Fundamentos de analítica 2: Modelos de pronósticos basados en series de tiempo.

Diego Agudelo

UNIVERSIDAD ICESI

Temas del día

- Introducción
- Modelo ARIMA
- Evaluación de pronósticos
- Consideraciones finales

Introducción a las series de tiempo

- Ya hemos visto como encontrar un pronóstico por medio de:
  - Media móvil
  - Suavización Exponencial
  - Simple
  - Holt
  - Holt-Winters
  - Tendencias + variables dummy estacionales

- Ahora veremos otra manera de explicar el comportamiento de una serie.

Diego Agudelo
UNIVERSIDAD ICESI

Modelos ARIMA

- Como lo habíamos mencionado…..
- En algunas ocasiones la historia puede afectar el valor presente y futuro de una serie …
- Esta idea fue popularizada por Box-Jenquins a principios de los 70.

Diego Agudelo
UNIVERSIDAD ICESI

Modelos ARIMA

- Pero históricamente estos modelos no han sido ampliamente empleados en la empresas

Porque:

1. Dificultad en la identificación del modelo
2. El reto de explicar como se obtiene el modelo definitivo.

El primer problema se ha resuelto automatizando los cálculos

El segundo es inherente al método

Diego Agudelo

UNIVERSIDAD ICESI

ARIMA Vs Suavización Exponencial

Similitudes

- son métodos de extrapolación
- Más peso a los datos recientes
- Prácticos al momento de generar pronósticos
- En algunos casos generan pronósticos similares

ARIMA Vs Suavización Exponencial

Diferencias

- Los modelos ARIMA se basan en autocorrelaciones
- Suavización Exponencial estima la tendencia
- ARIMA trata de eliminar la tendencia antes de modelar las autocorrelaciones

Procesos ARIMA – Introducción

Estos modelos son construidos para procesos estacionarios

Es decir:

- Con media constante
- Con varianza constante
- Autocovarianza no depende del tiempo

Procesos ARIMA – Introducción

Pero la mayoría de las series empleadas en las empresas son no estacionarias porque:

- Tienen tendencia o Cambios de niveles aleatorios

Procesos ARIMA – Introducción

Por tanto las series son transformadas para convertirlas en estacionarias:

- Lo más común es sacar diferencias (quita la tendencia) $X_t - X_{t-1}$
- Calcular el logaritmo de la serie (Quita la varianza no constante)

Procesos ARIMA – Introducción

Por tanto las series son transformadas para convertirlas en estacionarias:

- Lo más común es sacar diferencias (quita la tendencia) $X_t - X_{t-1}$
- Calcular el logaritmo de la serie (Quita la varianza no constante)

Procesos ARIMA – Introducción

$$y_t = \delta + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Procesos ARIMA – Introducción

Cuadro 8.1 Características de los procesos $AR(p)$ y $MA(q)$.

	$AR(p)$ $\phi_p(L)\, y_t = \varepsilon_t$	$MA(q)$ $y_t = \theta_q(L)\, \varepsilon_t$
Estacionariedad	Todas las raíces de $\phi_p(z) = 0$ están por fuera del circulo unitario	No existe restricción
ACF	$\rho_s \to 0$ exponencialmente u oscilante	$\rho_s \neq 0$ para $s \leq q$ $\rho_s = 0$ para $s > q$
PACF	$rho_s^* \neq 0$ para $s \leq p$ $\rho_s^* = 0$ para $s > p$	$\rho_s^* \to 0$ exponencialmente u oscilante
Invertivilidad	No existe restricción para procesos $AR(p)$ estacionarios	Todas las raíces de $\theta_q(z) = 0$ están por fuera del circulo unitario

Procesos ARIMA – Introducción

Ahora podemos usar la fuerza bruta:

