right] \end{aligned}$$

Which gives, using the fact that $x_t = u_t + (\phi - \theta)\sum_{j=1}^{\infty} \phi^{j-1} u_{t-j}$ and $x_t - 1 = u_t - 1 + (\phi - \theta)\sum_{j=1}^{\infty} \phi^{j-1} u_{t-1-j}$ from earlier,

$$\begin{aligned} h = 0 : \gamma_x(0) &= \phi\gamma_x(1) + E\left[u_t x_t\right] - \theta E\left[u_{t-1} x_t\right] \\ &= \phi\gamma_x(1) + E\left[u_t^2 + (\phi - \theta)\sum_{j=1}^{\infty} \phi^{j-1} u_{t-j} u_t\right] - \theta E\left[u_{t-1} u_t + (\phi - \theta)\sum_{j=1}^{\infty} \phi^{j-1} u_{t-j} u_{t-1}\right] \\ \gamma_x(0) &= \phi\gamma_x(1) + \sigma^2 - \theta(\phi - \theta)\sigma^2 \end{aligned}$$

$$\begin{aligned} h = 1 : \gamma_x(1) &= \phi\gamma_x(0) + E\left[u_t x_{t-1}\right] - \theta E\left[u_{t-1} x_{t-1}\right] \\ &= \phi\gamma_x(0) + E\left[u_t u_{t-1} + (\phi - \theta)\sum_{j=1}^{\infty} \phi^{j-1} u_{t-1-j} u_t\right] - \theta E\left[u_{t-1} u_{t-1} + (\phi - \theta)\sum_{j=1}^{\infty} \phi^{j-1} u_{t-1-j} u_{t-1}\right] \\ \gamma_x(1) &= \phi\gamma_x(0) + 0 - \theta\sigma^2 \end{aligned}$$

$$\begin{aligned} h \geq 2 : \gamma_x(h) &= \phi\gamma_x(h-1) + E\left[u_t x_{t-h}\right] - \theta E\left[u_{t-1t-h}\right] \\ \gamma_x(h) &= \phi\gamma_x(h-1) \end{aligned}$$

Giving the autocorrelation function,

$$\rho_x(h) = \begin{cases} 1 & h = 0 \\ \frac{(\phi-\theta)(1-\phi\theta)}{1+\theta^2-2\phi\theta} & h = 1 \\ \phi\rho_x(h-1) & h = 2 \end{cases}$$

# ARDL(p,r)

$$y_t = \underbrace{\phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p}}_{AR(p)} + \underbrace{\mu}_{constant} + \underbrace{\gamma_0 x_t + \gamma_1 x_{t-1} + \ldots + \gamma_r x_{t-r}}_{DL(r)} + \underbrace{\varepsilon_t}_{error}$$

## Static Model

$$y_t = \mu + \gamma_0 x_t + \varepsilon_t$$

We call this model static since all parameters are contemporaneous.

If $\varepsilon_t$ iid, $E[\varepsilon_t \mid x_t, x_{t-1}, \ldots] = 0$, and $\operatorname{Var}(\varepsilon_t \mid x_t, x_{t-1}, \cdots) = \sigma_u^2$ then we can use standard inference.

But if $\varepsilon_t$ is temporally dependent (depends on their past values - hence are not iid) then while estimator is still consistent, variance gets messy,

We can either:

(1) Correct for Temporal Dependence,

(2) Model it (Dynamic Model).

## Dynamic Model

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \gamma_0 x_t + \gamma_1 x_{t-1} + \ldots + \gamma_r x_{t-r} + \varepsilon_t$$

Often adding just one lag is enough for the errors to now be iid ( $\varepsilon_t$ iid). If this is the case and if it is also the case that $E[\varepsilon_t \mid x_t, x_{t-1}, \ldots] = 0$ and $\operatorname{Var}(\varepsilon_t \mid x_t, x_{t-1}, \ldots) = \sigma_u^2$ then we get standard inference.

Note that although the same term $\varepsilon_t$ is used in these two models it is capturing different errors.

## Lag Form

$$\begin{aligned}(1 - \gamma_1 L - \ldots - \gamma_p L^p) y_t &= \mu + (\beta_0 + \beta_1 L + \cdots + \beta_r L^r) x_t + \varepsilon_t \\ C_p(L) y_t &= \mu + B_r(L) x_t + \varepsilon_t\end{aligned}$$

*Stablility*: The ARDL model is stable iff the roots of $C_p(L)$ are outside of the unit circle, and hence $C_p(L)$ is invertible.

$$\frac{B_r(L)}{C_p(L)} = D_\infty(L)$$

is convergent if the model is stable.

### DL Representation

$$\begin{aligned}y_t &= \frac{\mu}{C_p(L)} + \frac{B_r(L)}{C_p(L)} x_t + \frac{1}{C_p(L)} \varepsilon_t \\ y_t &= \alpha + D_\infty(L) x_t + u_t \\ y_t &= \alpha + \sum_{j=0}^{\infty} \delta_j x_{t-j} + u_t\end{aligned}$$

Note that the new error term $u_t = \frac{\varepsilon_1}{C_p(L)}$ is autocorrelated.

Also the constant $\alpha = \frac{\mu}{C_p(L)} = \frac{\mu}{C_p(1)}$ is finite.

## ECM Representation

$$\Delta y_t = (\gamma - 1)\left(y_{t-1} - \theta x_{t-1} - \tau\right) + \beta \Delta x_t + \varepsilon_t$$

**Example: ARDL(1,1)**

$$\begin{aligned}
y_t &= \mu + \gamma_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t \\
y_t - y_{t-1} &= \mu + (\gamma_1 - 1)\, y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t \\
\Delta y_t &= \mu + (\gamma_1 - 1)\, y_{t-1} + \beta_0 x_t - \beta_0 x_{t-1} + \beta_0 x_{t-1} + \beta_1 x_{t-1} + \varepsilon_t \\
\Delta y_t &= \mu + (\gamma_1 - 1)\, y_{t-1} + \beta_0 \Delta x_t + (\beta_0 + \beta_1)\, x_{t-1} + \varepsilon_t \\
\Delta y_t &= (\gamma_1 - 1)\left[y_{t-1} + \frac{(\beta_0 + \beta_1)}{(\gamma_1 - 1)} x_{t-1} + \frac{\mu}{(\gamma_1 - 1)}\right] + \beta_0 \Delta x_t + \varepsilon
\end{aligned}$$

