Series de Tiempo

Fundamentos de Econometría

Juan Palomino[1]

[1]Magister en Economía Aplicada con Mención Estudios Regionales
juan.palominoh@pucp.pe

Departamento de Economía

Índice

Introducción

- En esta parte del curso veremos análisis de series de tiempo y procesos estocásticos.
  - Una **serie de tiempo** es un conjunto de observaciones x_t, donde t indica el tiempo de ocurrencia de la observación $(t = 1, 2, 3, ..., N)$.
  - El tiempo puede ser medido en cualquier unidad (ej. minutos, días, meses, años, etc.)
  - **Serie de tiempo discreta**: cuando el conjunto de observaciones se hace en un conjunto discreto de periodos, por ejemplo, en intervalos fijos de tiempo (diarias, mensuales, trimestrales, anuales, etc.).
  - **Serie de tiempo continua:** cuando las observaciones se realizan sobre un intervalo continuo de tiempo, por ejemplo, $[0, T]$. Las observaciones en sí pueden seguir siendo discretas.

Introducción

- Las observaciones sucesivas en el tiempo son, en general, no independientes. Esto significa que las observaciones pasadas pueden ser usadas para **predecir** futuras observaciones.
- **Serie deterministica:** Si, dadas las obsevaciones $x_1, ..., x_{t-1}$, la observación x_t se puede predecir exactamente.
- **Serie estocástica:** Si las observaciones futuras no se pueden predecir exactamente.
- En series estocásticas, las observaciones futuras tendrán una distribución de probabilidad. Si las observaciones son dependientes, entonces esta distribución de probabilidad será dependiente de las observaciones pasadas.

Componentes de una serie temporal

- Se divide en tres componentes
- **Componente tendencia:** representa el componente de largo plazo que representa el crecimiento o disminución en la serie sobre un periodo amplio.
- **Componente estacional:** es un patrón de cambio que se repite a sí mismo tras cierto periodo. (año, mes, etc).
- **Componente aleatorio:** este componente no responde a ningún patrón de comportamiento. Este mide la variabilidad de la serie después de retirar los otros componentes.

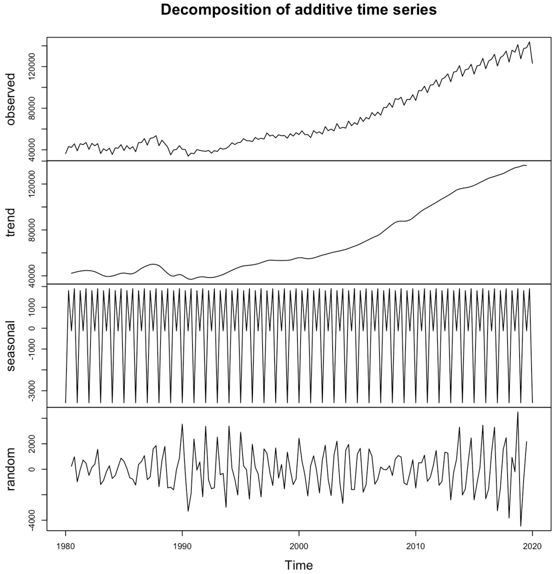
Componentes de una serie temporal

- De estos tres componentes los dos primeros son componentes determinísticos, mientras que la última es aleatoria.

$$X_t = T_t + E_t + I_t$$

- donde T_t es la tendencia, E_t es el componente estacional, I_t es el componente aleatorio.

Componentes de una serie temporal



- **Series Estacionarias:** una serie es estacionaria cuando es estable a lo largo del tiempo, es decir, cuando la media y varianza son constantes en el tiempo. Graficamente, se observa que los valores de la serie tienden a oscilar alrededor de una media constante y la variabilidad con respecto a esa media también permanece constante en el tiempo.
- **Series No Estacionarias:** son series en las cuales la tendencia y/o variabilidad cambian en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante.

