Modelado Estocástico

Clase 7

ESTACIONARIEDAD: Un proceso y_t es (débilmente) estacionario si

- $E(y_t) = \mu < \infty$ para todo t
- $Var(y_t) = \sigma^2 < \infty$ para todo t
- $Cov(y_t, y_{t-j}) = \gamma_j$ para todo t y $j = \pm 1, \pm 2, \pm 3, \ldots$

En otras palabras, un proceso es débilmente estacionario si tiene media y varianza constantes y las autocovarianzas γ_j dependen solamente de la diferencia de períodos "j". Cuando un proceso es estacionario, decimos que es integrado de orden cero, y se denota I(0).

Por otro lado, decimos que un proceso o serie de tiempo es fuertemente estacionario (o estrictamente estacionario) si reemplazamos la condición (c) anterior por la siguiente:

- la distribución conjunta de $\{y_t, y_{t+j_1}, y_{t+j_2}, \dots, y_{t+j_n}\}$ depende solo de los intervalos separando las fechas $(j_1, j_2, \dots, j_n)$ y no depende del momento t.

Si un proceso es estrictamente estacionario y tiene segundos momentos finitos, entonces también es débilmente estacionario.

Cuando un proceso no es estacionario, solemos tomar diferencias y considerar la serie en diferencias, $\Delta y_t = y_t - y_{t-1}$. Si Δy_t es estacionaria, entonces decimos que y_t es integrada de orden uno y se denota I(1). Un proceso I(1) es no estacionario. En general, casi todas las series de tiempo suelen ser I(0) ó I(1). Si Δy_t no fuera estacionaria, uno debería proceder a tomar nuevamente diferencias, es decir, tomar diferencias de las diferencias, es decir, $\Delta(\Delta y_t)$ para analizar si ésta es estacionaria.

La estacionariedad se testea con tests de raíces unitarias y el más tradicional es el estadístico de Dickey Fuller o Dickey Fuller aumentado (ADF) que muchos software estadísticos/econométricos calculan.

Cuando trabajamos con series de tiempo, lo primero que debemos revisar es el orden de integración de cada variable (es decir, si es o no es estacionaria la serie en niveles). La idea detrás de esto es que queremos relacionar variables con el mismo orden de integración.

Test de estacionariedad de Dickey-Fuller Aumentado (ADF):

Sea y_t un proceso AR(1), queremos testear si el *rho* es 1 (en cuyo caso tendríamos un *random walk*)

$y_t = \rho y_{t-1} + \varepsilon_t; \quad H_0: \rho = 1, \quad H_A: \rho < 1$

Esto es equivalente a decir que (resto y_{t-1} a ambos lados del "=")

$$y_t - y_{t-1} = \rho y_{t-1} + \varepsilon_t - y_{t-1}$$
$$\Delta y_t = (\rho - 1) y_{t-1} + \varepsilon_t; \quad (*)$$

De modo que,

$$H_0: \rho - 1 = 0, H_A: \rho - 1 < 0$$

Como bajo H_0 nuestra serie de tiempo es no estacionaria, esto ya no se testea con un estadístico "t" convencional. Dickey y Fuller mostraron que los valores críticos para testear esta hipótesis nula son mayores a los de un test "t" convencional. Los valores críticos deben aumentarse en una cantidad que depende del tamaño de la muestra. Los distintos software reportan directamente los valores críticos y el *p-value* de este test.

Como el test es a una cola, los valores críticos son siempre negativos.

Table II.5.1 Critical values of the Dickey–Fuller distribution[a]

Sample size	Significance level		
	1%	5%	10%
25	−3.75	−3.00	−2.62
50	−3.58	−2.93	−2.60
100	−3.51	−2.89	−2.58
250	−3.46	−2.88	−2.57
500	−3.44	−2.87	−2.57
∞	−3.43	−2.86	−2.57

[a]We only give the critical values for the case where the Dickey–Fuller regression includes a constant but no time trend. As explained in Section II.5.3.3, deterministic trends are rarely present in financial asset returns.

Un problema que tiene el test de Dickey Fuller es que si hubiera autocorrelación en los residuos de la regresión de Dickey Fuller (∗), estos valores críticos ya no estarían bien calculados. Por esta razón, Dickey y Fuller propusieron incluir rezagos de la variable dependiente para remover esta potencial autocorrelación, y a este test se lo llama *Augmented Dickey Fuller* o ADF.

