UNIVERSIDAD NACIONAL DE INGENIERÍA

ESCUELA CENTRAL DE POSGRADO

Facultad de Ingeniería Económica, Estadística y CCSS

Unidad de Posgrado - FIEECS

Curso: Forecasting for Data Science

Docente: Mg. Omar A. Chincaro Del C.

Forecasting for Data Science

1.- Introducción

2.- Componentes de una serie temporal

3.- Suavizado de un serie

4.- Series estacionarias:

- Modelos Autorregresivos (AR(p))
- Modelos Medias Móviles (MA(q))
- Modelos Mixtos: ARMA(p,q)

5.- Series no estacionarias

- Modelos Mixtos: ARIMA(p,d,q)

6.- Series estacionales: SMA, SAR, SARI, SIMA, SARIMA

7.- Series heterecedásticas: ARCH y GARCH

8.- Pronósticos

SERIES ESTACIONARIAS

¿Qué es la estacionariedad en serie temporal? ¿Por qué es estacionaria una serie?

Una serie estacionaria cuando la media y varianza de la serie se mantienen constante a lo largo del tiempo, es decir tanto la media como la varianza no dependen del tiempo.

MODELOS PARA UNA SERIE DE TIEMPO

A) FORMA DE MEDIA MOVIL INFINITO

Esta forma se fundamenta en el hecho de que toda la serie se origina como una suma infinita de choques aleatorios:

$$Z_t = \theta_0 + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3} + \theta_4\varepsilon_{t-4} + \ldots\ldots\ldots\ldots\ldots.$$

$$Z_t = \theta_0 + \sum_{j=1}^{\infty} \theta_j \varepsilon_{t-j}$$

$$Z_t = \Theta_0 + \theta(B)\,\epsilon_t$$

Donde θ_j son los parámetros del modelo, ε_t las perturbaciones del modelo, B es operador de desface y t es el tiempo.

¿Qué es la estacionariedad en serie temporal? ¿Por qué es estacionaria una serie?

B) FORMA AUTOREGRESIVA E INVERTIDA

Es cuando el modelo se puede expresarse en su forma autorregresiva, es decir, como una suma ponderada de sus valores pasados mas un choque aleatorio:

$$Z_t = \phi_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + \phi_4 Z_{t-4} + \dots\dots\dots\dots\dots.$$

$$Z_t = \phi_0 + \sum_{j=1}^{\infty} \phi_j Z_{t-j}$$

$$Z_t = \phi_0 + \phi(B)\, Z_t$$

Puede observarse que los polinomios $\theta(B)$ **y** $\phi(B)$ son inversos y determinan la invertibilidad y estacionariedad de los procesos respectivos.

CONDICIONES DE ESTACIONARIEDAD E INVERTIBILIDAD

1.- ESTACIONARIEDAD
Una serie de tiempo es estacionaria si tienen medias y varianzas finitas, no dependan del tiempo. Es decir cuando las raíces del polinomio que la generan tengas raíces mayores que la unidad.

2.- INVERTIBILIDAD
Una serie de tiempo es invertible cuando es posible expresar en su forma autorregresiva, implicando que la ecuación característica que lo generan tenga raíces mayores que la unidad

3.- CORRELOGRAMA
El análisis estadístico y la inspección gráfica visual de los valores muestrales tanto de las funciones de autocorrelación como los funciones de autocorrelación parcial constituyen lo que se llama "Análisis del correlograma"

3.1.- FUNCION DE AUTOCORRELACIÓN (FAC)

Esta función es la más usada que se utiliza para describir la estacionariedad o no estacionariedad de una serie de tiempo, consecuentemente en ella se determinan el orden de las Medias Móviles. Se definen de la siguiente manera:

$$\mathrm{CORR}(K) = \frac{COV(K)}{COV(0)}$$

3.2.- FUNCION DE AUTOCORRELACIÓN PARCIAL (FACP)

La función de autocorrelación parcial se define como:

$$\Phi(kk) = \mathrm{CORR}[(Z(t) - E(Z(t) / Z(t-1), Z(t-2), \ldots\ldots, Z(t-k+1)) ; Z(t-k) - E(Z(t-k) / Z(t-1), Z(t-2), \ldots\ldots, Z(t-k+1))]$$

y sirve para determinar el orden de la componente autoregresiva de la serie temporal.

