

9
24



**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

FACULTAD DE CIENCIAS

**INTRODUCCION A LA DESESTACIONALIZACION DE
SERIES DE TIEMPO APLICADO A MODELOS
ECONOMICOS.**

T E S I S

QUE PARA OBTENER EL TITULO DE:
A C T U A R I O
P R E S E N T A :
RICARDO SAUL AVENDAÑO GUTIERREZ



MEXICO, D. F.

1996

**TESIS CON
FALLA DE ORIGEN**



**TESIS CON
FALLA DE ORIGEN**



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL
AVENIDA DE
MEXICO

M. en C. Virginia Abrín Batule
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo de Tesis: INTRODUCCION A LA
DESESTACIONALIZACION DE SERIES DE TIEMPO APLICADO A MODELOS ECONOMICOS.

realizado por RICARDO SAUL AVENDAÑO GUTIERREZ

con número de cuenta 8233809-1 , pasante de la carrera de ACTUARIA

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis	M. en C. JUAN GONZALEZ HERNANDEZ.
Propietario	
Propietario	DR. VICTOR MANUEL GUERRERO GUZMAN.
Propietario	DR. ANTONIO ARPAZ RAMONET.
Suplente	DRA. MARIA EDITH PACHECO GOMEZ
Suplente	M. en C. JOSE LOPEZ ESTRADA.

Consejo Departamental de Matemáticas

M. en C. ALEJANDRO BRAVO HOJICA.

CON TODO MI CARIÑO, DEDICO ESTE ESFUERZO A LAS DOS
PRINCIPALES MOTIVACIONES DE MI VIDA:

SRA. MA. DOLORES GUTIERREZ G.

Y

SR. JESUS AVENDAÑO SANCHEZ S.

GRACIAS, PADRES.

AGRADEZCO ESPECIALMENTE A MI AMIGO Y DIRECTOR
DE TESIS

M. EN C. JUAN GONZALEZ HERNANDEZ

POR LA PACIENCIA Y ATENCION PRESTADA.

GRACIAS POR SU AYUDA A: MIS HERMANOS, MI NOVIA,
MIS AMIGOS Y MIS SINODALES.

INDICE

INTRODUCCION	1
CAPITULO 1. INTRODUCCION A LA PROBABILIDAD, ESTADISTICA Y SERIES DE TIEMPO.	3
1-1 PROBABILIDAD.	4
1-1.1 ESPACIOS DE PROBABILIDAD.	4
1-1.2 PROBABILIDAD CONDICIONAL E INDEPENDENCIA.	5
1-1.3 VARIABLES ALEATORIAS.	6
1-1.3.1 VARIABLE ALEATORIA DISCRETA.	6
1-1.3.2 VARIABLES ALEATORIAS CONTINUAS.	7
1-1.3.3 DISTRIBUCION NORMAL.	7
1-1.4 ESPERANZA.	8
1-1.4.1 MOMENTOS DE ORDEN r	9
1-1.4.2 VARIANZA.	9
1-1.5 TEOREMA CENTRAL DEL LIMITE.	10
1-2 ESTADISTICA.	11
1-2.1 MEDIDAS DESCRIPTIVAS.	11
1-2.1.1 MEDIDAS DE TENDENCIA CENTRAL.	11
1-2.1.2 MEDIDAS DE DISPERSION.	12
1-2.2 ESTIMACION PUNTUAL.	12
1-2.2.1 PROPIEDADES DE LOS ESTIMADORES.	13
1-2.2.2 METODOS PARA OBTENER ESTIMADORES.	14
1-2.3 ESTIMACION POR INTERVALOS.	15
1-2.3.1 METODO PIVOTAL.	15
1-3 SERIES DE TIEMPO.	17
3.1 COMPONENTES DE UNA SERIE DE TIEMPO.	17
1-3.2 TIPOS DE ESTIMACIONES.	20
1-3.3 METODOS DE ESTIMACION.	21
1-3.3.1 METODOS CUALITATIVOS.	21
1-3.3.2 METODOS CUANTITATIVOS.	21
CAPITULO 2. SERIES DE TIEMPO CON TENDENCIA E IRREGULARIDAD.	24
2-1 ANALISIS DE REGRESION LINEAL MULTIPLE.	26
2-1.1 ENFOQUE MATRICIAL DEL MODELO.	26
2-1.2 SUPUESTOS DEL MODELO DE REGRESION LINEAL.	27
2-1.3 ESTIMACION POR MINIMOS CUADRADOS.	27
2-1.4 MATRIZ DE VARIANZA COVARIANZA.	29
2-1.5 COEFICIENTE DE DETERMINACION r^2	30
2-1.6 INTERVALOS DE CONFIANZA.	32
2-1.7 PRUEBAS DE HIPOTESIS.	32
2-1.8 EJEMPLO.	35

2-2 SUAVIZAMIENTO EXPONENCIAL SIMPLE.....	42
2-2.1 APROXIMACION POR SUAVIZAMIENTO EXPONENCIAL.....	42
2-2.2 DETERMINACION DE LA CONSTANTE DE SUAVIZAMIENTO.....	44
2-2.3 INSTRUMENTACION DEL PROCESO.....	44
2-2.4 ANALISIS DE LOS ERRORES DE ESTIMACION.....	45
2-2.4.1 PROCESOS DE CONTROL ADAPTABLE.....	46
2-2.5 EJEMPLO.....	48
2-3 SUAVIZAMIENTO EXPONENCIAL DOBLE.....	52
2-3.1 APROXIMACION POR SUAVIZAMIENTO EXPONENCIAL.....	52
2-3.2 INTERVALOS DE CONFIANZA.....	56
2-3.3 EJEMPLO.....	58
CAPITULO 3. SERIES DE TIEMPO ESTACIONALES.....	65
3-1 MEDIAS MOVILES.....	67
3-1.1 MEDIAS MOVILES CENTRADAS.....	67
3-1.2 MEDIA MOVIL $M/A(3 \times 3)$	67
3-1.3 MEDIAS MOVILES DE SPENCER Y HENDERSON.....	68
3-2 METODO DE DESCOMPOSICION MULTIPLICATIVO.....	71
3-2.1 ESTIMACION DE LAS COMPONENTES.....	71
3-2.2 PRONOSTICOS E INTERVALOS DE CONFIANZA.....	74
3-2.3 EJEMPLO.....	75
3-3 METODO DE DESCOMPOSICION CENSO II.....	80
3-3.1 AJUSTE POR DIAS HABLES.....	80
3-3.2 AJUSTE ESTACIONAL INICIAL.....	81
3-3.3 AJUSTE ESTACIONAL FINAL.....	83
3-3.4 PRUEBAS ESTADISTICAS.....	84
3-3.5 RESUMEN.....	85
3-3.6 EJEMPLO.....	87
3-4 METODO DE DESCOMPOSICION WINTERS.....	92
3-4.1 ACTUALIZACION DE LOS ESTIMADORES.....	92
3-4.2 LOS ESTIMADORES INICIALES.....	93
3-4.3 CUANDO NO EXISTE TENDENCIA.....	96
3-4.4 INTERVALOS DE CONFIANZA.....	96
3-4.5 EJEMPLO.....	98
CAPITULO 4. METODOLOGIA BOX-JENKINS.....	106
4-1 PASOS BASICOS DE LA METODOLOGIA BOX-JENKINS.....	108
4-1.1 IDENTIFICACION DEL MODELO.....	108
4-1.1.1 SERIES DE TIEMPO ESTACIONARIAS Y NO ESTACIONARIAS.....	108
4-1.1.2 AUTOCORRELACION Y AUTOCORRELACION PARCIAL.....	110
4-1.1.3 IDENTIFICACION DE UN MODELO PARTICULAR DE ESTACIONARIDAD.....	113
4-1.2 ESTIMACION DE LOS PARAMETROS DEL MODELO.....	115

4-1.3 ADECUACION DEL MODELO	116
4-1.4 OBTENCION DE PRONOSTICOS	117
4-1.5 EJEMPLO.	119
4-2 MODELOS BOX-JENKINS MAS COMUNES.	133
4-2.1 MODELOS DE PROMEDIOS MOVILES.	133
4-2.2 MODELOS AUTORREGRESIVOS.	135
4-2.3 MODELOS AUTORREGRESIVOS Y DE PROMEDIOS MOVILES.	138
4-2.4 RESUMEN DE MODELOS BOX-JENKINS NO ESTACIONALES.	139
4-2.5 MODELOS ESTACIONALES.	142
4-2.5.1 DIFERENTES TIPOS DE ESTACIONARIDAD.	149
CONCLUSIONES.	154
BIBLIOGRAFIA	157

INTRODUCCION.

El presente trabajo es una introducción a la descomposición de series de tiempo económicas, fundado básicamente en la idea de que las series de tiempo están constituidas por varias componentes. En el análisis de series de tiempo se encuentran básicamente dos categorías:

- a) El análisis en el dominio del tiempo
- b) El análisis espectral o en el dominio de las frecuencias

El análisis espectral requiere un grado avanzado en el conocimiento de las matemáticas y por lo tanto su uso no es muy generalizado como el primero. Este trabajo se basa en el análisis en el dominio del tiempo.

Es importante la descomposición de series de tiempo ya que cada una de las componentes de una serie es causada por distintos fenómenos. En lo que respecta a la componente estacional existen varias causas: el hecho que los meses tengan más o menos días en el calendario o bien que ciertas festividades estén fijadas en determinadas fechas del calendario; otra causa es que las dependencias gubernamentales tienden a fijar fechas dentro del año para realizar ciertas actividades; una causa más es el clima, que de acuerdo a cada estación del año, fija ciertos patrones a seguir (de aquí el término estacionalidad).

Esta tesis está organizada en cuatro capítulos. El capítulo 1, es simplemente una introducción a conceptos básicos de estadística, probabilidad y series de tiempo. El capítulo 2 presenta los métodos de regresión lineal múltiple y el método de suavizamiento exponencial para pronosticar series de tiempo descritas por las componentes de tendencia e irregularidad. Con estos métodos se hacen estimaciones del índice de empleo en el sector manufacturero. En el capítulo 3 se describen los métodos de descomposición multiplicativo, Censo II y Winters para pronosticar series de tiempo descritas por las componentes de tendencia, estacionalidad e irregularidad. En esta parte se hacen ejercicios con dos series de tiempo que por naturaleza son estacionales: el índice de producción

pesquera y el índice de ventas de electricidad. El capítulo 4 es un análisis de la metodología desarrollada por Box y Jenkins: identificación del modelo, estimación de los parámetros del modelo, adecuación del modelo y obtención de pronósticos. Además de la presentación de los modelos más comunes de esta metodología.

Este trabajo consiste en la presentación matemática de las series de tiempo, así que los ejemplos que se dan de series económicas deben verse más como la explicación de la técnica, que la búsqueda de un sentido estrictamente económico.

CAPITULO I.

INTRODUCCION A LA PROBABILIDAD, ESTADISTICA Y SERIES DE TIEMPO.

La probabilidad y la estadística tienen que ver de alguna manera en diversas facetas de la vida. Es fácil enterarse por medio de los noticieros, del pronóstico del tiempo o encontrar artículos en el periódico con aplicaciones estadísticas, como por ejemplo, las predicciones que se hacen del índice nacional de precios, resultados electorales, índices delictivos, estudios sobre efectos de la contaminación atmosférica en las personas, etc. Cada artículo reporta alguna información, pronóstico o conclusiones sobre la organización y análisis de datos numéricos.

Por otro lado, a las predicciones de eventos futuros se les llama pronósticos y al proceso de llegar a estas se le conoce como pronosticar. Existen diversas formas de pronosticar los eventos futuros de cualquier variable, una técnica estadística importante, es el análisis de series de tiempo. Aunque en los pronósticos el interés principal radica en el futuro, el análisis de series de tiempo comienza examinando el pasado, por lo tanto, no se puede pasar por alto las regularidades y patrones apreciables en las series históricas. Si los pronósticos se basan en estas regularidades y patrones, naturalmente se está expresando la confianza en que el futuro se deriva del pasado con cierto grado de consistencia y que lo que ha sucedido en el pasado seguirá sucediendo, en mayor o menor grado, en el futuro.

En el presente capítulo se mencionará, a manera de recordatorio, algunos conceptos de probabilidad y estadística, que sirven como bases en el desarrollo de este trabajo. Así como las nociones preliminares acerca de las series de tiempo.

1-1 PROBABILIDAD.

En algunos libros de texto, se dice que la probabilidad nace y alcanza un notable desarrollo con la observación del juego de dados, cuando el Caballero de Mere (Un apasionado jugador de dados), planteó un problema de división de apuestas a Pascal y este mando su solución a Fermat. Esto no es estrictamente cierto, ya que el Triángulo de Pascal que resuelve el problema ya era conocido desde tiempo atrás. Los Hindúes 200 años antes de cristo ya tenían conocimiento de él. Bhaskara en 1178 en su tratado Lilavati (Laiencia noble) descubre métodos para calcular permutaciones y combinaciones. Narayana en el siglo XIV, Chu Shih Chien en el precioso Espejo de los Cuatro Elementos en 1303, Michael Stiffel (1486-1567) en su libro Aritmética Integra en 1544 son algunas de las referencias sobre el triángulo de Pascal.

En el siglo XIX Laplace publica su "Teoría analítica de la probabilidad". En el siglo XX el cálculo de probabilidades se enriquece con las ideas fundamentales del análisis, de la teoría de conjuntos y su medida y el análisis funcional. En este siglo destacan en la investigación de la teoría y aplicaciones, fundamento y lógica de la probabilidad: Kolmogorov, Chebychev, Markov, Pearson, Fisher, Neymann, R. Von Neumann.

El concepto de probabilidad no estaba rigurosamente definido. Se debe a Borel su definición rigurosa basada en la medida de conjuntos (la teoría de la medida fue desarrollada en buena parte por Borel y Lebesgue). De esta definición se llega al establecimiento de su axiomática: Kolmogorov (Berlín, 1933).

1-1.1 ESPACIOS DE PROBABILIDAD.

La teoría de probabilidades estudia los métodos de análisis que son comunes en el tratamiento de fenómenos aleatorios, cualquiera que sea el área en que éstos se presenten. La probabilidad es, pues, la ciencia de los fenómenos aleatorios, en el sentido de que estudia las propiedades de estos fenómenos que dependen esencialmente del concepto de aleatoriedad y no de otros aspectos particulares.

Fenómeno Aleatorio. Un fenómeno aleatorio es un fenómeno empírico que se caracteriza por la propiedad de que, al observarlo bajo determinado conjunto de condiciones, no siempre se obtiene el mismo resultado (de manera que no existe regularidad determinista) sino que los diferentes resultados ocurren con regularidad estadística. Esto quiere decir que existen números entre 0 y 1 que representan el límite de las frecuencias relativas en algún sentido con la que se observan los diferentes resultados en una serie de repeticiones independientes del fenómeno.

Espacio Muestral. Al conjunto formado por todos los diferentes resultados de un fenómeno aleatorio se le llama espacio muestral del fenómeno. Generalmente se denota por Ω .

Evento. Un evento es simplemente un subconjunto del espacio muestral Ω .

Un evento ocurre si y sólo si el resultado del fenómeno aleatorio es un elemento del evento. La unión de dos eventos es un evento y el complemento de un evento es un evento.

Definición: Un σ -campo \mathcal{A} , sobre un conjunto no vacío Ω , es una familia no vacía de subconjuntos de Ω tal que:

- i) Si $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
- ii) Si $A_n \in \mathcal{A}, n = 1, 2, \dots \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

Desde el punto de vista matemático, un evento es un elemento de la familia \mathcal{A} . El conjunto vacío \emptyset , se llama el evento imposible. El espacio muestral Ω se llama el evento seguro.

Definición: El σ -campo \mathcal{B} generado por la familia $\mathcal{A} = \{(-\infty, x], x \in \mathbb{R}\}$ de intervalos infinitos de la forma $(-\infty, x] = \{r \in \mathbb{R} : x \leq r\}$, se llama σ -campo de Borel de la recta \mathbb{R} .

Medida de Probabilidad. Una medida de probabilidad (o simplemente una probabilidad), sobre un σ -campo \mathcal{A} , de subconjuntos de un conjunto no vacío Ω es una función real $P: \mathcal{A} \rightarrow \mathbb{R}$ definida sobre \mathcal{A} tal que:

- i) $P(\Omega) = 1$
- ii) $P(A) \geq 0 \quad \forall A \in \mathcal{A}$
- iii) Si A_n con $n = 1, 2, \dots$ son ajenos (esto es $A_i \cap A_j = \emptyset$ si $i \neq j$)
 $\Rightarrow P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{i=1}^{\infty} P(A_n)$.

Espacio de Probabilidad. Un espacio de probabilidad denotado por (Ω, \mathcal{A}, P) , es un conjunto $\Omega \neq \emptyset$, un σ -campo \mathcal{A} y una medida de probabilidad P , definida en \mathcal{A} .

1-1.2 PROBABILIDAD CONDICIONAL E INDEPENDENCIA.

Probabilidad Condicional. Sean A y B dos eventos tal que $P(A) > 0$. Entonces la probabilidad condicional de B dado A , se escribe $P(B|A)$, se define como:

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Si $P(A) = 0$, $P(B|A)$ no está definida.

Independencia. Sean A y B dos eventos definidos en el espacio de probabilidad (Ω, \mathcal{A}, P) . Se dice que los eventos A y B son independientes si:

$$P(A \cap B) = P(A)P(B)$$

Dos eventos que no son independientes, se llaman dependientes.

1-1.3 VARIABLES ALEATORIAS.

Desde el punto de vista intuitivo, una variable aleatoria describe una característica observable de una población.

Variable Aleatoria. Una variable aleatoria, sobre un espacio de probabilidad (Ω, \mathcal{A}, P) , es una función real $X: (\Omega, \mathcal{A}, P) \rightarrow \mathbb{R}$, definida sobre (Ω, \mathcal{A}, P) , tal que para todo número real x , el conjunto $\{X \leq x\} = \{\omega \in \Omega \mid X(\omega) \leq x\}$ es un evento.

Distribución de Probabilidad. La distribución de probabilidad de una variable aleatoria $X: (\Omega, \mathcal{A}, P) \rightarrow \mathbb{R}$, es la función real $P_X: \mathcal{B} \rightarrow \mathbb{R}$ definida sobre el σ -campo de Borel \mathcal{B} , tal que

$$P_X(B) = P(X \in B) = P(\{\omega \in \Omega \mid X(\omega) \in B\})$$

1-1.3.1 VARIABLE ALEATORIA DISCRETA.

Una variable aleatoria X , definida sobre un fenómeno aleatorio (Ω, \mathcal{A}, P) , es una variable aleatoria discreta (v.a.d.), si su conjunto de valores $X(\Omega) = \{X(\omega) : \omega \in \Omega\}$, es finito o numerable. Es decir, si el conjunto $X(\Omega)$ puede escribirse como una sucesión de números reales:

$$X(\Omega) = \{x_1, x_2, \dots\}$$

Función de Masa de Probabilidad. La función de variable real $f: \mathbb{R} \rightarrow \mathbb{R}$, definida por:

$$f(x) = P(X = x)$$

se llama función de masa de probabilidad, o simplemente función de masa de la variable aleatoria X .

En general, una función real de variable real, $f: \mathbb{R} \rightarrow \mathbb{R}$ es una función de masa si:

- i) $f(x) \geq 0$
- ii) $\{x \in \mathbb{R} : f(x) \neq 0\}$
- iii) Sea $\{x \in \mathbb{R} : f(x) \neq 0\} = \{x_1, x_2, x_3, \dots\}$. Entonces $\sum_i f(x_i) = 1$

$$\sum_i f(x_i) = 1$$

Función de Densidad. A la función $f: \mathbb{R} \rightarrow \mathbb{R}$ que cumple con la condición:

$$P_X(B) = P(X \in B) = \int_B f(x) dx$$

se le llama función de densidad de la variable aleatoria X y tiene las siguientes propiedades:

$$f(x) \geq 0, x \in \mathbb{R}$$

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

1-1.3.2 VARIABLES ALEATORIAS CONTINUAS.

Variable Aleatoria Continua. Una variable aleatoria X es de tipo continuo si la distribución de probabilidad P_X es absolutamente continua y está especificada por una función de densidad.

Función de Distribución. La función de distribución de una variable aleatoria $X: (\Omega, \mathcal{A}, P) \rightarrow \mathbb{R}$, es la función $F: \mathbb{R} \rightarrow \mathbb{R}$, real de variable real, definida por:

$$F(x) = P(X \leq x)$$

Frecuentemente se denota por F_x la función de distribución de X .

1-1.3.3 DISTRIBUCION NORMAL.

La distribución normal juega un papel muy importante tanto en la teoría, como en las aplicaciones. Su gran importancia radica en el hecho de que muchas distribuciones tienen aproximadamente una distribución normal.

Se define la función de densidad de una variable aleatoria X normal con parámetros (μ, σ^2) , como la función:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$$

En consecuencia la función de densidad de una variable aleatoria que tiene una distribución normal, con parámetros $\mu = 0, \sigma^2 = 1$ es:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, -\infty < x < \infty$$

La función $\phi: \mathbb{R} \rightarrow \mathbb{R}$ definida arriba, se llama función de densidad normal estándar. La gráfica de la función de densidad normal, es una curva en forma de campana. La función de distribución correspondiente a ϕ , se denota por $\Phi(x)$.

$$\phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Su gráfica es una curva en forma de S. Una propiedad muy útil de la función de distribución normal estándar es la siguiente:

$$\Phi(-x) = 1 - \Phi(x), \quad x \in \mathbb{R}$$

Se puede escribir la función de distribución $F(x)$ de una variable aleatoria $X \sim N(\mu, \sigma^2)$ normal con parámetros (μ, σ^2) en términos de la función de distribución normal estándar ϕ .

$$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

$$P(a \leq X \leq b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

1-1.4 ESPERANZA.

Para conocer el comportamiento de una variable aleatoria X es conveniente conocer su valor promedio.

En general si X es una variable aleatoria simple, que toma los valores x_1, x_2, \dots, x_n la esperanza de X es:

$$E(X) = \sum_{i=1}^n x_i f(x_i)$$

La esperanza de X , también recibe los nombres de valor medio, la media, el valor esperado de X .

ESPERANZA (CASO DISCRETO). Sea X una variable aleatoria discreta cuyos valores posibles x_1, x_2, \dots . Se dice que X tiene esperanza finita si:

$$\sum |x_j| f(x_j) < \infty$$

En este caso, la esperanza de X , esta definida por:

$$E(X) = \sum_{j=1}^{\infty} x_j f(x_j)$$

Si $\sum_{j=1}^{\infty} |x_j| f(x_j) = \infty \Rightarrow X$ no tiene esperanza finita y $E(X)$ no está definida.

ESPERANZA (CASO CONTINUO). Sea X una variable aleatoria cuya distribución está especificada por una función de densidad $f: \mathbb{R} \rightarrow \mathbb{R}$. Entonces, X tiene esperanza finita si:

$$\int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

En este caso la esperanza de X , es:

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx < \infty$$

1-1.4.1 MOMENTOS DE ORDEN r .

El r -ésimo momento ($r = 1, 2, 3, \dots$) de una variable aleatoria X , es la esperanza de su r -ésima potencia:

en el caso continuo

$$E(X^r) = \int_{-\infty}^{\infty} x^r f_X(x) dx$$

en el caso discreto

$$E(X^r) = \sum_{x} x^r f_X(x)$$

1-1.4.2 VARIANZA.

Sea X variable aleatoria con segundo momento finito. Entonces la varianza de X , denotada por $V(X)$ está definida por:

$$V(X) = E[(X - E(X))^2]$$

A menudo se denota la $E(X)$ por μ y $V(X)$ por σ^2 . Y a la desviación estándar por $\sigma = \sqrt{V(X)}$.

A la cantidad $E[(X - E(X))(Y - E(Y))]$ se le llama covarianza de X y Y , y se escribe $COV(X, Y)$.

1-1.5 TEOREMA CENTRAL DEL LIMITE.

En las aplicaciones de la teoría de probabilidades a fenómenos reales, hay un resultado que es muy destacado, se conoce este resultados como el teorema central del límite.

Sean X_1, X_2, \dots variables aleatorias independientes e idénticamente distribuidas con media μ y varianza finita σ^2 y $S_n = X_1 + X_2 + \dots + X_n$. Entonces:

$$\lim P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \Phi(x)$$

Donde $\Phi(x)$ se distribuye $N(0, 1)$

Este teorema es de fundamental importancia, ya que justifica el uso de métodos de curva normal en una amplia gama de problemas; se aplica a poblaciones infinitas y también a poblaciones donde n , pese a ser grande, constituye en cambio una pequeña parte de la población. Es difícil decir con exactitud cuán grande debe ser n de manera que se aplique el teorema central del límite, pero $n = 30$ suele considerarse suficientemente grande [7].