Cuando incluimos constante y tendencia, la ecuación anterior $(*)$ nos queda

$$\Delta y_t = c + (\rho - 1)y_{t-1} + dt + rezagos + \varepsilon_t; \quad (**)$$

Las hipótesis nula y alternativa son las mismas:

$$H_0 : \rho - 1 = 0, H_A : \rho - 1 < 0$$

A este test $(**)$ se lo llama “*Augmented Dickey Fuller*” test (ADF test). No está explícito por una cuestión de espacio, pero en el lado derecho de la expresión $(**)$ se incluyen también los rezagos $\Delta y_{t-1}, \Delta y_{t-2}, etc$, tantos como sean necesarios para remover la autocorrelación que pudiera haber en los residuos.

Table II.5.2 Critical values of the augmented Dickey–Fuller distribution[a]

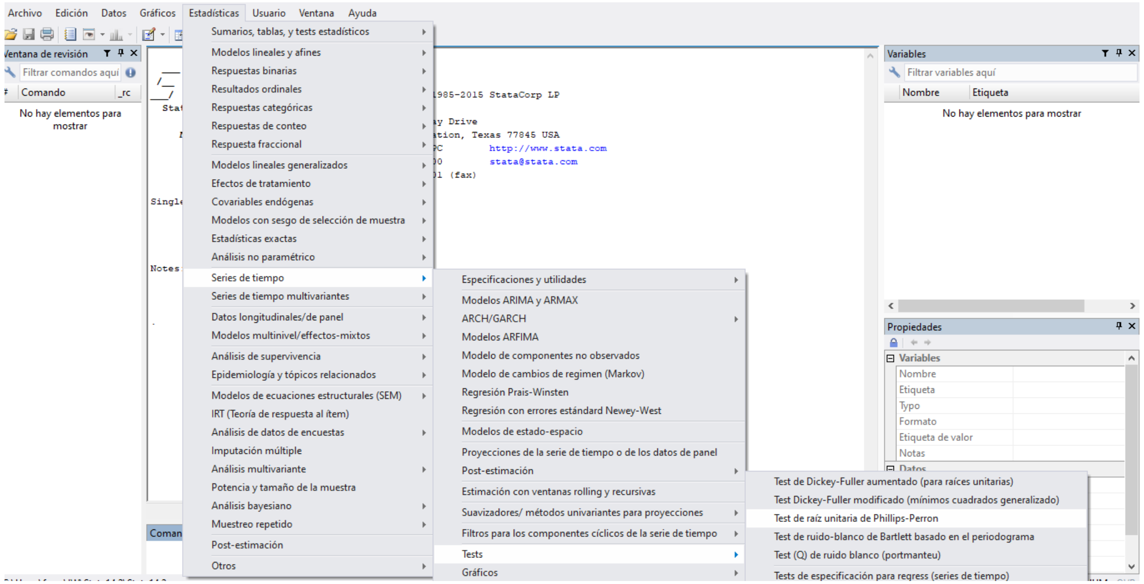
Number of lags	Significance Level		
	1%	5%	10%
1	−3.43	−2.86	−2.57
2	−3.90	−3.34	−3.05
3	−4.30	−3.74	−3.45
4	−4.65	−4.10	−3.81
5	−4.96	−4.42	−4.13

[a] Again we only give the critical values for the case where the augmented Dickey–Fuller regression includes a constant but no time trend.

En Stata, el test ADF se testea con el comando: `dfuller` *`varname`*, `lags`(# rezagos). Este comando por defecto supone que no hay tendencia determinística, pero permite que se incluya una tendencia determinística, con la opción `trend`.

El test ADF supone que los errores en la regresión (∗∗) son i.i.d.. Hay otro test de raíces unitarias, el de Phillips-Perron, que relaja este supuesto.

El test de raíces unitarias de Phillips-Perron, que también está disponible en Stata (comando `pperron`), permite que los errores sean dependientes y con varianzas heterocedásticas. Como los retornos de activos financieros suelen presentar heterocedasticidad condicional, el test de raíces unitarias de Phillips-Perron suele ser más apropiado para el análisis de datos en finanzas.



Tendencias Determinísticas

Recién, cuando agregamos el término "dt" en $(**)$, lo que agregamos era una tendencia determinística o "*time trend*". Desde el punto de vista gráfico, muchas veces es extremadamente difícil diferenciar entre una serie de tiempo con una tendencia determinística y una serie de tiempo con una raíz unitaria.