- *criterio de Akaike (Hirotugu AKAIKE, 1981)* (AIC, por su sigla en inglés):

$$AIC(p,q) = \ln\left(\hat{\sigma}^2\right) + \frac{2(p+q)}{T}$$

- *criterio Bayesiano de Schwarz (Schwarz, 1978)* (SBC, por su sigla en inglés):

$$SBC(p,q) = \ln\left(\hat{\sigma}^2\right) + \frac{\ln(T)\,2(p+q)}{T}$$

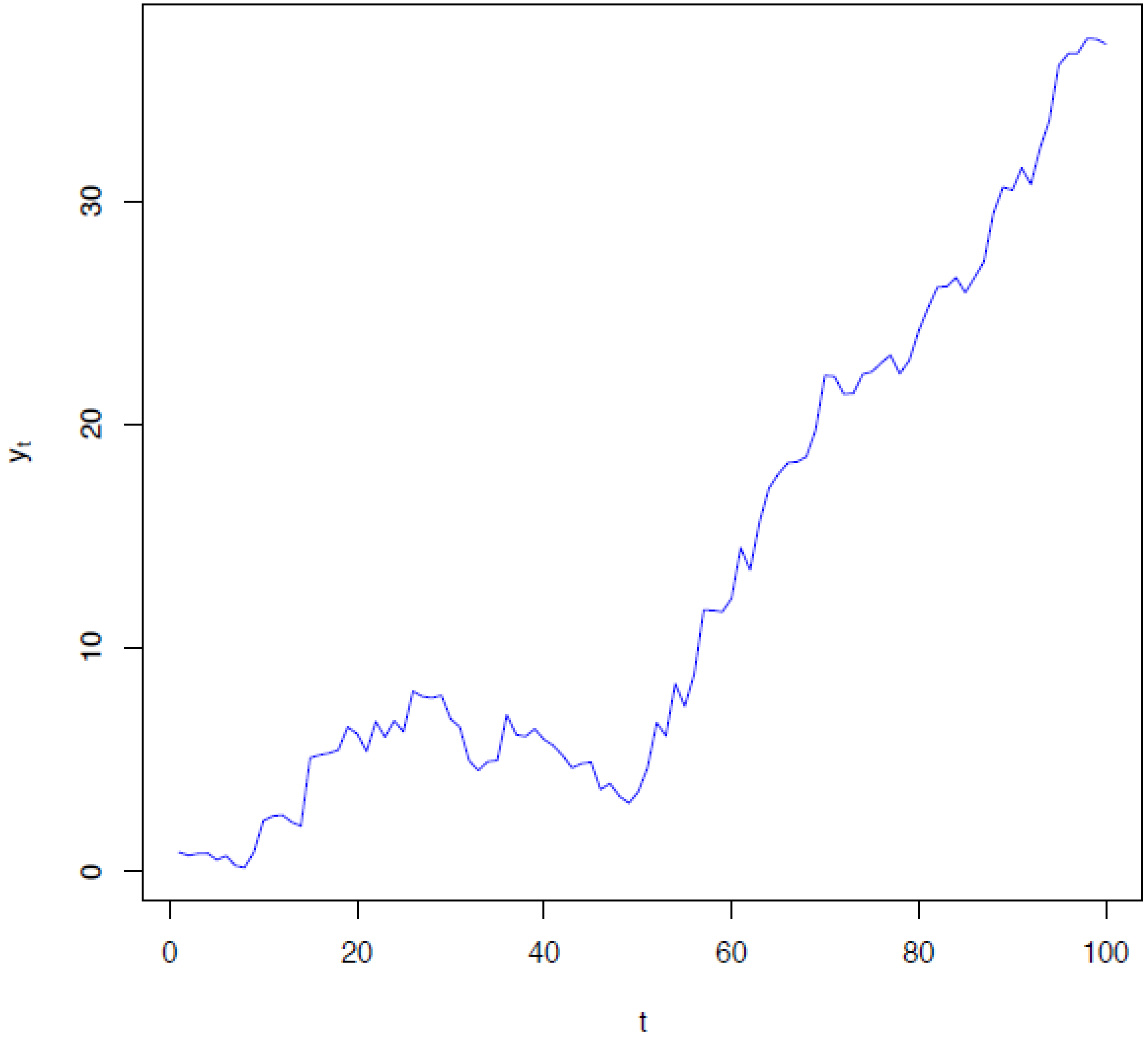
Procesos ARIMA – Introducción

Pero nada garantiza que los errores cumplan los supuestos:

- No Autocorrelación
- No comportamiento GARCH o ARCH
- Normalidad?