1-2 ESTADISTICA.

La información sobre un tópico se adquiere frecuentemente en forma de datos numéricos. un análisis de estos datos se hace para tener una mejor comprensión del fenómeno bajo estudio y permitir extraer conclusiones. De estas últimas se hacen generalizaciones cuya validez se checa por medio de alguna investigación posterior.

La estadística es la metodología que ha sido desarrollada para interpretar y extraer conclusiones de los datos.

Algunas veces es posible obtener valores de todos los individuos de la población y en estos casos se dice que se censó a la población. Desgraciadamente no siempre es posible hacer un censo ya que puede ser que no se tenga el dinero o el tiempo para hacerlo o porque el individuo se destruya al observarlo, en estos casos es necesario considerar sólo a un subconjunto de la población y en base a este obtener información sobre toda la población, es claro que para que esta información sea válida dicho subconjunto debe representar adecuadamente a la población; a este subconjunto se le conoce con el nombre de muestra.

Cuando se observa sólo una muestra, el extraer conclusiones sobre la población a partir de ella involucra cierta incertidumbre la cual es inherente al proceso de generalizar a partir de unos cuantos a muchos. La metodología desarrollada para pasar de observaciones de la muestra a afirmaciones acerca de la población es llamada inferencia estadística.

1-2.1 MEDIDAS DESCRIPTIVAS.

A las características numéricas de la población se les conoce como parámetros. Las dos clases más importantes de parámetros, aunque no las únicas son aquellas que localizan el centro y describen la dispersión de la población, tales medidas son llamadas de tendencia central y de variación, respectivamente

1-2.1.1 MEDIDAS DE TENDENCIA CENTRAL.

Las medidas descriptivas numéricas que localizan el centro de la distribución de mediciones, son llamadas medidas de tendencia central. A continuación se citan las más conocidas y utilizadas

Moda. Es el valor que ocurre con mayor frecuencia dentro de un conjunto de mediciones.

Mediana. Es el valor que divide a la mitad el conjunto de observaciones, de tal manera que las observaciones en una mitad son menores o iguales que el valor de la mediana y en la otra mitad son mayores o iguales.

Otras medidas en la familia a la que pertenece la mediana son los cuartiles, deciles, etc., que son casos especiales de los valores conocidos como cuantiles de orden p ($0 < p < 1$). Un cuantil es un valor Q_p tal que satisface las siguientes dos condiciones:

- a) La proporción de observaciones que es menor a Q_p es menor o igual que p .
- b) La proporción de observaciones que excede Q_p es menor o igual que $(1-p)$.

La media. La media de un conjunto de datos es el promedio aritmético de ellos.

Las tres medias que se han presentado son las de mayor uso en la práctica. Para decidir cual de ellas usar en un problema particular, se debe analizar la situación específica para así efectuar una buena elección. Una recomendación general, que está relacionada con la forma de la distribución de las observaciones es que si ésta es asimétrica, la mediana es una mejor medida de localización que la media en ese caso.

1-2.1.2 MEDIDAS DE DISPERSION.

Son medidas que indican el grado de variabilidad de los datos. Existen varias medidas de dispersión de las cuales se mencionan las siguientes.

Rango. Sea una muestra de tamaño n , $x_{(1)}$ el valor mínimo y $x_{(n)}$ el valor máximo. El rango se calcula como:

$$r = x_{(n)} - x_{(1)}$$

Por construcción el rango es número positivo que indica cual es la distancia máxima que se puede observar entre dos valores muestrales cualesquiera.

Desviación estándar. En este caso nuevamente se utiliza un parámetro de centralidad \bar{x} y se calcula como:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

Esta estadística obtiene un promedio de las desviaciones que tienen las observaciones con respecto a la media muestral, pero al elevarlas al cuadrado, "amplifica" las diferencias mayores que uno en valor absoluto mientras que disminuye la contribución de aquellas que son menores que uno en valor absoluto. Al obtener la raíz se tiene una medida en las mismas unidades que el parámetro de centralidad.

1-2.2 ESTIMACION PUNTUAL.

El problema de estimación puntual es determinar, en base a los resultados de un experimento, el posible valor de dichos parámetros. En general dicho problema se puede formular de la siguiente manera.

Considere una variable aleatoria X , que puede tomar valores discretos o continuos, cuya distribución está dada por una función de densidad, $f(x; \theta)$ que depende de un parámetro desconocido θ . Supóngase que se toma una muestra de dicha variable aleatoria x_1, x_2, \dots, x_n .

El problema básico es determinar una función $\hat{\theta} = \hat{\theta}(x_1, \dots, x_n)$ que estime de la "mejor manera posible" el valor del parámetro θ de $f(x; \theta)$. Dicha función $\hat{\theta} = \hat{\theta}(x_1, \dots, x_n)$ se conoce como el estimador del parámetro θ de $f(x; \theta)$ y al valor específico que tome para ciertos valores observados $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ se le llamará estimación o estimada.

Una estadística es una función de la muestra que no depende de parámetros desconocidos. Un estimador es una estadística.

1-2.2.1 PROPIEDADES DE LOS ESTIMADORES.

Insesgamiento. Sea $\hat{\theta}$ un estimador para θ en $f(x; \theta)$, se dice que $\hat{\theta}$ es insesgado si:

$$E[\hat{\theta} | \theta] = \theta \quad \forall \theta$$

Eficiencia. Este criterio proporciona una medida de la "bondad" de un estimador insesgado en cuanto a su varianza. Para definir (y utilizar) este criterio es necesario hacer una suposición bastante fuerte, esta es que se conoce el estimador insesgado óptimo, o sea el de varianza mínima, δ^* .

Entonces se efectúa la comparación entre un estimador insesgado cualquiera $\hat{\theta}$ en términos de la varianza definiendo:

$$eficiencia(\hat{\theta}) = \frac{Var(\delta^*)}{Var(\hat{\theta})}$$

que es la eficiencia de $\hat{\theta}$ respecto a δ^* .

Si un estimador tiene una *eficiencia* $(\hat{\theta}) = 0.8$ $\hat{\theta}$ es 80% eficiente.

Consistencia. Sea $\hat{\theta}$ un estimador para θ en $f(x; \theta)$, se dice que $\hat{\theta}$ es consistente si:

$$\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta \quad \text{y} \quad \lim_{n \rightarrow \infty} Var(\hat{\theta}) = 0$$

Suficiencia. Un estimador es suficiente si contiene la información necesaria para estimar el parámetro de interés. Sea X una variable aleatoria cuya función de densidad está dada por $f(x; \theta)$. Considerese una muestra x_1, x_2, \dots, x_n y una estadística $t = t(x_1, \dots, x_n)$.

Se dice que la estadística t es suficiente para el parámetro θ si la distribución condicional de la muestra x_1, \dots, x_n dada la estadística t , no depende del valor de θ , esto es:

$$P(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3, \dots, X_n \leq x_n | \theta)$$

no depende de θ .

1-2.2.2 METODOS PARA OBTENER ESTIMADORES.

Hasta el momento, se han revisado las principales propiedades de los estimadores y ahora el problema que se plantea es el de proponer estimadores que puedan cumplir con algunas, sino es que con todas, estas propiedades. Así es que ahora es de interés proponer ciertos métodos para encontrar estimadores.

Método de momentos. Este es el método más antiguo para generar estimadores de un parámetro desconocido, es muy fácil de usar en la mayoría de los casos y genera estimadores bastante razonables. Sin embargo, no tiene todas las propiedades de los estimadores máxima verosimilitud como se verá más adelante.

Dada una variable aleatoria X , con función de densidad $f(x; \theta)$, si se desea estimar los parámetros de la distribución de X , se debe expresar cada uno de tales parámetros como función de los momentos poblacionales, y la estimación se realiza al sustituir en esa función los momentos poblacionales por los valores de los momentos muestrales. Esto es, dado un parámetro θ y sí:

M_i es el i -ésimo momento poblacional

$$M_i = E(X^i)$$

m_i es el i -ésimo momento muestral

$$m_i = E(X^i)$$

se tiene que expresar $\theta = H(M_1, \dots, M_k)$ y estimar θ mediante $\hat{\theta} = H(m_1, \dots, m_k)$.

A este método se le conoce también como de analogía.

Método de máxima verosimilitud. Este método es de los más antiguos y aunque fue empleado por Gauss en algunos problemas aislados, su descripción formal como método de estimación ocurrió hasta este siglo y se debe a R. A. Fisher. El método es interesante desde el punto de vista intuitivo pues elige como estimador de un parámetro θ , aquel valor que maximiza la probabilidad de obtener una muestra como la que se ha obtenido y en la cual se basa la estimación.

Sea X una variable aleatoria cuya densidad esta dada por $f(x; \theta)$, considere una muestra aleatoria X_1, \dots, X_n y evalúe la función de densidad conjunta de la muestra en los valores específicos de las observaciones muestrales con la notación:

$$f(X; \theta) = \prod_{i=1}^n f(X_i; \theta) = L(\theta)$$

$L(\theta)$ se le llama función de verosimilitud. Para los valores fijos de las X_i , la función de verosimilitud es, exclusivamente una función del parámetro θ . Entonces el método máxima verosimilitud (para abreviar M.V.) consiste en estimar θ por un valor $\hat{\theta}(X)$ tal que:

$$L(\hat{\theta}) = \sup L(\theta)$$

1-2.3 ESTIMACION POR INTERVALOS.

Los métodos de intervalos de confianza están pensados para dar una idea del valor numérico real que el parámetro puede tener y también una medida de cuanto confianza se puede tener en el estimador como una aproximación de tal valor. Estos métodos se basan en la construcción de intervalos aleatorios. En general un intervalo donde por lo menos uno de sus límites es una variable aleatoria se denomina intervalo aleatorio.

La idea de estimar por intervalos es la de construir, en base a una muestra aleatoria, un intervalo que permita estimar el verdadero valor del parámetro desconocido y además medir, en términos de probabilidad, la exactitud de la estimación.

Sea X_1, X_2, \dots, X_n una muestra aleatoria con función de densidad $f(x; \theta)$. Si el intervalo $I(X) = [\underline{\theta}(X), \bar{\theta}(X)]$ en donde θ es un parámetro real del valor desconocido y $\underline{\theta}(X), \bar{\theta}(X)$ son dos funciones reales de X es tal que la probabilidad de que $I(X)$ cubra el verdadero valor de θ es igual a un valor prefijado γ , entonces se dice que $I(X)$ es un intervalo de confianza para θ con un coeficiente de confianza γ .

1-2.3.1 METODO PIVOTAL.

Suponga que X_1, \dots, X_n es una muestra aleatoria de la distribución $F(X; \theta)$. Suponga además que es posible encontrar una función $g(X_1, \dots, X_n | \theta)$ tal que:

- i) $g(X_1, \dots, X_n | \theta)$ esté definida $\forall \theta$ en un intervalo (θ_1, θ_2) que contiene a θ_0 el verdadero valor del parámetro y para todo punto (X_1, \dots, X_n) en el espacio muestral
- ii) $g(X_1, \dots, X_n | \theta)$ es continua y monótona en θ (creciente o decreciente) y
- iii) la función de distribución de $g(X_1, \dots, X_n | \theta)$ no depende de θ ni de ningún otro parámetro desconocido.

Conociendo la distribución de g será posible encontrar un intervalo (g_1, g_2) para el cual $P(g_1 < g < g_2) = \alpha$. Entonces si θ_0 es el verdadero valor de θ , las soluciones $\underline{\theta}$ y $\bar{\theta}$ (con $\underline{\theta} < \bar{\theta}$) respectivamente de las ecuaciones $g_1 = g(X_1, \dots, X_n | \theta)$ y $g_2 = g(X_1, \dots, X_n | \theta)$ existen y el intervalo de confianza para θ es $I(X) = [\underline{\theta}(X), \bar{\theta}(X)]$ con un coeficiente de γ

Debe notarse que g_1, g_2 se pueden escoger de un número infinito de formas, en general se preferiría que fueran valores simétricos, cuando se pueda, y que el valor esperado de la longitud del intervalo sea lo más pequeño posible.

La idea de intervalos de confianza se puede generalizar al problema de estimación de un parámetro vectorial, en cuyo caso se habla de regiones de confianza, las cuales pueden tener formas geométricas diversas. Y la "optimalidad" de los mismos deberá ser determinada según criterios adecuados a cada caso [6].

1-3 SERIES DE TIEMPO.

Para iniciar el tema de las series de tiempo es necesario tener una idea de lo que esto significa. Se define inicialmente serie de tiempo como un conjunto de datos numéricos, observados a intervalos iguales de tiempo, que son el resultado de un fenómeno determinado. Invariablemente se denotará por $\{Z_t\}$ o, para ser explícito, el número de observaciones de la serie, también se podrá escribir como $\{Z_1, Z_2, \dots, Z_n\}$. Por ejemplo el índice inflacionario que publica mensualmente el Banco de México desde enero de 1969 hasta la fecha. Como puede observarse al referirnos a una serie de tiempo, debe especificarse que fenómeno representa, el período que separa a observaciones consecutivas, el número de datos y el origen (el cual se localiza en el período al que pertenece la primer observación).

Desde el punto de vista probabilístico una serie de tiempo es la sucesión de observaciones generadas por un proceso estocástico, cuyo conjunto índice se toma en relación al tiempo discreto.

Esto es: $\{Z_t; t \in T\}$ donde: T es el conjunto índice (Tiempo) y las Z_t son las variables aleatorias [4,5].

3.1 COMPONENTES DE UNA SERIE DE TIEMPO.

Al observar una serie de tiempo se supone la presencia de varias componentes, ya sea por el conocimiento que se tenga de la misma o por una simple inspección preliminar. Una serie de tiempo no solamente está formada por componentes deterministas, sino que también tiene la presencia de una componente aleatoria. Karl Pearson (1919) fue quien distinguió por vez primera la existencia de cuatro componentes dentro de una serie de tiempo. Dichas componentes comúnmente se utilizan para caracterizar los distintos movimientos que pueden presentar una serie como son: la tendencia, el ciclo, la estacionalidad y la fluctuación irregular.

Tendencia. Movimientos crecientes o decrecientes que caracterizan una serie de tiempo en un período mucho mayor en relación a la unidad de tiempo. Esta tendencia refleja un crecimiento a largo plazo o un declive en la serie.

La tendencia puede presentarse por diversos factores como son:

- i) Cambios tecnológicos en la Industria.
- ii) Cambios en los gustos del consumidor.
- iii) Incrementos en el ingreso per capita.
- iv) Incrementos en la población total.
- v) Mercado creciente.
- vi) Inflación o deflación.

Ciclo. Oscilación casi-periodica alrededor del nivel de la tendencia caracterizada por periodos alternantes de expansión y contracción. Esta fluctuación puede tener una duración de dos a diez años, o una medida uniforme de pico a pico o de valle a valle.

Las fluctuaciones cíclicas no necesariamente son causadas por cambios en factores económicos. Puede deberse a factores políticos o sociales, en México hacia el final del período presidencial siempre hay una contracción económica.

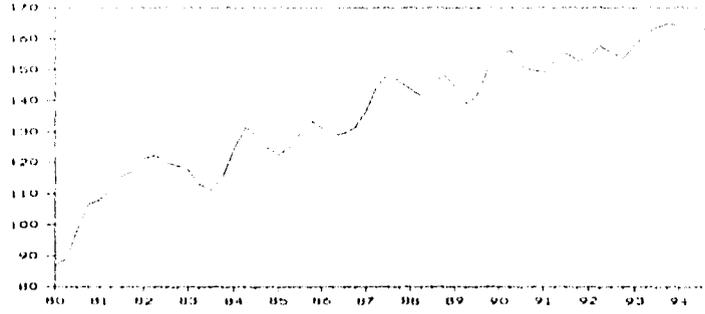
Estacionalidad. Fluctuaciones periódicas que recurren cada año con cierta regularidad, pero pueden evolucionar, aproximadamente en las mismas fechas y con la misma intensidad y las cuales, lo que es más importante, pueden medirse y eliminarse de la serie de tiempo en consideración. El fenómeno es causado principalmente por fuerzas no-económicas, exógenas al sistema económico, y no pueden controlarse o modificarse por los tomadores de decisiones. Por ejemplo en el mes que contiene la Semana Santa, los días laborados, en general, son inferiores a los otros meses. Otro ejemplo de efectos estacionales es la producción de azúcar que se ve disminuida en la zafra.

Fluctuación Irregular. Movimientos erráticos e imprevisibles relacionados con eventos de toda clase que siguen las series de tiempo y además tiene la característica de ser básicamente aleatoria. Cada movimiento representa un sobrante en la serie de tiempo que la tendencia, ciclo y variación estacional deben explicar. Es causada usualmente por eventos que no pueden ser pronosticados, también por errores en el análisis.

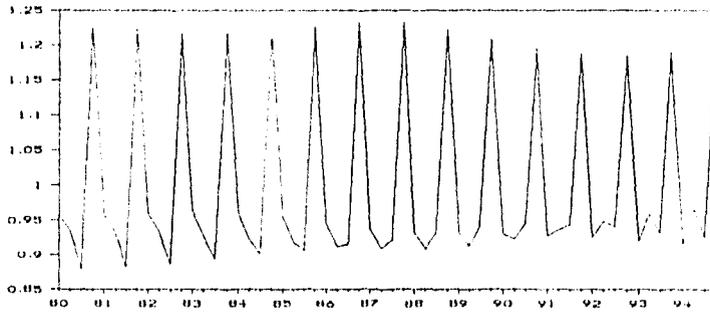
Quizá la mejor manera de apreciar el comportamiento de cada una de las componentes, sea mediante su observación visual; con esta finalidad, la gráfica 1.1 presenta un ejemplo de las componentes de tendencia-ciclo (debido a que no hay interés específico en el ciclo, por lo común no se realiza ningún esfuerzo para tratar de separar estas componentes, y se presentan conjuntamente), de estacionalidad y de irregularidad.

Gráfica 1.1

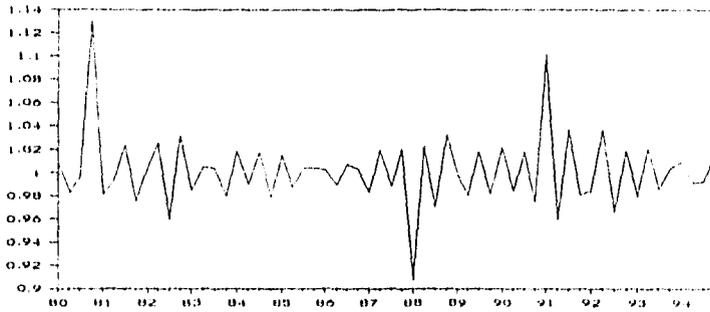
TENDENCIA-CICLO



ESTACIONALIDAD

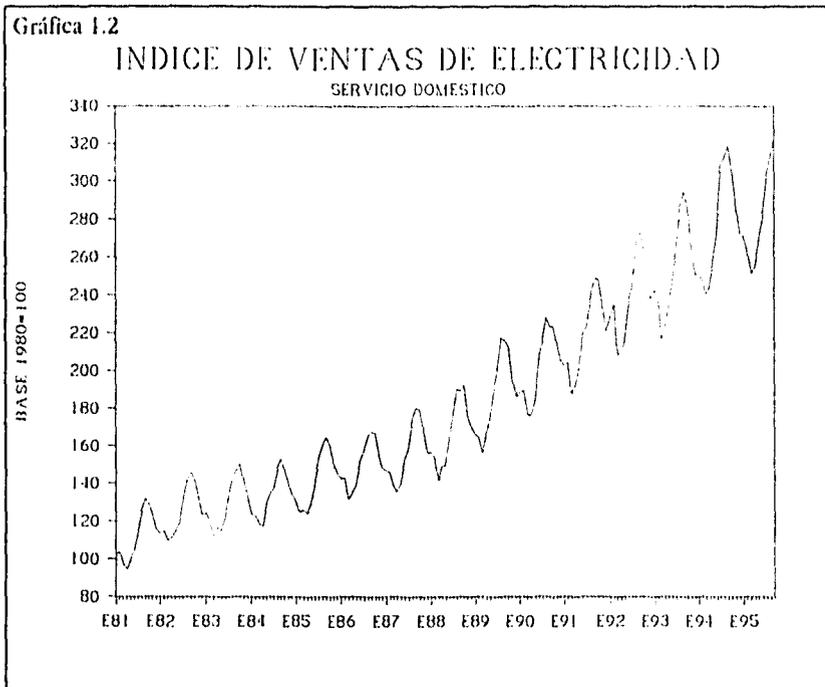


IRREGULARIDAD



1-3.2 TIPOS DE ESTIMACIONES.

Las componentes no siempre se presentan solas, pueden ocurrir en una combinación o todas al mismo tiempo. La gráfica 1.2, como se puede observar, muestra una serie que presenta todas las componentes al mismo tiempo. Por esta razón no existe una buena técnica de estimación, es decir, una técnica usada para estimar una serie de tiempo caracterizada por solo una componente, no es apropiada para una serie basada en una combinación de componentes. Hasta ahora uno de los problemas más importantes al hacer estimaciones de series de tiempo es escoger la técnica de estimación apropiada.



Se consideran dos tipos de estimaciones: puntual y por intervalos. Una estimación puntual es un número que representa "la mejor predicción" (de acuerdo al método de estimación usado) del valor de la variable en cuestión, dado un punto en el tiempo, es la mejor suposición para el valor futuro de la variable a estimar. Un intervalo de confianza es un rango de valores calculados a un porcentaje de confianza en el cual se asegura que el valor real de la variable estimada estará en dicho intervalo, muchas veces es útil desde el punto de vista de la planeación.

Desafortunadamente cualquier estimación que se haga de una serie de tiempo envuelve algún grado de incertidumbre. Se reconoce este factor para incluirlo en la componente irregular, la presencia de esta componente representa una inexplicable o imprevisible fluctuación en los datos por medio del cual algún error en la estimación se puede esperar.

1-3.3 METODOS DE ESTIMACION.

Como se mencionó antes, existen varias técnicas de estimación dependiendo del tipo de información con que se cuenta. Entre otras cosas, las distinciones respecto al tipo y volumen de los datos disponibles, conducen a efectuar una división dentro de las técnicas de estimación. Estos métodos pueden dividirse en dos tipos básicos: métodos cualitativos y métodos cuantitativos.

1-3.3.1 METODOS CUALITATIVOS.

Generalmente usan la opinión de expertos para predecir subjetivamente eventos futuros, éstos modelos muchas veces se requieren cuando no se dispone de la serie histórica, se dispone de pocos datos o no son muy confiables. Existen varias técnicas de estimación cualitativas, una de éstas técnicas se refiere a un ajuste subjetivo de una curva. Otro modelo común es el llamado modelo Delphi, una tercer técnica de estimación cualitativa es el cambio tecnológica¹.

Método Delphi. Se basa en una serie de cuestionarios, aplicados a un grupo de expertos en el área, para obtener pronósticos basados en la misma información, aunque no sean unánimes².

Cambio Tecnológico. Se basa en el hecho de que los cambios en la tecnología se presentan gradualmente y aparecen primero en una rama industrial aunque después se transmiten a otras, con lo cual se pretende tener en cuenta los impactos de diversos eventos en el futuro, así como las probabilidades de ocurrencia.

1-3.3.2 METODOS CUANTITATIVOS.

Estas técnicas, más objetivas, involucran el análisis de la serie histórica en un intento por predecir valores futuros de la variable de interés, utilizando herramientas matemáticas. Se agrupan en:

- a) Análisis de Series de Tiempo
- b) Análisis Causal.

La técnica más común es llamada serie de tiempo, en éstos modelos la serie histórica de la variable a pronosticar es analizada en un intento por identificar el patrón que siguen los datos. Asumiendo que se continuará en el futuro, éste patrón es extrapolado para producir pronósticos. Hay que observar que el modelo de serie

¹ Existen otras técnicas, como el método de impactos cruzados o el método de investigación morfológica

² Véase: Daklkey, Norman C. "The Delphi Method. An Experimental Study of Group Opinion" Corporation, Santa Monica, Calif 1969

de tiempo genera predicciones que son basadas solamente en el patrón histórico de la variable a ser pronosticada, así cualquier decisión de manejo implementada en el futuro no alterará predicciones generadas por un modelo de series de tiempo. Por lo tanto las series de tiempo son útiles cuando las condiciones son las mismas, pero no siempre son útiles al pronosticar cambios en políticas administrativas.

El uso del análisis causal involucra la identificación de otras variables que están relacionadas con la variable a ser predecida, ésta relación debe estar dada por un modelo estadístico que describa tal relación, la cual se usa para pronosticar la variable de Interés.

Factores a considerar al escoger una técnica de pronóstico:

1. El pronóstico deseado
2. El tiempo de realización
3. El patrón de los datos
4. El costo del pronóstico
5. La precisión deseada
6. La disponibilidad de los datos
7. La facilidad de operación y comprensión.

