Capítulo 2

**2. MARCO TEÓRICO**

* 1. **Series de Tiempo**

Una serie de tiempo es una realización de un proceso estocástico.

 

 **Figura 2.1 Serie de Tiempo**

Toda serie temporal debe estar constituida por cuatro componentes:

**Tendencia.-** Componente dominante a través del tiempo a largo plazo.

**Componente cíclica.-** Comportamientos periódicos observados en períodos largos.

**Componente oscilatoria.-** Comportamientos irregulares que se desarrollan en períodos pequeños.

**Componente aleatoria.-** La parte estocástica del proceso.

Para poder emplear un modelo de análisis de series temporales es necesario referirnos a los procesos estocásticos, ya que la secuencia de observaciones se registra a través del tiempo de acuerdo a una ley de probabilidad.

* 1. **Procesos Estocásticos**

Un Proceso Estocástico se define como una familia de variables aleatorias (definidas en una espacio probabilístico (Ω,Λ, p) ) que corresponden a momentos sucesivos del tiempo. Se nota por Y(t, ω) donde t es el tiempo y ω∈Ω, donde ω es la variable aleatoria.

El análisis de un proceso estocástico puede hacerse de 2 formas:

1. Mediante las funciones de distribución conjunta.
2. A partir de los momentos.

Notaremos ahora Y(t, ω) por Yt.[[1]](#footnote-2)

Para t = ti la función de distribución de la variable aleatoria Yti es F(Yti).

Para t = ti y t= tj, la función de distribución conjunta es F(Yti,Ytj).

En general para un conjunto finito de valores en el tiempo t1,t2,.....,tn la distribución conjunta es F(Yt1, Yt2,.......,Ytn).

Se dice que un proceso estocástico está perfectamente caracterizado cuando se pueden determinar las funciones de distribución conjunta para cualquier conjunto finito de variables t del proceso, es decir, para cada valor finito de n se puede determinar:

 F(Yt1, Yt2, ........., Ytn)

Este procedimiento en sí es muy complicado, por lo que se acostumbra a utilizar el método de los momentos.

En una distribución de probabilidades se puede calcular momentos de diverso orden, aunque los más utilizados son los de 1er orden y 2do orden.

En un proceso estocástico, la media o momento de 1er orden se define por:

Como momentos de segundo orden con respecto a la media, además de la varianza, se considera las covarianzas entre variables aleatorias referidas a distintos instantes de tiempo (autocovarianzas).

 Cuando t =s tenemos que:

 Los coeficientes de autocorrelación son:

Es preferible usar las autocorrelaciones pues estas carecen de unidades, es decir son medidas relativas a las que no afectan las unidades de medida.

La caracterización de un proceso estocástico por medio de los momentos de primero y segundo orden es en principio más incompleta que cuando se hace con funciones de distribución conjunta. Ahora si el proceso es normal (gaussiano) este queda perfectamente caracterizado a través de los dos primeros momentos gaussianos.

En una serie temporal se dispone de una observación para cada periodo de tiempo, por lo que se la puede considerar como una muestra de tamaño 1 tomada en periodos sucesivos de tiempo en un proceso estocástico.

A diferencia de un muestreo aleatorio simple donde cada extracción es independiente a las demás, en una serie temporal, el dato extraído para un período de tiempo, no será en general independiente de los datos de los períodos anteriores.

Si se dispone de n datos de una serie temporal con ellos hay que estimar n medias y n varianzas sin contar con las autocovarianzas, entonces tal como se plantea el problema, no tiene solución.

Para poder, a partir de una realización, efectuar inferencias sobre un proceso estocástico es preciso imponer restricciones a dicho proceso, las cuales son: que sea estacionario y ergódico.

* 1. **Procesos Estacionarios**

Se dice que un proceso estocástico es estacionario en sentido estricto cuando al realizar un mismo desplazamiento en el tiempo de todas las variables aleatorias de cualquier distribución conjunta finita, resulta que la distribución no varía, es decir:

 F(Yt1, Yt2, Yt3,..........,Ytk) = F(Yt1+n, Yt2+n,....................,Ytk+n)

Se dice que un proceso estocástico es estacionario de 1er orden o en media si : t E[Yt ] =

Es decir, la media permanece constante en el tiempo.