Supongamos que $\{Y_t\}_{t=-\infty}^{\infty}$ representa la suma de una constante μ más un ruido blanco Gaussiano $\{\varepsilon_t\}_{t=-\infty}^{\infty}$,

$$Y_t = \mu + \varepsilon_t$$

En este caso, note que

$$E(Y_t) = \mu + E(\varepsilon_t) = \mu + 0 = \mu$$

En cambio, si Y_t fuera una tendencia determinística más un ruido blanco Gaussiano

$$Y_t = \beta t + \varepsilon_t$$

su esperanza será, $E(Y_t) = \beta t$. (Esta es la esperanza no condicional, como estuvimos viendo hasta ahora). Esta esperanza depende ahora del tiempo. Y la varianza de Y_t será:

$$Var(Y_t) = E(Y_t - E(Y_t))^2 = E(\varepsilon_t)^2 = \sigma_\varepsilon^2$$

Esta serie de tiempo es no estacionaria ya que su esperanza depende del tiempo (no es igual a una constante μ para todo "t" como se indicó en la definición de estacionariedad). En finanzas no solemos prestarle mucha atención a este caso, pero es importante distinguir entre una serie de tiempo que es I(1) y otra serie de tiempo I(0) + una tendencia determinística. El punto crucial es comprender que el procedimiento para transformar un proceso o el otro en estacionario es completamente diferente: en el caso del proceso I(1) tomamos primeras diferencias para convertirlo en estacionario; por ejemplo, si el logaritmo del precio de un activo financiero es I(1), el retorno logarítmico de ese activo es I(0).

Mientras que en el caso del proceso I(0) + tendencia determinística, el procedimiento para convertirlo en estacionario consiste en tomar desvíos de la tendencia determinística estimada; es decir, removerle la tendencia determinística, y para lograr esto, habría que correr una regresión de la serie de tiempo en una constante y en el tiempo (esto es tomar como variable independiente el tiempo “t”, o sea, una tendencia determinística), y los residuos que se obtengan serán estacionarios.

¿Cuántos lags usamos en el ADF?

Ng & Perron (1995) sugieren:

- Elegir una cota superior para el número de rezagos, p_{max}
- Estimar el estadístico ADF con el número de lags p_{max}
- Si el valor absoluto del estadístico "t" asociado al último rezago de Δy_t (la variable dependiente en la regresión del ADF) es mayor o igual a 1.6, entonces elijo $p = p_{max}$. Si no, reduzco en uno la cantidad de rezagos y repito estos últimos dos pasos.

Este no es un punto menor porque si arbitrariamente elegimos un p chico, no logramos eliminar la correlación serial y los valores críticos no serían válidos, mientras que, por el otro lado, si elegimos un p demasiado grande, el poder estadístico del test se ve afectado.

Schwert (1989) sugiere usar como p_{max} el valor:

$$p_{max} = \left[12 \left(\frac{T}{100} \right)^{1/4} \right]$$

donde los corchetes representan la parte entera de lo que está adentro. Stata permite también usar otros criterios para seleccionar el número de rezagos (Akaike y Schwarz) con el comando:

```
dfgls varname
```

https://www.stata.com/manuals/tsdfgls.pdf

Este comando asume por defecto una tendencia determinística, y es otro test de raíces unitarias, similar al ADF, pero basado en GLS (*Generalized Least Squares*). Es posible que la cantidad óptima de lags que debe incluirse sea diferente si hay o no tendencia determinística. Para correrlo sin tendencia determinística debe usarse la opción `notrend`.

```
dfgls varname, notrend
```

Lag Operator:

El *lag operator* es un operador que todo lo que multiplica, lo rezaga: por ejemplo,

$$Lx_t = x_{t-1}$$

También, $L^2 x_t = LLx_t = Lx_{t-1} = x_{t-2}$

Y también, $L^k x_t = x_{t-k}$

Por ejemplo, una serie en diferencias podemos representarla también usando este operador:

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

Otro ejemplo: si y_t es un proceso $AR(p)$, entonces podemos representarlo de esta manera:

$$y_t\left(1 - \rho_1 L - \rho_2 L^2 - \cdots - \rho_p L^p\right) = \varepsilon_t$$

O bien como $y_t A(L) = \varepsilon_t,$

con $A(L) = 1 - \rho_1 L - \rho_2 L^2 - \cdots - \rho_p L^p$

Note también que

$$A(1) = 1 - \rho_1 - \rho_2 - \cdots - \rho_p$$

Y que un proceso AR(1) podemos escribirlo como

$$y_t(1 - \rho L) = \varepsilon_t$$

Y la condición para que un AR(1) sea débilmente estacionario es $|\rho| < 1$, o sea, la raíz de $1 - \rho z = 0$ sea mayor a uno.