El primer factor a considerar es que tipo de estimación se desea: puntual o por intervalos. En algunas situaciones un solo valor puede ser suficiente, mientras en otras se puede requerir todo un intervalo de confianza. Además, que algunas técnicas de estimación producen intervalos de confianza técnicamente correctos, mientras que con otros métodos no es así.

El segundo factor que influye al escoger un método de estimación es la periodicidad de los datos. Las estimaciones generalmente son puntos en el tiempo, por ejemplo, días, semanas, meses, trimestres o años. A la longitud del período se le llama comúnmente en inglés "time frame" (marco del tiempo). Difícilmente un "time frame" grande producirá estimaciones exactas, conforme aumenta el "time frame" es más conveniente utilizar técnicas de estimación cualitativas.

A menudo la técnica para estimar un fenómeno dado, la puede definir el patrón de los datos. Si estos presentan tendencia, estacionalidad o una componente cíclica o una combinación de estas, se escogerá un método u otro.

Existen varios costos a considerar. Primero, el costo de desarrollo de los procesos a seguir en el método seleccionado, la complejidad y por lo tanto el costo de esos procesos varía de técnica a técnica. Segundo, el costo derivado de la obtención y mantenimiento de la información, algunos métodos requieren mucha información mientras otros solamente la elemental, reduciendo este costo. Por último, el costo real de operación de cada método de estimación.

Otro factor muy importante es la exactitud que se desea del pronóstico, en algunas situaciones, una estimación que tiene un error del 20%, puede ser aceptable; en otras situaciones una estimación que tiene un error del 1% puede ser desastrosa.

La disponibilidad de la información es un factor que puede determinar el método de estimación a usar, ya que varios métodos requieren diferentes cantidades de datos históricos.

Por último el conocimiento y la facilidad para operar un método de estimación por un experto es muy importante, ya que él es el responsable de las predicciones hechas.

CAPITULO 2.

SERIES DE TIEMPO CON TENDENCIA E IRREGULARIDAD.

En este capítulo se discutirá la estimación de series de tiempo que presentan en su comportamiento las componentes de tendencia e irregularidad. Esto es, modelos de la forma:

$$y_t = T_t + \epsilon_t$$

donde: T_t es la tendencia en el tiempo t y ϵ_t es la irregularidad en el tiempo t .

Especialmente se examinarán estimaciones de series de tiempo para las cuales T_t puede asumir cualquiera de las siguientes formas:

1.-

$$T_t = \beta_0$$
$$\Rightarrow y_t = T_t + \epsilon_t$$
$$\therefore y_t = \beta_0 + \epsilon_t$$

En este caso se supone que no existe tendencia. Esto significa que la serie fluctúa aleatoriamente alrededor de un nivel medio β_0 .

2.-

$$T_t = \beta_0 + \beta_1 t$$
$$\Rightarrow y_t = T_t + \epsilon_t$$
$$\therefore y_t = \beta_0 + \beta_1 t + \epsilon_t$$

En este caso se supone que existe tendencia lineal. Es decir, la serie de tiempo fluctúa aleatoriamente alrededor de un nivel medio siguiendo una función lineal. Donde la pendiente es β_1 y la ordenada al origen es β_0 . Si $\beta_1 > 0$ entonces el nivel medio de la serie se incrementa conforme transcurre el tiempo. Si $\beta_1 < 0$ entonces el nivel medio de la serie decrece conforme transcurre el tiempo.

En este capítulo se discutirán los procesos de regresión lineal y el de suavizamiento exponencial y se usarán para pronosticar series de tiempo con tendencia y sin tendencia. Antes de continuar se deberá tener presente que los modelos y métodos discutidos en esta parte suponen que los $\epsilon_1, \epsilon_2, \dots$ en el modelo:

$$y_t = T_t + \epsilon_t$$

son no correlacionados (este concepto se explica en el capítulo 3). Si las observaciones son solo débilmente correlacionadas entonces los métodos discutidos aquí aun producirán probablemente pronósticos bastante exactos, pero, si las observaciones son fuertemente correlacionadas, la metodología Box-Jenkins discutida en el capítulo 4, utiliza esa dependencia para producir pronósticos que probablemente son más exactos que los pronósticos producidos por los métodos del capítulo 3.

2-1 ANALISIS DE REGRESION LINEAL MULTIPLE.

En este punto se introducirá una técnica de estimación cuantitativa; el análisis de regresión lineal múltiple, la cual es una herramienta importante en el análisis y pronóstico de series de tiempo. Se presenta en notación de álgebra matricial porque proporciona un método condensado para manejar los modelos de regresión con cualquier número de variables; una vez formulado el modelo múltiple y resuelto en notación matricial, la solución se puede aplicar a una, dos o más variables [3].

2-1.1 ENFOQUE MATRICIAL DEL MODELO.

Supongase que y es una función lineal de las variables x_1, x_2, \dots, x_p y un error ϵ . Entonces el modelo de regresión lineal múltiple es:

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_p x_{tp} + \epsilon_t$$

donde: y_t es la variable dependiente; p representa el número de variables independientes usadas en el modelo; $x_{t1}, x_{t2}, \dots, x_{tp}$ representan el valor de esas p variables independientes; $\beta_0, \beta_1, \dots, \beta_p$ son los parámetros desconocidos que relacionan a la variable dependiente con las variables independientes y ϵ_t es el error aleatorio que describe la influencia sobre y_t de todos los otros factores distintos a las p variables independientes.

Supongase que se cuenta con n observaciones, se puede resumir el modelo de regresión, escribiendo una serie de n ecuaciones, como sigue:

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_p x_{1p} + \epsilon_1$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_p x_{2p} + \epsilon_2$$

.....

.....

$$y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_p x_{np} + \epsilon_n$$

Entonces, la formulación matricial del modelo es:

$$Y = X\beta + \epsilon$$

donde:

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

2-1.2 SUPUESTOS DEL MODELO DE REGRESION LINEAL.

Las suposiciones clásicas del modelo de regresión lineal múltiple son:

- i) La especificación del modelo esta dada por $Y = X\beta + \epsilon$.
- ii) los elementos de X son fijos y tienen varianza finita. Además el rango de X es p el cual es menor que el número de observaciones n .
- iii) ϵ se distribuye normalmente con esperanza cero y varianza constante. Los errores son independientes y por lo tanto no correlacionados.

La suposición de que X tiene rango p garantiza la no colinealidad (es decir, que una de las columnas de X no sea combinación lineal del resto de columnas). La suposición de los errores, posiblemente la más fuerte, ya que ella garantiza las propiedades estadísticas, así como las aritméticas del proceso de estimación por mínimos cuadrados. Adicionalmente con la normalidad se supone que cada uno de los errores tiene esperanza cero, las varianzas todas son constantes y las covarianzas son cero.

2-1.3 ESTIMACION POR MINIMOS CUADRADOS.

Para hallar el estimador mínimo cuadrado $\hat{\beta}$ de β , primero se denota la estimación de y_i como:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_p x_{ip}$$

la cual puede escribirse de manera condensada en notación matricial como:

$$\hat{Y} = X\hat{\beta}$$

El objetivo es encontrar un vector con parámetros $\hat{\beta}$, el cual minimice la suma de errores al cuadrado, esto es:

$$\hat{e}'\hat{e} = (\hat{e}_1 \hat{e}_2 \dots \hat{e}_n) \begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \vdots \\ \hat{e}_n \end{pmatrix} = \hat{e}_1^2 + \hat{e}_2^2 + \dots + \hat{e}_n^2 = \sum \hat{e}_i^2$$

ahora, como $Y = X\beta + e$ y $\hat{Y} = X\hat{\beta}$ se obtiene:

$$\begin{aligned} \hat{e} &= Y - \hat{Y} \\ &= Y - X\hat{\beta} \end{aligned}$$

Por lo tanto:

$$\begin{aligned} \hat{e}'\hat{e} &= (Y - X\hat{\beta})'(Y - X\hat{\beta}) \\ &= Y'Y - \hat{\beta}'X'Y - Y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} \\ &= Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta} \end{aligned}$$

Para determinar los estimadores mínimo cuadrados, se minimiza la suma de errores al cuadrado usando las reglas de derivación matricial como sigue:

$$\frac{\partial \sum \hat{e}_i^2}{\partial \hat{\beta}} = -2X'Y + 2X'X\hat{\beta}$$

Igualando a cero la ecuación anterior, se tiene:

$$-2X'Y + 2X'X\hat{\beta} = 0$$

$$\hat{\beta} = (X'X)^{-1}X'Y$$

donde: $(X'X)^{-1}$ es la inversa de la matriz $X'X$

$\hat{\beta}$ es un estimador insesgado de β . Esto es $E(\hat{\beta}_i) = \beta_i$, para toda $i = 0, 1, 2, \dots, p$

Los estimadores mínimo cuadrados de β se distribuyen normalmente, ya que $\epsilon_i \sim N(0, \sigma^2)$ con $i = 1, 2, \dots, n$, entonces $\hat{\beta}_i$ es normal.

1 La matriz $(X'X)$ tiene inversa por la suposición ii

2-1.4 MATRIZ DE VARIANZA COVARIANZA.

El método matricial permite calcular tanto la varianza de $\hat{\beta}_i$, como las covarianzas de dos elementos de $\hat{\beta}$, por ejemplo: $\hat{\beta}_i$ y $\hat{\beta}_j$. Estas varianzas y covarianzas se necesitan para la inferencia estadística.

Para llegar a la matriz de varianza y covarianza, se sustituye $Y = X\beta + e$ en $\hat{\beta} = (X'X)^{-1}X'Y$ entonces, se obtiene:

$$\begin{aligned}\hat{\beta} &= (X'X)^{-1}X'(X\beta + e) \\ &= \beta + (X'X)^{-1}X'e \\ \therefore \hat{\beta} - \beta &= (X'X)^{-1}X'e\end{aligned}$$

Por definición:

$$\begin{aligned}\text{var}(\hat{\beta}) &= E\{(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\} \\ &= E\{[(X'X)^{-1}X'e]['(X'X)^{-1}X'e]\} \\ &= E\{(X'X)^{-1}X'e e'X(X'X)^{-1}\}\end{aligned}$$

Recordando que las X son fijas, al calcular la esperanza, se obtiene:

$$\begin{aligned}\text{var}(\hat{\beta}) &= (X'X)^{-1}X'E(e e')X(X'X)^{-1} \\ &= (X'X)^{-1}X'\sigma^2 I X(X'X)^{-1} \\ &= \sigma^2 (X'X)^{-1}\end{aligned}$$

Notar que para llegar al resultado anterior se hizo uso del supuesto $E(e e') = \sigma^2 I$.

Un estimador insesgado para σ^2 esta dado por la fórmula:

$$\begin{aligned}\hat{\sigma}^2 &= \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - (p + 1)} \\ \therefore \hat{\sigma} &= \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - (p + 1)}}\end{aligned}$$

Más fácilmente, las varianzas de las $\hat{\beta}_i$, están dadas por:

$$\text{Var}(\hat{\beta}_i) = \hat{\sigma}^2 C_{ii}$$

donde C_{ii} es el i -ésimo elemento de la diagonal de la matriz $(X'X)^{-1}$.

Las covarianzas están dadas por:

$$\text{Cov}(\hat{\beta}_i, \hat{\beta}_j) = \hat{\sigma}^2 C_{ij}$$

2-1.5 COEFICIENTE DE DETERMINACION r^2 .

Para introducir r^2 supongase que se conocen los valores de y_b pero no se conocen los valores de $x_{i1}, x_{i2}, \dots, x_{ip}$ con los cuales se pronostica y_b . En este caso un pronóstico razonable sería:

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

El error de estimación sería entonces: $(y_i - \bar{y}_i)$.

Sin embargo, se tienen los valores $x_{i1}, x_{i2}, \dots, x_{ip}$ para pronosticar y_b . El pronóstico de y_i es:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_p x_{ip}$$

Entonces el error de estimación sería: $(y_i - \hat{y}_i)$.

La cantidad $\sum_{i=1}^n (y_i - \bar{y}_i)^2$ se llama variación total y es igual a la suma de los errores al cuadrado que se obtienen si no se usan las variables independientes.

La cantidad $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ se llama variación no explicada y es igual a la suma de los errores al cuadrado que se obtienen si se usan las variables independientes.

La cantidad $\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2$ se llama variación explicada ya que:

$$\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2 = \sum_{i=1}^n (y_i - \bar{y}_i)^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

representa la reducción en la suma de errores al cuadrado que se ha realizado al usar las variables independientes al hacer pronósticos.

Se define el coeficiente de determinación r^2 por la ecuación:

$$r^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

r^2 es ampliamente usado como una medida de la bondad del ajuste en una regresión. Es decir, el r^2 mide la proporción o porcentaje de la variación total en Y explicada por el modelo de regresión. Sus propiedades más importantes son:

- i) Es una cantidad no negativa.
- ii) $0 \leq r^2 \leq 1$

Para $r^2 = 1$ quiere decir ajuste perfecto, mientras que para $r^2 = 0$ quiere decir que no hay relación entre la variable dependiente y las variables explicatorias.

Otro camino para expresar r^2 es:

$$\begin{aligned} r^2 &= \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} = \frac{\text{variación explicada}}{\text{variación total}} \\ &= \frac{\sum (y_i - \bar{y})^2 - \sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2} \\ &= 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2} \\ &= 1 - \frac{\text{variación no explicada}}{\text{variación total}} \end{aligned}$$

En notación matricial:

$$r^2 = \frac{\hat{\beta} X' Y - n \bar{Y}^2}{Y' Y - n \bar{Y}^2}$$

2-1.6 INTERVALOS DE CONFIANZA

Usando la inferencia estadística [3] se puede mostrar que un intervalo al $(100 - \alpha)\%$ de confianza para \hat{y}_t es:

$$\hat{y}_t \pm t_{\left(\frac{\alpha}{2}, n-(p+1)\right)} \cdot \hat{\sigma} \cdot f_t$$

donde: $t_{\left(\frac{\alpha}{2}, n-(p+1)\right)}$ es el valor crítico de la distribución t asociado a un nivel de significancia de $\alpha/2$ con $(n - (p + 1))$ grados de libertad. $\hat{\sigma}$ ya se definió anteriormente y

$$f_t = \sqrt{1 + X'_t (X'X)^{-1} X_t}$$

f_t es una medida de la incertidumbre que resulta del desconocimiento de los valores verdaderos de los parámetros $\beta_0, \beta_1, \dots, \beta_p$ y $X'_t = (1, x_{t1}, x_{t2}, \dots, x_{tp})$.

El intervalo es una función de t y del nivel de confianza $(100 - \alpha)\%$, mientras más grande es el nivel de confianza, más grande es el intervalo. El nivel de confianza $(100 - \alpha)\%$ puede ser arbitrario, usualmente $\alpha = 5$, en tal caso se calcularía un intervalo al 95% de confianza.

Un intervalo al $(100 - \alpha)\%$ de confianza para β_i , está dado por:

$$\hat{\beta}_i \pm t_{\left(\frac{\alpha}{2}, n-(p+1)\right)} \hat{\sigma} \sqrt{C_{ii}}$$

2-1.7 PRUEBAS DE HIPOTESIS.

Considérese ahora, como decidir, por medio de pruebas de hipótesis, si la variable independiente x_{it} tiene importancia adicional significativa sobre combinaciones lineales de las variables independientes

$$x_{i1}, \dots, x_{i,i-1}, x_{i,i+1}, \dots, x_{ip}$$

con relación al modelo:

$$y_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_{i-1} x_{i-1,t} + \beta_i x_{it} + \beta_{i+1} x_{i+1,t} + \dots + \beta_p x_{pt} + \epsilon_t$$

Es razonable decidir que la variable independiente x_{it} tiene importancia adicional significativa si se rechaza la hipótesis nula:

$$H_0: \beta_i = 0$$

la cual nos dice que ese cambio x_{it} no afecta a y_{it} en contra de la hipótesis alternativa

$$H_1: \beta_i \neq 0$$

la cual nos dice que ese cambio x_{it} produce cambios sobre y_{it} .

En la teoría clásica de pruebas de hipótesis, la regla para decidir cuando rechazar H_0 se basa frecuentemente en especificar una probabilidad de cometer un Error del Tipo I. Un Error del Tipo I, se comete si se rechaza H_0 dado que H_0 es verdadera. A la probabilidad de cometer un Error del Tipo I se le denota como α . Aunque se desea que α sea pequeña, haciendo α muy pequeña se puede cometer el Error del Tipo II. Un Error del Tipo II, se comete si no se rechaza H_0 dado que H_0 es falsa. Generalmente $\alpha = 0.05$, esto es, se debe establecer una regla para decidir cuando rechazar H_0 basada en la probabilidad igual a 0.05 de cometer el Error del Tipo II. En general, se puede mostrar [3] que si la regla de decisión basada en la probabilidad de cometer el Error del tipo I es igual a α , entonces la regla que se debe usar para rechazar la hipótesis nula

$$H_0: \beta_i = 0$$

en contra de la hipótesis alternativa

$$H_1: \beta_i \neq 0$$

y por lo tanto decidir que la variable independiente x_{it} tiene importancia adicional significativa, si:

$$\left| \frac{\hat{\beta}_i}{\hat{\sigma} \sqrt{C_{ii}}} \right| > t_{\left(\frac{\alpha}{2}, n-(p-1)\right)}$$

A las pruebas de hipótesis de la forma:

$$H_0: \beta_i = \beta_j \quad vs \quad H_1: \beta_i > \beta_j$$

$$H_0: \beta_i = \beta_j \quad vs \quad H_1: \beta_i < \beta_j$$

$$H_0: \beta_i = \beta_j \quad vs \quad H_1: \beta_i \neq \beta_j$$

se les conoce como: prueba de cola superior, prueba de cola inferior y prueba de dos colas respectivamente.

El estadístico de prueba es:

$$t_{\beta_i} = \frac{\hat{\beta}_i - \beta_i}{\hat{\sigma} \sqrt{C_{ii}}}$$

y la región de rechazo depende de la prueba:

$$\{t_{\beta_i} > t_{\frac{\alpha}{2}, n-(p+1)}\}$$

Para verificar cualquier hipótesis particular sobre β_i , se sustituye el valor hipotético de β_i en t_{β_i} y si el valor resultante de t_{β_i} se encuentra en la región de rechazo, se desecha la hipótesis.

2-1.8 EJEMPLO.

Desde 1973 el Banco de México elabora una encuesta de tipo cualitativa, llamada "Encuesta Mensual de Coyuntura", sobre la opinión y las expectativas de los empresarios en el sector manufacturero del país, con respecto a la producción, ventas, inventarios precios y empleo. El aspecto cualitativo se refiere al tipo de información que se recaba del empresario. A éste se le pregunta para cada una de las variables (por ejemplo: el número de trabajadores de su empresa) si, aumentó, permaneció igual o disminuyó con relación al mes anterior (opinión) y con relación al mes de captación expectativa. La información captada por la encuesta tiene una regularidad mensual, los resultados se ponderan por el número de personal ocupado de cada empresa y se contabilizan de la siguiente manera:

- a) Porcentaje ponderado de respuestas que indican que la variable aumentó.
- b) Porcentaje ponderado de respuestas que indican que la variable permaneció igual.
- c) Porcentaje ponderado de respuestas que indican que la variable disminuyó.

De esta manera se generan series de tiempo con dichos porcentajes para cada una de las variables de la encuesta.

En este ejemplo, se desea encontrar un modelo de regresión lineal múltiple para hacer pronósticos acerca de la variación mensual del índice de empleo en el sector de manufacturas, base 1985=100.

La variable dependiente denotada por y_t , es igual a la variación mensual del índice de empleo y se cree subjetivamente que está en función de las variables dependientes x_{1t} y x_{2t} donde:

x_{1t} =Porcentaje de respuestas de la Encuesta de Coyuntura que indican que el número de trabajadores en el sector manufacturero aumentó.

x_{2t} =Porcentaje de respuestas de la Encuesta de Coyuntura que indican que el número de trabajadores en el sector manufacturero disminuyó.

Un posible modelo de regresión lineal múltiple puede ser:

$$y_t = \beta_0 + \beta_1 x_{1t} + x_{2t} + \epsilon_t$$

La tabla 2.1 muestra los datos para llevar a cabo los cálculos de la regresión.

TABLA 2.1 Índice de empleo en el sector manufacturero y porcentajes de la Encuesta de Coyuntura

	t	Índice de empleo	Variación mensual	empleo aumentó	empleo disminuyó	producción aumentó	producción disminuyó
1992	1	97.3	0.77	15.20	19.20	44.20	30.10
f	2	97.7	0.39	18.60	16.40	36.60	35.50
nt	3	98.0	0.36	15.90	16.00	52.00	18.80
a	4	97.4	-0.64	11.80	15.70	31.80	41.20
m	5	97.0	-0.36	16.90	18.30	48.90	27.60
i	6	96.9	-0.15	14.30	21.60	41.80	29.40
j	7	96.6	-0.25	15.10	21.00	37.50	29.20
a	8	95.9	-0.75	12.00	25.90	22.30	41.60
s	9	95.3	-0.63	15.50	28.40	36.80	29.70
o	10	95.1	-0.23	11.50	30.80	47.50	23.20
n	11	94.5	-0.60	13.20	30.00	22.90	42.20
d	12	93.1	-1.47	9.50	37.70	29.70	46.10
1993	13	91.9	-1.30	17.60	29.40	41.70	40.80
f	14	92.2	0.27	16.40	27.70	41.60	30.90
m	15	92.1	-0.04	19.70	26.60	61.50	19.90
a	16	91.5	-0.64	18.40	32.20	24.80	53.00
m	17	91.5	-0.10	14.70	37.20	47.10	30.30
i	18	90.7	-0.81	20.60	30.20	36.20	41.30
j	19	90.0	-0.82	21.00	32.90	36.80	40.00
a	20	89.3	-0.77	18.50	33.80	38.60	36.10
s	21	88.4	-0.93	17.80	32.90	36.90	38.60
o	22	87.6	-0.90	17.90	30.90	48.80	27.70
n	23	87.2	-0.48	21.50	27.40	40.00	38.90
d	24	86.3	-1.04	11.80	35.80	36.60	46.30
1994	25	85.9	-0.47	24.60	33.50	45.30	38.30
f	26	86.5	0.64	25.30	27.30	39.50	39.60
m	27	86.6	0.16	26.50	28.20	55.20	27.50
a	28	86.6	-0.02	25.90	23.10	46.80	35.10
m	29	86.8	0.28	23.30	28.70	42.50	37.80
i	30	86.7	-0.15	22.50	32.40	49.40	30.60
j	31	86.4	-0.38	19.60	33.90	37.60	46.30
a	32	86.0	-0.41	20.90	30.60	51.50	27.10
s	33	85.9	-0.12	19.60	31.50	52.00	51.40
o	34	86.0	0.15	27.50	28.30	52.00	30.40
n	35	86.0	-0.04	19.08	25.30	44.60	39.50
d	36	85.2	-0.90	15.00	29.90	32.00	54.40

Fuente: Banco de México.

$$Y = \begin{pmatrix} 0.77 \\ 0.39 \\ \cdot \\ \cdot \\ \cdot \\ 0.90 \end{pmatrix} \quad X = \begin{pmatrix} 1 & 15.20 & 19.20 \\ 1 & 18.60 & 16.40 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & 15.00 & 29.90 \end{pmatrix}$$

por tanto

$$(X'X)^{-1} = \begin{pmatrix} 1.08526 & 0.02343 & -0.02249 \\ -0.02343 & 0.00137 & 0.00006 \\ -0.02249 & -0.00006 & 0.00084 \end{pmatrix} \quad X'Y = \begin{pmatrix} -13.96 \\ -215.45 \\ -432.86 \end{pmatrix}$$

Entonces los estimadores mínimo cuadrados de β_0 , β_1 y β_2 son:

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (X'X)^{-1} X'Y = \begin{pmatrix} -0.36893 \\ 0.05539 \\ -0.03660 \end{pmatrix}$$

Un intervalo al 95% de confianza para y_t es:

$$\hat{y}_t \pm t_{\left(\frac{\alpha}{2}, n-(p+1)\right)} \cdot \hat{\sigma} \cdot f_t$$

donde:

$$\hat{y}_t = -0.36893 + 0.05539 X_{it} - 0.03660 X_{it}^2$$

y aceptando que $X'_{37} = (1 \quad 15.40 \quad 47.50)$

$$t_{\left(\frac{\alpha}{2}, 36-(2+1)\right)} = t_{\left(\frac{\alpha}{2}, 33\right)} \approx 1.96$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{36} (y_i - \hat{y}_i)^2}{36 - (2 + 1)}} = \sqrt{\frac{4.97542}{33}} = 0.38829$$

$$f_{37} = \sqrt{1 + X'_{37} (X'X)^{-1} X_{37}} = \sqrt{1 + 0.36130} = 1.16674$$

$$\hat{y}_{37} = -0.36893 + 0.05539 \times 15.40 - 0.03660 \times 47.50 = -1.25452$$

Por lo tanto el intervalo al 95% de confianza para y_t con $t = 37$ es:

$$-1.25452 \pm (1.96)(0.38790)(1.16674)$$

$$-1.25452 \pm 0.88705$$

Ahora el cálculo de r^2 es:

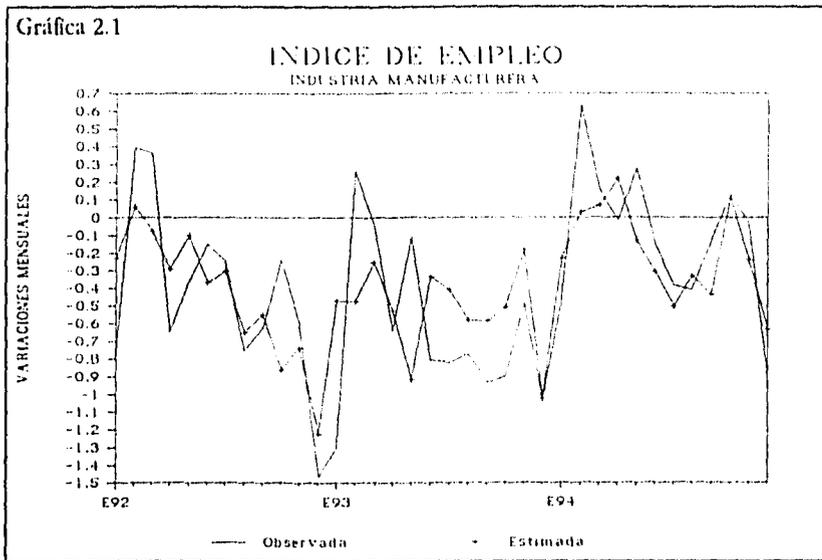
$$r^2 = \frac{\sum_{i=1}^{36} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{36} (y_i - \bar{y})^2} = \frac{3.64572}{8.62114} = 0.42288$$

Lo cual quiere decir que el modelo de regresión:

$$Y_t = \beta_0 + \beta_1 X_{1t} + X_{12} + \epsilon_t$$

explica el 42,28% de la variación total en las 36 observaciones.