Clasificación de las series temporales

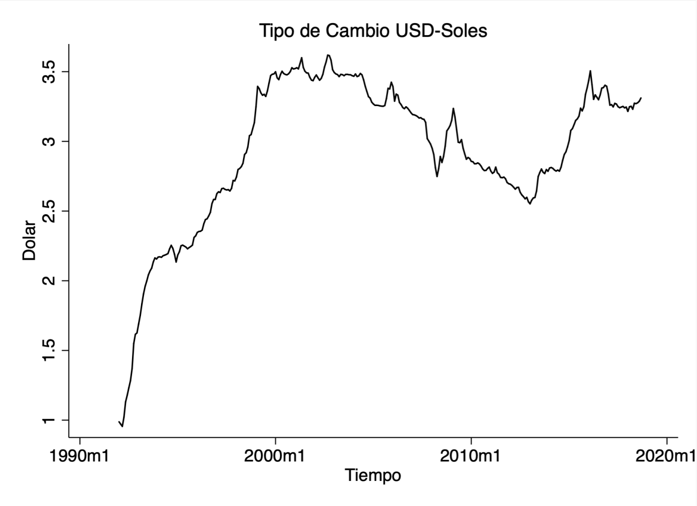


Figura: Tipo de Cambio

Clasificación de las series temporales

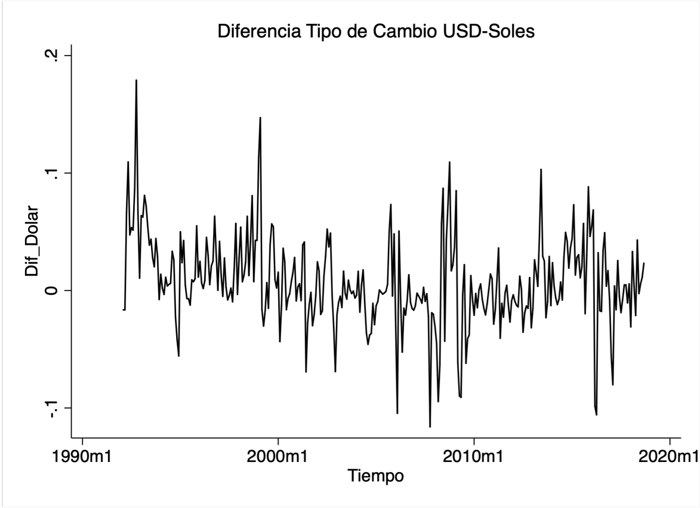


Figura: Diferencia Tipo de Cambio

Clasificación de las series temporales

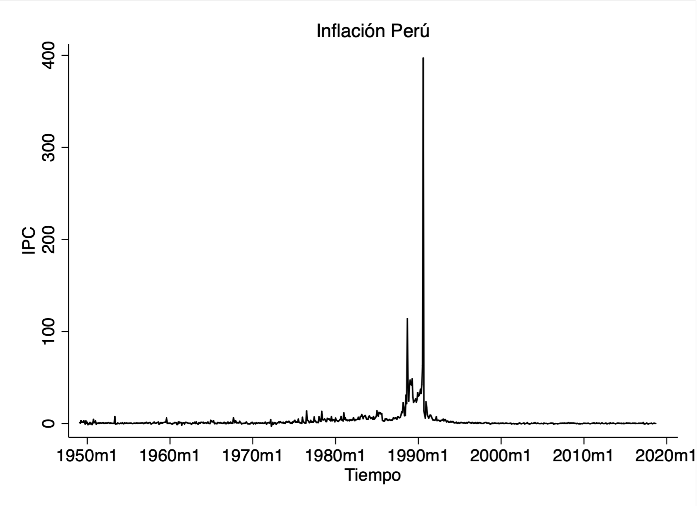


Figura: Inflación

Clasificación de las series temporales

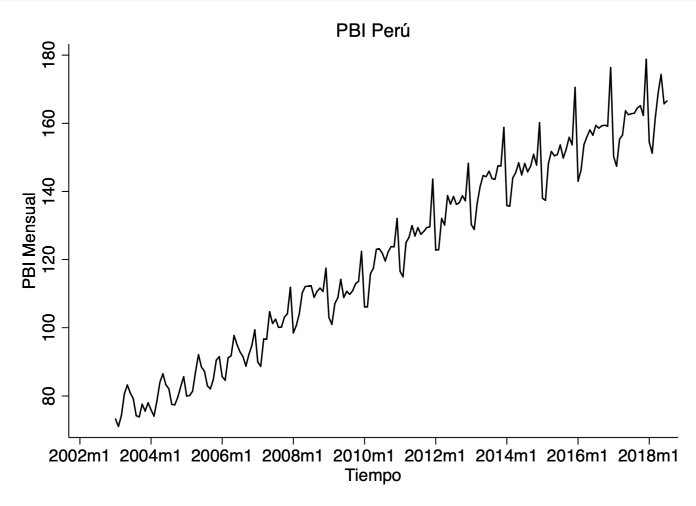


Figura: PBI

Clasificación de las series temporales