Proceso AR(2): suponga que

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \varepsilon_t$$

Este proceso autorregresivo puede reescribirse como

$$y_t(1 - \rho_1 L - \rho_2 L^2) = \varepsilon_t$$

Un proceso AR(2) es débilmente estacionario si las raíces de

$$1 - \rho_1 z - \rho_2 z^2 = 0$$

están afuera del círculo unitario.

Ejemplo: considere el proceso AR(2):

$$y_t = 0.75 y_{t-1} - 0.25 y_{t-2} + \varepsilon_t$$

Lo reescribimos como:

$$y_t(1 - 0.75L + 0.25L^2) = \varepsilon_t$$

Calculamos las raíces de

$$1 - 0.75z + 0.25z^2 = 0$$

$$a = \frac{1}{4}, b = -\frac{3}{4}, c = 1$$

$$\frac{\frac{3}{4} \pm \sqrt{\frac{9}{16} - 4\frac{1}{4}}}{\frac{2}{4}} = \frac{3}{2} \pm 2\sqrt{\frac{-7}{16}} = \frac{3}{2} \pm i\sqrt{\frac{7}{4}}$$

donde $i = \sqrt{-1}$

El módulo de estas dos soluciones, es decir, la distancia desde el origen hasta cada una de las raíces es

$$\sqrt{\left(\frac{3}{2}\right)^2+\left(\sqrt{\frac{7}{4}}\right)^2}=\sqrt{\frac{9}{4}+\frac{7}{4}}=2$$

Por lo tanto, este proceso AR(2) es estacionario.

En general, si y_t es un proceso AR(p),

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \cdots + \rho_p y_{t-p} + \varepsilon_t$$

Lo podemos reescribir como

$$y_t\left(1 - \rho_1 L - \rho_2 L^2 - \cdots - \rho_p L^p\right) = \varepsilon_t$$

Si las raíces de

$$1 - \rho_1 z - \rho_2 z^2 - \cdots - \rho_p z^p = 0$$

están afuera del círculo unitario, entonces se puede demostrar que este proceso AR(p) va a ser débilmente estacionario.

Si el proceso es débilmente estacionario y todas las raíces fueran reales, entonces las raíces deberán ser todas mayores a uno en valor absoluto.

Estacionariedad en procesos MA:

Un proceso de medias móviles puede representarse como la suma de términos de distintos rezagos de un mismo proceso i.i.d..

Por ejemplo, si y_t sigue un MA(1),

$$y_t = \theta\varepsilon_{t-1} + \varepsilon_t \;\text{ con }\; \varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$$

Este proceso es siempre estacionario si θ es finito. Y en general, un proceso de medias móviles de orden q,

$$y_t = \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q} + \varepsilon_t$$

es estacionario si $\theta_1, \theta_2, \dots, \theta_q$ son todos finitos.

Estacionariedad en procesos ARMA:

Considere el proceso ARMA(p,q),

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \cdots + \rho_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Este proceso puede reescribirse como

$$y_t A(L) = B(L)\varepsilon_t$$

donde $A(L) = 1 - \rho_1 L - \rho_2 L^2 - \cdots - \rho_p L^p$ y
$B(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$.

Este proceso ARMA(*p,q*) será débilmente estacionario si y solo si los coeficientes de medias móviles $\theta_1, \theta_2, \dots, \theta_q$ son todos finitos y si todas las raíces del polinomio $1 - \rho_1 z - \rho_2 z^2 - \cdots - \rho_p z^p = 0$ se encuentran fuera del círculo unitario.

Cálculo de la media y varianza:

Suponga que tenemos un proceso AR(1):

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

Para calcular la media no condicional, suponiendo que este proceso es estacionario, vamos a suponer que

$$E(y_t) = E(y_{t-1}) = \mu.$$

Por lo tanto,

$$\mu = c + \rho\mu \ + E(\varepsilon_t)$$

$$\mu = c + \rho\mu$$

$$\mu(1-\rho) = c$$

$$\mu = \frac{c}{1-\rho}$$

La varianza ya la calculamos para el AR(1) sin constante, y ahora, acá, el resultado va a ser el mismo, porque resulta ser la misma expresión:

$$\gamma_0 = Var(y_t) = E(y_t - \mu)^2 = E\left(\sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j}\right)^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

Y todas las autocovarianzas también van a ser las mismas que para un AR(1) sin constante, porque son, cada una, las mismas expresiones en ambos casos.