The LR expectation of this given that $y_t$ and $x_t$ are stationary must be that:

$$\begin{aligned}
&E\left[y_{t-1} - \theta x_{t-1} - \tau\right] = 0 \\
&\text{or} \\
&E\left[y_{t-1} - \theta x_{t-1}\right] = \tau
\end{aligned}$$

**Example: ARDL(2,2)**

ECM Form

$$\begin{aligned}
y_t &= \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + u_t \\
y_t - y_{t-1} &= (\gamma_1 - 1)\, y_{t-1} + \gamma_2 y_{t-2} + \beta_0 x_t - \beta_0 x_{t-1} + \beta_1 x_{t-1} + \beta_0 x_{t-1} + \beta_2 x_{t-2} + u_t \\
\Delta y_t &= (\gamma_1 - 1)\, y_{t-1} + (\beta_1 + \beta_0)\, x_{t-1} + \gamma_2 y_{t-2} + \beta_0 \Delta x_t + \beta_2 x_{t-2} + u_t \\
\Delta y_t &= (\gamma_1 - 1)\left[y_{t-1} + \frac{(\beta_0 + \beta_1)}{(\gamma_1 - 1)} x_{t-1}\right] + \gamma_2 y_{t-2} + \beta_0 \Delta x_t + \beta_2 x_{t-2} + u_t
\end{aligned}$$

Assume $y(t)$ and $x(t)$ are stationary processes and let $E[x(t)] = \mu$. Given that $E[x_t] = \mu$ and $E[u_t] = 0$, and the fact that $x_t$ and $y_t$ are stationary, then it must be the case that,

$$E\left[\Delta y_t\right] = E\left[\Delta x_t\right] = 0 \text{ and } E\left[y_{t-i}\right] = E\left[y_t\right] \text{ and } E\left[x_{t-i}\right] = E\left[x_t\right] \text{ for all } i$$

$$\begin{aligned}
E\left[\Delta y_t\right] &= (\gamma_1 - 1)\left[E\left[y_{t-1}\right] + \frac{(\beta_0 + \beta_1)}{(r_1 - 1)} E\left[x_{t-1}\right]\right] + \gamma_2 E\left[y_{t-2}\right] + \beta_0 E\left[\Delta x_t\right] + \beta_2 E\left[x_{t-2}\right] + E\left[u_t\right] \\
0 &= (\gamma_1 - 1)\left[E\left[y_t\right] + \frac{(\beta_0 + \beta_1)}{(\gamma_1 - 1)} \mu\right] + \gamma_2 E\left[y_t\right] + \beta_0 (0) + \beta_2 \mu + (0) \\
0 &= (\gamma_1 - 1)\, E\left[y_t\right] + (\beta_0 + \beta_1)\, \mu + \gamma_2 E\left[y_t\right] + \beta_2{}^{\mu} \\
E\left[y_y\right] &= \frac{-(\beta_0 + \beta_1 + \beta_2)}{\gamma_2 + \gamma_1 - 1} \mu = \frac{(\beta_0 + \beta_1 + \beta_2)}{(1 - \gamma_2 - \gamma_1)} \mu
\end{aligned}$$

## Multipliers

Recall for this that we are considering an ARDL(p,r),

$$\begin{aligned} y_t &= \frac{\mu}{C_p(L)} + \frac{B_r(L)}{C_p(L)} x_t + \frac{1}{C_p(L)} \varepsilon_t \\ &= \frac{\mu}{C_p(L)} + D_\infty(L) x_t + \frac{1}{C_p(L)} \varepsilon_t \end{aligned}$$

Where,

$$\begin{aligned} C_p(L) &= 1 - \gamma_1 L - ... - \gamma_p L^p \\ B_r(L) &= \beta_0 + \beta_1 L + \cdots + \beta_r L^r \\ D_\infty(L) &= \sum_{j=0}^{\infty} \delta_j L^j = \delta_0 + \delta_1 L + \delta_2 L^2 + ... \end{aligned}$$

And further recall that,

$$\begin{aligned} A_q(L) &= \left(1 + \alpha_1 L + \alpha_2 L^2 + \ldots + \alpha_q L^q\right) \\ A_q(0) &= \left(1 + \alpha_1 0 + \alpha_2 0^2 + \ldots + \alpha_q 0^q\right) = 1 \\ A_q(1) &= (1 + \alpha_1 + \alpha_2 + \ldots + \alpha_q) \end{aligned}$$

### Impact (Contemporaneous) Multiplier

'How does today's $x_t$ influence today's $y_t$'

$$m_0 = \frac{\delta y_t}{\delta x_t} = D_\infty(0) = \delta_0 = \frac{B_r(0)}{C_p(0)} = \beta_0$$

### J-th lag Multiplier

'How does $x_t$ j days ago influence today's $y_t$'

$$m_j = \frac{\delta y_t}{\delta x_{t-j}} = \delta_j \neq \beta_j$$

### Total/Long-run Multiplier

'Total effect'

$$m_{total} = \sum_{j=0}^{\infty} m_j = D(1) = \sum_{j=0}^{\infty} \delta_j = \frac{B_r(1)}{C_p(1)}$$

## Transmission Effects

(Assume that $\delta_j \geq 0$)

### Mean Lag

'How concentrated (or diluted) the effect of $x_t$ on $y_t$ is'.

Earlier lags get a higher weight (remember $t-4$ is earlier than $t$).

$$Meanlag = \frac{\sum_{j=0}^{\infty} j\delta_j}{\sum_{j=0}^{\infty} \delta_j} = \frac{D'(1)}{D(1)} = \frac{B'(1)}{B(1)} - \frac{C'(1)}{C(1)}$$

Where

$$D'(1) = \left.\frac{dD(L)}{dL}\right|_{L=1}$$

### Median Lag

'The time when $y_t$ has accumulated 50% of the total effect'

$$Medianlag = \min_q \left\{ \frac{\sum_{j=0}^{q} \delta_j}{\sum_{j=0}^{\infty} \delta_j} \geq 0.5 \right\}$$

If you add $\delta_0, \delta_1, \delta_2$ and get an answer $> 0.5$ then the median is 2

# Asymptotics & Estimation: AR(p)

Here we will consider a stationary AR(1), that is a model,

$$y_t = \phi y_{t-1} + u_t$$

Where $|\phi| < 1$, $u_t \sim iid(0, \sigma_u^2)$, $E[u_t^4] = \mu_4 < \infty$, $E[y_0] = 0$, $\text{Var}(y_0) = \frac{\sigma^2}{1-\phi^2}$ and where $y_0$ is independent of all $u_t$.

All this reasoning should extend to the AR(p).