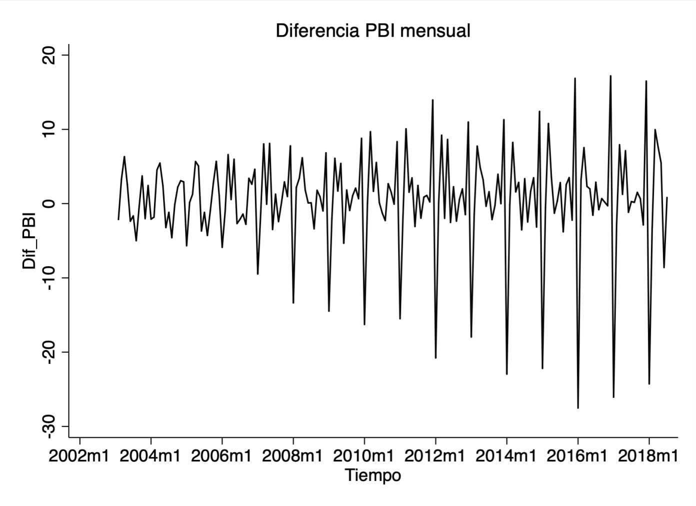


Figura: Diferencia PBI

Clasificación de las series temporales

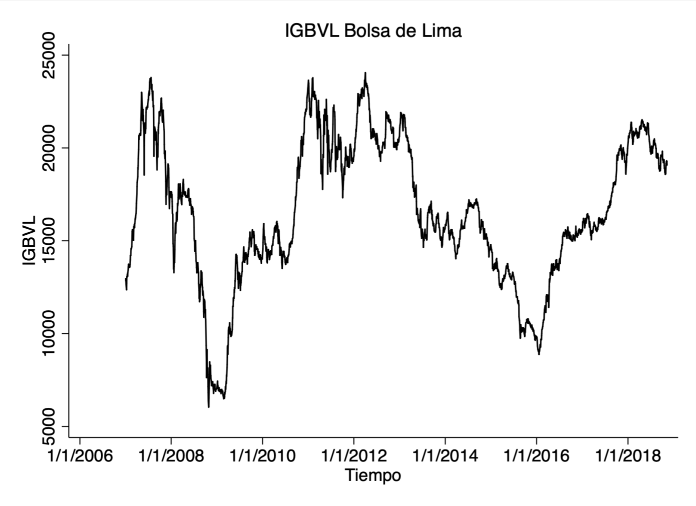


Figura: IGBVL

Clasificación de las series temporales

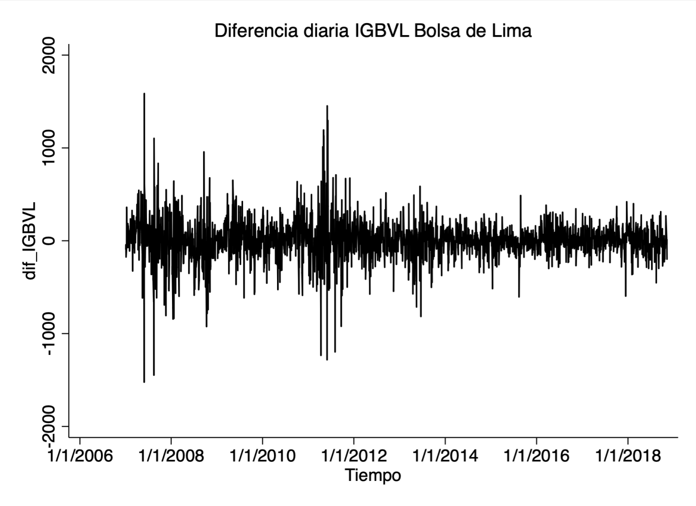


Figura: Diferencia IGBVL

Procesos Estacionarios

- En los gráficos presentados además de la serie misma, vimos también la diferencia en la variable, o el cambio de la variable de un periodo a otro