La gráfica 2.1 puede ser de utilidad para determinar la funcionalidad de la regresión. Otro camino es atendiendo al valor de r^2 . Sin embargo las pruebas de hipótesis pueden ser de mayor utilidad para construir un mejor modelo.



Ahora supongase que :

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{12} + \beta_3 X_{13} + \beta_4 X_{14} + \epsilon_t$$

donde X_{1t} y X_{12} ya se han definido anteriormente y

X_{13} = Porcentaje de respuestas de la Encuesta de Coyuntura que indican que la producción en el sector manufacturero aumentó.

X_{14} = Porcentaje de respuestas de la Encuesta de Coyuntura que indican que la producción en el sector manufacturero disminuyó.

Los datos de X_{13} y X_{14} también se muestran en la tabla 2.1

Usando el estadístico t_{β_i} se puede evaluar la importancia adicional de la constante β_0 y las variables independientes X_{i1} , X_{i2} , X_{i3} y X_{i4} en el modelo de regresión.

TABLA 2.2 Estadísticos t_{β_i} para el modelo:

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \epsilon_i$$

Variable independiente	i	$\hat{\beta}_i$	σ_{β_i}	$t_{\beta_i} = \frac{\hat{\beta}_i - 0}{\sigma_{\beta_i}}$
cte	0	-0.27191	0.09041	-0.39384
X_{i1}	1	0.04740	0.01696	2.79149
X_{i2}	2	-0.02879	0.01220	-2.35933
X_{i3}	3	0.00551	0.01167	0.47446
X_{i4}	4	-0.01113	0.01134	-0.98145

Ya que $p = 4$ y $n = 36$ se tiene:

$$t_{\left(\frac{5}{2}, 31\right)} \approx 1.96$$

En la tabla 2.2 se nota que t_{β_1} y t_{β_2} son mayores en valor absoluto que 1.96, por lo tanto se puede concluir que X_{i1} y X_{i2} tiene importancia adicional significativa. Sin embargo, se cree subjetivamente que X_{i4} se puede incluir en el modelo, ya que t_{β_4} se aproxima un poco mas al valor de $t_{\left(\frac{5}{2}, 31\right)}$. Por lo tanto, únicamente se excluyen del modelo la constante β_0 y X_{i3} por ser menos significativas.

Entonces considérese el modelo:

$$y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i4} + \epsilon_i$$

TABLA 2.3 Estadísticos t_{β_i} para el modelo:

$$y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \epsilon_t$$

Variable independiente	i	$\hat{\beta}_i$	σ_{β_i}	$t_{\beta_i} = \frac{\hat{\beta}_i - 0}{\sigma_{\beta_i}}$
X_{1t}	1	0.05147	0.01101	4.67484
X_{2t}	2	-0.02786	0.01042	-2.67363
X_{3t}	3	-0.01508	0.00702	-2.14678

Nuevamente la tabla 2.3 muestra los valores de t_{β_i} . Ahora $p = 3$ y $n = 36$ se tiene:

$$t_{\left(\frac{5}{2}, 32\right)} \pm 1.96$$

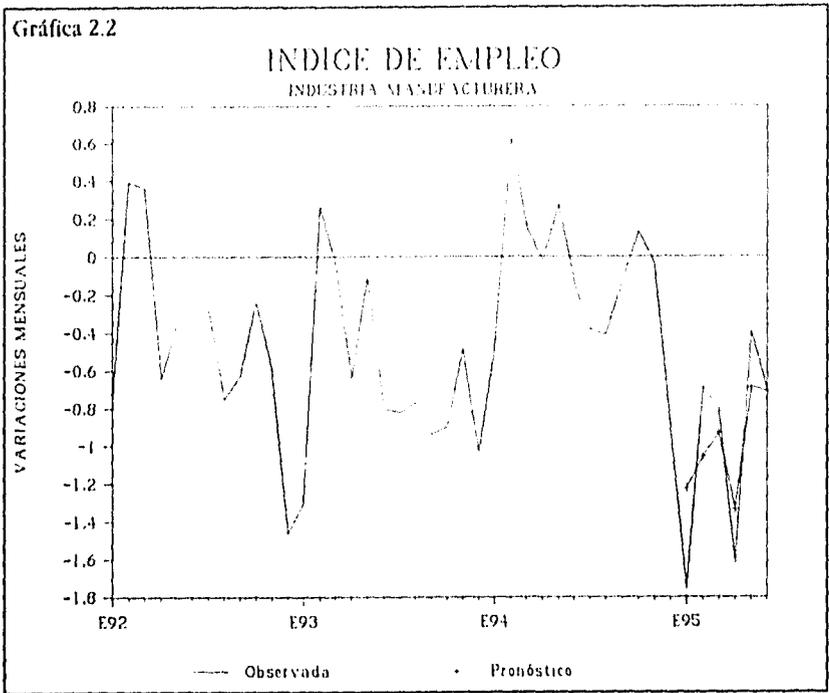
De este análisis se concluye que las variables X_{1t} , X_{2t} y X_{3t} son altamente significativas, por lo tanto:

$$y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \epsilon_t$$

puede ser un buen modelo para hacer pronósticos acerca de la variación mensual del índice de empleo.

La gráfica 2.2 muestra los pronósticos para el año 1995, usando el modelo más apropiado, atendiendo al estadístico t_{β_i} .

Gráfica 2.2



2-2 SUAVIZAMIENTO EXPONENCIAL SIMPLE.

En esta sección se discutirá la situación en la cual se supone que el nivel medio de la serie no cambia a través del tiempo, es decir se asume que no existe tendencia. Un modelo apropiado puede ser

$$Y_t = \beta_0 + \epsilon_t$$

En este caso, la serie está descrita por un valor o nivel medio β_0 el cual no cambia en el tiempo, combinada con fluctuaciones aleatorias ϵ_t las cuales provocan que las observaciones se desvíen del nivel medio.

2-2.1 APROXIMACION POR SUAVIZAMIENTO EXPONENCIAL.

Supongase que se tiene una serie de tiempo en la cual la última observación se localiza en el período $n - 1$, esto es, Y_1, Y_2, \dots, Y_{n-1} una serie de tiempo con $n - 1$ observaciones. Dadas estas $n - 1$ observaciones se pretende estimar β_0 (el nivel medio de la serie). El estimador mínimo cuadrado de β_0 en el período $n - 1$ es:

$$\hat{\beta}_0(n-1) = \bar{y} = \sum_{t=1}^{n-1} \frac{Y_t}{n-1}$$

Dado este estimador, la predicción para un valor futuro $n - 1 + \tau$, donde τ es un entero positivo, sigue siendo

$$\hat{\beta}_0(n-1) = \bar{y} = \sum_{t=1}^{n-1} \frac{Y_t}{n-1}$$

Ahora supongase que se obtiene una observación más reciente Y_n y se desearía incorporar esta nueva observación en el estimador β_0 . Esto es, se desea obtener un estimador más actual de β_0 que se hace en la nueva observación Y_n y en las anteriores Y_1, Y_2, \dots, Y_{n-1} .

Análogamente como antes, el estimador mínimo cuadrado de β_0 en el período n es

$$\hat{\beta}_0(n) = \bar{y} = \sum_{t=1}^n \frac{Y_t}{n}$$

Otro camino para incorporar esta nueva observación dentro del estimador de β_0 es el método conocido como "Suavizamiento Exponencial Simple" [8].

Para esto considérese la diferencia entre la última observación y_n y la estimación de β_0 en el periodo $n - 1$, esto es:

$$e_n = y_n - \hat{\beta}_0(n-1)$$

Ahora incorpórese $\alpha \in (0, 1)$ de tal forma que la siguiente ecuación se cumpla

$$\begin{aligned} \hat{\beta}_0(n) &= \hat{\beta}_0(n-1) + \alpha e_n \\ \therefore \hat{\beta}_0(n) &= \hat{\beta}_0(n-1) + \alpha [y_n - \hat{\beta}_0(n-1)] \end{aligned}$$

Así el nuevo estimador esta basado en el estimador anterior.

Para simplificar la notación se define $S_n = \hat{\beta}_0(n)$ de modo que la ecuación anterior se pueda reescribir como:

$$\begin{aligned} S_n &= S_{n-1} + \alpha(y_n - S_{n-1}) \\ S_n &= S_{n-1} + \alpha y_n - \alpha S_{n-1} \\ S_n &= \alpha y_n + (1 - \alpha)S_{n-1} \end{aligned}$$

Esta ecuación define el proceso llamado "Suavizamiento Exponencial Simple", donde S_n es el estimador (o estadístico) de suavizamiento y la fracción α se llama constante de suavizamiento.

Notese que S_n puede escribirse como una combinación lineal de todas las observaciones anteriores al tiempo n . Para ver esto considérese el estimador de suavizamiento en el periodo $n - 1$

$$S_{n-1} = \alpha y_{n-1} + (1 - \alpha)S_{n-2}$$

sustituyendo S_{n-1} en S_n se tiene:

$$\begin{aligned} S_n &= \alpha y_n + (1 - \alpha)[\alpha y_{n-1} + (1 - \alpha)S_{n-2}] \\ &= \alpha y_n + \alpha(1 - \alpha)y_{n-1} + (1 - \alpha)^2 S_{n-2} \end{aligned}$$

Ahora considérese el estimador de suavizamiento en el periodo $n - 2$

$$S_{n-2} = \alpha y_{n-2} + (1 - \alpha)S_{n-3}$$

sustituyendo S_{n-2} en S_n se tiene:

$$S_n = \alpha y_n + \alpha(1-\alpha)y_{n-1} + (1-\alpha)^2[\alpha y_{n-2} + (1-\alpha)S_{n-1}]$$

$$= \alpha y_n + \alpha(1-\alpha)y_{n-1} + \alpha(1-\alpha)^2 y_{n-2} + (1-\alpha)^3 S_{n-1}$$

Sustituyendo recursivamente para $S_{n-1}, S_{n-1}, \dots, S_2, y S_1$ se obtiene:

$$S_n = \alpha y_n + \alpha(1-\alpha)y_{n-1} + \alpha(1-\alpha)^2 y_{n-2} + \dots + \alpha(1-\alpha)^{n-1} y_1 + (1-\alpha)^n S_0$$

En la práctica S_0 se obtiene calculando el promedio aritmético de las primeras observaciones de la serie.

Los coeficientes $\alpha, \alpha(1-\alpha), \alpha(1-\alpha)^2, \dots, \alpha(1-\alpha)^{n-1}$, miden la aportación que las observaciones $y_n, y_{n-1}, y_{n-2}, \dots, y_1$ hacen al estimador S_n (la observación más reciente contribuye con mayor peso que las anteriores), además que dichos coeficientes decrecen geométricamente, por tal razón el nombre de "Suavizamiento Exponencial".

Ahora considérese el problema de hacer pronósticos, sea $S_n = \hat{\beta}_0(n)$, se desea pronosticar la serie de tiempo para el periodo $n + \tau$. Ya que el modelo es $y_t = \beta_0 + \epsilon_t$ el pronóstico es simplemente $\hat{y}_{n+\tau}(n) = S_n$ que es el actual estimador de β_0 . La notación $\hat{y}_{n+\tau}(n)$ se usa para enfatizar que el pronóstico es para el periodo $n + \tau$ y fue hecho en el periodo n .

2-2.2 DETERMINACION DE LA CONSTANTE DE SUAVIZAMIENTO.

Al estimar por medio del proceso de "Suavizamiento Exponencial Simple" uno mismo puede determinar el valor de la constante α . En general, cuando la serie es bastante inestable, es decir, la componente Irregular tiene varianza "grande", se elegirá un α pequeña. Para una serie más estable, en la cual la componente irregular tiene varianza "pequeña", se elegirá un α más grande. Otro camino para escoger α es el uso de la simulación, este proceso consiste en simular un conjunto de estimaciones usando diferentes valores de α . Esto es, para cada valor distinto de α se genera una estimación, esta es comparada con el valor verdadero de la serie, generando de esta manera un conjunto de errores para cada estimación. El valor de α que produce la "mejor estimación" es el que se usará para generar pronósticos para futuras observaciones de la serie de tiempo.

El conjunto de observaciones que tiene la suma de errores al cuadrado más pequeña, se considera la "mejor estimación".

2-2.3 INSTRUMENTACION DEL PROCESO.

En este punto se esbozará la secuencia de pasos involucrados en la instrumentación de un proceso de "Suavizamiento exponencial simple".

Primero, se calculará un estimador inicial de β_0 , usando las primeras observaciones (por ejemplo seis), después comenzando con este estimador inicial, se utiliza la ecuación S_n para simular toda la información disponible, utilizando un valor particular de α . Este proceso de simulación es repetido para otros valores de α y se calcula la suma de errores al cuadrado (SEC) para cada una de ellas, el valor de α que minimice la SEC, es el valor que se escogerá para hacer predicciones. Ya que se ha determinado α , el proceso se puede utilizar para predecir valores futuros de la serie, para esto se calcula un nuevo estimador inicial de β_0 usando toda la información. El nuevo estimador y la constante α óptima son usados para realizar la operación de suavizamiento exponencial simple sobre todos los datos, este último proceso produce el mejor estimador de β_0 .

Es importante notar que existen dos aspectos arbitrarios en este proceso, para examinarlos, sea n el número de observaciones. El primer aspecto es la elección de n_1 , el cual se definió como el número de observaciones escogidas del total n para calcular el estimador S_0 y a su vez poder determinar la constante α , generalmente n_1 es igual a la mitad del número de observaciones. El segundo aspecto es la selección de n_2 que es el número de observaciones usadas para calcular el valor de S_0 , después que se ha determinado el valor de α , es razonable tomar $n_2 = n$.

2-2.4 ANALISIS DE LOS ERRORES DE ESTIMACION.

Ya que un sistema de estimación nunca producirá estimaciones perfectas, a continuación se discutirán algunos métodos que pueden usarse para determinar cuando algo anda mal en el sistema de estimación. Se desearía determinar si los errores de estimación son mayores que los producidos por un sistema más "exacto". Para esto supongase que se tienen n errores de estimación $e_1(\alpha), e_2(\alpha), \dots, e_n(\alpha)$ donde α denota la constante de suavizamiento con la cual dichos errores fueron obtenidos. Definase la siguiente suma de errores:

$$Y(\alpha, n) = \sum_{i=1}^n e_i(\alpha)$$

es obvio que:

$$Y(\alpha, n) = Y(\alpha, n-1) + e_n(\alpha)$$

Ahora se define la desviación media absoluta como:

$$D(\alpha, n) = \frac{\sum_{i=1}^n |e_i(\alpha)|}{n}$$

entonces el $JND(\alpha, n)$ se define como:

$$IND(\alpha, n) = \left| \frac{Y(\alpha, n)}{D(\alpha, n)} \right|$$

Si $IND(\alpha, n)$ es "grande" significa que $Y > D$, lo cual dice a su vez que el proceso está produciendo errores que son ya sea consistentemente grandes o consistentemente pequeños. Esto implica que el sistema de estimación está produciendo estimaciones más pequeñas o más grandes que los valores de la serie de tiempo, por lo tanto un valor grande de $IND(\alpha, n)$ indica que el proceso no es exacto. En la práctica si $IND(\alpha, n)$ excede de un límite de control denotado por K , para dos o más periodos consecutivos, esto es un fuerte indicador de que los errores son mayores que los que produciría un sistema más "exacto", generalmente K se toma entre cuatro y seis.

Si $IND(\alpha, n) > 6$ indica que el proceso no se está realizando correctamente y sugiere alguna corrección, en este caso existen varias posibilidades: una opción es que el modelo necesite cambios, es decir, puede adicionarse o eliminarse una o más variables, otra posibilidad es que el modelo no necesite cambios pero la constante de suavizamiento sí.

2-2.4.1 PROCESOS DE CONTROL ADAPTABLE.

A menudo se usan procesos que cambian automáticamente el valor de la constante de suavizamiento, estos se conocen como "Procesos de Control Adaptable".

Se han desarrollado varias técnicas al respecto, ya que ellas permiten que la constante de suavizamiento se adapte a cambios en los parámetros de la serie. Aquí se examinará el método elaborado por Chow², el cual introduce a parte de la constante α otros dos valores α_U y α_L tal que:

$$\alpha_U = \alpha + d$$

$$\alpha_L = \alpha - d$$

donde d es una constante positiva (generalmente se propone $d = 0.05$).

Ahora se generan tres estimaciones, una para cada α y se calcula $D(\alpha_L, n)$, $D(\alpha, n)$ y $D(\alpha_U, n)$, si sucede que:

$$D(\alpha, n) < \begin{cases} D(\alpha_L, n) \\ D(\alpha_U, n) \end{cases} \quad \text{Los errores generados al estimar con } \alpha \text{ son más pequeños que los errores generados al usar } \alpha_U \text{ y } \alpha_L \text{ por lo tanto es razonable seguir usando } \alpha \text{ para generar pronósticos.}$$

² Vease: Chow, W "Adaptive Control of the Exponential Smoothing Constant" Journal of Industrial Engineering 1965

$D(\alpha_L, n) < \begin{cases} D(\alpha, n) \\ D(\alpha_U, n) \end{cases}$ Se debe reducir α (entre α_L y α) y sumar y restar cl para fijar nuevos límites.

$D(\alpha_U, n) < \begin{cases} D(\alpha, n) \\ D(\alpha_L, n) \end{cases}$ Se debe incrementar α (entre α y α_U) y sumar y restar cl para fijar nuevos límites.

Chow reporto buenos resultados con este método, sin embargo, una desventaja, es que se deben hacer tres estimaciones³.

³ Johnson y Montgomery [8], presenta otro proceso de control adaptable.

2-2.5 EJEMPLO.

Ahora considérese el método de suavizamiento exponencial simple para hacer pronósticos de la serie de tiempo del ejemplo anterior (variación mensual del índice de empleo). Para iniciar el proceso supóngase $\alpha = 0.1$.

TABLA 2.4 Estadísticos de suavizamiento exponencial simple, con $\alpha = 0.1$ y $S_0 = -0.42923$

	t	y_t	S_t
E92	1	-0.77	-0.46372
F	2	0.39	-0.37803
M	3	0.36	-0.30401
A	4	-0.64	-0.33804
M	5	-0.36	-0.34043
J	6	-0.15	-0.32146
J	7	-0.25	-0.31390
A	8	-0.75	-0.35774
S	9	-0.63	-0.38488
O	10	-0.23	-0.36988
N	11	-0.60	-0.39262
D	12	-1.47	-0.49989
E93	13	-1.30	-0.57988
F	14	0.27	-0.49527
M	15	-0.04	-0.44976
A	16	-0.64	-0.46906
M	17	-0.10	-0.43253
J	18	-0.81	-0.46981
J	19	-0.82	-0.50483
A	20	-0.77	-0.53155
S	21	-0.93	-0.57171
O	22	-0.90	-0.60462
N	23	-0.48	-0.59225
D	24	-1.04	-0.63673
E94	25	-0.47	-0.62011
F	26	0.64	-0.49457
M	27	0.16	-0.42947
A	28	-0.02	-0.38844
M	29	0.28	-0.32197
J	30	-0.15	-0.30458
J	31	-0.38	-0.31239
A	32	-0.41	-0.32221
S	33	-0.12	-0.30229
O	34	0.13	-0.25882
N	35	-0.04	-0.23665
D	36	-0.90	-0.30306

El estimador β_0 se obtiene promediando las primeras $\frac{n}{2}$ observaciones. Así $S_0 = -0.42923$

$$\begin{aligned}
 S_1 &= \alpha y_1 + (1 - \alpha) S_0 \\
 &= (0.1)(-0.77) + (1 - 0.1)(-0.42923) \\
 &= -0.46372
 \end{aligned}$$

$$\begin{aligned}
 S_2 &= \alpha y_2 + (1 - \alpha)S_1 \\
 &= (0.1)(0.39) + (1 - 0.1)(-0.46372) \\
 &= -0.37803
 \end{aligned}$$

Continuando de esa manera se puede completar la tabla 2.4

Para determinar la constante de suavizamiento óptima se realiza un proceso de simulación. Esto es, generar distintas estimaciones de y_t para diferentes valores de $\alpha \in (0, 1)$ con incrementos de 0.02 y determinar el α que minimice la suma de errores al cuadrado.

TABLA 2.5 Suma de errores al cuadrado para diferentes valores de α

α	Suma de errores al cuadrado
0.02	8.79878
0.04	8.88340
0.06	8.93993
0.08	8.97675
0.10	9.00197
0.12	9.02164
0.14	9.03958
0.16	9.05796
0.18	9.07784
0.20	9.09962

De la tabla 2.5 se concluye que la constante α que minimiza la suma de errores al cuadrado es 0.02. Esta α es la que se utilizará para generar pronósticos de y_t .