Se dice que un proceso estocástico es estacionario de 2 orden

(o en sentido amplio) si se verifica :

2) La autocovarianza entre dos períodos distintos de tiempo viene afectada únicamente por el lapso de tiempo transcurrido entre estos dos períodos.

En un proceso estacionario en sentido amplio las autocorrelaciones están dadas por :

La representación gráfica de ρk para valores de k = 0, 1, 2 .......... se denomina Correlograma.

 

 **Figura 2.2.**

Cuando el proceso estocástico es estacionario en sentido amplio, puede estimarse los parámetros , γo, γ1, γ2.......... a partir de una sola realización.

A más de la estacionariedad es necesario que el proceso estacionario sea ergódico.

Si un proceso estacionario es ergódico entonces:

Si para valores altos de k, ρk también es alto, se tiene que al aumentar el tamaño de la muestra se añade poca información nueva, en consecuencia los estimadores obtenidos no serían consistentes.

* 1. **Procesos Lineales**

Son procesos estocásticos estacionarios en sentido amplio y ergódicos que consisten en combinaciones lineales de variables aleatorias

* Proceso puramente aleatorio: “ruido blanco”

* Procesos Autorregresivos de orden p: AR(p)

* Procesos de medias móviles de orden q MA (q)

* Procesos ARMA(p,q), combinación de los dos anteriores.

* 1. **Ecuaciones en Diferencias Finitas**

**Operador de diferencias: Δ**


#### Operador de retardos: L

El operador L se puede usar para expresar un modelo con retardos:

 Equivale a:



 Operador polinomial

 de retardos: φ(L) 🡺 φ(L)Yt = ∈t

#####  Ecuaciones de 1er orden

 Asumamos que la solución es de la forma:

 Si t = 0

Una condición suficiente y necesario para que el proceso estocástico sea estacionario es que :

 Puesto que:

Alternativamente, se puede usar la ecuación polinomial de retardos:

 Como sabemos que:


##  Ecuaciones de 2do orden

En una ecuación homogénea de 2do orden, cuando las raíces son reales y diferentes la condición de estacionariedad implica que:

 Si es raíz única:

Si son imaginarias o complejas se tiene que r<1, es decir las raíces deben caer dentro del círculo unidad.

Como alternativa a la ecuación característica, se puede usar la ecuación polinomial de retardos:

En este caso las raíces L1 y L2 deben caer fuera del círculo unidad.

* 1. **Modelos Lineales**
		1. **Modelos Autorregresivos (AR)**

Un modelo de proceso estocático será un AR de orden p, si se cumple que los valores presentes se los representa como p valores pasados.

**Modelo AR (1)**

Una condición necesaria y suficiente para que el proceso estocástico sea estacionario es que |1| < 1.

 Si |1| < 1 entonces las autocovarianzas serían:



 Las autocorrelaciones estarían dadas por:

 **Modelo AR(2)**

 Para que un modelo AR(2) sea estacionario se tiene que dar:

 Las autocovarianzas están dadas por:

 Las autocorrelaciones están dadas por:



 **Modelo AR(p)**

ó



Para que el proceso sea estacionario se requieren que las raíces de la ecuación φ(L)=0 estén fuera del círculo unidad.

Si tenemos que:

Y si multiplicamos esta ecuación por Yt-k y a su vez tomamos esperanzas, tenemos que:

 Si dividimos la última ecuación para γo tenemos:

Tomando ρo,ρ1,ρ2,........,ρp-1 como condicione iniciales determinadas a partir de φo,φ1,φ2,........,φp (conocidos), la ecuación anterior nos permite calcular ρk para cualquier valor de k ≥p.

A la inversa, si se conocen ρ1,ρ2,........,ρp utilizando la última ecuación para k=1,2,......p se puede calcular φ1,φ2,........,φp.

Matricialmente



* + 1. **Modelos de Medias Móviles (MA)**

Modelo representado en términos de un ruido blanco correspondiente a ese tiempo y q valores pasados de ese ruido blanco.

 **Modelo MA(1)**



Las autocovarianzas están dadas por:

Las autocorrrelaciones están dadas por:

**Modelo MA(2)**

Las autocorrelaciones están dadas por:

Las autocorrelaciones están dadas por:

**Modelo MA(q)**

 ó

 Las autocovarianzas están dadas por:

 Las autocorrelaciones están dadas por:

* + 1. **Modelos Mixtos Autorregresivos-Medias Móviles**

Modelos ARMA (p,q) es una combinación de las dos anteriores.