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

- Notamos que hay algunas series que al restarle a cada observación el valor del periodo, la serie resultante tiende a estar acotada entre dos valores, sin embargo para otras no parece tan claro.
- ¿Qué serie representa variables estacionarias y cuales son no estacionarias?

Procesos Estacionarios

- Se dice que y_t es estacionario si su esperanza y varianza son constante a lo largo del tiempo, y además si la covarianza entre dos valores de la serie depende sólo del largo de tiempo que separa esos dos valores (y no del momento que es evaluado).
- En otras palabras, la serie de tiempo y_t es estacionaria si para todo t
  - $E(y_t) = \mu$ (media constante)
  - $var(y_t) = \sigma^2$ (varianza constante)
  - $cov(y_t, y_{t+s}) = \gamma_s$ (covarianza depende de s, no de t)

Índice

Modelo AR(1)

- Supongamos que la serie y_t es una variable económica que observamos en el tiempo. En línea con la mayoría de las variables económicas, asumimos que y_t tiene naturaleza estocástica porque no podemos predecirla perfectamente.
- Las series de tiempo univariadas son ejemplos de procesos estocásticos donde la variable y está relacionada con sus valores pasados y valores actual y pasados del error (no tenemos variables explicativas $x's$)
- El modelo autorregresivo de orden 1, AR(1), es un modelo de series de tiempo univariado útil para explicar la diferencia entre estacionario y no estacionario. Viene dado por:

$$y_t = \rho y_{t-1} + v_t \quad |\rho| < 1$$

- Donde los errores v_t son independientes, con media cero y varianza σ_v^2
- Veremos que $|\rho| < 1$ implica que y_t es estacionario.

Modelo AR(1)

- Podemos expresar el modelo de la siguiente manera:

$$
\begin{aligned}
y_t &= \rho y_{t-1} + v_t \\
&= \rho(\rho y_{t-2} + v_{t-1}) + v_t \\
&= \rho^t y_0 + \sum_{i=0}^{t-1} \rho^i v_{t-i}
\end{aligned}
$$

- Esperanza

$$
\begin{aligned}
E(y_t) &= E(\rho^t y_0 + \sum_{i=0}^{t-1} \rho^i v_{t-i}) \\
&= \rho^t y_0 + \sum_{i=0}^{t-1} \rho^i E(v_{t-i}) \\
&= \rho^t y_0 = 0
\end{aligned}
$$

- Si t es suficientemente grande (o pensar que la serie no tiene un $t = 0$, sino que sigue infinitamente a $t = -\infty$).

Modelo AR(1)

- Varianza

$$
\begin{aligned}
var(y_t) &= var(\rho^t y_0 + \sum_{i=0}^{t-1} \rho^i v_{t-i}) \\
&= \sum_{i=0}^{t-1} \rho^{2i} var(v_{t-i}) \\
&= \sigma_v^2 \sum_{i=0}^{t-1} \rho^{2i}
\end{aligned}
$$

- y si recordamos que $\sum_{i=0}^{t-1} \alpha^i = \frac{1-\alpha^t}{1-\alpha}$, luego:

$$
var(y_t) = \sigma_v^2 \frac{1-\rho^{2t}}{1-\rho^2} = \sigma_v^2 \frac{1}{1-\rho^2}
$$

- Si t es suficientemente grande y $|\rho| < 1$ (sino no converge).