## Asymptotics: Sample Mean

$$\bar{y}_t = \frac{1}{T}\sum_{t=1}^{T} y_t = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{\infty}\phi^i u_{t-i}$$

### Consistency

We will use the Covariance-Stationary LLN in this case, which requires that we check that the autocovariances are absolutely summable.

$$\begin{aligned}
\gamma_h &= \phi^{|h|}\frac{\sigma_u^2}{1-\phi^2} \\
\sum_{h=0}^{\infty}|\gamma_h| &= \sum_{h=0}^{\infty}\left|\phi^{|h|}\frac{\sigma_u^2}{1-\phi^2}\right| = \sum_{h=0}^{\infty}|\phi|^{|h|}\frac{\sigma_u^2}{|1-\phi^2|} \\
&= \frac{\sigma_u^2}{|1-\phi^2|}\sum_{h=0}^{\infty}|\phi|^h = \frac{\sigma_u^2}{|1-\phi^2|}\left(1 + |\phi|^1 + |\phi|^2 + \ldots\right) \\
&= \frac{\sigma_u^2}{|1-\phi^2|}\frac{1}{1-|\phi|} < \infty
\end{aligned}$$

Having shown that they are, hence we can use the Covariance-Stationary LLN,

$$\bar{y}_t = \frac{1}{T}\sum_{t=1}^{T} y_t \xrightarrow{P} \mu = E\left[y_t\right] = 0$$

### Asymptotic Normality

Here we will use the Wold Decomposition CLT on the Wold Decomposition of the AR(1),

$$y_t = \sum_{i=0}^{\infty}\phi^i u_{t-i}$$

Given we have already assumed $u_t \sim iid(0, \sigma_u^2)$, we just need to check the condition that the coefficients on the error term are absolutely summable, that is,

$$\sum_{i=0}^{\infty}|\psi_i| < \infty$$

In our case,

$$\sum_{i=0}^{\infty} |\psi_j| = \sum_{i=0}^{\infty} |\phi^i| = 1 + |\phi| + |\phi^2| + |\phi^3| + ...$$
$$= \frac{1}{1-|\phi|} < \infty$$

Hence the Wold Decomposition CLT applies and so,

$$\sqrt{T}\big(\bar{y}_t - \mu\big) \xrightarrow{D} N\bigg(0,\ \sum_{h=-\infty}^{\infty} \gamma_h\bigg)$$

Where $\mu = 0$ and where $\sum_{h=-\infty}^{\infty} \gamma_h = \sigma_u^2 \Psi^2(1)$. Recall that in this case,

$$\Psi(L) = \frac{1}{1-\phi L}$$
$$\Psi(1) = \frac{1}{1-\phi}$$

Therefore,

$$\sum_{h=-\infty}^{\infty} \gamma_h = \sigma_u^2 \Psi^2(1) = \sigma_u^2 \bigg(\frac{1}{1-\phi}\bigg)^2$$

Using all of this and the Wold Decomposition CLT,

$$\sqrt{T}\big(\bar{y}_t\big) \xrightarrow{D} N\bigg(0,\ \sigma_u^2 \Big(\frac{1}{1-\phi}\Big)^2\bigg)$$

## Estimation: OLS

We estimate $\phi$ in the AR(1) model by OLS,

$$\widehat{\phi}_T = argmin \sum_{t=1}^{T} u_t^2 = argmin \sum_{t=1}^{T} \big(y_t - \phi y_{t-1}\big)$$

Giving the FOC,

$$0 = -2 \sum_{t=1}^{T} y_{t-1}\big(y_t - \widehat{\phi}_T y_{t-1}\big)$$
$$0 = \sum_{t=1}^{T} y_{t-1} y_t - \widehat{\phi}_T \sum_{t=1}^{T} y_{t-1}^2$$

Hence,

$$\widehat{\phi}_T = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2} = \frac{\sum_{t=1}^{T} y_{t-1}\left(\phi y_{t-1} + u_t\right)}{\sum_{t=1}^{T} y_{t-1}^2} = \phi + \frac{\sum_{t=1}^{T} y_{t-1} u_t}{\sum_{t=1}^{T} y_{t-1}^2}$$

## Asymptotics: OLS Estimator

### Consistency

$$\widehat{\phi}_T - \phi = \frac{T^{-1}\sum_{t=1}^T y_{t-1}u_t}{T^{-1}\sum_{t=1}^T y_{t-1}^2}$$

Consider the **numerator** to start, and notice that, where $\mathcal{I}_{t-1}$ is the information set we have at $t-1$, $E[y_{t-1}u_t \mid \mathcal{I}_{t-1}] = y_{t-1}E[u_t \mid \mathcal{I}_{t-1}] = 0$. We have a Martingale Difference Series (mds), so we will apply the mds LLN.

We just need to check condition $(b)$ of mds LLN for $p=1$ and $m_t = y_{t-1}u_t$,

$$\begin{aligned}
\lim_{T\to\infty}\max_{t\le T} E\left|y_{t-1}u_t\right|^2 &= \lim_{T\to\infty}\max_{t\le T} E\left[{y_{t-1}}^2{u_t}^2\right] \\
&= \lim_{T\to\infty}\max_{t\le T} = E\left[y_{t-1}^2 E\left(u_t^2 \mid \mathcal{I}_{t-1}\right)\right] \\
&= \frac{\sigma^2}{1-\phi^2}\sigma^2 = \frac{\sigma^4}{1-\phi^2} < \infty
\end{aligned}$$

The condition holds, so,

$$T^{-1}\sum_{t=1}^T y_{t-1}u_t \xrightarrow{P} 0$$

Now considering the **denominator**,

$$\begin{aligned}
T^{-1}\sum_{t=1^T}^T y_{t-1}^2 &= T^{-1}\left(y_0^2 + y_1^2 + \ldots + y_{T-1}^2\right) = T^{-1}\left(y_0^2 + y_1^2 + \ldots + y_{T-1}^2 + y_T^2 - y_T^2\right) \\
&= T^{-1}\sum_{t=1}^T y_t^2 + T^{-1}\left(y_0^2 - y_T^2\right)
\end{aligned}$$

And we can work out, simply by squaring the AR(1) model,

$$\begin{aligned}
\left\{y_t\right\}^2 &= \left\{\phi y_{t-1} + u_t\right\}^2 \\
y_t^2 &= \phi^2 y_{t-1}^2 + u_t^2 + 2\phi y_{t-1}u_t
\end{aligned}$$

