Series de Tiempo Descomposición

Orlando Belli Hesse

Test de estacionariedad

- Utilizaremos el test de Dickey Fuller aumentado
  - H0 : La serie no es estacionaria
  - H1 : La serie es estacionaria

https://www.statsmodels.org/dev/generated/statsmodels.tsa.stattools.adfuller.html

Si p_value < 5 % entonces se rechaza H0, se concluye que la serie es estacionaria

P_value es la probabilidad de que ocurra H0

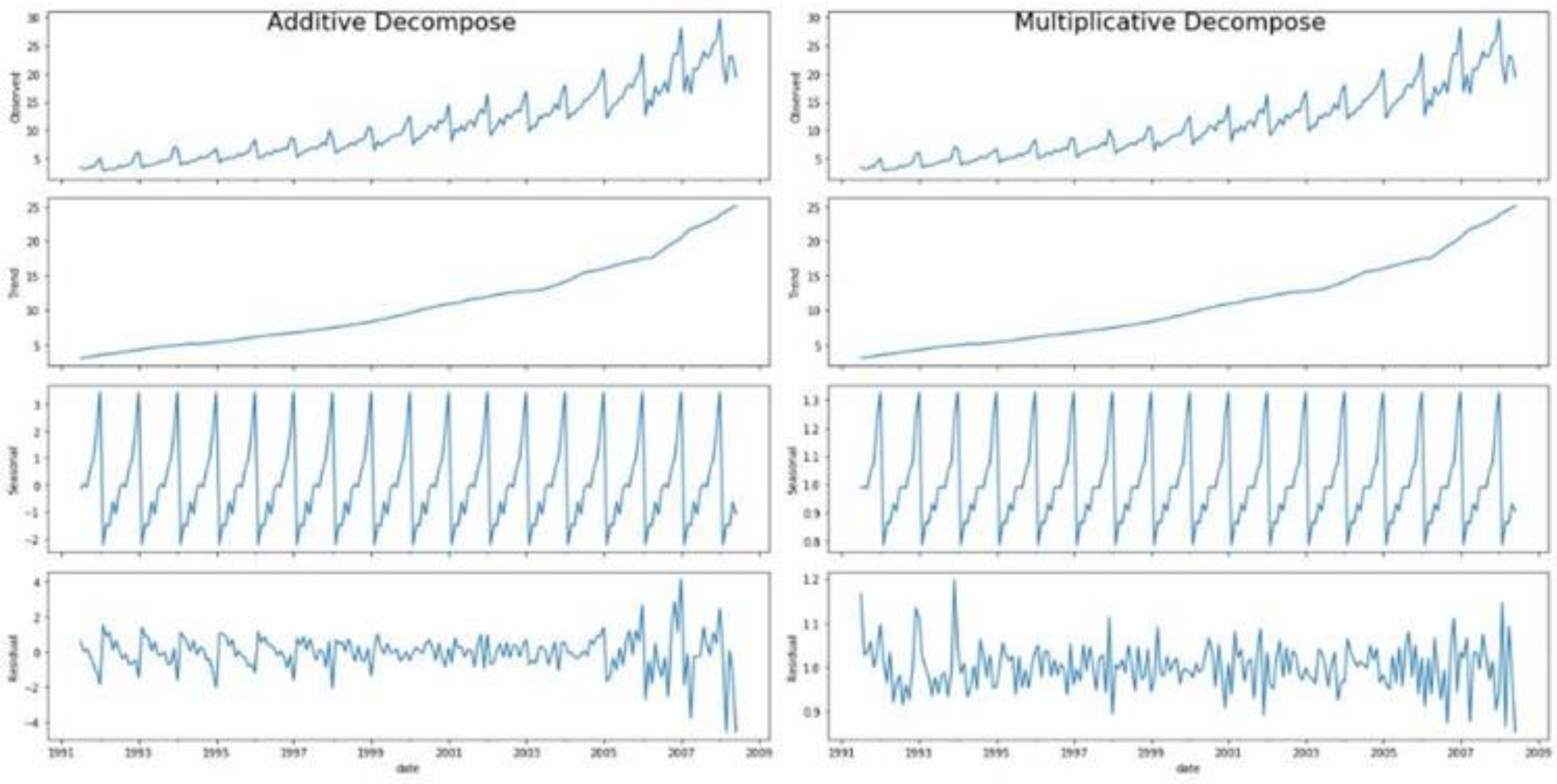
```
sts.adfuller(df.market_value)
```

```
(-1.7369847452352454,
 0.4121645696770613,
 18,
 5002,
 {'1%': -3.431658008603046,
  '10%': -2.567077669247375,
  '5%': -2.862117998412982},
 39904.880607487445)
```

Estacionalidad

Descompone la serie en 3 efectos:

- Tendencia (Tt)
- Estacional (St)
- Residual (et)