¿Por qué es importante que la serie de tiempos sea estacionaria?

En la mayoría de los modelo de series temporales es indispensable que la serie sea estacionario

¿Cómo conseguimos la estacionariedad en varianza?: Aplicando log neperiano a la serie y/o haciendo uso de las transformaciones de box – cox.

¿Cómo conseguimos la estacionariedad en media? Tomando la primera diferencia de la serie

¿Raíces unitarias?

Una raíz unitaria es una tendencia estocástica de la serie temporal. Existen dos formas de remover la tendencia de una serie: 1) Hacer las primeras diferencias y 2) Regresión tendencial.

Los test de raíces unitarias puede ser utilizado para determinar si una serie tienen tendencias:

a) **El test de Dicker – Fuller Aumentado (ADF),** elimina la autocorrelación e indica si una serie es estacionario o no
b) **El test de Phillips - Perron (PP),** corrige la autocorrelación y heterocedasticidad de los errores.

PROCESO DE MODELACIÓN

METODOLOGÍA DE BOX JENKINS

ESPECIFICACIÓN

Analizar la gráfica de la serie (análisis situacional y estadístico).
Identificar tendencias y volatilidades en la serie.

IDENTIFICACIÓN

Analizar la estacionariedad en varianza y en media
Usar test de raíz unitaria
Transformación de Box Cox
Usar las FAC y FCP para identificar posibles modelos.

ESTIMACIÓN

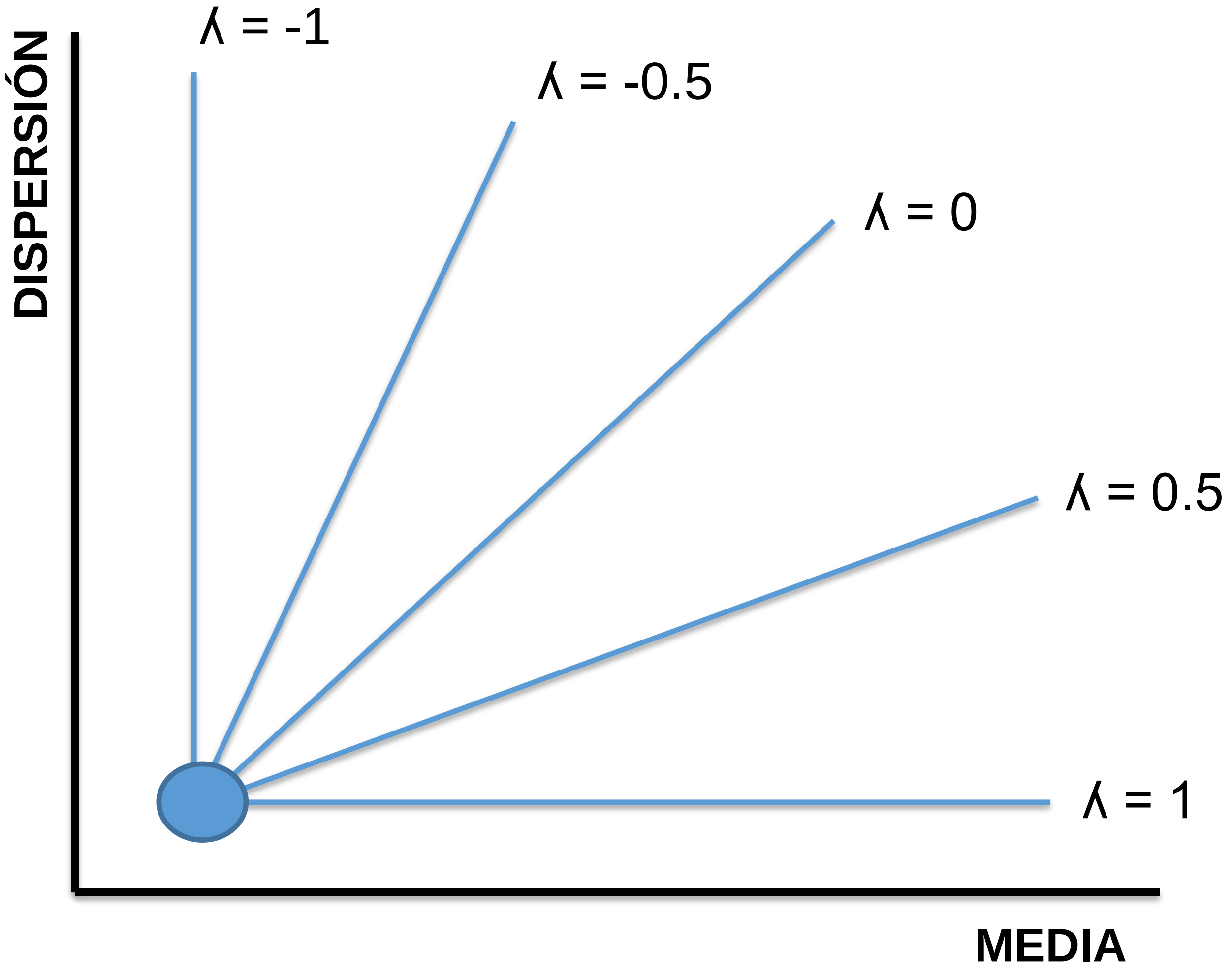
Estimar los parámetros del modelo con métodos estadísticos.
Analizar la significancia de los parámetros estimados.