Now substituting this in,

$$\begin{aligned}
T^{-1}\sum_{t=1}^T y_{t-1}^2 &= T^{-1}\sum_{t=1}^T\left(\phi^2 y_{t-1}^2 + u_t^2 + 2\phi y_{t-1}u_t\right) + T^{-1}\left(y_0^2 - y_T^2\right) \\
&= \phi^2 T^{-1}\sum_{t=1}^T y_{t-1}^2 + T^{-1}\sum_{t=1}^T\left(u_t^2 + 2\phi y_{t-1}u_t\right) + T^{-1}\left(y_0^2 - y_T^2\right) \\
T^{-1}\sum_{t=1}^T y_{t-1}^2 - \phi^2 T^{-1}\sum_{t=1}^T y_{t-1}^2 &= T^{-1}\sum_{t=1}^T u_t^2 + 2\phi T^{-1}\sum_{t=1}^T y_{t-1}u_t + T^{-1}\left(y_0^2 - y_T^2\right) \\
(1-\phi^2)T^{-1}\sum_{t=1}^T y_{t-1}^2 &= T^{-1}\sum_{t=1}^T u_t^2 + 2\phi T^{-1}\sum_{t=1}^T y_{t-1}u_t + T^{-1}\left(y_0^2 - y_T^2\right) \\
T^{-1}\sum_{t=1}^T y_{t-1}^2 &= \frac{1}{1-\phi^2}\left[T^{-1}\sum_{t=1}^T u_t^2 + 2\phi T^{-1}\sum_{t=1}^T y_{t-1}u_t + T^{-1}\left(y_0^2 - y_T^2\right)\right]
\end{aligned}$$

Now analysising each term,

$$T^{-1}\sum_{t=1}^{T}u_t^2 \xrightarrow{P} \sigma^2 \ (iid\ LLN)$$

$$T^{-1}\sum_{t=1}^{T}y_{t-1}u_t \xrightarrow{P} 0 \ (mds\ LLN)$$

$$T^{-1}\big(y_0^2 - y_T^2\big) \xrightarrow{P} 0$$

Hence the denominator,

$$T^{-1}\sum_{t=1}^{T}y_{t-1}^2 \xrightarrow{P} \frac{\sigma^2}{(1-\phi^2)}$$

So **overall**,

$$\widehat{\phi}_T = \phi + \frac{T^{-1}\sum_{t=1}^{T}y_{t-1}u_t}{T^{-1}\sum_{t=1}^{T}y_{t-1}^2} \xrightarrow{P} \phi + \frac{0}{\sigma^2/\,(1-\phi^2)} = \phi$$

**Asymptotic Normality**

$$\sqrt{T}\left(\widehat{\phi}_T - \phi\right) = \frac{T^{-1/2}\sum_{t=1}^{T}y_{t-1}u_t}{T^{-1}\sum_{t=1}^{T}y_{t-1}^2}$$

Starting with the **denominator**, from the argument above,

$$T^{-1}\sum_{t=1}^{T}y_{t-1}^2 \xrightarrow{P} \frac{\sigma^2}{1-\phi^2}$$

Then the **numerator**,

mds CLT applies (argumentation is tedious)

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}y_{t-1}u_t \xrightarrow{D} N\left(0, E\left[y_{t-1}^2u_t^2\right]\right) = N\left(0,\ \frac{\sigma^4}{1-\phi^2}\right)$$

And so **overall**,

$$\sqrt{T}\big(\widehat{\phi}_T - \phi\big) = \frac{T^{-1/2}\sum_{t=1}^{T}y_{t-1}u_t}{T^{-1}\sum_{t=1}^{T}y_{t-1}^2} \xrightarrow{D} \frac{N\left(0,\ \frac{\sigma^4}{1-\phi^2}\right)}{\frac{\sigma^2}{1-\phi^2}} = N\left(0,\ \frac{\frac{\sigma^4}{1-\phi^2}}{\left(\frac{\sigma^2}{1-\phi^2}\right)^2}\right) = N\Big(0,\ (1-\phi^2)\Big)$$

# Asymptotics, Estimation & Selection: ARMA(p,q)

## Asymptotics: Sample Mean

### Consistency

Use Covariance-Stationary LLN for a stationary ARMA model.

### Asymptotic Normality

Use the Wold Decomposition (the MA representation).

## Estimation

We get our estimates $\widehat{\phi}$, $\widehat{\theta}$, and $\widehat{\sigma}$, in the ARMA(1,1) model from the Yule-Walker equations,

$$\begin{aligned} h = 0 &: \gamma_x(0) = \phi\gamma_x(1) + \sigma^2 - \theta(\phi - \theta)\sigma^2 \\ h = 1 &: \gamma_x(1) = \phi\gamma_x(0) + 0 - \theta\sigma^2 \\ h \geq 2 &: \gamma_x(h) = \phi\gamma_x(h-1) \end{aligned}$$

The ARMA(p,q) is estimated by ML.

## Selection

How many lags should we use in the model? Having many lags implies a more fleixble model with less bias, but fewer lags means a lower variance...

(1) Stepwise Testing Down Procedure

- Start with some $p$ lags
  (1) Perform a t-test $H_0 \ : \ \phi_p = 0$
  (2) If we accept $H_0 \mid \phi_p = 0$ then repeat $p-1$ lags until we reject $H_0$ (as long as stationary and weakly dependent $\widehat{\phi}_i$ is asymptotically normal).
- Problem: $\widehat{\phi}_i$ could be significant by chance.

(2) Information Criteria

$$\min \ IC(k) = \log\left(\widehat{\sigma}_k\right) + k\frac{P(T)}{T}$$

- Where:
  - There are $\bar{k}$ alternative models, $M_1, \ldots, M_{\bar{k}}$ where $k = 1, \ldots, \bar{k}$ represent the number of parameters in the model,
  - $\widehat{\sigma}_k$ is the variance of the residuals of model $M_k$,
  - T is the sample size,
  - P(T) is a penalty for including too many lags.

# Asymptotics & Estimation: ARDL(p,r)

## Estimation

### With iid Errors

$$y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + \alpha y_{t-1} + \varepsilon_t$$

Where $\varepsilon_t \sim iid\ N\left(0, \sigma^2\right)$ , $x_t$ is stationary AR(1) and $\varepsilon_t$ of all past $x_t,\ x_{t-1},\ x_{t-2}, \ldots$ and past $y_{t-1},\ y_{t-2}, \ldots$

Use OLS to estimate,

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

### With Autocorrelated Errors

Consider ARDL $(1, 0)$

$$y_t = \beta_0 x_t + \gamma_1 y_{t-1} + \varepsilon_t$$

With $|\gamma_1| < 1$ and where $\varepsilon_t$ is autocorrelated.