Para que el modelo sea estacionario se requiere que las raíces de (L) = 0 caiga fuera del círculo unidad.

 **Modelo ARMA(1,1)**

Las autocovarianzas están dadas por:

* + 1. **Modelo ARIMA(p,d,q)**

A un proceso integrado Yt se le denomina proceso ARIMA(p,d,q) si tomando diferencias de orden d, se obtiene un proceso estacionario Wt del tipo ARMA(p,q).

La I de la palabra ARIMA significa integrado.

Sea Yt el modelo original (no estacionario) tomando diferencias de orden d se tiene un modelo estacionario Wt del tipo ARMA(p,q).

La representación de un proceso ARIMA(p,d,q) es:

Una clase considerablemente amplia de series se puede modelizar por medio de los procesos ARIMA gracias a que aplicando transformaciones de tipo no lineal, muchas series pasan a ser representables por medio de estos modelos.

Hay series en las que a lo largo de un periodo extenso de tiempo la varianza puede estar afectada por una tendencia la cual no desaparece al tomar diferencias.

Si se presenta esta característica la transformación adecuada puede consistir en la toma de logaritmos.

* 1. **Elaboración de Modelos**

## Fase de Identificación

En la 1era etapa de esta fase se procede a efectuar un análisis de estacionariedad. En el caso en que la serie no sea estacionario se aplican las transformaciones adecuadas con objeto de convertirla en estacionaria.

Así, si se dispone de una serie Yt, se eligirá valores para y tales que sea estacionaria la serie:

Tanto a Yt, como a Wt se les puede centrar (restar el valor de la media) si se considera necesario.

En la 2da etapa, si procede a determinar el orden de la parte autorregresiva (es decir p) y el de la parte de medias móviles (es decir q) del proceso ARMA que se considere generador de la serie estacionaria Wt.

## Fase de Estimación

##  En esta fase se obtienen valores estimados para los parámetros:

Una vez concluida esta fase, se tiene conocimiento de un proceso que hipotéticamente, ha podido generar la serie temporal transformada Wt, a partir de la cual se puede obtener la serie original Yt.

En consecuencia, la serie Wt, a partir de la expresión anterior, y conociendo los parámetros se puede expresar como:

Se puede determinar el estimador de ∈t, y si el modelo fuera adecuado, debería ser una serie que se aproxima al ruido blanco.

### Fase de Validación

Esta fase va dirigida a establecer si se produce o no esta adecuación entre datos y modelos.

### Fase de Predicción

Aquí se analizan pronósticos en términos probabilísticos de valores futuros de la variable. Esta fase es la prueba de fuego del modelo.

* 1. **Identificación de Modelos Estacionarios**

Para la identificación de los modelos estacionarios se utilizará principalmente la función de autocorrelación estimada (FACE) y la función de autocorrelación parcial estimada (FACPE).

* + 1. **Función de Autocorrelación Estimada (FACE)**

A partir de una muestra de tamaño N de valores de una serie estacionaria Wt se puede calcular un coeficiente de autocorrelación muestral de orden k.

**Observaciones:**

1. El proceso original Yt se convierte en un proceso estacionario Wt al tomar diferencias. En cada diferencia que se toma se pierde una observación al comienzo de la serie (es decir en cada aplicación del operador Δ).

Si para obtener Wt se han tomado d diferencias, entonces la muestra original de tamaño T se ha reducido a una muestra de tamaño N = T-d.

1. rk es el estimador de ρk de menor sesgo.
2. Para rN-1 se dispone de un solo sumando para el numerador, en cambio rN no se puede calcular. Al disminuir el número de sumandos en el numerador se pierde eficiencia en la estimación, por lo que no se recomienda calcular coeficientes de autocorrelación para tamaños de k superiores a 1/3 o 1/4 de la muestra.

La secuencia de valores rk para k = 1,2,3,......... constituye el correlograma estimado rk ( que es una forma gráfica de la FACE).

* + 1. **Función de Autocorrelación Parcial Estimada (FACPE)**

La función de autocorrelación parcial estimada se la puede calcular de dos formas:

1. **Mediante ecuaciones de Yule-Walker**

El inconveniente de esta forma de cálculo es la obtención de la matriz inversa que implica un gran número de operaciones.