De nuevo, el estimador β_0 se obtiene promediando las primeras $\frac{n}{2}$ observaciones. Así $S_0 = -0.42923$ y con $\alpha = 0.02$, se tiene:

$$\begin{aligned}
 S_1 &= \alpha y_1 + (1 - \alpha)S_0 \\
 &= (0.02)(-0.77) + (1 - 0.02)(-0.42923) \\
 &= -0.43613 \\
 S_2 &= \alpha y_2 + (1 - \alpha)S_1 \\
 &= (0.02)(0.39) + (1 - 0.02)(-0.43613) \\
 &= -0.41955
 \end{aligned}$$

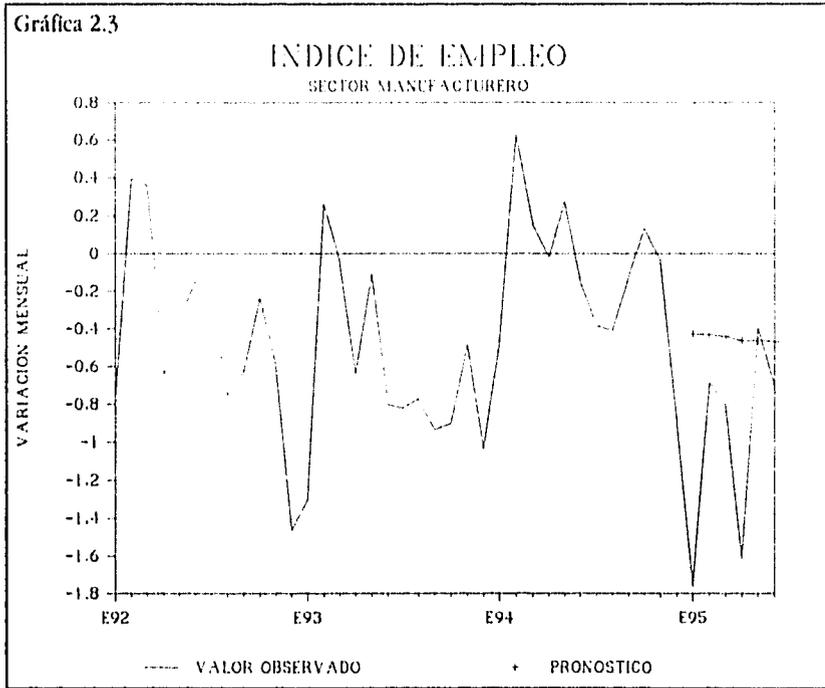
Continuando de esa manera se puede completar la tabla 2.6

TABLA 2.6 Estadístico de suavizamiento exponencial simple con $\alpha = 0.02$ y $S_0 = -0.42923$

	t	\mathcal{I}_t	S_t	S_{t-1}	$e^{-\lambda(t-1)}$	$(\alpha - \lambda(t-1))$
E92	1	-0.77	-0.43613	-0.42923	-0.54493	0.11898
F	2	0.49	-0.41955	-0.43613	0.82937	0.68785
M	3	0.36	-0.40391	-0.41955	0.78170	0.61105
A	4	-0.64	-0.40872	-0.40391	-0.24043	0.05780
M	5	-0.36	-0.40778	-0.40872	0.04677	0.00219
J	6	-0.15	-0.40264	-0.40778	0.25704	0.06607
J	7	-0.25	-0.39951	-0.40264	0.15681	0.02459
A	8	-0.75	-0.40656	-0.39951	-0.35284	0.12450
S	9	-0.65	-0.41102	-0.40656	-0.22253	0.04952
O	10	-0.23	-0.40749	-0.41102	0.17615	0.03103
N	11	-0.60	-0.41129	-0.40749	-0.18985	0.03604
D	12	-1.47	-0.43237	-0.41129	-1.05400	1.11092
E93	13	-1.30	-0.44972	-0.43237	-0.86744	0.75244
F	14	0.27	-0.43540	-0.44972	0.71595	0.51259
M	15	-0.04	-0.42749	-0.43540	0.39521	0.15619
A	16	-0.64	-0.43180	-0.42749	-0.21528	0.04634
M	17	-0.10	-0.42524	-0.43180	0.32802	0.10760
J	18	-0.81	-0.43284	-0.42524	-0.38007	0.14445
J	19	-0.82	-0.44059	-0.43284	-0.58722	0.14994
A	20	-0.77	-0.44721	-0.44059	-0.33143	0.10984
S	21	-0.93	-0.45693	-0.44721	-0.48591	0.23611
O	22	-0.90	-0.46581	-0.45693	-0.44385	0.19701
N	23	-0.48	-0.46611	-0.46581	-0.01513	0.00023
D	24	-1.04	-0.47753	-0.46611	-0.57098	0.32602
E94	25	-0.47	-0.47739	-0.47753	0.00697	0.00005
F	26	0.64	-0.45514	-0.47739	1.11273	1.23817
M	27	0.16	-0.44291	-0.45514	0.61159	0.37405
A	28	-0.02	-0.43443	-0.44291	0.42370	0.17952
M	29	0.28	-0.42022	-0.43443	0.71067	0.50506
J	30	-0.15	-0.41478	-0.42022	0.27213	0.07405
J	31	-0.38	-0.41413	-0.41478	0.03210	0.00103
A	32	-0.41	-0.41406	-0.41413	0.00350	0.00001
S	33	-0.12	-0.40824	-0.41406	0.29106	0.08471
O	34	0.13	-0.39743	-0.40824	0.54066	0.29231
N	35	-0.04	-0.39022	-0.39743	0.36033	0.12984
D	36	-0.90	-0.40043	-0.39022	-0.51055	0.26067

La gráfica 2.3 muestra la variación mensual observada del índice de empleo y sus pronósticos para los siguientes seis meses.

Gráfica 2.3



2-3 SUAVIZAMIENTO EXPONENCIAL DOBLE.

En esta sección se discutirá la estimación puntual y por intervalos de series de tiempo caracterizadas por una tendencia lineal, utilizando el método conocido como suavizamiento exponencial doble.

Ahora considérese la situación en la cual el nivel medio de la serie de tiempo cambia a través del tiempo, específicamente supongase que cambia siguiendo una función lineal. Así un modelo apropiado para esta serie puede ser

$$y_t = \beta_0 + \beta_1 t + \epsilon_t$$

ya que la expresión $\beta_0 + \beta_1 t$ indica una relación lineal entre el nivel medio y el tiempo. La pendiente de esta relación es β_1 y la ordenada al origen es β_0 . Entonces la serie de tiempo queda explicada por la tendencia implícita en esta línea recta, combinada con fluctuaciones aleatorias, las cuales provocan que las observaciones se desvien de la tendencia.

Si $\beta_1 > 0$, esto implica que el nivel medio de la serie de tiempo se incrementa conforme avanza el tiempo; si $\beta_1 < 0$, esto implica que el nivel medio de la serie de tiempo decrece conforme avanza el tiempo.

2-3.1 APROXIMACION POR SUAVIZAMIENTO EXPONENCIAL.

Como en el caso simple, sean y_1, y_2, \dots, y_{n-1} , dadas estas $n-1$ observaciones, se desea encontrar los estimadores de β_0 y β_1 . Los estimadores mínimo cuadrados de β_0 y β_1 los cuales se denotan por $\hat{\beta}_0(n-1)$ y $\hat{\beta}_1(n-1)$ respectivamente, están dados por:

$$\hat{\beta}_1(n-1) = \frac{(n-1) \sum_{t=1}^{n-1} t y_t - \left(\sum_{t=1}^{n-1} t \right) \left(\sum_{t=1}^{n-1} y_t \right)}{(n-1) \sum_{t=1}^{n-1} t^2 - \left(\sum_{t=1}^{n-1} t \right)^2}$$
$$\hat{\beta}_0(n-1) = \frac{\sum_{t=1}^{n-1} y_t}{n-1} - \hat{\beta}_1(n-1) \left(\frac{\sum_{t=1}^{n-1} t}{n-1} \right)$$

Dados estos estimadores, el pronóstico en el periodo $(n-1)$ para un valor futuro $(n-1) + \tau$, donde τ es un entero positivo, es:

$$y_{n-1+\tau}(n-1) = \hat{\beta}_0(n-1) + \hat{\beta}_1(n-1)(n-1 + \tau)$$

Ahora incorpórese una nueva observación y_n , los estimadores mínimo cuadrados $\beta_0(n)$ y $\beta_1(n)$ usando las n observaciones, se encuentran reemplazando $(n-1)$ por n en las fórmulas anteriores.

Otra aproximación usada frecuentemente para determinar $\hat{\beta}_0$ y $\hat{\beta}_1$ se conoce con el nombre de "Suavizamiento exponencial doble" [8]. Este método proporciona una estimación de β_1 en el periodo n dada por:

$$\hat{\beta}_1(n) = \frac{\alpha}{1-\alpha} (S_n - S_n^{[2]})$$

donde: S_n es el estimador de suavizamiento simple y la expresión $S_n^{[2]}$ se conoce como el estadístico de suavizamiento doble⁴ y se calcula como sigue:

$$S_n^{[2]} = \alpha S_n + (1-\alpha) S_{n-1}$$

La constante $\alpha \in (0, 1)$ es otra vez la ya mencionada constante de suavizamiento.

El proceso de suavizamiento exponencial doble ofrece un estimador de β_0 en el periodo n y esta dada por:

$$\hat{\beta}_0(n) = 2S_n - S_n^{[2]} - n\hat{\beta}_1(n)$$

sustituyendo $\hat{\beta}_1(n)$ en la ecuación anterior:

$$\hat{\beta}_0(n) = 2S_n - S_n^{[2]} - n \left[\frac{\alpha}{1-\alpha} (S_n - S_n^{[2]}) \right]$$

La obtención de las ecuaciones para $\hat{\beta}_0$ y $\hat{\beta}_1$ no son del todo intuitivas y se omite aquí su desarrollo.

Ahora supongase que se cuenta con todos los datos de la serie de tiempo, incluyendo el periodo n , la ecuación para hacer un pronóstico para $(n+\tau)$ es

$$\begin{aligned} \hat{y}_{n+\tau}(n) &= \hat{\beta}_0(n) + \hat{\beta}_1(n)(n+\tau) \\ &= [\hat{\beta}_0(n) + \hat{\beta}_1(n)n] + \hat{\beta}_1(n)\tau \\ &= \alpha_0(n) + \hat{\beta}_1(n)\tau \end{aligned}$$

⁴ Notar que $S_n^{[2]}$ no es el cuadrado de S_n , es más bien el estadístico de suavizamiento doble.

donde:

$$\alpha_0(n) = \hat{\beta}_0(n) + \hat{\beta}_1(n)n$$

Sustituyendo $\hat{\beta}_0(n)$ en la ecuación anterior se tiene:

$$\begin{aligned}\alpha_0(n) &= \hat{\beta}_0(n) + \hat{\beta}_1(n)n \\ &= [2S_n - S_n^{[2]} - n\hat{\beta}_1(n)] + \hat{\beta}_1(n)n \\ &= 2S_n - S_n^{[2]}\end{aligned}$$

Ahora sustituyendo $\alpha_0(n)$ y $\hat{\beta}_1(n)$ en $\hat{y}_{n+\tau}(n)$ se tiene:

$$\begin{aligned}\hat{y}_{n+\tau} &= \alpha_0(n) + \hat{\beta}_1(n)\tau \\ &= 2S_n - S_n^{[2]} + \frac{\alpha}{1-\alpha}(S_n - S_n^{[2]})\tau \\ &= \left(2 + \frac{\alpha\tau}{1-\alpha}\right)S_n - \left(1 + \frac{\alpha\tau}{1-\alpha}\right)S_n^{[2]}\end{aligned}$$

Para empezar el proceso de suavizamiento exponencial doble se debe disponer de los valores iniciales de S_0 y $S_0^{[2]}$. Estos estadísticos no son del todo intuitivos, la asignación directa de sus valores no es inmediata, pero se pueden encontrar aplicando una regresión a los datos y así obtener estimadores de los coeficientes β_0 y β_1 .

Se sabe que para el periodo n :

$$\hat{\beta}_1(n) = \frac{\alpha}{1-\alpha}(S_n - S_n^{[2]})$$

Así para el periodo $n = 0$ se tiene:

$$\hat{\beta}_1(0) = \frac{\alpha}{1-\alpha}(S_0 - S_0^{[2]})$$

También se sabe que para el periodo n :

$$\hat{\beta}_0(n) = 2S_n - S_n^{[2]} - n \frac{\alpha}{1-\alpha}(S_n - S_n^{[2]})$$

Así para el periodo $n = 0$ se tiene:

$$\hat{\beta}_0(0) = 2S_0 - S_0^{[2]}$$

Por lo tanto, de $\hat{\beta}_0(0)$ y $\hat{\beta}_1(0)$ se puede obtener S_0 y $S_0^{[2]}$ en términos de los estimadores iniciales $\hat{\beta}_0(0)$ y $\hat{\beta}_1(0)$

$$S_0 = \hat{\beta}_0(0) - \left(\frac{1-\alpha}{\alpha}\right)\hat{\beta}_1(0)$$

$$S_0^{[2]} = \hat{\beta}_0(0) - 2\left(\frac{1-\alpha}{\alpha}\right)\hat{\beta}_1(0)$$

Las soluciones a estas ecuaciones producen el valor inicial de los estadísticos S_0 y $S_0^{[2]}$ necesarios para el proceso de suavizamiento. Si no se dispone de los estimadores iniciales de β_0 y β_1 obviamente no se puede determinar los valores de S_0 y $S_0^{[2]}$, en tal caso es común en la práctica asignar el valor de la primera observación en la serie de tiempo a ambos estadísticos S_0 y $S_0^{[2]}$.

Dados los estimadores iniciales de S_0 y $S_0^{[2]}$ se puede entonces calcular las ecuaciones de suavizamiento:

$$S_n = \alpha y_n + (1-\alpha)S_{n-1} \quad \text{y} \quad S_n^{[2]} = \alpha S_n + (1-\alpha)S_n^{[2]}$$

para todos y cada uno de sus periodos.

Ahora ya se pueden hacer pronósticos para $(n + \tau)$, usando la siguiente ecuación:

$$\hat{y}_{n+\tau}(n) = \left(2 + \frac{\alpha\tau}{1-\alpha}\right)S_n - \left(1 + \frac{\alpha\tau}{1-\alpha}\right)S_n^{[2]}$$

Para encontrar la "mejor" constante de suavizamiento α , igual que en el caso simple, se puede hacer un proceso de simulación, esto es, generar un conjunto de pronósticos para un conjunto de α 's. El valor de α que minimice la suma de errores al cuadrado es el valor que se utilizará en el proceso de suavizamiento exponencial.

Otra vez las primeras n observaciones en la serie de tiempo (por ejemplo $6 \leq n \leq \frac{n}{2}$) se usan para obtener los estimadores iniciales de β_0 y β_1 vía análisis de regresión y entonces poder encontrar los valores iniciales de S_0 y $S_0^{[2]}$.

2-3.2 INTERVALOS DE CONFIANZA.

Considérese el siguiente modelo de regresión:

$$y_t = \beta_0 + \beta_1 f_1(t) + \beta_2 f_2(t) + \dots + \beta_p f_p(t) + \epsilon_t$$

donde: $f_1(t), f_2(t), \dots, f_p(t)$, son funciones que dependen del tiempo y no de variables de otras series de tiempo; este modelo de regresión se caracteriza porque el término:

$$\beta_0 + \beta_1 f_1(t) + \beta_2 f_2(t) + \dots + \beta_p f_p(t)$$

es una función lineal con parámetros desconocidos $\beta_0, \beta_1, \dots, \beta_p$. Más aun, hay un método llamado método general de suavizamiento directo⁵, el cual actualiza los estimadores de $\beta_0, \beta_1, \dots, \beta_p$ a la luz de nuevas observaciones obtenidas de periodo a periodo. Sin embargo, es importante mencionar que existe una teoría estadística respaldada en este método general, y se ha usado para obtener un intervalo al $(100 - \alpha)\%$ de confianza para observaciones futuras $y_{n+\tau}$ o para la suma de observaciones futuras:

$$\sum_{\tau=1}^T y_{n+\tau}$$

En particular, el método general emplea una constante de suavizamiento y supone que el origen es n y que $\beta_0, \beta_1, \dots, \beta_p$ son los estimadores de $\beta_0, \beta_1, \dots, \beta_p$. Entonces

$$\hat{y}_{n+\tau} = \hat{\beta}_0(n) + \hat{\beta}_1(n) f_1(n+\tau) + \hat{\beta}_2(n) f_2(n+\tau) + \dots + \hat{\beta}_p(n) f_p(n+\tau)$$

es la estimación de $y_{n+\tau}$. Más aún, sea:

$$\Delta(n) = \frac{\sum_{t=1}^n |y_t - \hat{y}_t(t-1)|}{n}$$

y sea $z_{\alpha/2}$ el punto en la escala de la curva normal con media 0 y varianza 1, tal que hay una área de $(100 - \alpha)\%$ bajo esta curva entre $-z_{\alpha/2}$ y $z_{\alpha/2}$. Entonces se puede mostrar⁶ que existen constantes c_1 y c_2 , las cuales no dependen de n , tal que:

⁵ Vease: Brown, R. "Smoothing, Forecasting and Prediction of Discrete Time Series" Prentice-Hall, Inc. Englewood Cliffs, New Jersey 1962

⁶ Idem

1.- El intervalo al $(100 - \alpha)\%$ de confianza para $\gamma_{n,\tau}$ esta dado por:

$$\hat{\gamma}_{n,\tau}(n) \pm z_{\alpha/2} d_{\tau} \Delta(n)$$

2.- El intervalo al $(100 - \alpha)\%$ de confianza para $\sum_{i=1}^n \gamma_{n,\tau}$ esta dado por:

$$\sum_{i=1}^n \hat{\gamma}_{n,\tau}(n) \pm z_{\alpha/2} q_{\tau} \Delta(n)$$

Las constantes d_{τ} y q_{τ} son funciones de τ y α y están dadas por las siguientes ecuaciones:

$$d_{\tau} = 1.25 \left[\frac{1 + \frac{\alpha}{(1+v)^3} [(1+4v+5v^2) + 2\alpha(1+3v)\tau + 2\alpha^2\tau^2]}{1 + \frac{\alpha}{(1+v)^3} [(1+4v+5v^2) + 2\alpha(1+3v) + 2\alpha^2]} \right]^{1/2}$$

$$q_{\tau} = 1.25 \left[\frac{\tau + \frac{\alpha\tau^2}{2(1+v)^3} [5(1+2v+\tau^2) + 4(1-\tau^2)\tau + \alpha^2\tau^2]}{1 + \frac{\alpha}{(1+v)^3} [(1+4v+5v^2) + 2\alpha(1+3v) + 2\alpha^2]} \right]^{1/2}$$

donde α es la constante de suavizamiento empleada en el proceso y $v = 1 - \alpha$.

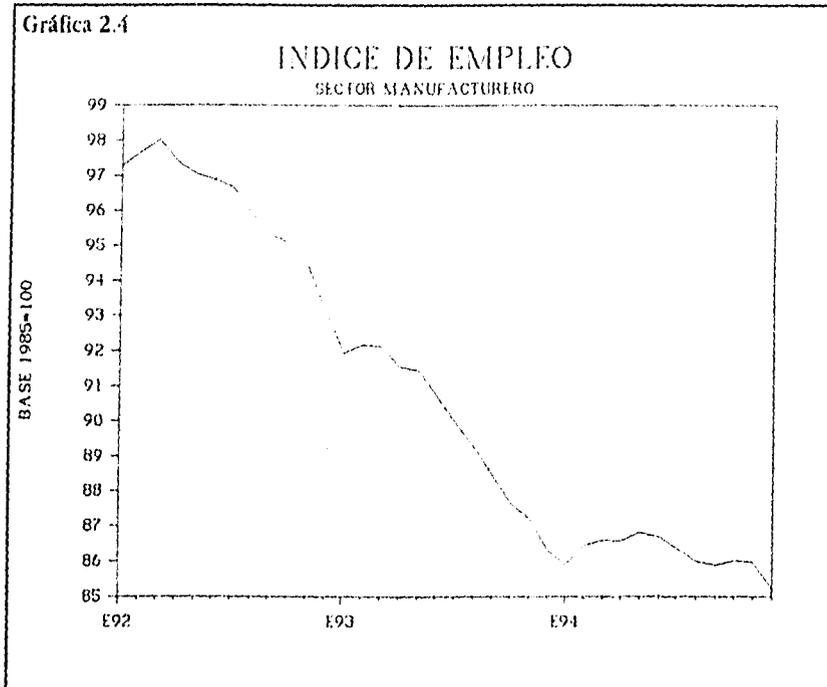
Usando la teoría estadística que hay detrás del método general de suavizamiento directo, se puede mostrar que para el proceso exponencial simple:

$$d_{\tau} = 1.25$$

$$q_{\tau} = 1.25 \left[\frac{\alpha(\tau^2 - \tau) + 2\tau}{2} \right]^{1/2}$$

2-3.3 EJEMPLO.

Ahora se intenta pronosticar el índice de empleo del sector manufacturero, gráfica 2.4 (no la variación mensual como en los ejemplos anteriores), utilizando el método de suavizamiento exponencial doble. El primer paso es la selección de la constante de suavizamiento α , esta selección se puede hacer mediante un proceso de simulación con los valores del índice de la tabla 2.1 y para distintos valores de α ; el valor de α que minimice la suma de errores al cuadrado será la seleccionada para llevar a cabo el método de suavizamiento.



Para empezar el proceso de simulación se deben estimar primero los parámetros de β_0 y β_1 en el modelo:

$$y_t = \beta_0 + \beta_1 t + \epsilon_t$$

por medio del análisis de regresión y empleando las primeras $\frac{3}{2}$ observaciones del índice de empleo manufacturero. Se encuentra que los estimadores mínimo cuadrados de β_0 y β_1 son:

$$\hat{\beta}_0 = 99.09855 \quad \text{y} \quad \hat{\beta}_1 = -0.46133$$

Ahora estos estimadores se pueden usar para encontrar los valores iniciales de los estadísticos de suavizamiento S_0 y $S_0^{(2)}$ y de esta manera empezar el proceso de simulación.

Se ilustrarán los cálculos comprendidos en el proceso de simulación considerando $\alpha = 0.1$

$$\begin{aligned} S_0 &= \hat{\beta}_0(0) - \left(\frac{1-\alpha}{\alpha}\right)\hat{\beta}_1(0) \\ &= 99.09855 - \left(\frac{1-0.1}{.01}\right)(-0.46133) \\ &= 103.2505 \end{aligned}$$

$$\begin{aligned} S_0^{(2)} &= \hat{\beta}_0(0) - 2\left(\frac{1-\alpha}{\alpha}\right)\hat{\beta}_1(0) \\ &= 99.09855 - 2\left(\frac{1-0.1}{0.1}\right)(-0.46133) \\ &= 107.4026 \end{aligned}$$

Dados estos estimadores iniciales se puede encontrar la estimación para enero de 1992 (período 1), calculada en el período 0, usando la ecuación:

$$\begin{aligned} \hat{y}_{t,t}(t) &= \left(2 + \frac{\alpha t}{1-\alpha}\right)S_t - \left(1 + \frac{\alpha t}{1-\alpha}\right)S_t^{(2)} \\ \hat{y}_{0,1}(0) &= \left(2 + \frac{\alpha(1)}{1-\alpha}\right)S_0 - \left(1 + \frac{\alpha(1)}{1-\alpha}\right)S_0^{(2)} \\ \hat{y}_1(0) &= \left(2 + \frac{0.1(1)}{1-0.1}\right)(103.2505) - \left(1 + \frac{0.1(1)}{1-0.1}\right)(107.4026) \\ &= 98.63721 \end{aligned}$$

Con estos estimadores iniciales se obtienen los siguientes valores de los estadísticos de suavizamiento:

$$\begin{aligned} S_1 &= \alpha \gamma_1 + (1 - \alpha) S_0 \\ &= (0.1)(97.3) + (1 - 0.1)(103.2505) \\ &= 102.6536 \end{aligned}$$

$$\begin{aligned} S_1^{[2]} &= \alpha S_1 + (1 - \alpha) S_0^{[2]} \\ &= (0.1)(102.6536) + (1 - 0.1)(107.1026) \\ &= 106.9277 \end{aligned}$$

Ahora se puede obtener la estimación para febrero de 1992 (periodo 2), calculada en el periodo 1, usando la ecuación:

$$\begin{aligned} \hat{y}_{1-1}(1) &= \left(2 + \frac{\alpha(1)}{1 - \alpha}\right) S_1 - \left(1 + \frac{\alpha(1)}{1 - \alpha}\right) S_1^{[2]} \\ \hat{y}_2(1) &= \left(2 + \frac{0.1(1)}{1 - 0.1}\right) (102.6536) - \left(1 + \frac{0.1(1)}{1 - 0.1}\right) (106.9276) \\ &= 97.90477 \end{aligned}$$

De la misma forma se obtiene la estimación para marzo de 1992 (periodo 3), calculada en el periodo 2

$$\begin{aligned} S_2 &= \alpha \gamma_2 + (1 - \alpha) S_1 \\ &= (0.1)(97.7) + (1 - 0.1)(102.6536) \\ &= 102.1547 \end{aligned}$$

$$\begin{aligned} S_2^{[2]} &= \alpha S_2 + (1 - \alpha) S_1^{[2]} \\ &= (0.1)(102.1547) + (1 - 0.1)(106.9277) \\ &= 106.4504 \end{aligned}$$

y

$$\hat{y}_{2,t}(2) = \left(2 + \frac{\alpha(1)}{1-\alpha}\right) S_t - \left(1 + \frac{\alpha(1)}{1-\alpha}\right) S_t^{(2)}$$

$$\hat{y}_3(2) = \left(2 + \frac{0.1(1)}{1-0.1}\right) (102.4547) - \left(1 + \frac{0.1(1)}{1-0.1}\right) (106.4501)$$

$$= 97.38177$$

Continuando de la misma manera se generan las estimaciones para todo el periodo, así como el error y el error al cuadrado para cada mes. Todos esos resultados se presentan en la tabla 2.7.