Stability allows us to write,

$$y_t = \frac{\beta_0}{1-\gamma_1 L} x_t + \frac{\varepsilon_t}{1-\gamma_1 L}$$

Mean independence of $\varepsilon_t$ and regressors fails

$$\begin{aligned}
E\left(\varepsilon_p y_{t-1}\right) &= E\left[\varepsilon_t \frac{\beta_0}{1-\gamma_1 L} x_{t-1} + \varepsilon_t \frac{\varepsilon_{t-1}}{1-\gamma_1 L}\right] \\
&= E\left[\varepsilon_t \frac{\beta_0}{1-\gamma_1 L} x_{t-1}\right] + E\left[\varepsilon_t \frac{\varepsilon_{t-1}}{1-\gamma_1 L}\right] \\
&= 0 + E\left[\varepsilon_t \frac{\varepsilon_{t-1}}{1-\gamma_1 L}\right] = 0 + E\left[\varepsilon_t\left(\varepsilon_{t-1} + \gamma_1 \varepsilon_{t-2} + \gamma_1^2 \varepsilon_{t-3} + \ldots\right)\right] \neq 0
\end{aligned}$$

If the errors are AR(1), that is if $\varepsilon_t$ is autocorrelated such that $\varepsilon_t = \phi\varepsilon_{t-1} + u_t$ where $u_t \sim idd(0, \sigma_u^2)$, then we can solve this by adding more lags.

Here is our ARDL(1,0) model,

$$y_t = \beta x_t + \gamma_1 y_{t-1} + \varepsilon_t$$

Using the AR(1) autocorrelated error and by adding the lags $y_{t-2}$ and $x_{t-1}$,

$$\begin{aligned}
&\varepsilon_t = \phi\varepsilon_{t-1} + u_t \quad , \quad u_t \sim idd(0, \sigma_u^2) \\
&u_t = \varepsilon_t - \phi\varepsilon_{t-1} \quad , \quad \varepsilon_t = y_t - \beta x_t - \gamma_1 y_{t-1} \quad , \quad \varepsilon_{t-1} = y_{t-1} - \beta x_{t-1} - \gamma_1 y_{t-2} \\
&u_t = y_t - \beta x_t - \gamma_1 y_{t-1} - \phi\left[y_{t-1} - \beta x_{t-1} - \gamma_1 y_{t-2}\right]
\end{aligned}$$

Which means finally,

$$y_t = (\gamma_1 + \phi) y_{t-1} - \gamma_1 y_{t-2} + \beta x_t - \beta x_{t-1} + u_t$$

with $u_t \sim iid(0, \sigma_u^2)$.

If, however, the errors are MA(1) autocorrelated, that is if $\varepsilon_t = u_t + \phi u_{t-1}$ , $u_t \sim iid(0, \sigma_u^2)$ then no such trick can be used. Instead we can use IVs.

## Asymptotics

Study asymptotics with mds LLN and CLT.

# Time Trends

$$y_t = \alpha + \delta t + u_t \quad , \quad u_t \sim iid\left(0, \sigma_u^2\right)$$

## Stochastic Properties

### Expectation

$$E\left[y_t\right] = E\left[\alpha + \delta t + u_t\right] = \alpha + \delta t$$

### Variance

$$\mathrm{Var}\left(y_t\right) = E\left[\left\{y_t - E\left(y_t\right)\right\}^2\right] = E\left[u_t^2\right] = \sigma_u^2$$

### Covariance

$$\mathrm{Cov}\left(y_t, y_{t-h}\right) = E\left[\left\{y_t - E\left(y_t\right)\right\}\left\{y_{t-h} - E\left(y_{t-h}\right)\right\}\right] = E\left[u_t u_{t-h}\right] = 0$$

### Correlation

$$\mathrm{Corr}\left(y_t, y_{t-h}\right) = \frac{\mathrm{cov}\left(y_t, y_{t-h}\right)}{\sqrt{\mathrm{var}\left(y_t\right)}\sqrt{\mathrm{var}\left(y_{t-h}\right)}} = 0$$

## OLS Estimation

$$Y = XB + U \text{ where, } \underset{(T\times 1)}{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \end{pmatrix} \ , \ \underset{(T\times 2)}{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & T \end{pmatrix} \ , \ \underset{(2\times 1)}{B} = \begin{pmatrix} \alpha \\ \delta \end{pmatrix} \ , \ \underset{(T\times 1)}{U} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_t \end{pmatrix}$$

We can just use the usual OLS matrix estimator,

$$\underset{(2\times 1)}{\widehat{B}_T} = \left(X'X\right)^{-1} X'Y$$

And in our case,

$$\underset{(2\times 1)}{\widehat{B}_T} = \begin{pmatrix} \widehat{a}_T \\ \widehat{\delta}_T \end{pmatrix}$$

$$\left(\widehat{B}_T - B\right) = \begin{pmatrix} \widehat{a}_T - a \\ \widehat{\delta}_T - \delta \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^T 1 & \sum_{t=1}^T t \\ \sum_{t=1}^T & \sum_{t=1}^T t^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^T u_t \\ \sum_{t=1}^T t u_t \end{pmatrix}$$

Where,

$$X'X = \begin{bmatrix} \sum_{t=1}^T & \sum_{t=1}^T t \\ \sum_{t=1}^T t & \sum_{t=1}^T t^2 \end{bmatrix} = \begin{bmatrix} T & \frac{T(T+1)}{2} \\ \frac{T(T+1)}{2} & \frac{T(T+1)(2T+1)}{6} \end{bmatrix}$$

## Asymptotic Distribution

Need to first stabilise the denominator: $(X'X)^{-1}$

$$\gamma_T = \begin{pmatrix} T^{\frac{1}{2}} & 0 \\ 0 & T^{\frac{3}{2}} \end{pmatrix}$$

$$\gamma_T^{-1}(X'X)^{-1}\gamma_T^{-1} = \begin{pmatrix} T^{-1}T & T^{-2}\frac{T(T+1)}{2} \\ T^{-2}\frac{T(T+1)}{2} & T^{-3}\frac{T(T+1)(2T+1)}{6} \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix} := Q$$

Where $Q$ is invertible.