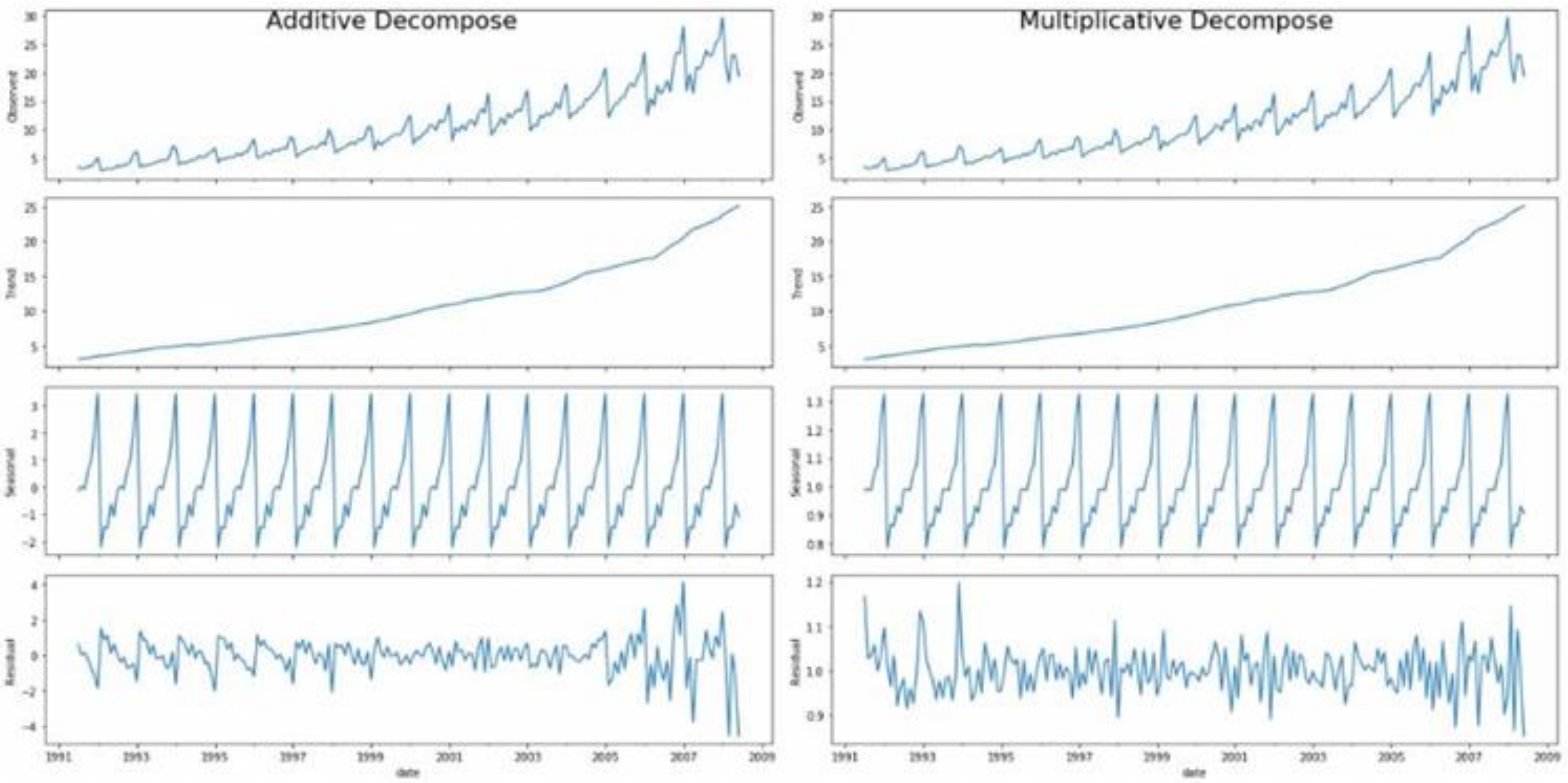
Descomposición Clásica

- Aditivo

$$X_t = S_t + T_t + e_t$$

- Multiplicativo

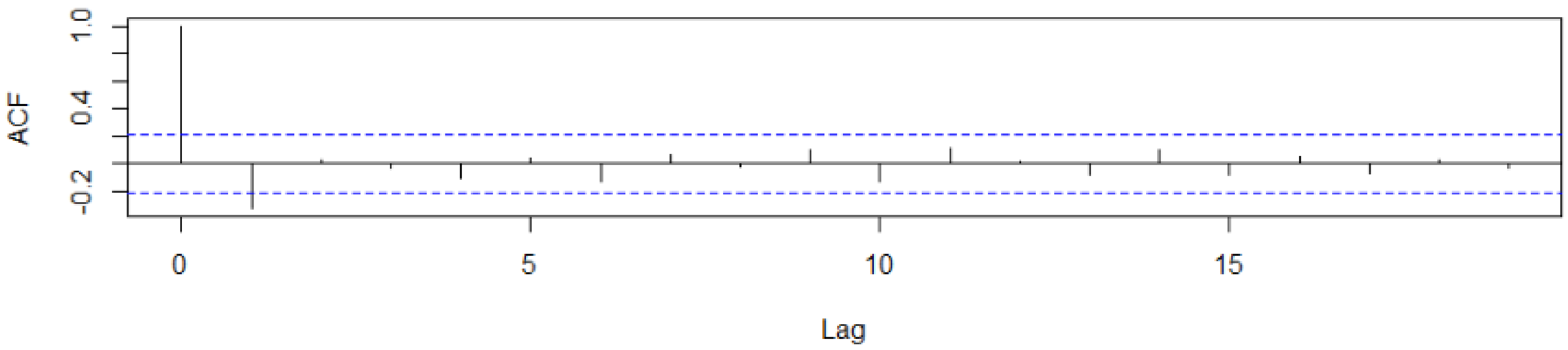
$$X_t = S_t \times T_t \times e_t$$



Autocorrelación simple ACF

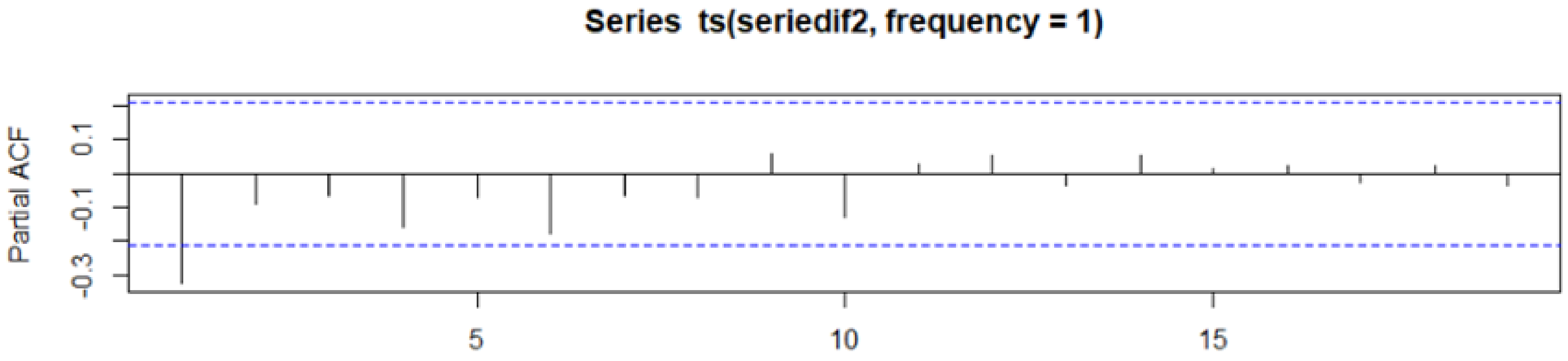
- Si las líneas verticales salen de la sombra azul, quiere decir que deben tener un coeficiente significativo