VERIFICACIÓN

Análisis de residuales del modelo: Aleatoriedad y Normalidad.
Validar los supuestos del modelo.

PRONÓSTICOS

TRANSFORMACIÓN DE BOX COX



$$y(\lambda) = \begin{cases} \frac{y^{(\lambda)}-1}{\lambda}, \lambda \neq 0 \\ ln(y), \lambda = 0 \end{cases}$$

LAMBDA	TRANSFOMRACIÓN
1	No existe transfomación a la serie
0.5	Raíz cuadrada a la serie
0.125	Raíz cuarta
0	Logaritmo neperiano
-0.125	Inversa de la raíz cuarta
-0.5	Inversa de la raiz cuadrada
-1	Inversa de la serie

MODELOS AR(P)

DEFINICIÓN

$$Z_t = \phi_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + \phi_4 Z_{t-4} + \dots\dots\dots\dots\dots + \phi_p Z_{t-p} + \varepsilon_t$$

ε_t ~ (0,Var(error)) Var(error) = cte ε_t = Aleatorios independientes

ESPERANZA

$$E(Z_t) = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \dots\dots\dots - \phi_P}$$

VARIANZA

$$VAR(Z_t) = \frac{Var(Error)}{1 - \phi_1\rho_1 + \phi_2\rho_2 + \dots\dots\dots + \phi_p\rho_P}$$

ρ_p Autocorrelación de la serie

CONDICIONES DE ESTACIONARIEDAD E INVERTIBILIDAD

1.- ESTACIONARIEDAD

Una serie de tiempo es estacionaria si tienen medias y varianzas finitas, no dependan del tiempo. Es decir cuando las raíces del polinomio que la generan tengas raíces mayores que la unidad.