Modelo AR(1)

- Autocovarianza

$$
\begin{aligned}
cov(y_t, y_{t+s}) &= cov(\rho^t y_0 + \sum_{i=0}^{t-1} \rho^i v_{t-i}, \rho^{t+s} y_0 + \rho^t y_0 + \sum_{j=0}^{t-1} \rho^j v_{t+s-j} \\
&= cov(\sum_{i=0}^{t-1} \rho^i v_{t-i}, \sum_{i=0}^{t+s-1} \rho^j v_{t+s-j}) \\
&= \rho^s cov(\sum_{i=0}^{t-1} \rho^i v_{t-i}, \sum_{i=-s}^{t-1} \rho^i v_{t-i}) \\
&= \rho^s cov(\sum_{i=0}^{t-1} \rho^i v_{t-i}, \sum_{i=0}^{t-1} \rho^i v_{t-i}) \\
&= \sigma_v^2 \rho^s \sum_{i=0}^{t-1} \rho^{2i} \\
&= \sigma_v^2 \frac{\rho^s}{1-\rho^2}
\end{aligned}
$$

- Observación: no depende de t, sino de la distancia s

Modelo AR(1)

- Podemos definir la función de autocovarianzas:

$$\gamma(y_t, y_{t+s} = cov(y_t, y_{t+s})$$

- en el caso donde sólo dependa de la distancia o lag s, $\gamma(s) = \gamma(y_t, y_{t+s})$ (fijarse que $\gamma(0) = var(y_t)$)
- También podemos definir la función de autocorrelación (ACF):

$$r(y_t, y_{t+s}) = \frac{cov(y_t, y_{t+s})}{\sqrt{var(y_t)var(y_{t+s})}}$$

- y nuevamente, en el caso que sólo dependa de s,

$$r(s) = r(y_t, y_{t+s}) = \frac{\gamma(s)}{\gamma(0)}$$

- Para AR(1), tenemos que $r(s) = \frac{\gamma(s)}{\gamma(0)} = \rho^s$.

Modelo AR(1) con constante

- En la realidad es difícil encontrarse con procesos de media cero.
- Podemos incluir un término μ de la siguiente manera

$$y_t - \mu = \rho(y_{t-1} - \mu) + v_t$$

- o expresado de otra forma

$$y_t = \alpha + \rho y_{t-1} + v_t \quad |\rho| < 1$$

- con $\alpha = \mu(1-\rho)$
- Extendiendo la serie, tenemos

$$\begin{aligned} y_t &= \alpha + \rho y_{t-1} + v_t \\ &= \alpha + \rho(\alpha + \rho y_{t-2} + v_{t-1}) + v_t \\ &= \rho^t y_0 + \alpha \sum_{i=0}^{t-1} \rho^i + \sum_{i=0}^{t-1} \rho^i v_{t-i} \end{aligned}$$

Modelo AR(1) con constante

- Esperanza. Si utilizamos la última expresión, rescatamos que

$$E(y_t) = \frac{\alpha}{1-\rho} = \mu$$

- Varianza

$$var(y_t) = var(\rho^t y_0 + \alpha \sum_{i=0}^{t-1} \rho^i + \sum_{i=0}^{t-1} \rho^i v_{t-i}) = \sigma_v^2 \frac{1}{1-\rho^2}$$

- Autocovarianza

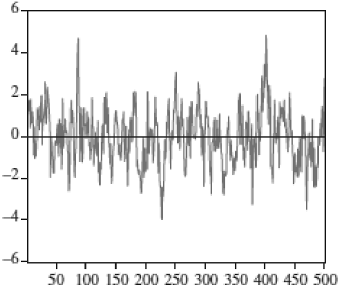
$$cov(y_t, y_{t+s}) = \sigma_v^2 \frac{\rho^2}{1-\rho^2} = \gamma(s)$$

- Autocorrelación (ACF)

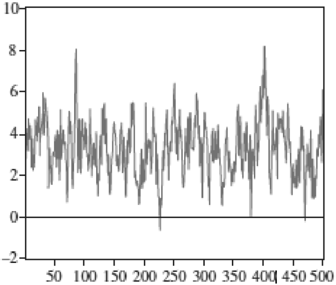
$$r(s) = \frac{\gamma(s)}{\gamma(0)} = \rho^s$$

- Observación: La ACF debiera ir decayendo a medida que aumentamos el lag s.