**Procedimiento de Durbin.-** Es un método recursivo que utiliza las ecuaciones:

El inconveniente de esta forma de cálculo es que al ser recursiva acumula errores de redondeo por lo que se debe utilizar algoritmos de doble precisión:

1. **Por Regresión**

La secuencia de valores estimados de φkk constituye el correlograma estimado φkk ( que es una forma gráfica de la FACPE).

* 1. **Modelos para Series de Tiempo Estacionales**

Estacionalidad es la tendencia a repetir un modelo de conducta cada período o estación de referencia, el cual es generalmente de un año. Las series estacionales se caracterizan por revelar una fuerte correlación en el período estacional.

En general, a estos modelos se los puede expresar como:

S es el período de la estacionalidad. Por ejemplo, si el período de estacionalidad es anual, S será 12; si el período de estacionalidad es trimestral, S será 4, etc

D = número de veces que se diferencia estacionalmente

d = número de veces que se diferencia estacionariamente

φp,φP,Θq,ΘQ son polinomios de grado p,P,q,Q con raíces de módulo superior a uno; ∈t es un ruido blanco.

Un proceso que satisface esta ecuación se lo denomina SARIMA[(p,d,q)x(P,D,Q)]S

El razonamiento que conduce a obtener un modelo SARIMA, consiste en aplicar en s series, obtenidas a partir de Yt, poniendo para cada serie los meses idénticos, la misma transformación:

Y suponer que la serie obtenida :

No tiene estacionalidades y es entonces modelizable por un ARIMA(p,d,q):

De la combinación de las dos últimas ecuaciones, obtenemos la ecuación general para estos modelos estacionales, detallada al principio.

Las series estacionales pueden detectarse analizando las funciones de autocorrelación y de autocorrelación parcial estimadas, pues ellas presentan grandes valores en módulo para los índices múltiplos de s.

La identificación de los parámetros P, D y Q de los factores estacionales se hace de forma similar al procedimiento utilizado para la identificación de los parámetros de los proceso ARIMA.

* 1. **Pronósticos para Series No Estacionales**

Como expresamos anteriormente , un proceso ARIMA, lo podemos definir como:

Ahora, una observación Yt+l generado por un proceso ARIMA, donde l ≥ 1, puede ser expresado de tres formas:

1. Directamente en términos de la ecuación diferencial

1. Como una suma de ponderaciones de errores en curso y previos ∈j

Donde Ψo = 1, y los pesos Ψ pueden ser obtenidos por la ecuación de coeficientes:

1. Como una suma de ponderaciones de observaciones previas, más un error aleatorio

* 1. **Pronósticos para Series Estacionales**

Una función de pronósticos para Series Estacionales, puede ser representada mediante una función de senos y cosenos:

Donde las b, son coeficientes adaptativos, y donde [s/2]= ½ S ,si s es par y [s/2]= ½(s-1) si s es impar.

* 1. **Elección del Mejor Modelo**

Seleccionar un modelo que capture de una gran forma la estructura de los datos de una serie, es sin lugar a dudas nuestro principal objetivo, por lo cual para llevar a cabo la elección de un buen modelo en el Análisis de Series de Tiempo que se detalla en el Capítulo 3, nos basaremos en dos criterios:

1. **Prueba de Student de significancia de parámetros.**

Esta prueba nos da a conocer si los parámetros del modelo son significativamente diferentes de cero, para lo cual se realiza el siguiente contraste de hipótesis:

 **Ho : θi = 0**

 **vs.**

 **H1: θi ≠ 0**

Si el valor p es menor que 0.05, se tiene evidencia estadística para afirmar que el parámetro es significativamente diferente de cero, con un nivel de confianza del 95%.

1. **Media Cuadrática del Error**

La Media Cuadratica del Error del Modelo sin lugar a dudas va a ser un factor preponderante para decidir cual modelo captura de una mejor manera la estructura de los datos.

Un buen modelo tiene que lograr reducirla, para así poder ser un gran medio de predicción.

1. *Véase Time Series Analysis, George P. Box, Gwilyn Jenkins, and Gregory Reinsel* [↑](#footnote-ref-2)