Tabla 2.7 Proceso de simulación para el índice de empleo manufacturero, usando el método de suavizamiento exponencial doble con $\alpha = 0.1$

	t	y_t	S_t	$S_t^{(2)}$	$\hat{y}_{t-1}(t)$	$(\hat{y}_t - \hat{y}_t(t-1))$	$(\hat{y}_t - \hat{y}_t(t-1))^2$
E92	1	97.28	102.65	106.93	98.64	1.36	1.84
F	2	97.66	102.15	106.45	97.90	0.24	0.06
M	3	98.02	101.74	105.98	97.38	-0.64	0.40
A	4	97.39	101.31	105.51	97.03	-0.35	0.13
M	5	97.03	100.88	105.05	96.63	-0.40	0.16
J	6	96.89	100.48	104.59	96.24	-0.64	0.41
J	7	96.65	100.10	104.14	95.91	-0.74	0.55
A	8	95.92	99.68	103.70	95.60	-0.32	0.10
S	9	95.32	99.24	103.25	95.22	-0.10	0.01
O	10	95.09	98.83	102.81	94.79	-0.31	0.09
N	11	94.53	98.40	102.37	94.41	-0.12	0.01
D	12	93.14	97.87	101.92	93.99	0.85	0.72
E93	13	91.93	97.28	101.45	93.38	1.45	2.09
F	14	92.18	96.77	100.99	92.64	0.46	0.21
M	15	92.14	96.31	100.52	92.08	-0.06	0.00
A	16	91.55	95.83	100.05	91.62	0.08	0.01
M	17	91.45	95.39	99.58	91.14	-0.31	0.10
J	18	90.72	94.92	99.12	90.73	0.02	0.00
J	19	89.97	94.43	98.65	90.26	0.29	0.09
A	20	89.28	93.91	98.17	89.74	0.46	0.21
S	21	88.44	93.37	97.69	89.18	0.74	0.54
O	22	87.65	92.79	97.20	88.56	0.91	0.83
N	23	87.23	92.24	96.71	87.90	0.67	0.45
D	24	86.52	91.65	96.20	87.27	0.95	0.90
E94	25	85.91	91.07	95.69	86.58	0.67	0.45
F	26	86.46	90.61	95.18	85.94	-0.52	0.27
M	27	86.60	90.21	94.68	85.53	-1.06	1.13
A	28	86.58	89.85	94.20	85.24	-1.34	1.79
M	29	86.82	89.54	93.73	85.01	-1.81	3.27
J	30	86.69	89.26	93.29	84.89	-1.80	3.25
J	31	86.36	88.97	92.85	84.78	-1.58	2.48
A	32	86.00	88.67	92.44	84.65	-1.35	1.83
S	33	85.90	88.39	92.03	84.49	-1.41	1.98
O	34	86.01	88.16	91.64	84.35	-1.66	2.75
N	35	85.98	87.94	91.27	84.28	-1.70	2.89
D	36	85.20	87.67	90.91	84.23	-0.97	0.95

De igual forma se simulan estimaciones para distintos valores de α . Cabe hacer mención que $\beta_0 = 99.09855$ y $\beta_1 = -0.46133$ son los mismos para cada simulación, aunque S_0 y $S_0^{(2)}$ difieren en cada caso ya que ellos dependen de α . La suma de errores al cuadrado

para todas las simulaciones se presentan en la tabla 2.8, en la cual se aprecia que el valor de la constante α que minimiza la suma de errores al cuadrado es 0.57, por lo tanto este valor es el seleccionado para hacer pronósticos.

TABLA 2.8 Suma de errores al cuadrado para diferentes valores de α

α	Suma de errores al cuadrado
0.10	32.95
0.20	18.74
0.30	13.96
0.40	12.11
0.50	11.35
0.57	11.22
0.60	11.24
0.70	11.61
0.80	12.34
0.90	13.41

Ahora se procede de la misma manera que en el proceso de simulación, solo que fijando $\alpha = 0.57$ y utilizando los mismos estimadores mínimo cuadrados $\hat{\beta}_0 = 99.09855$ y $\hat{\beta}_1 = -0.46133$. Los resultados se presentan en la tabla 2.9, tal que para diciembre de 1994 se observa que $S_{36} = 85.54570$ y $S_{36}^{[2]} = 85.75704$

TABLA 2.9 Estadísticos de suavizamiento exponencial doble con $\alpha = 0.57$ en la serie del índice de empleo.

	t	y_t	S_t	$S_t^{(2)}$	$\hat{y}_{t+1}(t)$
E92	1	97.28	98.21	98.89	98.64
F	2	97.66	97.90	98.43	96.63
M	3	98.02	97.97	98.12	96.91
A	4	97.39	97.64	97.85	97.61
M	5	97.03	97.29	97.53	97.15
J	6	96.89	97.06	97.26	96.74
J	7	96.65	96.83	97.01	96.59
A	8	95.92	96.31	96.61	96.39
S	9	95.32	95.75	96.12	95.61
O	10	95.09	95.37	95.69	94.88
N	11	94.53	94.89	95.24	94.63
D	12	93.14	93.89	94.47	94.09
E93	13	91.93	92.78	93.50	92.55
F	14	92.18	92.43	92.89	91.08
M	15	92.14	92.27	92.94	91.36
A	16	91.55	91.86	92.15	91.64
M	17	91.45	91.63	91.85	91.18
J	18	90.72	91.11	91.43	91.10
J	19	89.97	90.46	90.88	90.36
A	20	89.28	89.79	90.25	89.49
S	21	88.44	89.02	89.55	88.70
O	22	87.65	88.24	88.80	87.79
N	23	87.23	87.66	88.15	86.92
D	24	86.32	86.90	87.44	86.52
E94	25	85.91	86.34	86.81	85.64
F	26	86.46	86.41	86.58	85.24
M	27	86.60	86.51	86.54	86.00
A	28	86.58	86.35	86.55	86.43
M	29	86.82	86.70	86.64	86.56
J	30	86.69	86.70	86.67	86.86
J	31	86.36	86.50	86.57	86.75
A	32	86.00	86.22	86.37	86.34
S	33	85.90	86.04	86.18	85.86
O	34	86.01	86.02	86.09	85.70
N	35	85.98	86.00	86.01	85.86
D	36	85.20	85.55	85.76	85.91

Esas cantidades se pueden usar para generar pronósticos de la serie. El pronóstico para enero de 1995 (periodo 37), hecho en el periodo 36 es:

$$\hat{y}_{t+1}(t) = \left(2 + \frac{\alpha\tau}{1-\alpha}\right)S_t - \left(1 + \frac{\alpha\tau}{1-\alpha}\right)S_t^2$$

$$\hat{y}_{36+1}(36) = \left(2 + \frac{\alpha(1)}{1-\alpha}\right)S_{36} - \left(1 + \frac{\alpha(1)}{1-\alpha}\right)S_{36}^2$$

$$\hat{y}_{37}(36) = \left(2 + \frac{0.57(1)}{1-0.57}\right)(85.54570) - \left(1 + \frac{0.57(1)}{1-0.57}\right)(85.75704)$$

$$= 85.05423$$

Ahora supongase que en enero de 1995 el índice de empleo manufacturero observado fue 83.7, para encontrar el pronóstico para febrero de 1995 primero se debe calcular

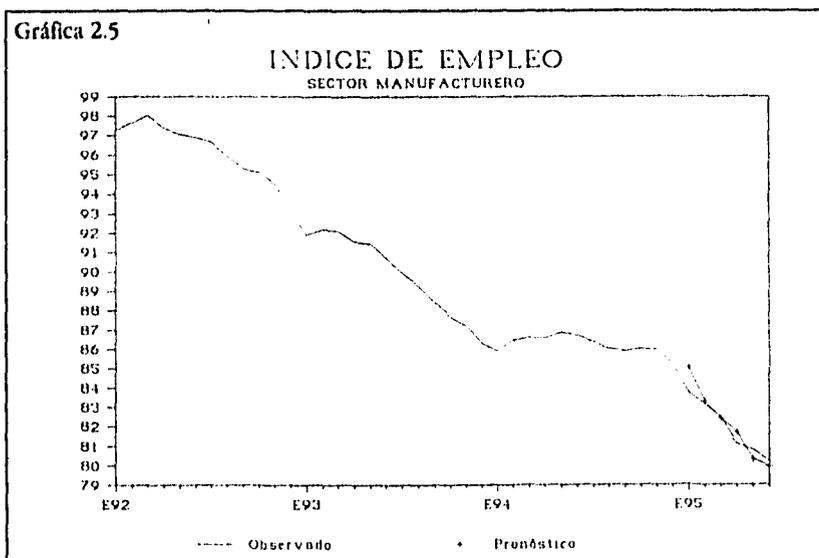
$$\begin{aligned}
 S_{37} &= \alpha \gamma_{37} + (1 - \alpha) S_{36} \\
 &= (0.57)(83.7) + (1 - 0.57)(85.54570) \\
 &= 84.49528
 \end{aligned}$$

$$\begin{aligned}
 S_{37}^{[2]} &= \alpha S_{37} + (1 - \alpha) S_{36}^{[2]} \\
 &= (0.57)(84.49528) + (1 - 0.57)(85.75704) \\
 &= 85.03783
 \end{aligned}$$

Por lo tanto el pronóstico para febrero de 1995 (periodo 38), calculado en el periodo 37 es:

$$\hat{y}_{37+1}(37) = \left(2 + \frac{\alpha(1)}{1-\alpha} \right) S_{37} - \left(1 + \frac{\alpha(1)}{1-\alpha} \right) S_{37}^2$$

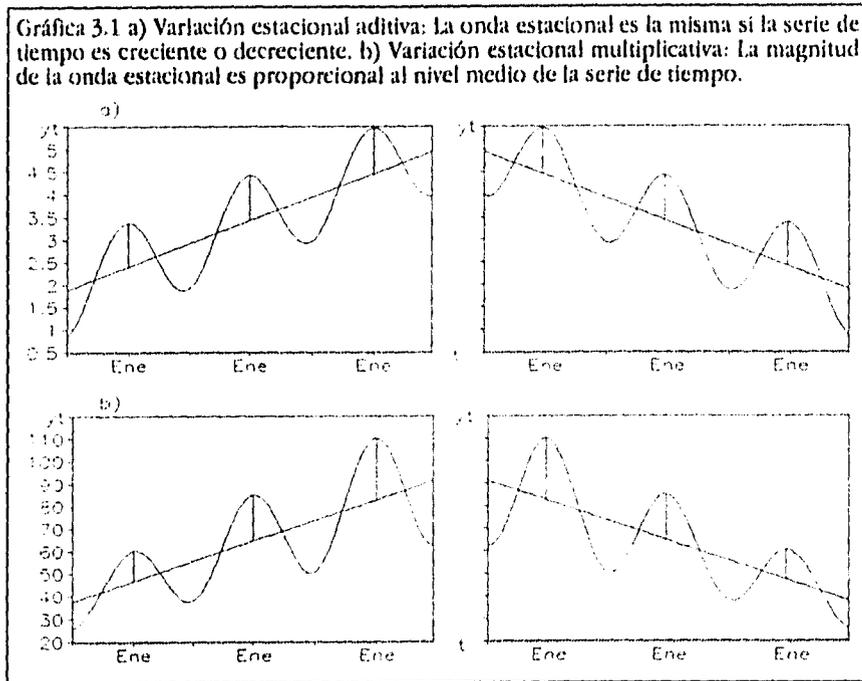
$$\hat{y}_{38}(37) = 83.23352$$



CAPITULO 3.

SERIES DE TIEMPO ESTACIONALES.

En este capítulo se estudiarán series de tiempo estacionales. Se definen dos tipos de variación estacional. El primer tipo es la variación estacional aditiva: si una serie de tiempo exhibe variación estacional aditiva, la magnitud de la onda estacional de la serie de tiempo es independiente del nivel medio determinado por la tendencia. El segundo tipo es la variación estacional multiplicativa: si una serie de tiempo exhibe variación estacional multiplicativa, la magnitud de la onda estacional de la serie de tiempo es proporcional al nivel medio determinado por la tendencia. Así, si el nivel medio de la serie de tiempo se esta incrementando, también la magnitud de la onda estacional, mientras que si el nivel medio de la serie de tiempo esta decreciendo, también decrece la magnitud de la onda estacional.



Muy pocas series de tiempo poseen variación estacional que sean precisamente aditivas o precisamente multiplicativas. Sin embargo, se debe intentar clasificar correctamente cada

serie de tiempo estacional, para entonces escoger un modelo apropiado y así poder analizar la serie de tiempo. Los modelos de series de tiempo estudiados en este capítulo son casos especiales del modelo:

$$Y_t = f(T_t, S_t) + \epsilon_t$$

donde: Y_t es el valor de la serie de tiempo al tiempo t .

T_t es el factor de la tendencia de la serie de tiempo al tiempo t .

S_t es el factor estacional de la serie de tiempo al tiempo t .

ϵ_t es el factor irregular de la serie de tiempo al tiempo t .

f es una función que relaciona el valor observado de la serie de tiempo con la tendencia y el factor estacional. Puede tomar una de las siguientes formas:
 $f(T, S) = T + S$ o $f(T, S) = T \times S$.

Se estudiará el método de descomposición multiplicativo y el método llamado Censo II, para lo cual será necesario introducir antes el concepto de media móvil y por último el método Winters, el cual es una extensión del proceso de suavizamiento exponencial para pronosticar series de tiempo con variación estacional.

3-1 MEDIAS MOVILES.

Las medias móviles se usan para eliminar estacionalidad y aleatoriedad en una serie de tiempo, en este sentido las medias móviles son la columna vertebral de los métodos de descomposición. El propósito de esta sección es revisar los diferentes tipos de medias móviles en general y las usadas en el método Censo II, el cual se describirá más adelante.

Una media móvil se calcula sumando las observaciones para un número de periodos en la serie de tiempo, dividiendo el resultado entre el número de periodos incluidos en la suma.

En notación, una media móvil de ancho m , denotada por $MA_t(m)$ (moving average), para un conjunto de observaciones $z_1, \dots, z_t, \dots, z_n$ se define como:

$$MA_t(m) = \sum_{j=-m'}^{m'} \frac{z_{t+j}}{2m'+1} = \sum_{j=-m'}^{m'} \frac{z_{t+j}}{m} \quad \text{con} \quad m' = \frac{m-1}{2}$$

Así se puede generar un conjunto de medias móviles a través del tiempo. Hay que notar, que para n observaciones quedarán solo $n - m + 1$ observaciones, es decir, se perderán $(m - 1) / 2$ medias por cada extremo de la serie.

3-1.1 MEDIAS MOVILES CENTRADAS.

Si las medias móviles originales se han basado en un número impar m de observaciones, la primer media móvil corresponde al periodo intermedio de esas m observaciones. Por el contrario, si la media móvil original se ha basado en un número par de observaciones, la primer media móvil corresponderá justamente al punto intermedio entre las dos observaciones centrales. Para definir a que periodo corresponde la primer media será necesario calcular una media móvil centrada, la cual resulta de calcular una $MA(2)$ a las medias móviles calculadas previamente. Ya que la mayoría de las series de tiempo económicas son mensuales o trimestrales, generalmente, el proceso de centrado será necesario.

3-1.2 MEDIA MOVIL $MA(3 \times 3)$.

Una media móvil de (3×3) , es una $MA(3)$ de una $MA(3)$. Esto se obtiene calculando una $MA(3)$ a los datos originales, como se muestra en las siguientes ecuaciones:

$$MA_2(3) = (z_1 + z_2 + z_3)/3$$

$$MA_3(3) = (z_2 + z_3 + z_4)/3$$

$$MA_4(3) = (z_3 + z_4 + z_5)/3$$

$$MA_5(3) = (z_4 + z_5 + z_6)/3$$

etc.

y a esta nueva serie se le calcula otra $MA(3)$, esto es:

$$MA'_3(3) = (MA_2(3) + MA_3(3) + MA_4(3))/3$$

sustituyendo se tiene:

$$MA'_3(3) = \left(\frac{z_1 + z_2 + z_3}{3} + \frac{z_2 + z_3 + z_4}{3} + \frac{z_3 + z_4 + z_5}{3} \right) / 3$$

$$MA'_3(3) = \frac{1}{9} (z_1 + 2z_2 + 3z_3 + 2z_4 + z_5)$$

Esta ecuación es equivalente a una $MA(5)$ ponderada con un peso de $\frac{1}{9}, \frac{2}{9}, \frac{3}{9}, \frac{2}{9}, \frac{1}{9}$ para el primero, segundo, tercero, cuarto y quinto término respectivamente.

De similar manera se puede obtener una $MA(3 \times 5)$ excepto que el número de términos promediados la primera vez es 5 y después 3. Esto da como resultado una $MA(7)$ ponderada, con un peso de $\frac{1}{15}, \frac{2}{15}, \frac{3}{15}, \frac{3}{15}, \frac{3}{15}, \frac{2}{15}, \frac{1}{15}$ para el primero, segundo, tercero, cuarto, quinto, sexto y séptimo término respectivamente.

3-1.3 MEDIAS MOVILES DE SPENCER Y HENDERSON.

Las medias móviles de Spencer y Henderson [10] son de orden mayor que las ya examinadas. La media móvil de Spencer es una $MA(5 \times 5 \times 4 \times 4)$ de orden cuatro, o equivalentemente es una $MA(15)$ ponderada y se obtiene calculando primero una $MA(4)$ a los datos originales:

$$MA_{2,5}(4) = (z_1 + z_2 + z_3 + z_4) / 4$$

$$MA_{3,5}(4) = (z_2 + z_3 + z_4 + z_5) / 4$$

$$MA_{4,5}(4) = (z_3 + z_4 + z_5 + z_6) / 4$$

$$MA_{5,5}(4) = (z_4 + z_5 + z_6 + z_7) / 4$$

etc.

a esta nueva serie se le calcula otra $MA(4)$, esto es:

$$MA'_{4}(4) = (MA_{2,5}(4) + MA_{3,5}(4) + MA_{4,5}(4) + MA_{5,5}(4)) / 4$$

sustituyendo se tiene:

$$MA'_{4}(4) = (z_1 + 2z_2 + 3z_3 + 4z_4 + 3z_5 + 2z_6 + z_7) / 16$$

una $MA(7)$ ponderada. Ahora usando esta media se le calcula una $MA(5)$, lo cual da como resultado:

$$MA''_{6}(5) = (MA'_{4}(4) + MA'_{5}(4) + MA'_{6}(4) + MA'_{7}(4) + MA'_{8}(4)) / 5$$

sustituyendo se tiene:

$$MA''_{6}(5) = (z_1 + 3z_2 + 6z_3 + 10z_4 + 13z_5 + 14z_6 + 13z_7 + 10z_8 + 6z_9 + 3z_{10} + z_{11}) / 80$$

una $MA(11)$ ponderada. Finalmente se aplica una $MA(5)$ a esta última media, lo cual da como resultado:

$$MA'''_{8}(5) = \frac{3}{4} (-MA''_{6}(5) + MA''_{7}(5) + MA''_{8}(5) + MA''_{9}(5) - MA''_{10}(5))$$

sustituyendo se tiene después de varios pasos algebraicos:

$$MA'''_{8}(5) = (-3z_1 - 6z_2 - 5z_3 + 3z_4 + 21z_5 + 46z_6 + 67z_7 + 74z_8 + 67z_9 + 46z_{10} + 21z_{11} + 3z_{12} - 5z_{13} - 6z_{14} - 3z_{15}) / 320$$

Por supuesto, los datos serán diferentes para medias móviles centradas en diferentes periodos. Por ejemplo:

$$MA'''_{9}(5) = (-3z_2 - 6z_3 - 5z_4 + \dots - 6z_{15} - 3z_{16}) / 320$$

Para series con n observaciones, $M.A.(15)$, se puede calcular para los periodos 8 a $n - 7$ y cada media incluye 15 términos.

Otra media móvil de Spencer comúnmente usada es la $M.A.(21)$ ponderada, la cual se puede obtener de manera similar. En adición a la media móvil de Spencer, la versión más nueva del Método Censo II también usa medias móviles ponderadas, llamadas de Henderson. Las hay de 5, 9, 13 y 25 términos. La selección de la media móvil específica se basa en la aleatoriedad presente en la serie (a mayor aleatoriedad, mayor número de términos en el promedio).

3-2 METODO DE DESCOMPOSICION MULTIPLICATIVO.

En esta sección se analizarán series de tiempo que tienen variación estacional multiplicativa. Ya que la variación estacional multiplicativa implica que la magnitud de la onda estacional de la serie de tiempo es proporcional al nivel medio determinado por la tendencia, se sigue que un modelo razonable para analizar una serie de tiempo con estas características es:

$$y_t = T \times S + \epsilon_t$$

donde el significado de la notación ya ha sido previamente discutido.

3-2.1 ESTIMACION DE LAS COMPONENTES.

El método de descomposición multiplicativo¹, desarrollado en los años veintes, fue por muchos años el método de descomposición más usado, además que es la base del Método de Descomposición Censo II, el cual se examinará más adelante. Este método supone un modelo multiplicativo de la forma:

$$y_t = S_t \times T_t \times C_t \times I_t$$

El cual, primero aísla la tendencia-ciclo de los datos, calculando una media móvil, cuyo número de términos sea igual al número de periodos en el año. Una media móvil de este ancho no contiene efectos estacionales y además cancela muy bien los efectos del factor irregular, por lo tanto es una buena estimación de la tendencia-ciclo y se representa por:

$$MA_t(m) = T_t \times C_t$$

donde m es el número de periodos en el año. Por otro lado, ya que:

$$y_t = S_t \times T_t \times C_t \times I_t$$

y

$$\hat{S}_t \times I_t = \frac{\hat{S}_t \times \hat{T}_t \times \hat{C}_t \times I_t}{\hat{T}_t \times \hat{C}_t}$$

la estimación de $S_t \times I_t$ se obtiene usando la fórmula:

$$\hat{S}_t \times I_t = \frac{y_t}{\hat{T}_t \times \hat{C}_t}$$

¹ Este método también es llamado método de descomposición clásico a proporción de medias móviles [9]

El siguiente paso es eliminar la aleatoriedad en $\hat{S}_t \times I_t$. Para esto se agrupan los valores de $\hat{S}_t \times I_t$ para los mismos periodos en toda la serie. Es decir, si la serie es mensual se agrupan enero con enero, febrero con febrero etc., si la serie es trimestral primeros trimestres con primeros trimestres. En seguida se promedian esos valores para cada agrupación. Con este proceso se espera remover la influencia del factor irregular, así las medias calculadas representan solamente los efectos del factor estacional. Se denotan estas medias como:

$$\bar{S}_t$$

Estos factores estacionales son ahora normalizados de tal manera que sumen el número de periodos en el año, sea L tal número. Esta normalización se realiza multiplicando cada valor de \bar{S}_t por la cantidad:

$$\frac{L}{\sum_{i=1}^L \bar{S}_t}$$

Así, la estimación de S_t , esta dada por:

$$\hat{S}_t = \bar{S}_t \left[\frac{L}{\sum_{i=1}^L \bar{S}_t} \right]$$

Notar que el proceso de normalización asegura que la suma de los factores estacionales suman L . Esto es porque:

$$\sum_{i=1}^L \hat{S}_t = \sum_{i=1}^L \bar{S}_t \left[\frac{L}{\sum_{i=1}^L \bar{S}_t} \right] = L \left[\frac{\sum_{i=1}^L \bar{S}_t}{\sum_{i=1}^L \bar{S}_t} \right] = L.$$

Ya que se han calculado los factores estacionales, se puede obtener una estimación de la tendencia. Esto se hace calculando primero los datos desestacionalizados; la observación desestacionalizada d_t se calcula:

$$d_t = \frac{y_t}{\hat{S}_t}$$

Esto se hace para eliminar los efectos estacionales, así, si a una observación le corresponde un factor estacional menor que uno, la observación desestacionalizada es mayor que la observación original, ($d_t > y_t$) y por lo tanto d_t se aproxima más al valor de la tendencia. De otro modo, si a una observación le corresponde un factor estacional mayor que uno, la observación desestacionalizada es menor que la observación original y otra vez d_t se aproxima más al valor de la tendencia.

Para encontrar el estimador de T_t , se parte de una inspección visual de los datos desestacionalizados y así poder ajustar la tendencia a una función constante, lineal o cuadrática según sea el caso ($T = \hat{\beta}_0$, $T = \hat{\beta}_0 + \hat{\beta}_1 t$, $T = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2$) y se calcula una regresión para estimar T_t usando el modelo deseado.

Cabe notar que esta regresión se puede realizar usando los datos de la serie original para la variable dependiente, sin embargo, este proceso puede distorsionar la estimación de la componente estacional.