Modelo AR(1) sin y con constante



(a) $y_t = 0{,}7y_{t-1} + v_t$



(b) $y_t = 1 + 0{,}7y_{t-1} + v_t$

Modelo AR(1) con constante y tendencia

- Otra extensión al modelo es considerar que AR(1) fluctua entorno a una linea recta, es decir

$$y_t - \mu - \delta t = \rho(y_{t-1} - \mu - \gamma(t-1)) + v_t \quad |\rho| < 1$$

- que también puede ser expresado de la siguiente manera

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

- donde $\alpha = \mu(1-\rho) + \rho\delta$ y $\lambda = \delta(1-\rho)$
- Esperanza

$$E(y_t) = \mu + \delta t$$

- Varianza

$$var(y_t) = \sigma_v^2 \frac{1}{1-\rho^2}$$

- ¿Es estacionario?

Modelo AR(1) con constante y tendencia

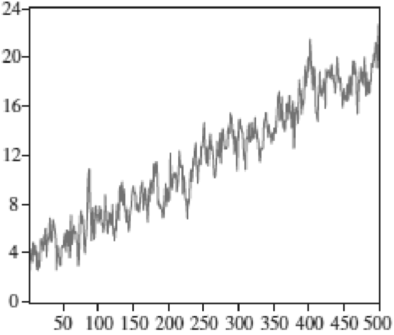


Figura: $1 + 0{,}01t + 0{,}7y_{t-1} + v_t$

Índice

Ruido Blanco ("white noise")

- Un ruido blanco es un caso simple de los procesos estocásticos, donde los valores son independientes e idénticamente distribuidos a lo largo del tiempo con media cero e igual varianza, se denota por ε_t.

$$\varepsilon_t \sim N(0, \sigma^2) \quad cov(\varepsilon_{t_i}, \varepsilon_{t_j}) = 0 \; \forall t_i \neq t_j$$

Ruido Blanco ("white noise")

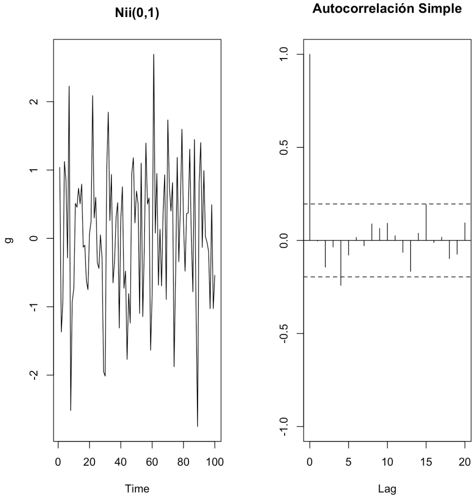


Figura: Ruido blanco: $\varepsilon_t \sim N(0, \sigma^2)$

Índice

Camino aleatorio ("Random Walk")

- Es un proceso estocástico y_t, donde la primera diferencia de este proceso estocástico es un ruido blanco y se cumple cuando $\rho = 1$:

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

  Esto es $\Delta y_t = v_t$. Su nombre proviene porque la serie pareciera caminar para arriba o para abajo sin seguir ningún patrón aparente.

- Expansión de la serie:

$$y_t = y_{t-1} + v_t = y_{t-2} + v_{t-1} + v_t = y_0 + \sum_{i=0}^{t-1} v_{t-i}$$

- Esperanza

$$E(y_t) = y_0$$

- Varianza

$$var(y_t) = t\sigma_v^2$$

Camino aleatorio ("Random Walk")