Procesos ARIMA – Introducción

En algunos casos las series no son estacionarias:

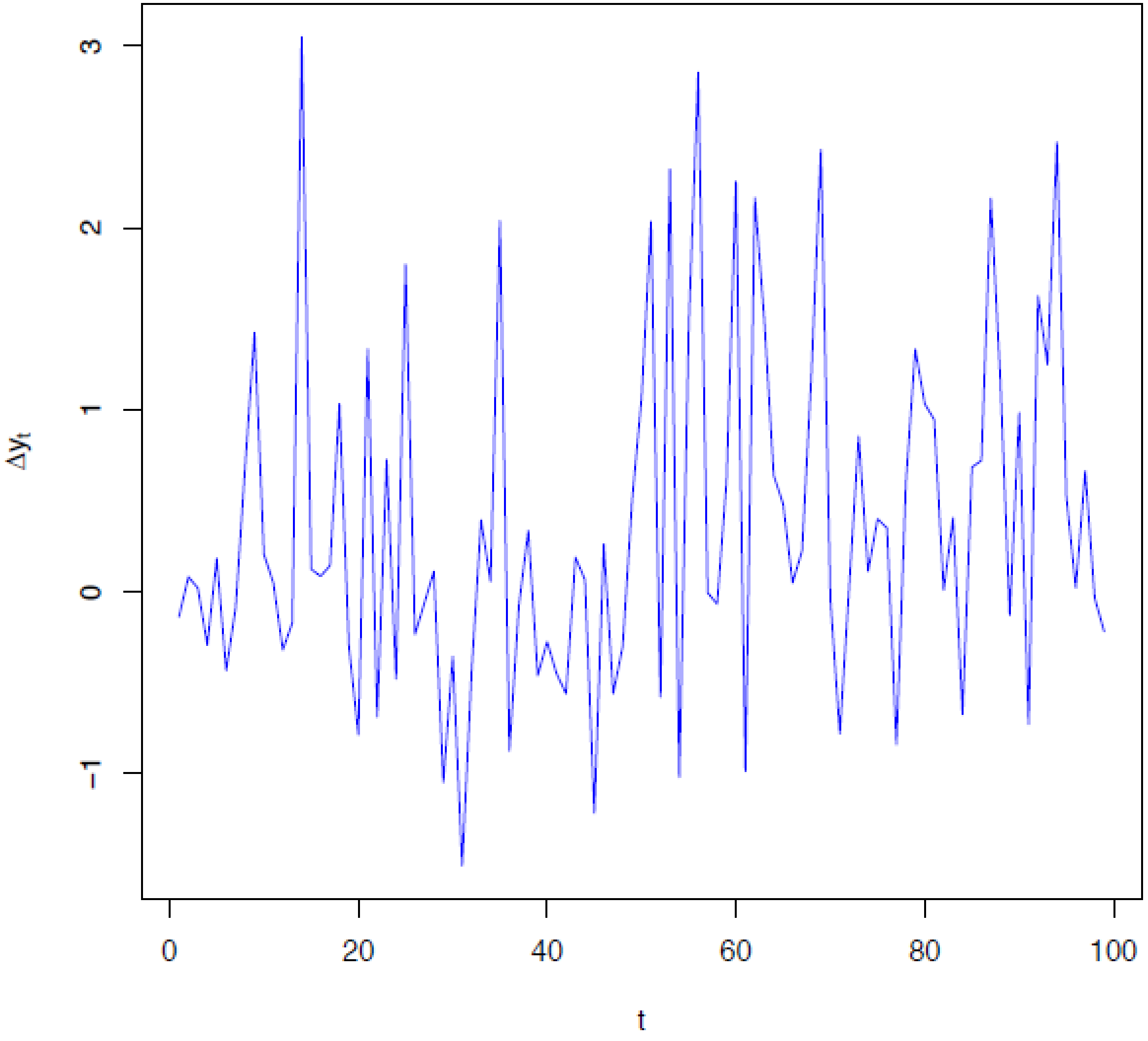


Diego Agudelo

UNIVERSIDAD ICESI

Procesos ARIMA – Introducción

Podemos integrar (l(1))



Diego Agudelo
UNIVERSIDAD ICESI

Procesos ARIMA – Introducción

Procesos ARIMA

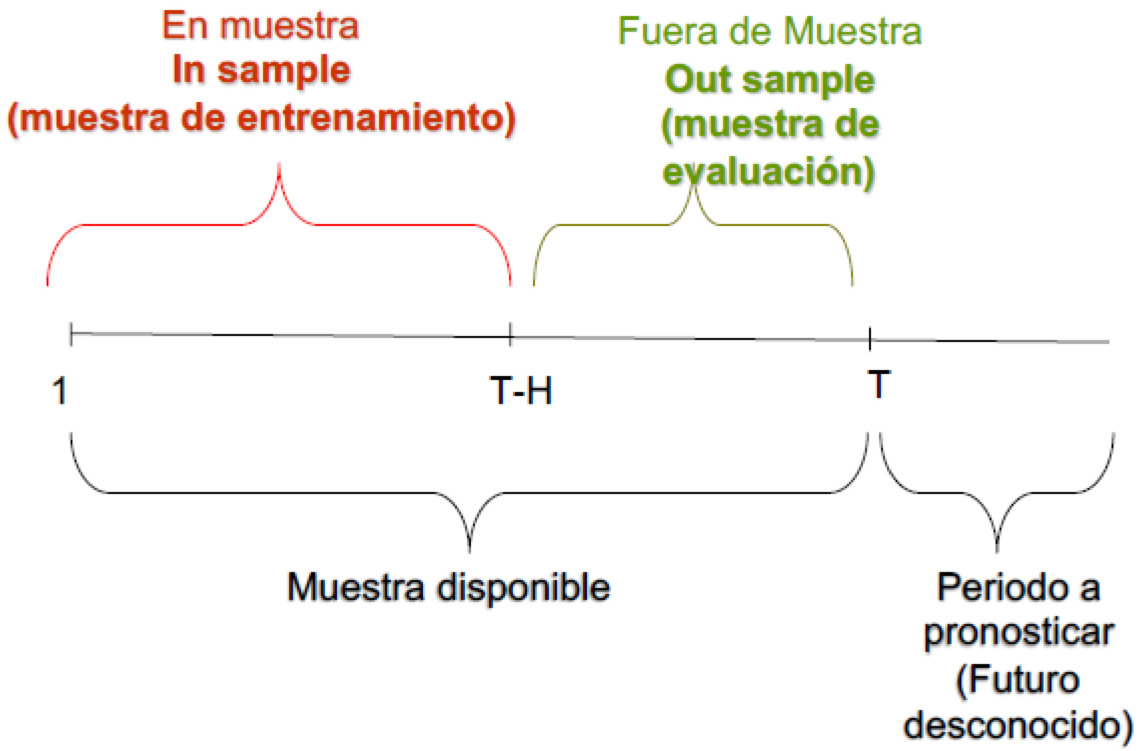
- ***ARIMA*** =
- Modelo compuesto por
- ***AR*** = procesos Autoregresivos
- ***MA*** =Procesos de media móvil
- ***I*** = # de veces que se diferencia la serie para obtener un proceso estacionario

$$\Delta^d y_t = \delta + \sum_{i=1}^{p} \phi_i \Delta^d y_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Diego Agudelo
UNIVERSIDAD ICESI