Now we need $\gamma_T$ 's to cancel out, this is how we do it,

$$\gamma_T\left(\widehat{B}_T - B\right) = \left\{\gamma_T^{-1}\left(X'X\right)\gamma_T^{-1}\right\}^{-1}\gamma_T^{-1}\left(X'U\right)$$

Hence we just need to consider,

$$\gamma_T(X'U) = \begin{bmatrix} T^{\frac{1}{2}} & 0 \\ 0 & T^{\frac{3}{2}} \end{bmatrix}^{-1}\begin{bmatrix} \sum_{t=1}^T u_t \\ \sum_{t=1}^T tu_t \end{bmatrix} = \begin{bmatrix} T^{-\frac{1}{2}}\sum_{t=1}^T u_t + 0\sum_{t=1}^T tu_t \\ 0\sum_{t=1}^T u_t + T^{-\frac{3}{2}}\sum_{t=1}^T tu_t \end{bmatrix} = \begin{bmatrix} T^{-\frac{1}{2}}\sum_{t=1}^T u_t \\ T^{-\frac{1}{2}}T^{-1}\sum_{t=1}^T tu_t \end{bmatrix}$$
$$= \begin{bmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^T u_t \\ \frac{1}{\sqrt{T}}\sum_{t=1}^T \frac{t}{T}u_t \end{bmatrix}$$

Row 1:

Given that $u_t \sim iid\left(0, \sigma^2\right)$ we can use the iid CLT so that,

$$\left(\frac{1}{\sqrt{T}}\right)\sum_{t=1}^T u_t \longrightarrow N\left(0, \sigma^2\right)$$

Row 2:

$$\text{Var}\left(\frac{t}{T}u_t\right) = \frac{t^2}{T^2}\sigma^2$$

hence we are dealing with heteroskedastic errors, which means we can't use the iid CLT but we can show this to be an mds instead,

$$E\left[\frac{t}{T}u_t \mid \mathcal{I}_{t-1}\right] = \frac{t}{T}E\left[u_t \mid \mathcal{I}_{t-1}\right] = 0$$

Hence we just need to check conditions $(i)$ and $(ii)'$ for the mds CLT, which hold (I'm not going to show it). For this row we get out as our answer,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^T \frac{t}{T}u_t \xrightarrow{D} N\left(0,\ \frac{\sigma^2}{3}\right)$$

Overall then,

$$\gamma_T(X'U) \xrightarrow{D} N(0,\ \sigma^2 Q)$$

NOT SURE IF THIS IS CORRECT??

# Unit Roots

## Hypothesis

We can test if the model,

$$y_t = \phi y_{t-1} + u_t$$

Has a unit root with the null hypothesis that it has a unit root against the alternative hypothesis that it is stationary,

$$H_0 : \phi = 1$$
$$H_1 : \phi < 1$$

Or equivalently,

$$\Delta y_t = \theta y_{t-1} + u_t \text{ for } \theta = \phi - 1$$
$$H_0 : \theta = 0$$
$$H_1 : \theta < 0$$

Note that to test for stationarity in the alternative we really need to test for $|\phi| < 1$ which implies $-1 < \phi < 1$ and therefore $-2 < \theta < 0$.

## Autocorrelation Functions of Random Walks

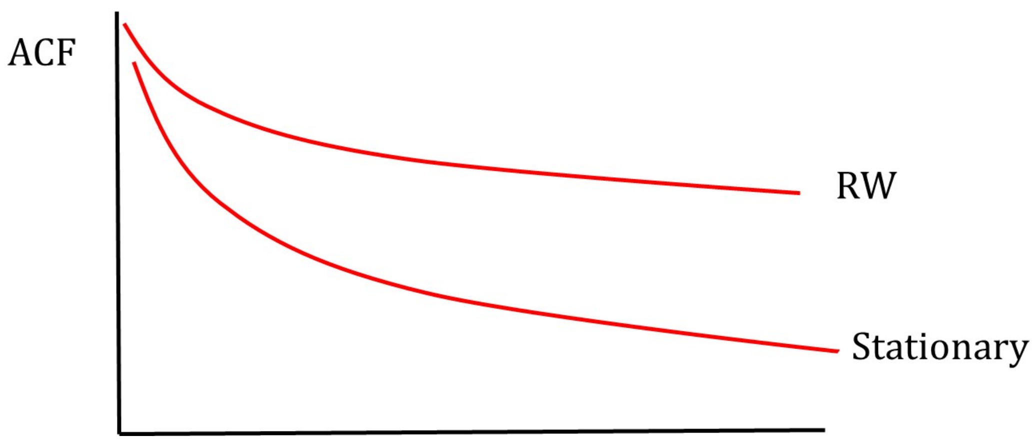



Figure 1: ACF for Random Walks vs Stationary AR(1)

Stationary trend reverts back to mean therefore ACF drops off quickly. Unit roots do not revert so this doesn't happen in the same way.

## Why a Random Walk is called a Unit Root: Lag Form

Random walk with drift: $y_t = \mu + y_{t-1} + u_t$

Lag Polynomial and its roots:

$$(1 - L)y_t = \mu + u_t$$
$$(1 - L) = 0$$
$$L = 1$$

Given non-stationary then roots of lag polynomial are on the unit circle, hence we have a 'unit root'.

## AR Representation

$$y_t = \mu + y_{t-1} + u_t \text{ with } u_t \sim \text{iid}\left(0, \sigma_u^2\right) \text{ and } y_0 = 0$$

## MA Representation

$$\begin{aligned}
y_t &= \mu + y_{t-1} + u_t \\
y_t &= \mu + (\mu + y_{t-2} + u_{t-1}) + u_t \\
y_t &= 2\mu + (\mu + y_{t-3} + u_{t-2}) + u_{t-1} + u_t \\
&\dots \\
y_t &= h\mu + y_{t-h} + u_{t-h+1} + u_{t-h+2} + \ldots + u_{t-1} + u_t \\
y_t &= h\mu + \sum_{j=0}^{h-1} u_{t-j} \\
\text{Letting } h = t \\
y_t &= t\mu + \sum_{j=0}^{t-1} u_{t-j} \\
y_t &= t\mu + u_t + u_{t-1} + \ldots + u_{t-t+2} + u_{t-t+1} \\
y_t &= t\mu + u_t + u_{t-1} + \ldots + u_2 + u_1 \\
y_t &= t\mu + \sum_{j=1}^{t} u_j
\end{aligned}$$

## Order of Integration

$y_t \sim I(1)$ : integrated of order 1 - difference once for stationarity.

$\Delta y_t \sim I(0)$ : stationary.

That is the first difference of an AR(1) random walk is stationary: $y_t - y_{t-1} = \Delta y_t = \mu + u_t$.