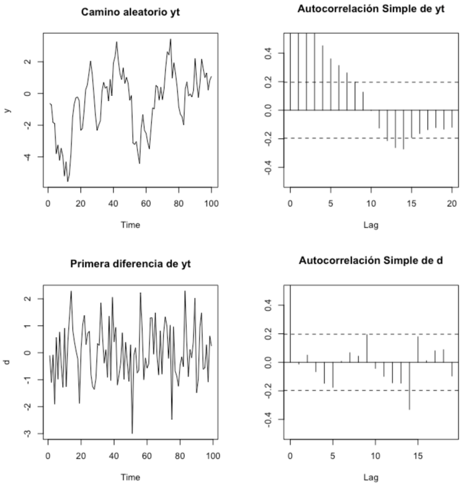


Figura: Random Walk

Regresiones Espúreas

- La principal razón de por qué es importante saber si una serie es estacionaria o no estacionaria antes de realizar un análisis de regresión, es que cuando las series usadas son **no estacionarias** existe la problemática de obtener resultados aparentemente significativos sobre data que no tiene relación. Estas regresiones se llaman **espúreas**.
  - En otras palabras, cuando series de tiempo no estacionarias son usadas en un análisis de regresión, el resultado puede indicar una relación significativa cuando en realidad no existe.
  - Estimadores MCO pierden sus propiedades y los estadísicos t no son confiables.
  - Como muchas variables económicas son **no estacionarias**, es importante tener presente cómo deben ser tratadas.
    - ¿Cómo podemos **testear** si una serie es estacionaria o no?
    - ¿Cómo podemos hacer regresiones cuando tenemos series **no estacionarias?**

Test de Estacionariedad

- Existen variados test para determinar si una serie es estacionaria o no.
- Discutiremos uno de los más populares en econometría, test de **Dickey-Fuller**.

1. Dickey-Fuller, sin constante y sin tendencia

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

2. Dickey-Fuller, con constante y sin tendencia

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

3. Dickey-Fuller, con constante y con tendencia.

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

Test de estacionariedad

1. Dickey-Fuller, sin constante y sin tendencia.

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

- o alternativamente

$$\Delta y_t = (\rho - 1) y_{t-1} + v_t = \beta y_{t-1} + v_t$$

- Luego, la hipótesis nula puede escribirse en términos de ρ o β

$$H_0 : \rho = 1 \Longleftrightarrow H_0 : \beta = 0$$
$$H_1 : \rho < 1 \Longleftrightarrow H_1 : \beta < 0$$

- La hipótesis nula es que la serie es **no estacionaria**. En otras palabras, si no rechazamos la nula, concluímos que no existe evidencia para rechazar que el proceso es no estacionario. Y, por el contrario, si rechazamos la nula, concluímos que la serie es estacionaria.

1. Dickey-Fuller, con constante y sin tendencia

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

- o alternativamente

$$\Delta y_t = \alpha + (\rho - 1) y_{t-1} + v_t = \alpha + \beta y_{t-1} + v_t$$

$$H_0 : \rho = 1 \iff H_0 : \beta = 0$$
$$H_1 : \rho < 1 \iff H_1 : \beta < 0$$

1. Dickey-Fuller, con constante y con tendencia

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

- o alternativamente

$$\Delta y_t = \alpha + \lambda t + (\rho - 1) y_{t-1} + v_t = \alpha + \lambda t + \beta y_{t-1} + v_t$$

$$H_0 : \rho = 1 \iff H_0 : \beta = 0$$
$$H_1 : \rho < 1 \iff H_1 : \beta < 0$$

Valores críticos del test Dickey-Fuller

- Un camino lógico para testear la hipótesis de DF, sería realizar la regresión por MCO y luego construír el estadístico $t = \hat{\beta}/\hat{se}(\hat{\beta})$
  - Desgraciadamente, el estadístico t ya no distribuye bajo una distribución t cuando la nula es cierta
    - **Problema**: cuando la nula es cierta, y_t es no estacionario y su varianza incrementa a medida que la muestra crece. Este incremento altera la distribución usual.
  - Para reconocer este hecho, usualmente a este estadístico se le llama τ (tau) y su valor debe ser comparado con valores especiales (los cuales cambian dependiendo del tipo de test DF que apliquemos).