- La gráfica de autocorrelación simple nos determina el modelo MA



Autocorrelación Parcial

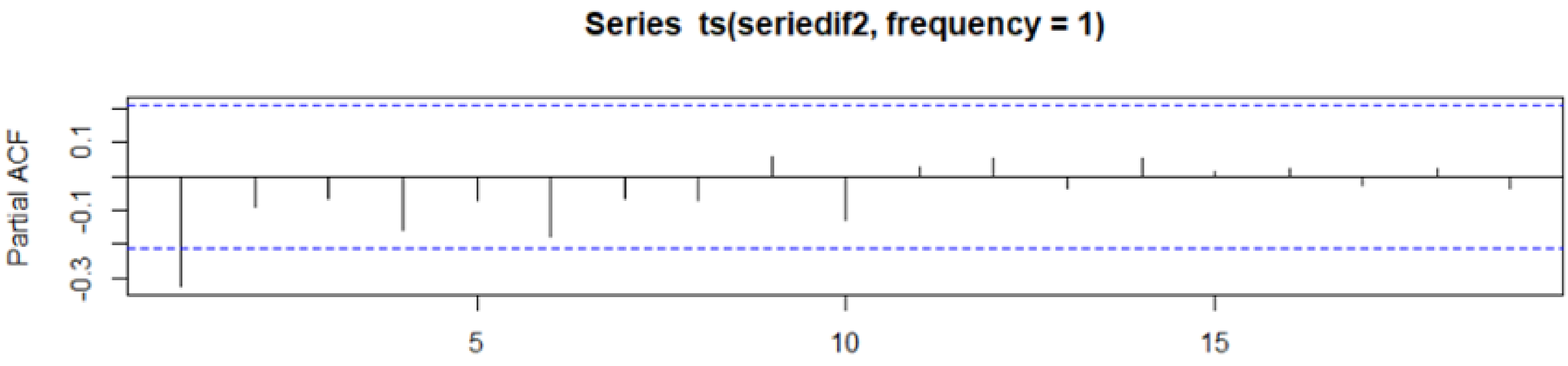
- Determina el impacto de autocorrelación que tiene el valor actual vs un primer, segundo, tercer, etc. Retraso, los que salen de los umbrales son significativos
- La gráfica de autocorrelación parcial nos determina el modelo AR