El cálculo de los factores cíclicos es posible ya que se han encontrado los factores de tendencia y estacionalidad. Se sabe que:

$$y_t = T_t \times S_t \times C_t \times I_t$$

y

$$\hat{C}_t \times I_t = \frac{\hat{T}_t \times \hat{S}_t \times \hat{C}_t \times I_t}{\hat{T}_t \times \hat{S}_t}$$

entonces se obtiene el estimador $\hat{C}_t \times I_t$ usando la ecuación

$$\hat{C}_t \times I_t = \frac{y_t}{\hat{T}_t \times \hat{S}_t}$$

En seguida, se está en condiciones de calcular el factor cíclico, esto se puede hacer promediando los valores de $\hat{C}_t \times I_t$. Se puede usar una $MA(3)$, ya que los movimientos irregulares raramente persisten más de dos o tres meses y con una $MA(3)$ se espera remueva esa influencia, además que no necesita centrarse.

Ya que se ha encontrado la estimación del factor cíclico, se puede obtener la estimación de los movimientos irregulares en la serie de tiempo. Esto se hace usando la fórmula:

$$I_t = \frac{\hat{C}_t \times \hat{I}_t}{\hat{C}_t}$$

3-2.2 PRONOSTICOS E INTERVALOS DE CONFIANZA.

El uso más común del método de descomposición incluye la generación de pronósticos usando solamente los factores de estacionalidad y de tendencia. Se limita únicamente a esas dos componentes porque las otras (la variación cíclica y los movimientos irregulares) no se pueden predecir con mucha precisión. Sin embargo, la variación cíclica en las series de tiempo se debería obtener de la misma forma que se obtuvieron los otros factores, pero en estos procesos aunque teóricamente existe su implementación, a menudo es difícil su uso, por varias razones. Primero, se puede usar solo si el ciclo está bien definido y si se repite varias veces con una duración constante, segundo, aun, si existe cada ciclo, no es fácil analizarlos por la escasez de datos. Para conseguir estimaciones más exactas deben existir varios ciclos, (por definición los ciclos tienen una duración de 2 a 10 años), por lo tanto, se requiere la información de más de 30 años para estimar este factor. Muchas veces no se tiene toda la información, por esa razón la variación cíclica en la serie de tiempo no se puede estimar.

Habiendo obtenido las estimaciones de T_t y S_t y suponiendo que los factores estacionales permanecen constantes para todos los periodos, la estimación de y_t es:

$$\hat{y}_t = \hat{T}_t \times \hat{S}_t$$

Ahora se definirá un intervalo de confianza para y_t . Ya que no hay teoría estadística detrás del método de descomposición, por medio del cual se obtuvo las estimaciones de T_t y S_t , aquí no hay intervalos de confianza para y_t estadísticamente correctos. Sin embargo, existe un método para encontrar intervalos de confianza para y_t , el cual trabaja razonablemente bien:

$$\hat{y}_t \pm INT \quad \text{con} \quad INT = t_{\left(\frac{\alpha}{2}, n-(p+1)\right)} \cdot \hat{\sigma} \cdot f_t$$

donde INT se obtiene al calcular un intervalo de confianza para la observación desestacionalizada d_t , suponiendo que d_t está descrita por el modelo de regresión:

$$d_t = T + e_t$$

donde T puede tomar cualquiera de estas formas: $T = \hat{\beta}_0$, $T = \hat{\beta}_0 + \hat{\beta}_1 t$, $T = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2$.

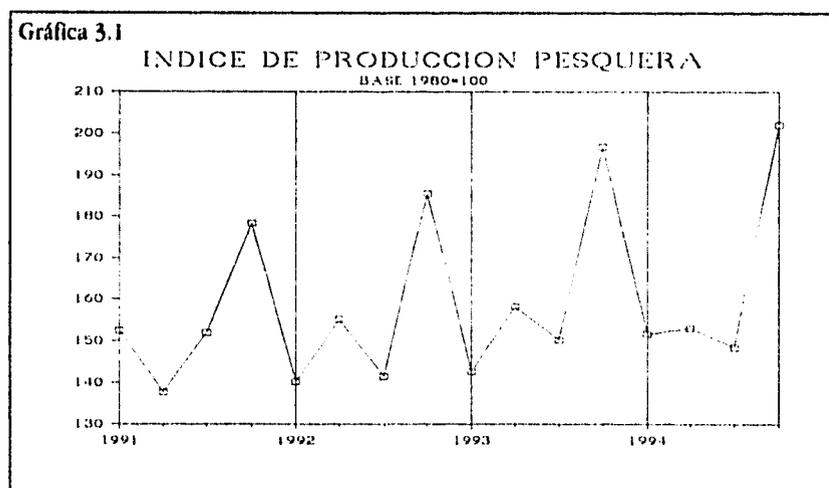
3-2.3 EJEMPLO.

Usando el método de descomposición multiplicativa se intentará analizar la serie de tiempo trimestral del índice de volumen de la producción pesquera, base 1980=100 de la tabla 3.1. Esta serie indiscutiblemente tiene un comportamiento estacional pues, al inicio del cuarto trimestre de cada año inicia la captura de camarón, reflejándose un notable incremento en ese trimestre y para el segundo trimestre aumenta la captura para hacer frente al consumo de semana santa.

Tabla 3.1 Índice de volumen de la producción pesquera. Base 1980 = 100

Año	Trim.	γ_t	Año	Trim.	γ_t
1991	1	152 4000	1993	1	142 5834
	2	137 7000		2	158 2917
	3	152 0000		3	150 2857
	4	178 4000		4	196 8537
1992	1	140 2000	1994	1	151 7087
	2	155 1880		2	153 0680
	3	141 5120		3	148 4823
	4	185 5360		4	202 1687

Fuente: Banco de México.



Para aislar la tendencia-ciclo, se debe calcular una media móvil de ancho 4 (por ser la serie trimestral) y posteriormente centrarla. Los resultados se muestran en la tabla 3.2

Tabla 3.2 Cálculo de las medias móviles.

Año	Trim.	y_t	MA(2x4)		$\hat{S}_t \times \hat{I}_t$
			MA(4)	$= \hat{I}_t \times \hat{C}_t$	
				$= \frac{y_t}{\hat{I}_t \times \hat{C}_t}$	
1991	1	152 4000			
	2	137 7000	155 125		
	3	152 0000	152 075	153 6000	0 9896
	4	178 4000	156 447	154 2610	1 1565
1992	1	140 2000	153 825	155 1560	0 9057
	2	155 1880	155 609	154 7170	1 0030
	3	141 5120	156 205	155 9069	0 9077
	4	185 5360	156 981	156 5928	1 1848
1993	1	142 5834	159 174	158 0775	0 9020
	2	158 2917	162 004	160 5889	0 9857
	3	150 2857	164 285	163 1443	0 9212
	4	196 8537	162 979	163 6320	1 2030
1994	1	151 7087	162 528	162 7536	0 9321
	2	153 0680	163 857	163 1926	0 9380
	3	148 4823			
	4	202 1687			

Con el fin de eliminar la aleatoriedad se agrupan los valores de $\hat{S}_t \times \hat{I}_t$ como se muestra abajo:

1er. Trim.	2o. Trim.	3er. Trim.	4o. Trim.
0 904	1 003	0 990	1 156
0 902	0 986	0 908	1 185
0 932	0 938	0 921	1 203
$\bar{S}_1 = 0.913$	$\bar{S}_2 = 0.976$	$\bar{S}_3 = 0.939$	$\bar{S}_4 = 1.181$

Ahora estos factores se normalizan usando la ecuación:

$$\frac{L}{\sum_{t=1}^L \bar{S}_t} = \frac{4}{\sum_{t=1}^4 \bar{S}_t} = \frac{4}{4.009106} = 0.997728$$

Así, el estimador de S_t está dado por:

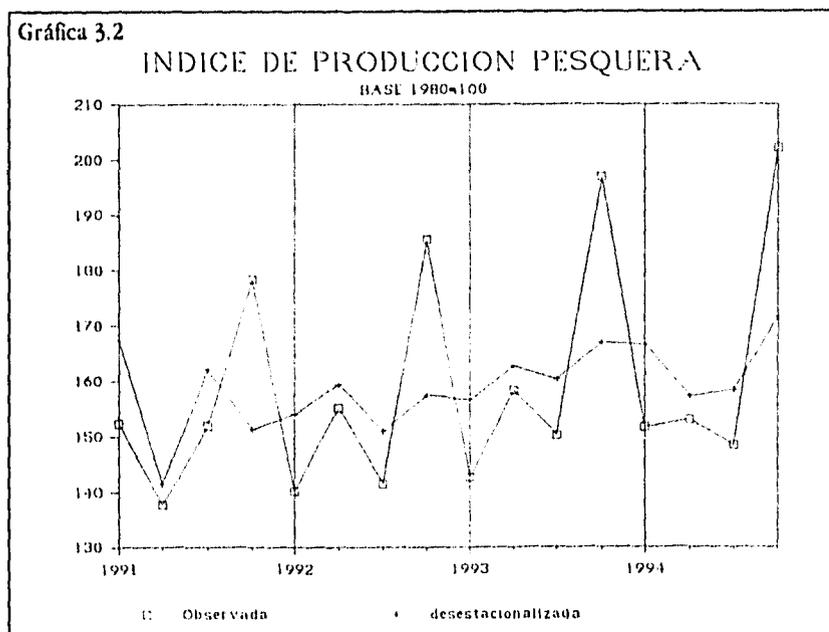
$$\hat{S}_t = \bar{S}_t \left[\frac{L}{\sum_{t=1}^L \bar{S}_t} \right] = \begin{cases} 0.911 & \text{si } t = 1 \\ 0.973 & \text{si } t = 2 \\ 0.937 & \text{si } t = 3 \\ 1.179 & \text{si } t = 4 \end{cases}$$

Con estos estimadores se puede obtener los datos desestacionalizados como se muestra en la tabla 3.3.

Tabla 3.3 Índice de volumen de la producción pesquera. Base 1980 = 100. Serie desestacionalizada.

Año	Trím.	Y_t	\hat{S}_t	$d_t = \frac{Y_t}{\hat{S}_t}$
1991	1	152 4000	0 911	167 373
	2	137 7000	0 973	141 470
	3	152 0000	0 937	162 160
	4	178 4000	1 179	151 345
1992	1	140 2000	0 911	153 974
	2	155 1880	0 973	159 437
	3	141 5120	0 937	150 971
	4	185 5360	1 179	157 399
1993	1	142 5834	0 911	156 592
	2	158 2917	0 973	162 626
	3	150 2857	0 937	160 331
	4	196 8537	1 179	167 000
1994	1	151 7087	0 911	166 614
	2	153 0680	0 973	157 259
	3	148 4823	0 937	158 407
	4	202 1687	1 179	171 509

Gráfica 3.2



Ya que la gráfica 3.2 indica que los datos desestacionalizados cambian siguiendo una función lineal, parece razonable suponer que:

$$\hat{T}_t = \beta_0 + \beta_1 t$$

Por lo tanto, por medio del análisis de regresión lineal, los datos desestacionalizados se ajustan a una línea recta, dando como resultado:

$$\hat{\beta}_0 = 146.178 \text{ y } \hat{\beta}_1 = 1.526.$$

Entonces la estimación de la tendencia esta dada por:

$$\hat{T}_t = \hat{\beta}_0 + \hat{\beta}_1 t = 146.178 + 1.526t$$

Ya que se han encontrado \hat{T}_t y \hat{S}_t se pueden calcular los factores cíclicos, realizando los siguientes pasos: Primero, calcular $\hat{C} \times \hat{I}$ como se muestra en la tabla 3.4. Segundo, los factores cíclicos se obtienen promediando la serie $\hat{C}_t \times \hat{I}_t$, usando una MA(3), tabla 3.5.

TABLA 3.4 Estimación de $\hat{C}_t \times \hat{I}_t$.

Año	Trim.	y_t	\hat{T}_t	\hat{S}_t	$\hat{T}_t \times \hat{S}_t$	$\hat{C}_t \times \hat{I}_t$ $= \frac{y_t}{\hat{T}_t \times \hat{S}_t}$
1991	1	152.4000	147.704	0.911	134.490	1.135
	2	157.7000	149.230	0.973	145.253	0.948
	3	152.0000	150.756	0.937	141.310	1.076
1992	1	178.4000	152.282	1.179	179.504	0.994
	2	140.2000	153.808	0.911	140.048	1.001
	3	155.1880	155.334	0.973	151.194	1.026
1993	1	141.5120	156.860	0.937	147.032	0.962
	2	185.5360	158.386	1.179	186.699	0.994
	3	142.5834	159.912	0.911	145.606	0.979
1994	1	158.2917	161.438	0.973	157.135	1.007
	2	150.2857	162.964	0.937	152.753	0.984
	3	196.8537	164.490	1.179	193.894	1.015
1994	1	151.7087	166.016	0.911	151.164	1.004
	2	153.0680	167.542	0.973	163.077	0.939
	3	148.4823	169.068	0.937	158.475	0.937
	4	202.1687	170.594	1.179	201.089	1.005

Por último, la componente irregular se encuentra usando la ecuación:

$$I_t = \frac{\hat{C}_t \times \hat{I}_t}{\hat{C}_t}$$

ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

TABLA 3.5 Estimación de $\hat{C}_t \times \hat{I}_t$

Año	Trim.	$\hat{C}_t \times \hat{I}_t$	\hat{C}_t	\hat{I}_t
1991	1	1.133		
	2	0.918	1.052	0.901
	3	1.076	1.006	1.069
	4	0.991	1.024	0.971
1992	1	1.001	1.007	0.994
	2	1.026	0.997	1.030
	3	0.962	0.994	0.968
	4	0.994	0.978	1.016
1993	1	0.979	0.993	0.986
	2	1.007	0.990	1.017
	3	0.984	1.002	0.982
	4	1.015	1.001	1.014
1994	1	1.004	0.986	1.018
	2	0.939	0.960	0.978
	3	0.937	0.960	0.976
	4	1.005		

Ahora se está en condiciones de hacer pronósticos y encontrar un intervalo de confianza para y_t con:

$$\hat{y}_t = \hat{T}_t \times \hat{S}_t$$

donde: $T = \beta_0 + \beta_1 t$ y $\hat{S}_1 = 0.911$, $\hat{S}_2 = 0.973$, $\hat{S}_3 = 0.937$ y $\hat{S}_4 = 1.179$

El pronóstico para y_{17} es:

$$\begin{aligned} \hat{y}_{17} &= \hat{T}_{17} \times \hat{S}_{17} \\ &= 172.120 \times 0.911 \\ &= 156.722 \end{aligned}$$

Más aun, después de calcular un intervalo al 95% de confianza para los datos desestacionalizados, el intervalo al 95% de confianza para y_{17} es:

$$\hat{y}_{17} \pm INT$$

donde:

$$INT = t_{(25, 14)} s^2 f_t = 2.145 \times 0.318 \times 3.269 = 2.228$$

Por lo tanto el intervalo al 95% de confianza para y_{17} es:

$$(\hat{y}_{17} - 2.228, \hat{y}_{17} + 2.228) \quad \hat{o} \quad (154.494, 158.950)$$

3-3 METODO DE DESCOMPOSICION CENSO II.

El método Censo II fué desarrollado por la Oficina de Censos de los Estados Unidos del Departamento de Comercio y ha sido muy usado por gran número de empresas y dependencias gubernamentales en los Estados Unidos. Julius Shiskin es considerado como el principal colaborador en el desarrollo de este método, el cual lleva varias modificaciones desde 1955, año en que aparece la primera versión [10].

Este proceso se basa en el método de descomposición multiplicativa, sin embargo, Censo II es mucho más elaborado; ofrece información adicional a través de pruebas empíricas diseñadas para revisar la adecuación de la descomposición, además aprovecha todas las ventajas de la computación hoy en día.

El método Censo II consiste de cuatro fases:

- 1) Se intenta ajustar los datos por la variación de días hábiles.
- 2) Estimación preliminar de factores estacionales y un ajuste preliminar estacional para la serie.
- 3) Refina el ajuste de tal forma que los factores estacionales se calculan con mayor exactitud. Adicionalmente se hace una estimación de la tendencia-ciclo y de la componente irregular.
- 4) Prepara un resumen estadístico que se puede usar para determinar que tan acertado ha sido el ajuste estacional y proporciona información necesaria para estimar la tendencia-ciclo con el propósito de hacer pronósticos.

3-3.1 AJUSTE POR DIAS HABILES.

El ajuste por días hábiles a menudo es necesario porque un mes dado puede tener distinto número de días laborados en diferentes años. En algunas industrias este factor es muy importante, ya que puede tener influencia significativa sobre su producción o ventas.

El primer paso es determinar el número de días hábiles para cada mes en los años de interés, en seguida se calcula el promedio sobre todos los años, de los días laborados en cada mes.

Sea d_j el número de días hábiles para cada mes y \bar{d}_j el promedio de días laborados en cada mes. El cociente:

$$\frac{d_j}{\bar{d}_j}$$

proporciona el factor de ajuste para cada mes.

Ahora para obtener la serie ajustada por días hábiles se multiplica la serie original por su factor correspondiente:

$$y_t = x_t \times \frac{d_t}{\bar{d}_t}$$

Estos datos son los que normalmente se usarán como entrada en Censo II y serán referidos como los datos originales ajustados por días hábiles.

3-3.2 AJUSTE ESTACIONAL INICIAL.

La segunda fase pretende separar preliminarmente la estacionalidad de la tendencia-ciclo, esto se hace calculando una media móvil centrada (el ancho de la media es igual al número de periodos que forman un año). Matemáticamente estos cálculos logran lo siguiente:

$$y_t = S_t \times T_t \times C_t \times I_t$$

$$M_t = T_t \times C_t$$

$$R_t = \frac{y_t}{M_t} = \frac{S_t \times T_t \times C_t \times I_t}{T_t \times C_t} = S_t \times I_t$$

donde M_t es la media móvil centrada.

Los valores de R_t contienen estacionalidad y aleatoriedad. Además, se debe mencionar que se pierden 6 datos al principio y 6 datos al fin de la serie por el uso de la media móvil.

Reemplazo de valores extremos. R_t puede incluir eventos inusuales como huelgas, paros, guerras, etc. El siguiente paso es excluir valores extremos antes de eliminar la aleatoriedad. Este proceso tiene dos etapas:

1. Calcula una $MA(3 \times 3)$ a R_t , digase R'_t . El propósito de este paso es eliminar la mayor aleatoriedad posible, sin embargo, calculando una $MA(3 \times 3)$ se pierden 2 datos al principio y 2 datos al final de la serie. Para evitar dicha pérdida, el Censo II estima estos valores como sigue: los dos valores perdidos al principio son iguales al promedio de los dos valores siguientes, los dos valores perdidos al final son iguales al promedio de los dos valores anteriores. Esto da cuatro valores tal que después de aplicar una $MA(3 \times 3)$ se tiene el mismo número de datos que antes.
2. Cálculo de la desviación estándar. Después de calcular la $MA(3 \times 3)$ se determina la desviación estándar para cada uno de los periodos, de la siguiente forma:

$$DS = \sqrt{\frac{\sum (R_t - R'_t)^2}{n/12}}$$

La desviación estándar se utiliza para construir límites de control que identifiquen valores extremos. Esto es

$$R'_t \pm 2 \times DS$$

Si algún dato se encuentra fuera de estos límites, se sustituye por la media de los dos valores más próximos a ese dato.

Factores estacionales preliminares. Después de haber reemplazado los valores extremos en R_t , sus datos son ajustados y usados para calcular los factores estacionales preliminares. El ajuste se efectúa de la siguiente manera:

1. Los 6 datos perdidos al principio y los 6 datos perdidos al final después de calcular R'_t , sus datos son reemplazados con los correspondientes valores de cada mes del año siguiente o precedente según sea el caso.
2. Se hace un siguiente ajuste a R'_t , tal que los promedios anuales sean 12, esto se hace calculando las sumas anuales y dividiendo cada una de ellas entre 12. El valor obtenido es la media de todos los meses para cada año, este valor divide a cada mes para el año apropiado dando una dando un promedio anual de 1.

Al momento se han hecho varios ajustes a R_t como parte de la etapa preliminar. Primero, son reemplazados los valores extremos, segundo, se estiman los datos perdidos y tercero, R_t se ajustó para que promediara 12, el objetivo fué eliminar el efecto de eventos inusuales. El siguiente paso es eliminar la aleatoriedad tomando un $MA(3 \times 3)$ para cada mes del año individualmente. Este promedio móvil es análogo al descrito en el paso 1 sobre el reemplazo de valores extremos excepto por la modificación de los datos por reemplazo de valores extremos y los ajustes utilizados.

El último paso de esta fase preliminar es dividir el dato original entre el factor preliminar estacional para obtener una serie ajustada preliminarmente. Esta serie constituye la base para después refinar las estimaciones de estacionalidad, tendencia, ciclo y aleatoriedad:

$$S_t' = \frac{Y_t}{S_t} = \frac{S_t \times T_t \times C_t \times I_t}{S_t} = T_t \times C_t \times I_t$$

donde: S_t' son los valores del ajuste estacional preliminar y S_t se obtuvo aplicando a R_t una $MA(3 \times 3)$.

3-3.3 AJUSTE ESTACIONAL FINAL.

En este paso de Censo II la serie ajustada preliminarmente por estacionalidad es procesada usando un promedio móvil para eliminar algún efecto estacional y/o irregular no detectado previamente.

La tendencia-ciclo. Usando los datos ajustados estacionalmente como punto de partida, se remueve la aleatoriedad aplicando una media móvil ponderada de Spencer. La razón de aplicar esta media, es que la serie S_t incluye tendencia-ciclo y aleatoriedad en sus componentes y esta media ofrece una curva suavizada que destaca la existencia de tendencia-ciclo en sus datos.

$$M'_t = T_t \times C_t$$

$$\frac{y_t}{M'_t} = \frac{S_t \times T_t \times C_t \times I_t}{T_t \times C_t} = S_t \times I_t$$

Este resultado da el factor estacional-irregular, donde: M'_t es la media móvil ponderada de Spencer.

Aplicando la media móvil ponderada de Spencer se pierden 7 datos al principio y 7 datos al final. Para evitar estas pérdidas cada dato perdido se reemplaza por un valor estimado: los primeros 7 valores son iguales a la media de las 4 siguientes observaciones y las 7 últimas son iguales a la media de las 4 observaciones anteriores.

Los factores de estacionalidad e irregularidad. Los factores de estacionalidad e irregularidad finales se calculan dividiendo los valores originales entre la media móvil ponderada de Spencer, en esta serie resultante se reemplazan los valores extremos y se hacen los ajustes para que los años promedien 12. Estos pasos son idénticos a los aplicados anteriormente. El resultado de este proceso son los factores de estacionalidad e irregularidad finales.

Los factores estacionales finales. Los factores estacionales finales se calculan aplicando una $MA(3 \times 3)$ (o una $MA(5 \times 5)$ si existiera aleatoriedad en la serie anterior). Estos factores se pronostican un año, multiplicando el factor del último renglón por tres, restando el factor del renglón anterior y dividiendo este resultado entre dos.

Matemáticamente este paso es equivalente a calcular el valor esperado para remover aleatoriedad que aun este presente.

$$e(S_t \times I_t) = S_t$$

donde S_t es el factor estacional final y e denota el valor esperado.

Ajuste estacional final. La serie ajustada estacionalmente se encuentra dividiendo los datos originales entre los factores estacionales finales. Al realizar este ajuste, las fluctuaciones en los datos originales causadas por la estacionalidad, se habrán removido completamente y solo la tendencia-ciclo y la componente irregular permanecerán. Ya que el ajuste estacional tiende a suavizar la serie, el resultado es una estimación mas refinada de la tendencia-ciclo y la componente irregular.

$$\frac{y_t}{c(S_t \times I_t)} = \frac{S_t \times T_t \times C_t \times I_t}{S_t} = T_t \times C_t \times I_t$$

La preparación del ajuste estacional final es la fase más importante, una característica fundamental de esta fase es que la tarea de separar la aleatoriedad y los factores estacionales no termina simultáneamente como en muchos métodos de descomposición. Esto agranda los requerimientos computacionales, pero generalmente se mejora la exactitud.

Antes de proceder con la fase final de Censo II, se necesitan dos conjuntos de valores: una estimación final de la tendencia ciclo y una estimación final de la componente aleatoria. El primero de estos cálculos se obtiene aplicando una media móvil ponderada de 15 meses a los datos de la serie ajustada por estacionalidad.

Matemáticamente este paso es equivalente a calcular el valor esperado de

$$c(T_t \times C_t \times I_t) = T_t \times C_t$$

Esta ecuación es mucho mejor estimación para la tendencia-ciclo que M_t o M'_t por que fué calculada usando $\frac{y_t}{c(S_t \times I_t)}$ a la cual se le aplica el ajuste estacional.

Finalmente, la ecuación $c(T_t \times C_t \times I_t)$ se puede dividir entre la ecuación $\frac{y_t}{c(S_t \times I_t)}$ para obtener la estimación de la componente irregular:

$$\frac{T_t \times C_t \times I_t}{T_t \times C_t} = I_t$$

3-3.4 PRUEBAS ESTADISTICAS.