Caso general ARMA(*p,q*):

$$y_t = c + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \cdots + \rho_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Este proceso puede reescribirse como

$$y_t A(L) = c + B(L)\varepsilon_t$$

con $A(L) = 1 - \rho_1 L - \rho_2 L^2 - \cdots - \rho_p L^p$ y $B(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$.

Si este proceso es estacionario (la constante "*c*" no cambia las condiciones para estacionariedad), entonces la media será

$$E(y_t) = c\left(1 - \sum_{j=1}^{p} \rho_j\right)^{-1}$$

Y la varianza,

$$\gamma_0 = Var(y_t) = E(y_t - \mu)^2 = \sigma_\varepsilon^2\left(1 - \sum_{j=1}^{p} \rho_j^2\right)^{-1}\left(1 + \sum_{i=1}^{q} \theta_i^2\right)$$

Los procesos estacionarios siempre vuelven a su media. Esta particularidad es relevante para el análisis de variables financieras porque implica que, si nos alejamos de la media, en algún momento deberíamos volver, y esto podría traducirse en una estrategia de inversión.

Por ejemplo, para un ARMA(*p,q*) estacionario,

$$y_t A(L) = c + B(L)\varepsilon_t$$

$A(L)$ tiene inversa y podemos reescribirlo como

$$y_t = cA(L)^{-1} + B(L)A(L)^{-1}\varepsilon_t$$

O sea,

$$y_t = c^* + G(L)\varepsilon_t$$

donde

$c^* = cA(L)^{-1}$ $\qquad$ $G(L) = B(L)A(L)^{-1}$

Note que $A(L)^{-1}$ es una serie infinita en L.

Si la parte de medias móviles también es invertible, es decir, las raíces de

$$1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q = 0$$

están afuera del círculo unitario, entonces $B(L)^{-1}$ existe y también podemos reescribir el proceso ARMA(p,q) como:

$$y_t A(L) B(L)^{-1} = \tilde{c} + \varepsilon_t$$

donde $\tilde{c} = B(L)^{-1} c$.

Note que $B(L)^{-1}$ es una serie infinita en L.

Es decir, el proceso ARMA(*p,q*) puede escribirse como,

$$y_t A(L) B(L)^{-1} = \tilde{c} + \varepsilon_t$$

Esto quiere decir que cuando $B(L)^{-1}$ existe, el proceso ARMA(*p,q*) puede reescribirse como un proceso autorregresivo de orden infinito.

Introducción de shocks:

Partamos de un proceso ARMA(*p,q*) estacionario e invertible ($B(L)^{-1}$ existe), sea $F(L) = A(L)B(L)^{-1}$, entonces el proceso ARMA(*p,q*) lo podemos escribir como un proceso autorregresivo de orden infinito

$$y_t F(L) = \tilde{c} + \varepsilon_t$$

Vamos a introducir un shock u_t de una unidad en el período t:

$$u_t = \begin{cases} 1 \; en \; el \; período \; t \\ 0 \qquad\qquad\qquad si \; no \end{cases}$$

De modo que,

$$y_t F(L) = \tilde{c} + \varepsilon_t + u_t$$

Notando que,

$$F(L)^{-1} = A(L)^{-1} B(L) = G(L) = 1 + g_1 L + g_2 L^2 + \cdots$$

Podemos decir que

$$y_t = F(L)^{-1}(\tilde{c} + \varepsilon_t + u_t)$$

$$y_t = G(L)(\tilde{c} + \varepsilon_t + u_t)$$
$$y_t = c^* + G(L)(\varepsilon_t + u_t)$$

Como $G(L) = 1 + g_1 L + g_2 L^2 + \cdots$, el shock de una unidad de u_t va a generar un cambio de la misma magnitud en el corto plazo en y_t pero también va a afectar a los valores futuros $y_{t+1}, y_{t+2}, y_{t+3}, \dots$. Los coeficientes $g_1,\ g_2,\ g_3, \dots$ pueden interpretarse como el efecto del shock sobre $y_{t+1}, y_{t+2}, y_{t+3}, \dots$. La suma de todos estos es $G(1) = 1 + g_1 + g_2 + \cdots$ es el efecto de largo plazo que el shock u_t genera sobre la serie y_t.

La función $g_s\,(s = 1,2,3, \dots)$ se llama **función de impulso respuesta** y mide el efecto que este shock u_t de una unidad genera sobre y_{t+s}.