Valores críticos del test Dickey-Fuller

Modelo	1 %	5 %	10 %
$\Delta y_t = \beta y_{t-1} + v_t$	-2.56	-1.94	-1.62
$\Delta y_t = \alpha + \beta y_{t-1} + v_t$	-3.43	-2.86	-2.57
$\Delta y_t = \alpha + \lambda t + \beta y_{t-1} + v_t$	-3.96	-3.41	-3.13
Normal estándar	-2.33	-1.65	-1.28

Cuadro: Valores críticos Dickey-Fuller

Dickey Fuller aumentado

- Una importante extensión del test de Dickey-Fuller permite la inclusión de modelos con el término del error autocorrelacionado. Esta autocorrelación es probable que aparezca si nuestro modelo no incorporó términos anteriores de la serie yt−s que capturen la dinámica de nuestro proceso.
  - Usando la versión de DF con constante, una extensión sería

$$\Delta y_t = \alpha + \beta y_{t-1} + \sum_{s=1}^{m} a_s \Delta y_{t-s} + v_t$$

  - con $\Delta y_{t-s} = y_{t-s} - y_{t-s-1}$. Se agregan tantos términos de retraso como sea necesario para asegurar que los residuos no están autocorrelacionados.
  - Este test, con variaciones de constante y tendencia respectiva, se conoce como **Dickey-Fuller aumentado**.
  - En la práctica siempre usamos el test DF aumentado.

Ejemplo test Dickey-Fuller Aumentado

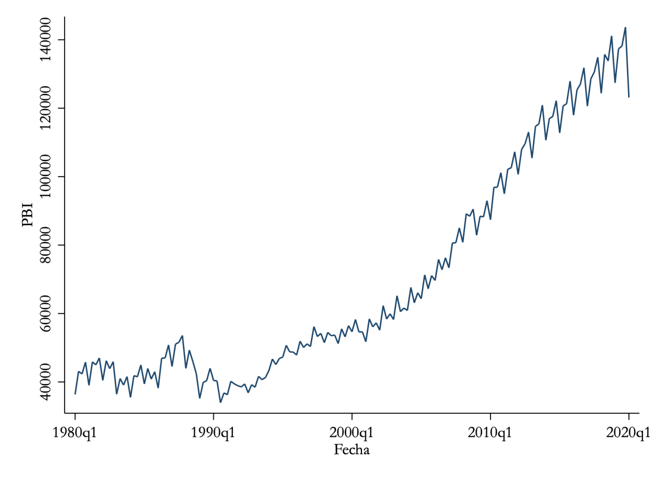


Figura: PBI

Ejemplo test Dickey-Fuller Aumentado ¿Cuál?

- $\Delta PBI_t = \beta PBI_{t-1} + v_t$

D.PBI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
PBI						
L1.	.0051563	.0053681	0.96	0.338	-.0054457	.0157583

- $\Delta PBI_t = \alpha + \beta PBI_{t-1} + v_t$

D.PBI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
PBI						
L1.	-.0076767	.012921	-0.59	0.553	-.0331968	.0178434
_cons	1073.838	983.5991	1.09	0.277	-868.8603	3016.537

- $\Delta PBI_t = \alpha + \lambda t + \beta PBI_{t-1} + v_t$

D.PBI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
PBI						
L1.	-.0825618	.0307306	-2.69	0.008	-.1432606	-.021863
_trend	56.25532	21.02992	2.68	0.008	14.71725	97.79338
_cons	1731.254	995.7844	1.74	0.084	-235.6087	3698.117

Ejemplo test Dickey-Fuller Aumentado

- $\Delta PBI_t = \beta PBI_{t-1} + v_t$

Dickey-Fuller test for unit root Number of obs = **160**

	Test Statistic	Interpolated Dickey-Fuller 1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	**0.961**	**-2.592**	**-1.950**	**-1.614**

- $\Delta PBI_t = \alpha + \beta PBI_{t-1} + v_t$

Dickey-Fuller test for unit root Number of obs = **160**

	Test Statistic	Z(t) has t-distribution 1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	**-0.594**	**-2.350**	**-1.655**	**-1.287**

p-value for Z(t) = **0.2766**

- $\Delta PBI_t = \alpha + \lambda t + \beta PBI_{t-1} + v_t$

Dickey-Fuller test for unit root Number of obs = **159**

	Test Statistic	Interpolated Dickey-Fuller 1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	**-25.583**	**-4.020**	**-3.442**	**-3.142**

MacKinnon approximate p-value for Z(t) = **0.0000**