Modelo AR(p)

$$X_t = \alpha + \varrho_1 X_{t-1} + \cdots + \varrho_p X_{t-p} + \varepsilon_t$$

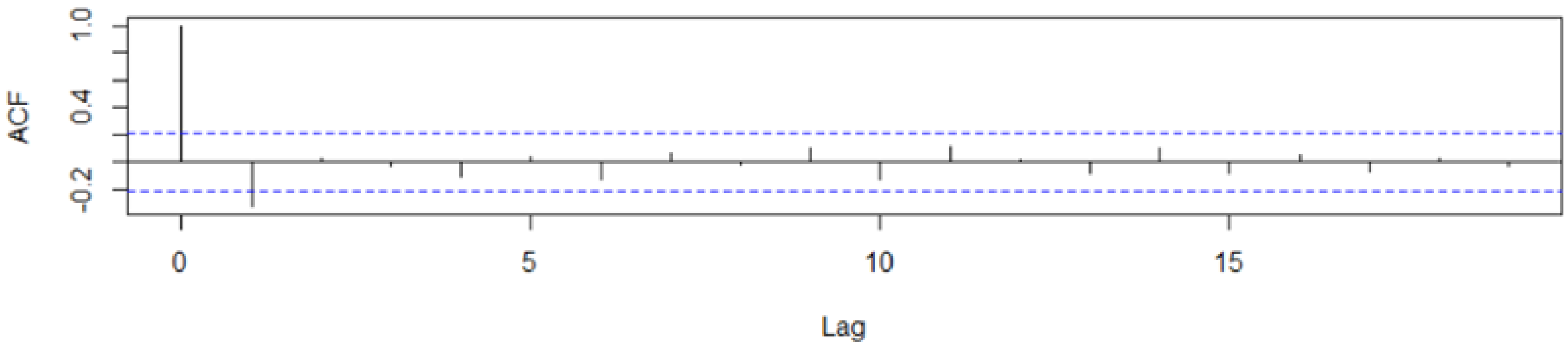
- Quiere decir que el valor actual depende de sus valores pasados:
- Si depende de un solo valor anterior , es una regresión lineal simple
- Si depende de varios valores anteriores es una regresión lineal múltiple



Modelo MA(q)

$$y_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \ldots + \theta_q a_{t-q}$$

- yt: valor actual
- µ : Valor constante sobre la cual se mueve la variable
- ϴ1 , ϴ2 …. , ϴn : Coeficientes a estimar
- at ,at-1 ,…., at-n: Errores observados de los periodos t, t-1, … , t-n
- Quiere decir que el valor actual depende de sus errores pasados
- q : parámetro de la media móvil para predecir y calcular errores



Modelos Mixtos ARMA(p,q)

La consideración de la parte estacional supone la modelización de la parte regular y la estacional, lo que daría lugar a la combinación de modelos ARMA(p,q) x SARMA(P,Q):

$$y_t = \mu + \underbrace{\phi_1 y_{t-1} + \ldots + \phi_p y_{t-p}}_{\text{AR(p)}} + \underbrace{\Phi_1 y_{t-s} + \ldots + \Phi_P y_{t-Ps}}_{\text{SAR(P)}} + a_t + \underbrace{\theta_1 a_{t-1} + \ldots + \theta_q a_{t-q}}_{\text{MA(q)}} + \underbrace{\Theta_1 a_{t-s} + \ldots + \Theta_Q a_{t-Qs}}_{\text{SMA(Q)}}$$

EJERCICIOS: Exprese, en forma algebraica, los siguientes modelos:

ARMA(1,2) $y_t = \mu + \phi_1 y_{t-1} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}$ SARMA(4,0) $y_t = \mu + \Phi_1 y_{t-4} + a_t$

ARMA(0,1) x SARMA(4,8) $y_t = \mu + \Phi_1 y_{t-4} + a_t + \theta_1 a_{t-1} + \Theta_1 a_{t-4} + \Theta_2 a_{t-8}$

ARMA(1,1) x SARMA(12,12) $y_t = \mu + \phi_1 y_{t-1} + \Phi_1 y_{t-12} + a_t + \theta_1 a_{t-1} + \Theta_1 a_{t-12}$

Estacionariedad en los modelos ARIMA

Como los modelos ARMA sólo pueden ser aplicados a series que no muestren ningún tipo de tendencia, será necesario recurrir a **métodos de eliminación de las tendencias (en media y varianza).**

a) **Tendencia en varianza** o heteroscedasticidad. Prácticamente todas las series procedentes del mundo socioeconómico presentan heteroscedasticiadad, en mayor o menor grado.