2.- INVERTIBILIDAD

Una serie de tiempo AR(p) es invertible cuando es posible expresar en su forma autorregresiva

3.- CORRELOGRAMA

El análisis estadístico y la inspección gráfica visual de los valores muestrales tanto de las funciones de autocorrelación como los funciones de autocorrelación parcial constituyen lo que se llama “Análisis del correlograma”

MODELOS MA(q)

DEFINICIÓN

$$Z_t = \theta_0 + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3} + \theta_4\varepsilon_{t-4} + \ldots\ldots\ldots\ldots\ldots + \varepsilon_{t-q} + \varepsilon_t$$

ε_t ~ (0,Var(error)) Var(error) = cte ε_t = Aleatorios independientes

ESPERANZA

$$E(Z_t) = \theta_0$$

VARIANZA

$$VAR(Z_t) = Var(Error) * (1 + \theta_1^2 + \theta_2^2 + \ldots\ldots\ldots\ldots.. + \theta_q^2)$$

CONDICIONES DE ESTACIONARIEDAD E INVERTIBILIDAD

1.- ESTACIONARIEDAD

La serie MA(q) es estacionario dado que no existen tendencia siempre y cuando la V(Error) sea constante en el tiempo.

2.- INVERTIBILIDAD

Una serie de tiempo MA(q) es invertible cuando $|\theta_j| < 1$

3.- CORRELOGRAMA

El análisis estadístico y la inspección gráfica visual de los valores muestrales tanto de las funciones de autocorrelación como los funciones de autocorrelación parcial constituyen lo que se llama "Análisis del correlograma"

MODELOS ARMA(p,q)

DEFINICIÓN

$$Z_t = \phi_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots\ldots. + \phi_p Z_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \varepsilon_t$$

ε_t ~ (0,Var(error)) Var(error) = cte ε_t = Aleatorios independientes

ESPERANZA

$$E(Z_t) = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \ \ldots\ldots\ldots - \phi_P}$$

SERIES NO ESTACIONARIAS

INTRODUCCIÓN

También llamados Procesos Arima, presentan en su mayoría series de tiempo con un grado de no estacionariedad en media y/o varianza. Consiste en realizar diferencia simples y/o realizar transformaciones sobre la serie observada, con la finalidad de obtener una serie estacionaria. Dentro de esta clase de Modelos se encuentran: ARIMA(p,d,q), ARI(p,d,0) y los IMA(0,d,q).

PROCESOS ARIMA(p,d,q)

$$\Delta^d Z_t = \varphi(B) Z_t + \omega(B)\varepsilon_t$$

$$Z_t = \phi_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots\ldots + \phi_p Z_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \varepsilon_t$$

ESPECIFICACIÓN DE UN MODELO ARIMA

PASO 1: No estacionariedad en varianza

a) **Gráfica de la serie:** Se da cuando en la gráfica de la serie existe mucha volatilidad con respecto a su valor promedio. Si esto ocurre hay que hacer transformaciones de BOX – COX
b) **Grafica de dispersión vs media:** Esta grafica es importante por que dá una idea que tipo de transformación hay que realizar a la serie.
c) **Determinar el valor optimo de (ʎ):** Con la finalidad de encontrara la transformación de BOX – COX más optima.

ESPECIFICACIÓN DE UN MODELO ARIMA

PASO 2: No estacionariedad en media

a) **Gráfica de la serie:** Se da cuando en la gráfica de la serie existe tendencia. Si esto ocurre hay que tomar tantas diferencias sea necesaria a la serie, hasta obtener un grafica de una serie sin tendencias.
b) **Gráfica de correlograma:** Cuando no hay un decaimiento rápido en las funciones de autocorrección parcia (facp) y autocorrelación (fac), entonces las serie no es estacionaria en media. Sin embargo el correlograma de la serie diferenciada si existe un decaimiento rápido, si esto ocurre la serie es estacionaria en media.

PASO 3: Determinación de los valores de "p" y "q"

a) **Valores de p:** Aquí hay que ver el decaimiento rápido de la facp del correlograma
b) **Valores de q:** Aquí hay que ver el decaimiento rápido de la fac del correlograma