## Expectation

$$
\begin{aligned}
E\left[y_t\right] &= E\left[\mu t + \sum_{j=1}^{t} u_j\right] = \mu t + \sum_{j=1}^{t} E\left[u_j\right] = \mu t \\
E\left[y_t\right] &= \mu t
\end{aligned}
$$

## Variance

$$
\begin{aligned}
\operatorname{Var}\left(y_t\right) &= \operatorname{Var}\left(\mu t + \sum_{j=1}^{\prime} u_j\right) = \operatorname{Var}\left(\sum_{j=1}^{\prime} u_j\right) = \sum_{j=1}^{\prime} \operatorname{Var}\left(u_j\right) = {\sigma_u}^2 t \\
\operatorname{Var}\left(y_t\right) &= {\sigma_u}^2 t
\end{aligned}
$$

## Covariance

### Method 1

$$
\begin{aligned}
\operatorname{Cov}\left(y_t, y_{t-h}\right) &= \operatorname{Cov}\left(\mu h + y_{t-h} + \sum_{j=0}^{h-1} u_{t-j}, y_{t-h}\right) \\
\operatorname{Cov}\left(y_t, y_{t-h}\right) &= 0 + \operatorname{var}\left(y_{t-h}\right) + 0 \\
\operatorname{Cov}\left(y_t, y_{t-h}\right) &= (t-h)\sigma_u^2
\end{aligned}
$$

To better explain the jump from the first to the second line,

$$
\operatorname{Cov}\left(y_t, y_{t-h}\right) = \operatorname{Cov}\left(\mu h, y_{t-h}\right) + \operatorname{Cov}\left(y_{t-h}, y_{t-h}\right) + \operatorname{Cov}\left(\sum_{j=0}^{h-1} u_{t-j}, y_{t-h}\right)
$$

Where the first covariance is zero because $\mu h$ is a constant, and the last covariance is zero because $\left\{u_{t-j}\right\}_{j=0}^{h-1}$ is temporally *after* period $t-h$, and $y_t$ (or in this case $y_{t-h}$) is only affected by $u_t$'s that come before it.

### Method 2

$$
\begin{aligned}
&\operatorname{Cov}\left(y_t, y_s\right) = E\left[\left(y_t - Ey_t\right)\left(y_s - Ey_s\right)\right] \\
&y_t - Ey_t = y_t - \mu t = \sum_{j=1}^{t} u_j \\
&\operatorname{Cov}\left(y_t, y_s\right) = E\left[\left(\sum_{j=1}^{t} u_j\right)\left(\sum_{i=1}^{s} u_j\right)\right] \\
&E\left[u_j u_i\right] \neq 0 \text{ for i = j, hence} \\
&\sum_{j=1}^{t}\sum_{i=1}^{s} E\left[u_j u_i\right] = {\sigma_u}^2 \min\{s, t\}
\end{aligned}
$$

# Unit Roots (Nonstandard) Asymptotics & Estimation

For this section we will consider a Random Walk without drift.

$$y_t = y_{t-1} + u_t \text{ where } y_0 = 0$$
$$y_t = \sum_{j=1}^{t} u_j$$

## Asymptotics: Sample Mean

$$\bar{y}_t = \frac{1}{T}\sum_{t=1}^{T} y_t = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t} u_j$$

### Consistency

The sample has the expectation,

$$E\left[\bar{y}_t\right] = \frac{1}{T}\sum_{t=1}^{T} E\left[y_t\right] = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{t} E\left[u_i\right] = 0 = E[y_t]$$

Hence to show consistency we just need to show by MSC that the variance tends to zero in the limit,

$$\operatorname{Var}\left(\bar{y}_t\right) = \operatorname{Var}\left(\frac{1}{T}\sum_{t=1}^{T} y_t\right) = \frac{1}{T^2}\operatorname{Var}\left(\sum_{t=1}^{T} y_t\right) = \frac{1}{T^2}\left[\sum_{t=1}^{T}\operatorname{Var}\left(y_t\right) + 2\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\operatorname{Cov}\left(y_t, y_s\right)\right]$$

Where the **first term**,

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T}\operatorname{Var}\left(y_t\right) &= \frac{1}{T^2}\sum_{t=1}^{T}\operatorname{Var}\left(\sum_{j=1}^{t} u_j\right) = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{j=1}^{t}\operatorname{var}\left(u_j\right)\\ &= \frac{1}{T^2}\sum_{t=1}^{T} t\sigma_u^2 = \frac{\sigma_u^2}{T^2}\sum_{t=1}^{T} t = \frac{\sigma_u^2}{T^2}\frac{T(T+1)}{2} = \frac{\sigma_u^2(T+1)}{2T}\end{aligned}$$

And the **second term**,

$$\begin{aligned}\frac{2}{T^2}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\operatorname{Cov}\left(y_t, y_s\right) &= \frac{2}{T^2}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\sigma_u^2\min\{s,t\} \;\text{ where } s = t+1 \Rightarrow s > t \text{ hence,}\\ &= \frac{2\sigma_t^2}{T^2}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T} t = \frac{2\sigma_t^2}{T^2}\sum_{t=1}^{T-1} t\sum_{s=t+1}^{T} 1\\ &\sum_{s=1}^{T} 1 = T, \text{ but we have } t \text{ less numbers than that,}\\ &\text{hence } \sum_{s=t+1}^{T} 1 = (T-t)\\ &= \frac{2\sigma_t^2}{T^2}\sum_{t=1}^{T-1} t(T-t) = \frac{2\sigma_u^2}{T^2}\left[T\sum_{t=1}^{T} t - \sum_{t=1}^{T} t^2\right]\end{aligned}$$

We can go from $T-1$ to $T$ since when $t=T$ the sum is 0 anyway

$$\frac{2}{T^2}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\mathrm{Cov}\left(y_t,y_s\right)=\frac{2\sigma_u^2}{T^2}\left[\sum_{t=1}^{T}t-\sum_{t=1}^{T}t^2\right]=\frac{2\sigma_u^2}{T^2}\left[T\frac{T(T+1)}{2}-\frac{T(T+1)(2T+1)}{6}\right]$$
$$=\sigma_u^2\frac{(T+1)(T-1)}{3T}$$

So **overall**,

$$\begin{aligned}\mathrm{Var}\left(\bar{y}_t\right)&=\sigma_u^2\frac{(T+1)}{2T}+\sigma_u^2\frac{(T+1)(T-1)}{3T}\\&=\sigma_u^2\left[\frac{3(T+1)}{6T}+\frac{2(T+1)(T-1)}{6T}\right]\\&=\sigma_u^2\frac{(T+1)}{6T}\left(3+2(T-1)\right)\\&=\sigma_u^2\frac{(T+1)(2T+1)}{6T}\end{aligned}$$

Which we can approximate when $T$ gets large as,

$$\mathrm{Var}\left(\bar{y}_t\right)\approx\sigma_u^2\frac{T}{3}$$

So the variance explodes as $T$ gets large, implying that the sample mean is not consistent.