Procesos ARIMA – Introducción

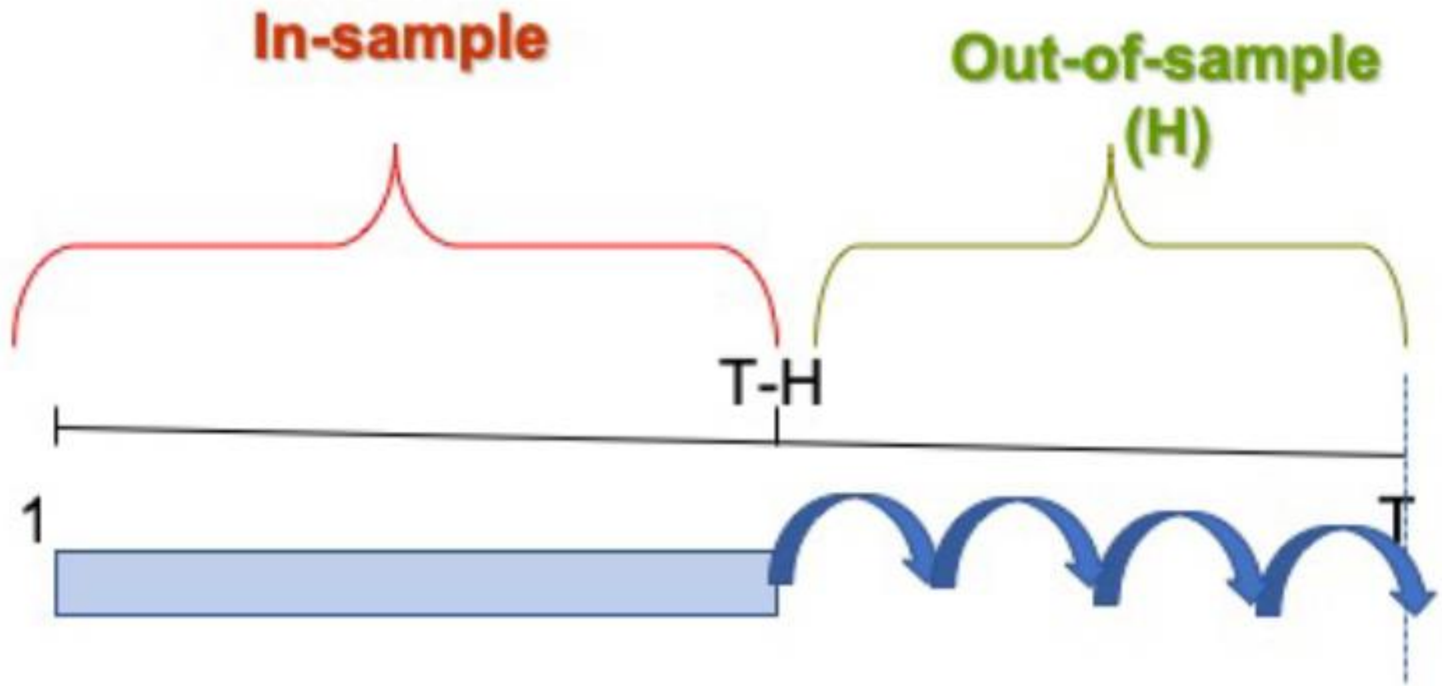
Como comparar la capacidad de predicción de varios modelos:



Diego Agudelo
UNIVERSIDAD ICESI

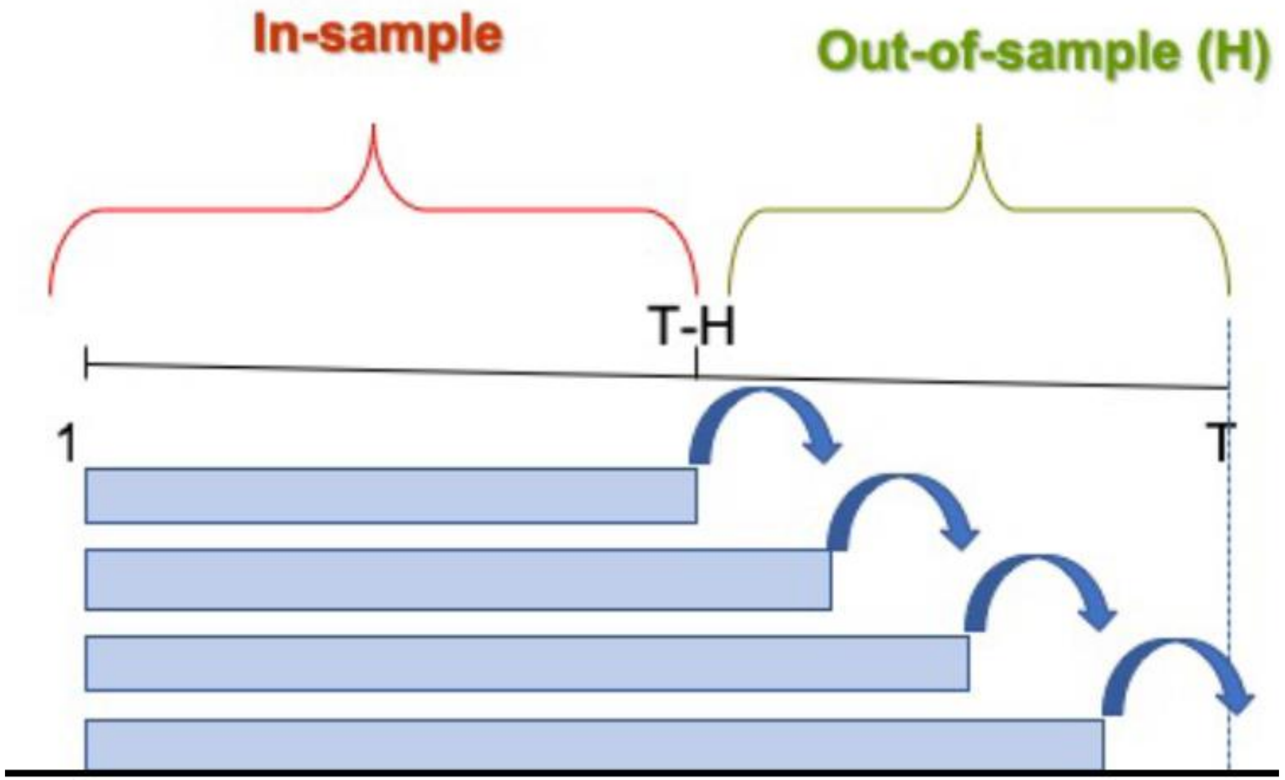
Procesos ARIMA – Introducción

Ventana Fija:



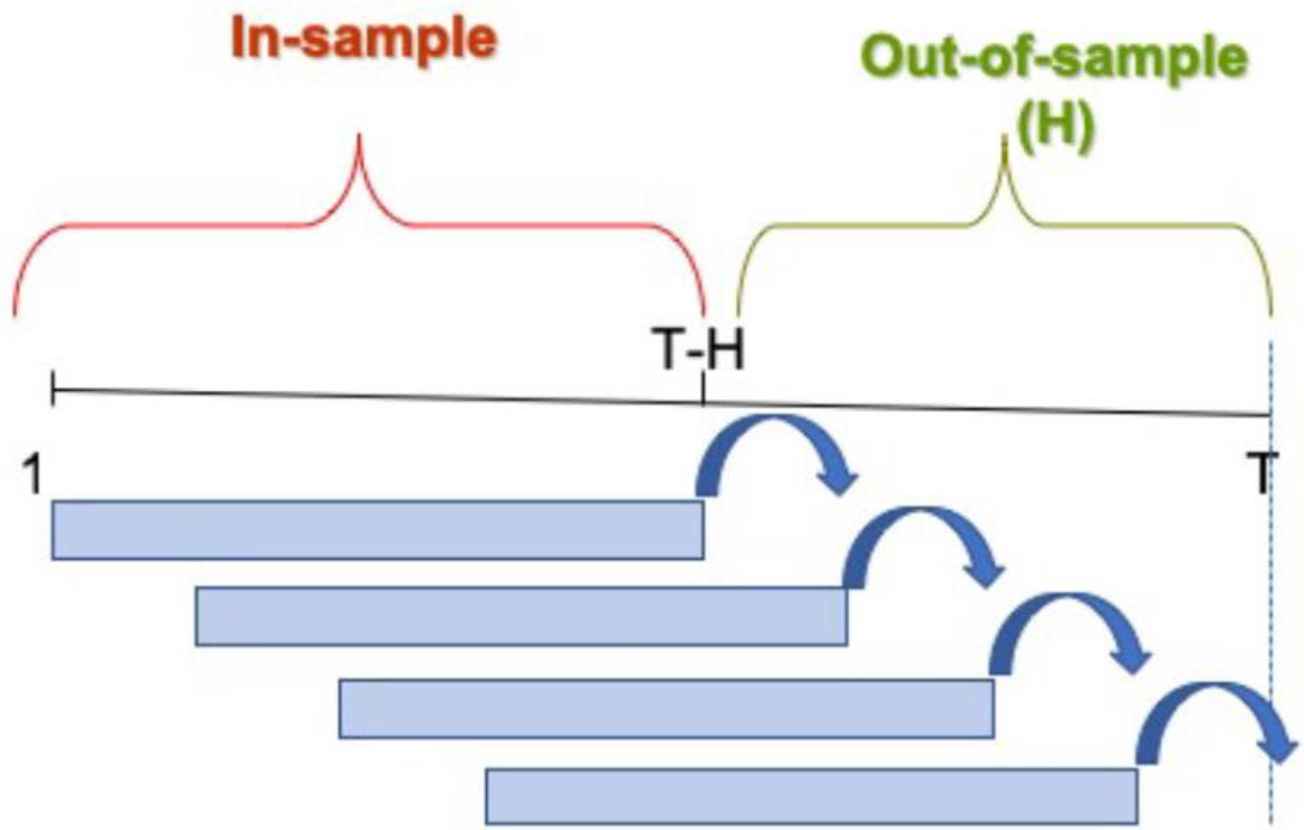
Procesos ARIMA – Introducción

Ventana Recursiva:



Procesos ARIMA – Introducción

Rolling window:



Diego Agudelo

Metricas

- *RMSE* (Raíz media cuadrada del error):

$$RMSE = \sqrt{\frac{1}{h}\sum_{t=T-H}^{T}(\hat{y}_t - y_t)^2}$$

- *MAE* (error absoluto promedio):

$$MAE = \frac{1}{h}\sum_{t=T-H}^{T}|\hat{y}_t - y_t|$$

- *MAPE* (error absoluto porcentual promedio):

$$MAPE = \frac{1}{h}\sum_{t=T-H}^{T}\frac{|\hat{y}_t - y_t|}{\hat{y}_t}$$

Diego Agudelo
UNIVERSIDAD ICESI

Consideraciones finales

Errores son ruido blanco

Autocorrelación	• Box-Pierce – Ljung-Box • Prueba de rachas • Gráficos ACF y PACF
No comportamiento ARCH o GARCH (homoscedasticidad)	• Ljung-Box (residuales al cuadrado) • Gráficos ACF y PACF (residuales al cuadrado)
Normalidad	• Jarque Bera y Shapiro-Wilk • q-q plot

Diego Agudelo
UNIVERSIDAD ICESI