Dado que tampoco existen contrastes univariantes de homoscedasticidad muy efectivos, en la práctica se opta por eliminar (o mitigar) esta tendencia en todas las series, sometiéndolas a algún tipo de transformación.

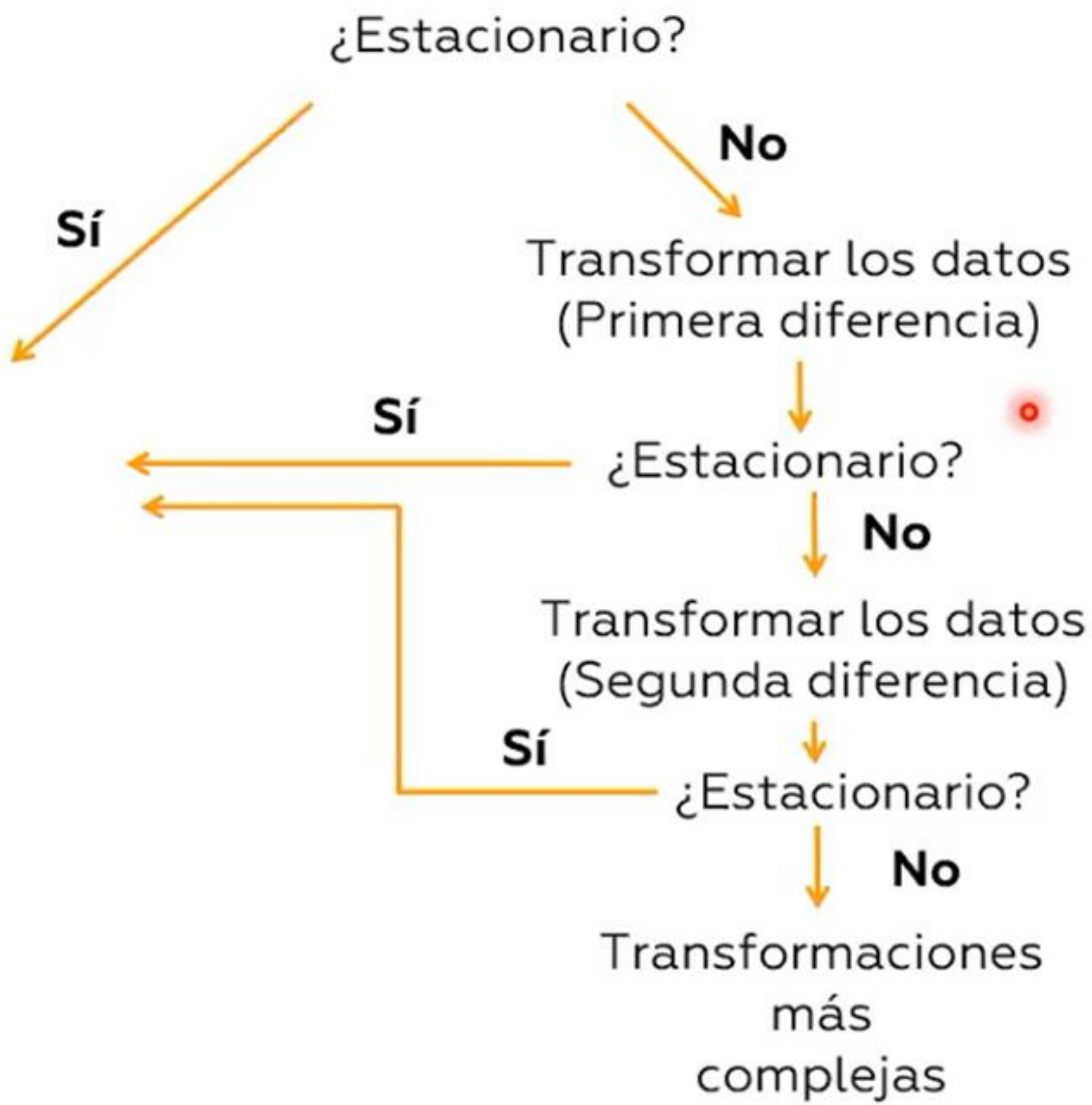
Las transformaciones más habituales son la **transformación logarítmica o cualquier otra perteneciente a la familia de Box-Cox**:

Estacionariedad en los modelos ARIMA

b) Tendencia en media o tendencia (sin más). Una de las formas más elementales de eliminar esta tendencia es proceder a calcular **diferencias sucesivas** de la misma. Así, si la serie y_t muestra tendencia en media, posiblemente ya no la tendrá con la siguiente transformación:

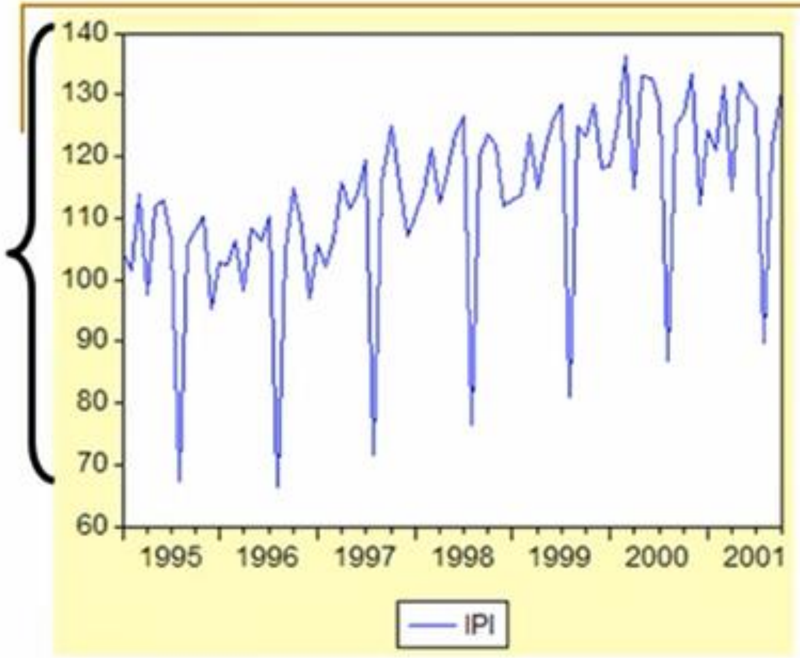
$$\Delta y_t = y_t - y_{t-1}$$

El proceso contrario a la **diferenciación** es la **integración**.
De este modo, las series que necesitan de **1 diferencia** para ser estancionarias en media se dice que son **integradas de orden 1:**

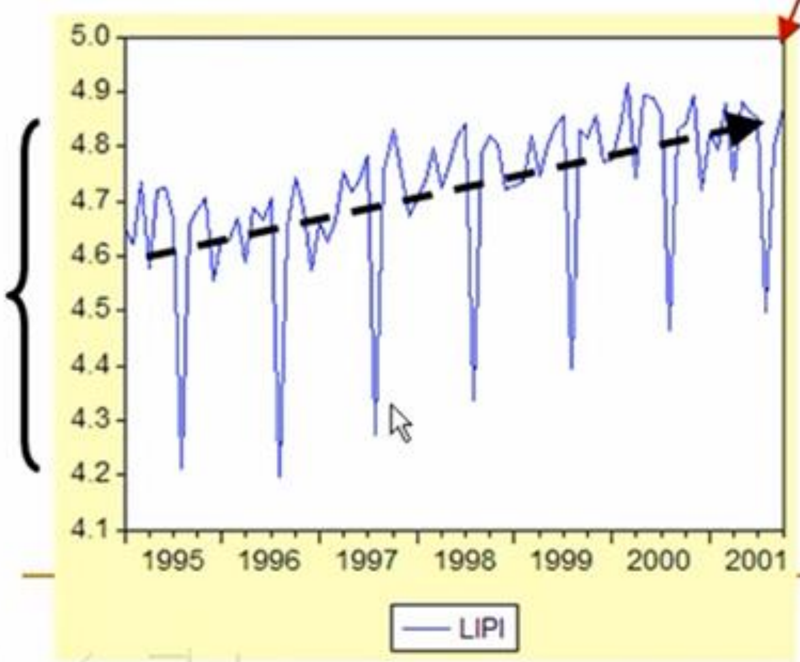

¿Estacionario?
Sí
No
Transformar los datos
(Primera diferencia)
Sí
¿Estacionario?
No
Transformar los datos
(Segunda diferencia)
Sí
¿Estacionario?
No
Transformaciones
más
complejas