**Asymptotic Normality**

If we **assume that the errors are normally distributed**, then using the fact that a linear combination of normally distributed RVs is normally distributed,

$$\begin{aligned}&u_t\sim\ iid\ N\left(0,\sigma_u^2\right)\\&\mathrm{var}\left(\frac{\bar{y}_t}{\sqrt{T}}\right)=\frac{1}{T}\,\mathrm{var}\left(\bar{y}_t\right)\approx\frac{\sigma_u^2}{3}\ ,\ E\left[\frac{\bar{y}_t}{\sqrt{T}}\right]=0\\&\frac{\bar{y}_t}{\sqrt{T}}=\frac{1}{T^{3/2}}\sum_{t=1}^{T}y_t=\frac{1}{T^{3/2}}\sum_{t=1}^{T}\sum_{j=1}^{t}u_j\sim N\left(0,\frac{\sigma_u^2}{3}\right)\end{aligned}$$

If we instead make the weaker assumption that the **errors are iid but not normal**

$$\begin{aligned}&u_t\sim i.i.d\left(0,\sigma_u^2\right)\\&\frac{\bar{y}_t}{\sqrt{T}}=\frac{1}{T^{3/2}}\sum_{t=1}^{T}\sum_{j=1}^{t}u_j\xrightarrow{D}N\left(0,\frac{\sigma_u^2}{3}\right)=\sigma_uN\left(0,\frac{1}{3}\right)\overset{D}{=}\sigma_uD_1\end{aligned}$$

In this case it can also be shown that,

$$\frac{1}{T^2}\sum_{t=1}^{T}y_t^2\xrightarrow{D}\sigma_u^2D_2$$

## Estimation: OLS

$$y_t=\phi y_{t-1}+u_t\ \text{ with } u_t\sim\ iid\ \left(0,\sigma^2\right)\ \text{ and } E\left[u_t^4\right]=\mu_4<\infty$$

OLS Estimator is given by,

$$\widehat{\phi}_T=\frac{\sum_{t=1}^{T}y_{t-1}y_t}{\sum_{t=1}^{T}y_{t-1}^2}=\frac{\sum_{t=1}^{T}y_{t-1}\left(\phi y_{t-1}+u_t\right)}{\sum_{t=1}^{T}y_{t-1}^2}=\phi+\frac{\sum_{t=1}^{T}y_{t-1}u_t}{\sum_{t=1}^{T}y_{t-1}^2}$$

## Asymptotics: OLS Estimator

### Consistency

Consistency of the OLS estimator can be shown in much the same way as it was for the stationary AR(1). The using the mds LLN it can be shown for the numerator that,

$$T^{-1}\sum_{t=1}^{T} y_{t-1}u_t \xrightarrow{P} 0$$

While it can be shown for the denominator that,

$$T^{-1}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{P} \frac{\sigma^2}{(1-\phi)^2}$$

So overall the estimator is consistent.

### Asymptotic Distribution: Dicky-Fuller Distribution

Recall that in the stationary case when $|\phi| < 1$,

$$\sqrt{T}\left(\widehat{\phi}_T - \phi\right) \xrightarrow{D} N(0,\ 1-\phi^2)$$

In the Unit Root case $\phi = 1$,

$$\widehat{\phi}_T - \phi = \frac{\sum_{t=1}^{T} y_{t-1}u_t}{\sum_{t=1}^{T} y_{t-1}^2}$$

$$\widehat{\phi}_T - 1 = \frac{\sum_{t=1}^{T} y_{t-1}u_t}{\sum_{t=1}^{T} y_{t-1}^2}$$

We can't use the usual normalisation $T^{1/2}$ because $T^{1/2}(\widehat{\phi}_T - 1) \xrightarrow{P} 0$, so instead we use $T$ and show that,

$$T(\widehat{\phi}_T - 1) \xrightarrow{D} DF_{OLS}$$

Taking first the **denominator**, recall from before that, although we didn't show why, it can be shown that,

$$\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1} \xrightarrow{D} \sigma_u^2 D_2$$

Considering then the **numerator**,

$$y_t^2 = {y_{t-1}}^2 + {u_t}^2 + 2y_{t-1}u_t \Rightarrow y_{t-1}u_t = \frac{1}{2}\left[{y_t}^2 - {y_{t-1}}^2 - {u_t}^2\right]$$

Then,

$$\frac{1}{T}\sum_{t=1}^{T} y_{t-1}u_t = \frac{1}{2T}\sum_{t=1}^{T}\left[{y_t}^2 - {y_{t-1}}^2 - {u_t}^2\right] = \frac{1}{2}\left[\frac{1}{T}\sum_{t=1}^{T}\left({y_t}^2 - {y_{t-1}}^2\right) - \frac{1}{T}\sum_{t=1}^{T}{u_t}^2\right]$$

For the first term of this new term we can show that,

$$\sum_{t=1}^{T}(y_t^2 - y_{t-1}^2) = \left(y_1^2 + +y_2^2 + ... + y_T^2 - \underbrace{y_0^2}_{=0 \text{ by assumption}} - y_1^2 - y_2^2 - ... - y_{T-1}^2\right) = y_T^2$$

Hence considering again the whole numerator,

$$\frac{1}{2}\left[\frac{1}{T}y_T^2 - \frac{1}{T}\sum_{t=1}^{T}u_t^2\right] = \frac{1}{2}\left[\left\{\frac{1}{\sqrt{T}}y_T\right\}^2 - \frac{1}{T}\sum_{t=1}^{T}u_t^2\right] \xrightarrow{D} \frac{1}{2}\left(\left[N(0,\sigma_u^2)\right]^2 - \sigma_u^2\right) = \frac{1}{2}\sigma_u^2(\mathcal{X}_1^2 - 1)$$

So **overall**,

$$T\left(\widehat{\phi}_T - 1\right) = \frac{T^{-1}\sum_{t=1}^{T} y_{t-1}u_t}{T^{-2}\sum_{t=1}^{T} y_{t-1}^2} \xrightarrow{D} \frac{\left(\mathcal{X}_1^2 - 1\right)}{2D_1} = DF_{OLS}$$

## Spurious Regression & Cointegration