A veces podemos estar interesados en conocer el tiempo que transcurre hasta que la mitad del efecto de largo plazo del shock u_t fue incorporado en el proceso y_t. A esto se lo conoce como el "*median lag*".

Otra medida interesante es el "*mean lag*", que se define como

$$\text{Mean lag} = \frac{\sum_{i=0}^{\infty} i\, g_i}{\sum_{i=0}^{\infty}\, g_i} = G(1)^{-1}G'(1)$$

Ejemplo: considere el proceso ARMA(2,1) (ejemplo II.5.2 de Carol Alexander)

$$y_t = 0.03 + 0.75y_{t-1} - 0.25y_{t-2} + \varepsilon_t + 0.5\varepsilon_{t-1}$$

Lo reescribimos como:

$$y_t(1 - 0.75L + 0.25L^2) = 0.03 + \varepsilon_t(1 + 0.5L)$$

De modo que,

$A(L) = 1 - 0.75L + 0.25L^2,$

$B(L) = 1 + 0.5L,$

$G(L) = A(L)^{-1}B(L)$

$G(L) = (1 - 0.75L + 0.25L^2)^{-1}(1 + 0.5L)$

Una manera de conocer los coeficientes $g_1,\ g_2,\ g_3, \ldots$ es plantear lo siguiente e igualar los coeficientes de L, L^2, L^3, etc :

$$(1 - 0.75L + 0.25L^2)G(L) = 1 + 0.5L$$

$$(1 - 0.75L + 0.25L^2)(1 + g_1 L + g_2 L^2 + \cdots) = 1 + 0.5L$$

Igualo los coeficientes que tengan L:

$$g_1 - 0.75 = 0.5 \Rightarrow g_1 = 1.25$$

Igualo los coeficientes que tengan L^2:

$$g_2 - 0.75g_1 + 0.25 = 0 \Rightarrow g_2 = \frac{11}{16}$$

Igualo los coeficientes que tengan L^3:

$$g_3 - 0.75g_2 + 0.25g_1 = 0 \Rightarrow g_3 = \frac{13}{64}$$

Igualo los coeficientes que tengan L^4:

$$g_4 - 0.75g_3 + 0.25g_2 = 0 \Rightarrow g_4 = -\frac{5}{256}$$

Y así sucesivamente. En el Excel pueden ver el gráfico de la función de impulso respuesta. El efecto de largo plazo es 3, el *mean lag* es 0.833 y el *median lag* es 0.4. En el Excel pueden ver un gráfico con la función de impulso-respuesta (los efectos del shock se ven por aproximadamente 10 períodos) y también la función de impulso-respuesta acumulada.

Cointegración:

Si las series y_t y x_t fueran ambas integradas de orden 1, es posible regresar y_t en x_t usando MCO.

Supongamos que estamos interesados en estimar el modelo

$$y_t = \alpha + \beta x_t + u_t$$

Si las series y_t y x_t son ambas I(1) y ***si los residuos de esta regresión por el método de MCO son estacionarios***, entonces podemos decir que y_t y

x_t están **cointegradas** y que $\boldsymbol{\hat{\beta}}$ **es el estimador de la relación de cointegración**. En otras palabras, los residuos $e_t = y_t - \hat{\alpha} - \hat{\beta}x_t$, si son estacionarios, lo que intuitivamente estamos obteniendo es que las series y_t y x_t guardan una tendencia estocástica común.

Solamente si ambas, y_t y x_t son I(1), es posible correr una regresión por MCO y que el análisis sea estadísticamente válido, si los residuos que se obtienen en esta regresión son estacionarios.

Cointegración es una medida de dependencia de largo plazo entre variables (por ejemplo, activos financieros). Si los precios de dos activos financieros estuvieran cointegrados, esto nos está diciendo que si bien no sabemos dónde pueden encontrarse estos precios en algún momento en el futuro, si conociéramos uno de ellos, entonces es posible determinar aproximadamente dónde se ubicará el otro; que los residuos de la relación de cointegración sean estacionarios, básicamente nos está diciendo que los residuos revierten a la media.

Pensemos en esos residuos como un *spread*, que, de verificarse estacionarios, podemos decir que ese *spread* será *mean reverting* (*spread* entre precios de activos financieros o entre diferentes tasas de interés). Una relación de cointegración es una relación de largo plazo entre las variables.

Referencias:

- Ng, S., & Perron, P. (1995). Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, *90*(429), 268-281.