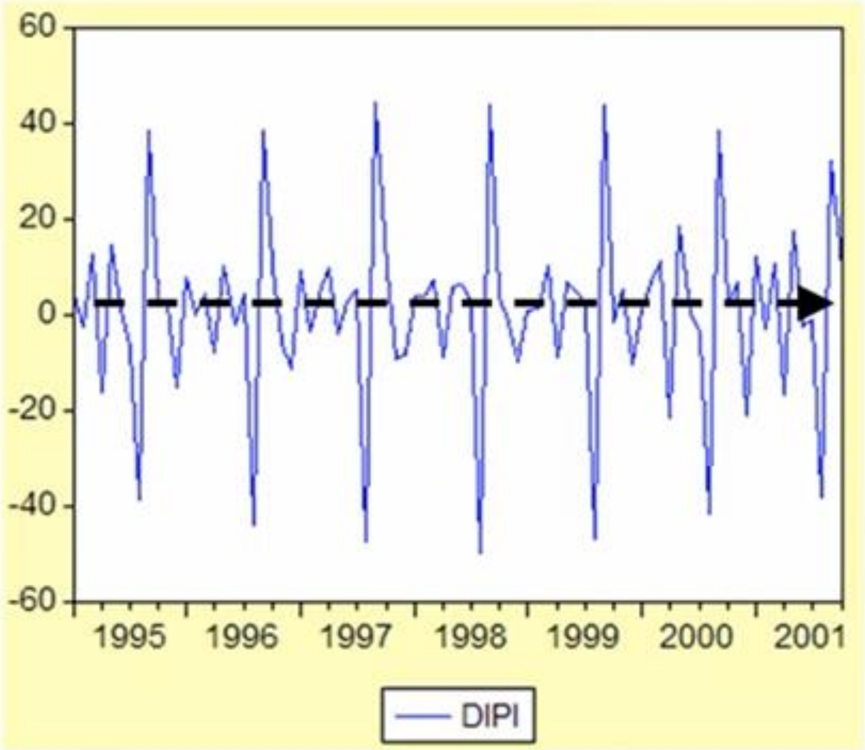
Ejemplo:

1) Serie original = IPI



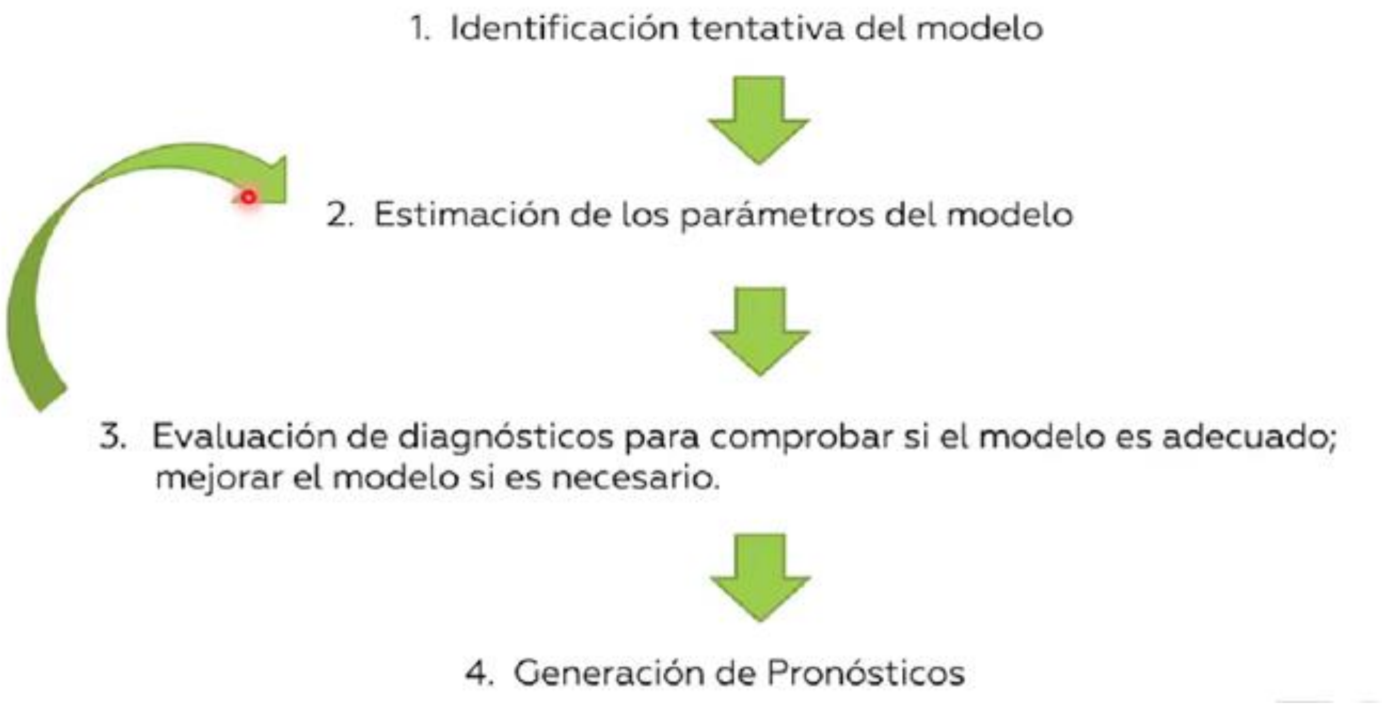
2) Transformación logarítmica: LIPI = LOG(IPI)





3) Transformación en primeras diferencias:**DIPI** = IPI – IPI(-1)

Esquema de escoger modelos series temp



Análisis de residuos

H0: Los residuos son ruido blanco
H1: no es ruido blanco

Ruido blanco significa que el Error:

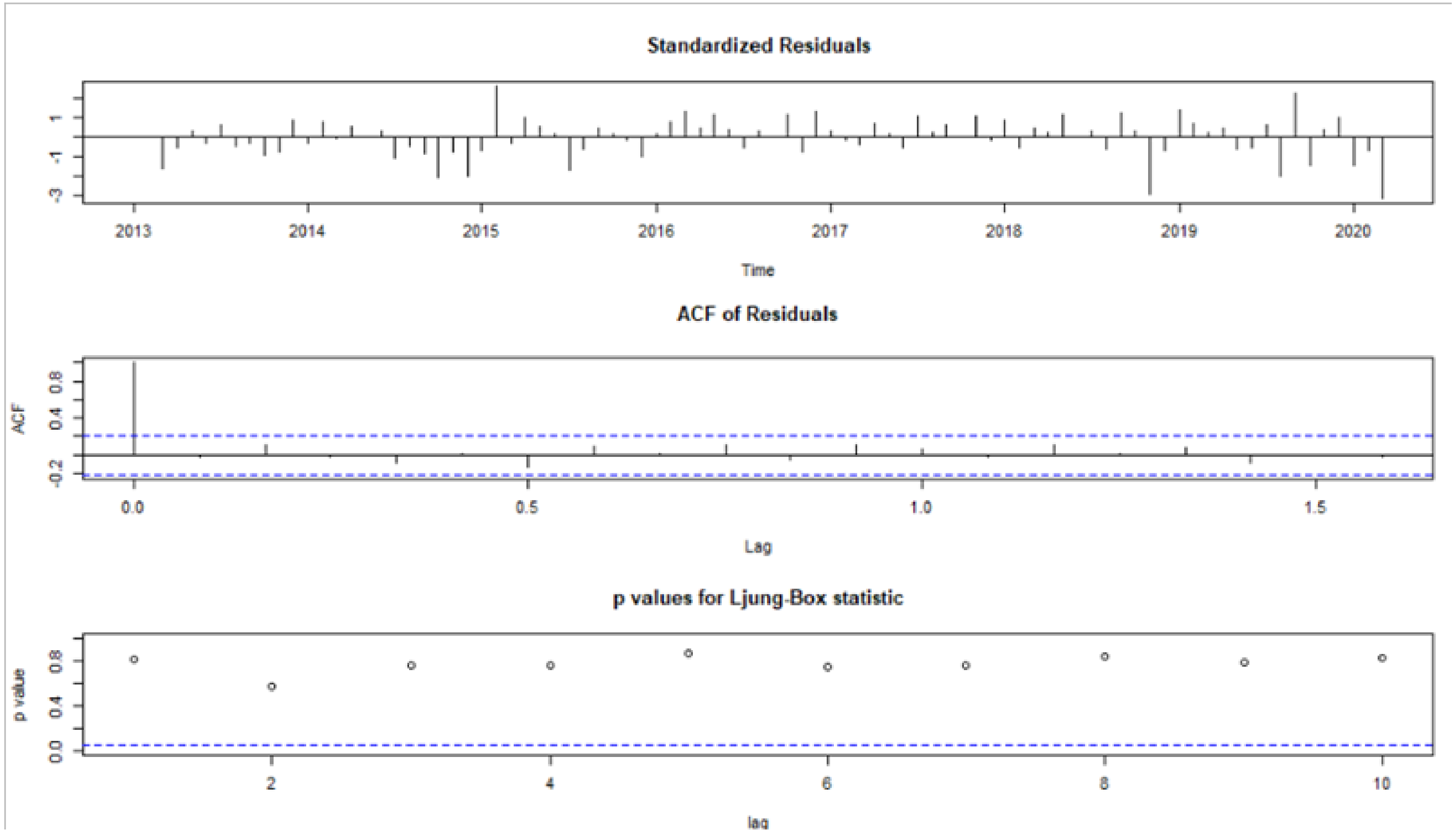
Media igual a cero
Varianza constante
No estar serialmente correlacionada

```
> Box.test(residuals(modelo1),type="Ljung-Box")

        Box-Ljung test

data:  residuals(modelo1)
X-squared = 0.055048, df = 1, p-value = 0.8145
```

Análisis de residuos



Modelos SARIMAX

SARIMAX (1, 0, 2) (2, 0, 1, 5)

$$y_t = c + \varphi_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \Phi_1 (y_{t-5} + \varphi_1 y_{t-6}) +$$

$$\Phi_2 (y_{t-10} + \varphi_1 y_{t-11}) + \Theta_1 (\varepsilon_{t-5} + \theta_1 \varepsilon_{t-6} + \theta_2 \varepsilon_{t-7}) + \varepsilon_t$$