Después que en la fase III se han estimado las componentes básicas de la serie de tiempo, en la fase IV se usan una serie de estadísticos para determinar si o no la descomposición ha sido correcta. Estas pruebas no son estadísticas en un sentido rigurosamente matemático, pero se basan en consideraciones intuitivas. Las pruebas mas comúnmente usadas son:

La prueba del mes adyacente. Se calculan variaciones para un mes dado, promediando el mes anterior y el mes siguiente, obteniendo así una indicación de cuanto varía un mes particular respecto al anterior y al siguiente. Si el dato es no estacional la variación será pequeña, por el contrario, si existe fuerte estacionalidad la variación será considerable. Si todos los años son promediados para cada mes, los resultados darán una mejor indicación de la variabilidad de la serie.

La prueba de los eneros. Si se dividen los datos de la serie ajustada estacionalmente (final) entre el enero correspondiente, se obtiene así un conjunto de valores estandarizados en base a los eneros. Examinando esos valores se puede identificar algún patrón constante, si tal patrón existe se sugiere que la estacionalidad no ha sido removida de los datos. La prueba de los eneros revela cualquier estacionalidad intra-anual existente, mientras que la prueba del mes adyacente revela estacionalidad inter-anual. Ambas pruebas deben usarse en combinación para asegurarse que la desestacionalización sea exitosa.

La prueba de igualdad. En esta prueba se calcula una $MA(12)$ a los datos originales y se dividen entre la $MA(12)$ de los valores ajustados estacionalmente. Los primeros eliminarán la estacionalidad sin alterar el volumen de los datos y los segundos además eliminarán la aleatoriedad. Las variaciones entre esas dos medidas pueden usarse para identificar algún sobre-ajuste existente al desestacionalizar. Si las variaciones son cercanas a 100 indican que no hay sobreajuste, sin embargo, si las variaciones están por debajo de 90 o arriba de 110, esto indica que el ajuste estacional puede tener algún sobre-ajuste.

3-3.5 RESUMEN.

El método Censo II ha sido desarrollado empíricamente por expertos y probado con miles de series, aunque puede tener algunas fallas estadísticas. Al contrario del método Box-Jenkins, el cual se examinará posteriormente, Censo II no está basado teóricamente.

Hay muchas versiones diferentes del método Censo II, la descrita en esta sección es la $X-9$. La versión $X-11$ es comúnmente más aceptada la cual es muy similar a $X-9$ excepto por los siguientes cinco puntos:

- 1.- El ajuste por días hábiles no se hace en la primera fase, pero es calculado más adelante usando análisis de regresión.
- 2.- El suavizamiento de eventos inusuales (guerras, huelgas, inundaciones, etc.) se hace excluyendo los valores extremos que caigan fuera de algún límite de control.
- 3.- Se usan medias móviles de Heanderson. El ancho de la media móvil depende del nivel medio de la componente aleatoria, una componente aleatoria grande incluye más términos en la media de Heanderson (la media de Heanderson reemplaza la fórmula de Spencer usada en la versión $X-9$).
- 4.- El usuario siempre escogerá entre un modelo aditivo o multiplicativo.

5.- El $X - 11$ tiene muchas más opciones disponibles que $X - 9$.

Esas diferencias hacen al $X - 11$ más elaborado pero también más complicado que la versión $X - 9$. Se describió la versión $X - 9$ por ser mucho más simple y la esencia de los resultados es la misma.

3-3.6 EJEMPLO.

En esta parte, se presenta un ejercicio con el paquete de computación X-11. Ya que los cálculos son muy laboriosos solo se mostrará los cuadros de resultados de dicho paquete. La serie de tiempo a considerar es la del ejemplo anterior, solo que ahora desde el primer trimestre del año de 1980.

X-11 DECOMPOSITION

ORIGINAL DATA SERIES

YEAR	1ST QUARTER	2ND QUARTER	3RD QUARTER	4TH QUARTER
1980	83.8	81.8	86.2	148.2
1981	101.8	103.5	104.4	140.0
1982	117.1	116.8	101.9	149.7
1983	111.8	105.4	99.5	137.7
1984	122.1	119.9	118.2	149.0
1985	119.0	113.3	118.1	164.7
1986	124.7	116.2	119.4	162.5
1987	126.3	134.3	135.8	183.8
1988	121.5	131.6	131.7	187.7
1989	134.3	124.4	135.8	178.2
1990	147.4	141.7	145.5	174.6
1991	152.4	137.7	152.0	178.4
1992	140.2	155.2	141.5	185.5
1993	142.6	158.3	150.3	196.9
1994	151.7	153.1	148.5	202.2

X-11 DECOMPOSITION

FINAL SEASONAL-IRREGULAR RATIOS

YEAR	1ST QUARTER	2ND QUARTER	3RD QUARTER	4TH QUARTER
1980	96.4	91.8	87.9	126.0
1981	94.1	92.6	90.1	119.2
1982	96.9	95.7	84.7	125.3
1983	95.5	93.6	89.1	118.9
1984	98.5	91.8	91.4	119.2
1985	97.0	90.7	91.2	123.5
1986	94.7	90.1	92.5	123.9
1987	91.8	92.2	91.3	125.7
1988	91.2	92.1	90.5	126.4
1989	93.0	88.9	95.5	119.4
1990	95.1	90.8	96.1	116.5
1991	95.5	90.8	98.3	116.0
1992	90.9	99.0	91.3	120.4
1993	90.0	98.4	91.9	119.1
1994	92.3	96.0	92.0	121.1

X-11 DECOMPOSITION
FINAL SEASONAL FACTORS

YEAR	1ST QUARTER	2ND QUARTER	3RD QUARTER	4TH QUARTER
1980	95.8	93.5	87.8	122.8
1981	95.9	93.3	88.1	122.3
1982	96.1	93.2	88.7	121.8
1983	96.4	92.8	89.2	121.8
1984	96.2	92.3	90.0	122.0
1985	95.7	91.7	90.7	122.8
1986	94.6	91.2	91.5	123.3
1987	93.8	90.9	92.2	123.3
1988	93.3	90.8	93.3	122.3
1989	93.1	91.3	94.1	121.0
1990	93.0	92.3	94.6	119.6
1991	92.8	93.6	94.4	118.9
1992	92.5	94.8	91.0	118.6
1993	92.0	96.0	93.1	119.0
1994	91.6	96.6	92.6	119.4

X-11 DECOMPOSITION
PROJECTED ONE YEAR AHEAD SEASONAL FACTORS

YEAR	1ST QUARTER	2ND QUARTER	3RD QUARTER	4TH QUARTER
1995	91.4	96.9	92.3	119.6

X-11 DECOMPOSITION
FINAL SEASONALLY ADJUSTED SERIES

YEAR	1ST QUARTER	2ND QUARTER	3RD QUARTER	4TH QUARTER
1980	87.5	87.5	98.1	120.7
1981	106.1	110.9	118.4	114.4
1982	121.8	125.4	114.9	122.9
1983	115.9	113.6	111.5	113.1
1984	126.9	130.0	131.3	122.1
1985	121.4	123.5	130.3	134.1
1986	131.8	127.1	130.5	131.8
1987	134.6	147.7	147.2	149.0
1988	130.2	144.9	141.1	153.4
1989	144.3	136.2	144.3	147.3
1990	158.5	153.6	153.8	146.0
1991	164.2	147.1	161.0	150.1
1992	151.6	163.7	150.6	156.1
1993	154.9	164.9	161.4	165.4
1994	165.7	158.4	160.4	169.3

X-11 DECOMPOSITION

FINAL TREND CYCLE

YEAR	1ST QUARTER	2ND QUARTER	3RD QUARTER	4TH QUARTER
1980	86.7	89.1	98.5	106.5
1981	108.1	111.6	115.7	117.2
1982	121.4	122.2	119.8	119.0
1983	117.7	113.0	111.1	115.4
1984	124.6	131.3	129.0	124.8
1985	122.5	125.1	129.7	133.5
1986	131.4	128.8	129.6	131.4
1987	137.0	144.8	149.1	146.0
1988	143.6	141.7	145.4	148.5
1989	144.6	139.0	141.8	150.1
1990	155.1	156.1	151.1	149.8
1991	149.1	153.4	155.1	153.0
1992	154.1	157.7	155.9	153.4
1993	158.2	161.5	163.8	164.8
1994	164.1	159.8	161.7	166.5

X-11 DECOMPOSITION

FINAL IRREGULAR FACTORS

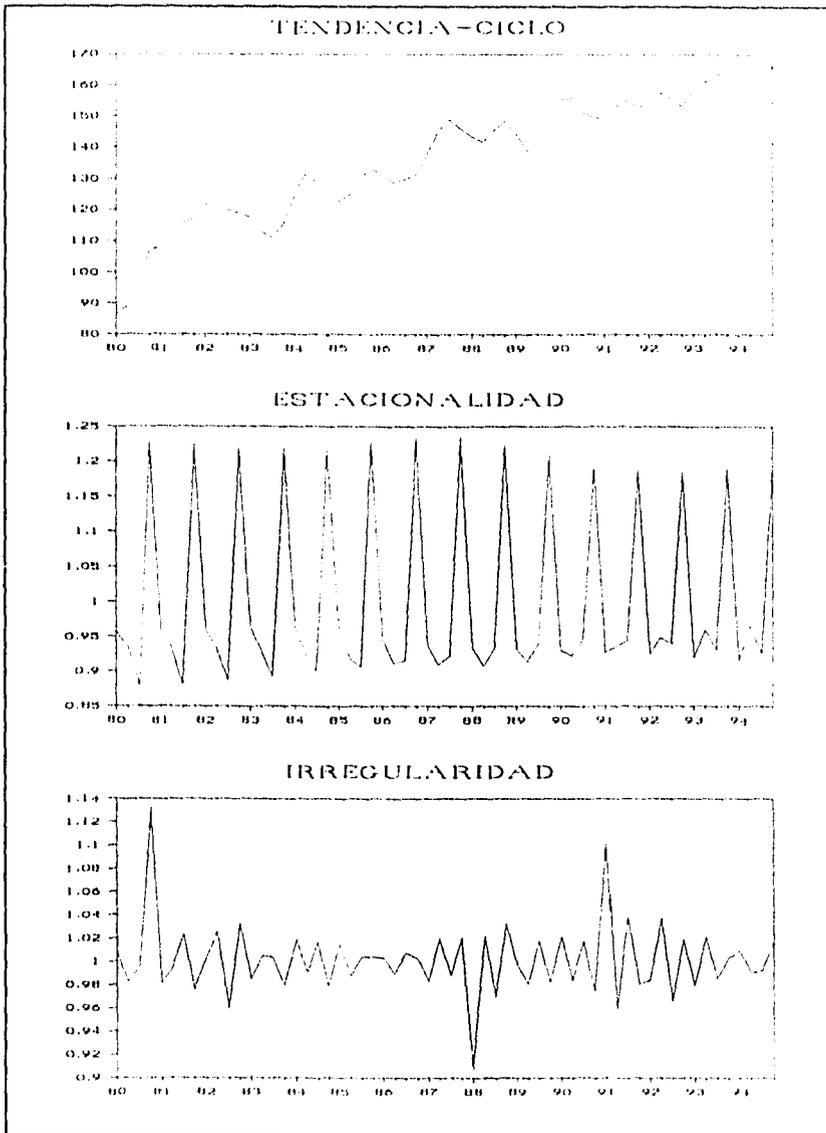
YEAR	1ST QUARTER	2ND QUARTER	3RD QUARTER	4TH QUARTER
1980	100.9	98.3	99.7	113.3
1981	98.1	99.4	102.4	97.6
1982	100.3	102.6	96.0	103.3
1983	98.5	100.5	100.3	98.0
1984	101.9	99.0	101.7	97.8
1985	101.5	98.7	100.4	100.4
1986	100.3	98.9	100.7	100.3
1987	98.3	102.0	98.8	102.1
1988	90.7	102.3	97.0	103.3
1989	99.8	98.0	101.8	98.1
1990	102.2	98.4	101.8	97.4
1991	110.2	95.9	103.8	98.1
1992	98.4	103.8	96.6	101.9
1993	97.9	102.1	98.5	100.3
1994	100.9	99.1	99.2	101.7

X-11 DECOMPOSITION

QUARTERS FOR CYCLICAL DOMINANCE MOVING AVERAGE QCD = 2

YEAR	1ST QUARTER	2ND QUARTER	3RD QUARTER	4TH QUARTER
1980		87.5	92.8	109.4
1981	113.4	108.5	114.7	116.4
1982	118.1	123.6	120.1	118.9
1983	119.4	114.8	112.6	112.3
1984	120.0	128.4	130.6	126.7
1985	123.3	124.0	126.9	132.2
1986	132.9	129.6	129.0	131.1
1987	133.2	141.2	147.5	148.1
1988	139.6	137.6	143.0	147.3
1989	148.8	140.2	140.3	145.8
1990	152.9	156.1	153.7	149.9
1991	155.1	155.7	154.1	155.5
1992	150.8	157.6	157.1	153.5
1993	155.7	159.9	163.2	163.4
1994	165.5	162.0	159.4	164.9

La siguiente gráfica muestra los factores finales de tendencia-ciclo, estacionalidad e irregularidad, producidos por el método Censo X-11.



Ya que se han obtenido los factores de tendencia ciclo finales se pueden hacer pronósticos de la serie de tiempo, multiplicando estos factores por el correspondiente factor estacional obtenido del cuadro PROJECTED ONE YEAR AHEAD SEASONAL FACTORS.

3-4 METODO DE DESCOMPOSICION WINTERS.

En esta sección se analizarán series de tiempo que tienen variación estacional multiplicativa. La técnica que se presenta son procesos de suavizamiento exponencial que se pueden usar para estimar series de tiempo con tendencia lineal y variación estacional multiplicativa. El método Winters [11] fué desarrollado por P. R. Winters y se puede usar para estimar la tendencia y los factores estacionales de una serie de tiempo, también se mostrará como esas estimaciones son usadas para obtener pronósticos de la serie, así mismo, se presentará un método intuitivo para calcular intervalos de confianza bastante exactos.

Este método supone que la serie de tiempo a estimar, se puede describir usando el modelo:

$$y_t = (\beta_0 + \beta_1 t) \times S_t \cdot c_t$$

3-4.1 ACTUALIZACION DE LOS ESTIMADORES.

Supóngase que al fin del periodo $n - 1$, se tienen estimadores para los parámetros del modelo β_0 , β_1 y S_t . En el tiempo n se obtiene una nueva observación y_n y se desea actualizar los estimadores considerando esta nueva observación. Se denotarán estos últimos estimadores como $\hat{\beta}_0(n)$, $\hat{\beta}_1(n)$ y $\hat{S}_t(n)$. Aquí $\hat{\beta}_0(n)$ es la ordenada al origen de la tendencia lineal. El estimador de la ordenada al origen usando la última observación como origen se denotará por $\alpha_0(n)$

Antes de continuar, se debe notar que hay varios factores estacionales, sea L el número de periodos que pueden ocurrir antes de que se vuelva a repita el primer periodo. Entonces para series de tiempo mensuales $L = 12$; y para series de tiempo trimestrales $L = 4$

Se definen los factores estacionales y sus estimaciones tal que la suma de ellos sea L . Esto es:

$$\sum_{i=1}^L S_i = L \quad \gamma \quad \sum_{i=1}^L \hat{S}_i(0) = L$$

Ahora, dada la observación y_n se desea obtener los estimadores $\alpha_0(n)$, $\hat{\beta}_1(n)$ y $\hat{S}_n(n)$ para actualizar $\alpha_0(n - 1)$, $\hat{\beta}_1(n - 1)$ que son los estimadores obtenidos en el periodo anterior, y $\hat{S}_n(n - L)$, el último estimador del factor estacional obtenido L periodos antes. Para actualizar el estimador de β_0 se usa la siguiente ecuación:

$$\alpha_0(n) = \alpha \frac{y_n}{\hat{S}_n(n-L)} + (1 - \alpha) [\alpha_0(n-1) + \hat{\beta}_1(n-1)]$$

donde: $\alpha_0(n)$ es el nuevo estimador de β_0 (comúnmente se le llama componente permanente) y α es una constante de suavizamiento entre 0 y 1. Aquí $[\alpha_0(n-1) \cdot \hat{\beta}_1(n-1)]$ es simplemente la estimación del nivel medio de la serie al tiempo n usando las estimaciones calculadas en el periodo anterior. Se divide \mathcal{Y}_n entre $\hat{S}_n(n-L)$ para desestacionalizar \mathcal{Y}_n .

Para obtener la estimación de β_1 , normalmente llamada tendencia, se utiliza la ecuación:

$$\hat{\beta}_1(n) = \beta[\alpha_0(n) - \alpha_0(n-1)] + (1 - \beta)\hat{\beta}_1(n-1)$$

donde β es una constante de suavizamiento entre 0 y 1.

La estimación del factor estacional se obtiene usando la ecuación:

$$\hat{S}_n(n) = \gamma \frac{\mathcal{Y}_n}{\alpha_0(n)} + (1 - \gamma)\hat{S}_n(n-L)$$

donde γ es una constante de suavizamiento entre 0 y 1. Hay que notar que se forzo a que los valores iniciales de los factores estacionales sumen L ; pero, después que el proceso de suavizamiento ha empezado, los factores estacionales no necesariamente deben sumar L .

Ya que se han obtenido los estimadores $\alpha_0(n)$, $\hat{\beta}_1(n)$ y $\hat{S}_n(n)$ se generan pronósticos para la serie usando la siguiente ecuación:

$$\hat{\mathcal{Y}}_{n+\tau}(n) = [\alpha_0(n) + \hat{\beta}_1(n)\tau] \hat{S}_{n+\tau}(n+\tau-L)$$

la cual produce pronósticos τ periodos adelante.

Hay que notar que para hacer un pronóstico para el periodo $n+\tau$, se usan los estimadores de los factores estacionales calculados en el periodo $n+\tau-L$, el cual se denota por $\hat{S}_{n+\tau}(n+\tau-L)$. Sin embargo, si un pronóstico se hace para un periodo mayor que L , el índice $n+\tau-L$ se refiere a un factor estacional que no se ha calculado, en tal caso el estimador más reciente se usara en el calculo del pronóstico.

3-4.2 LOS ESTIMADORES INICIALES.

Para empezar el proceso de suavizamiento, se requiere de los valores iniciales de los estimadores. Así se puede determinar $\alpha_0(0)$, $\hat{\beta}_1(0)$ y $\hat{S}_t(0)$ para $t = 1, 2, \dots, L$, donde L es el número de periodos diferentes.

Un camino para determinar los estimadores iniciales es usar el método de descomposición multiplicativa, así si, β_0 , β_t y \hat{S}_t con $t = 1, 2, \dots, l$ representan los estimadores obtenidos usando el método de descomposición multiplicativa, entonces $\alpha_0(0) = \hat{\beta}_0$, $\hat{\beta}_1(0) = \hat{\beta}_1$ y $\hat{S}_t(0) = \hat{S}_t$ para $t = 1, 2, \dots, l$.

Aunque el método de descomposición multiplicativa puede usarse para obtener los estimadores iniciales, el método que se describirá a continuación es usado más frecuentemente. Este método se presenta en Johnson y Montgomery y es similar al método presentado por Winters.

Supóngase que se dispone de datos para los últimos m años, sea \bar{y}_i la media de las observaciones en el i -ésimo año, donde $i = 1, 2, \dots, m$. Entonces el estimador Inicial de β_1 , esta dado por la ecuación:

$$\hat{\beta}_1(0) = \frac{\bar{y}_m - \bar{y}_1}{(m-1)l}$$

Así $\hat{\beta}_1(0)$ es simplemente el cambio en el nivel promedio por periodo de la media del año 1 a la media del año m .

El estimador Inicial de β_0 esta dado por:

$$\alpha_0(0) = \bar{y}_1 - \frac{l}{2}\hat{\beta}_1(0)$$

$\alpha_0(0)$ es, por lo tanto, el nivel medio de la serie a mitad del año 1 menos la cantidad que este nivel ha cambiado al empezar el año 1 hasta la mitad del mismo año.

La obtención de los estimadores Iniciales para los l factores estacionales se hace como sigue:

$$\hat{S}_t = \frac{y_t}{\bar{y}_1 - \left[\frac{(l+1)}{2} - j \right] \hat{\beta}_1(0)}$$

esta ecuación se puede calcular para cada periodo t para los m años. Aquí \bar{y}_1 es la media de las observaciones para el año al cual pertenece t , tal que si $1 \leq t \leq l$, entonces $i = 1$; si $l+1 \leq t \leq 2l$, entonces $i = 2$, etc. La letra j indica la posición del periodo t dentro de año. Si la serie de tiempo consiste de datos mensuales, entonces, para enero $j = 1$, para febrero $j = 2$, para marzo $j = 3$, etc.

La expresión \underline{S}_t es entonces la razón de observaciones en el periodo t para el nivel medio de la serie en ese periodo. Este es el factor por el cual se debe multiplicar el nivel medio para obtener la observación, por lo tanto \underline{S}_t representa factores que no contabilizan en el nivel de la serie ya que el nivel está determinado por una componente permanente y la tendencia, \underline{S}_t representa el factor estacional y el error ϵ_t .

La ecuación \underline{S}_t produce m estimaciones para cada factor estacional distinto, uno para cada año. Si la serie es mensual se obtienen m estimaciones de los factores estacionales para cada mes. Esas m estimaciones son entonces promediadas para producir una estimación para cada periodo diferente. Así se obtiene:

$$\bar{S}_t = \frac{1}{m} \sum_{k=0}^{m-1} \underline{S}_{t+kt} \quad \text{para } t = 1, 2, \dots, l$$

Tal que si la serie de tiempo es mensual se obtienen 12 factores estacionales, uno para cada mes. Se espera que esto remueva la influencia del término ϵ_t . Esto es razonable ya que se supone que el valor esperado de ϵ_t es cero, así los factores obtenidos representan solamente los efectos estacionales.

Finalmente, los factores son normalizados tal que sumen l , así el estimador inicial $\hat{S}_t(0)$ está dado por la fórmula:

$$\hat{S}_t(0) = \bar{S}_t \left[\frac{l}{\sum_{i=1}^l \bar{S}_i} \right] \quad \text{para } t = 1, 2, \dots, l$$

De esta manera, se obtienen los estimadores iniciales $\hat{\alpha}_0(0)$, $\hat{\beta}_1(0)$ y $\hat{S}_t(0)$ con $t = 1, 2, \dots, l$.

Se puede utilizar un proceso de simulación para determinar la combinación óptima de las constantes de suavizamiento α , β y γ . El conjunto (α, β, γ) se puede formar variando secuencialmente cada α , β y γ dentro de un intervalo $(0, 1)$ en incrementos de un tamaño dado.

Las estimaciones generadas por cada combinación son comparadas con los valores originales de la serie de tiempo, la combinación particular que minimiza la suma de errores al cuadrado es la que se usará para hacer pronósticos de la serie de tiempo. Este proceso de simulación determina la combinación óptima de α , β y γ para calcular los valores iniciales de $\hat{\alpha}_0(0)$, $\hat{\beta}_1(0)$ y $\hat{S}_t(0)$ con $t = 1, 2, \dots, l$. Estos valores iniciales se obtienen usando los primeros m_1 años (la selección de m_1 es un tanto arbitraria y depende más bien de la experiencia del observador), del total de los m años.

Nota, se puede mostrar que si $\alpha = \beta$, entonces el proceso acerca de la actualización es equivalente al proceso de suavizamiento exponencial doble [11].

3.4.3 CUANDO NO EXISTE TENDENCIA.

El método Winters también se usa para estimar series de tiempo que no tienen tendencia, en este caso las ecuaciones son:

$$\alpha_0(n) = \alpha \frac{y_n}{\hat{S}_n(n-L)} + (1-\alpha)\alpha_0(n-1)$$

$$\hat{S}_n(n) = \gamma \frac{y_n}{\alpha_0(n)} + (1-\gamma)\hat{S}_n(n-L)$$

Aquí la estimación inicial de la componente permanente $\alpha_0(0)$, es justamente el promedio de los valores de las observaciones en los m años. El estimador inicial de \hat{S}_1 se determina usando el método previamente discutido para la situación en la cual existe tendencia lineal, con la excepción de que \hat{S}_1 se calcula como sigue:

$$\hat{S}_1 = \frac{y_1}{\alpha_0(0)}$$

El método Winters, es entonces una modificación del método de suavizamiento exponencial simple o doble y se usa, como ya se dijo, para estimar series de tiempo con variación estacional y con tendencia o sin ella.

3.4.4 INTERVALOS DE CONFIANZA.

La teoría estadística del método Winters se basa en la teoría desarrollada para el método de suavizamiento exponencial. Aquí, teóricamente, no hay intervalos de confianza correctos para $y_{n+\tau}$. Sin embargo, existe un método intuitivo que produce intervalos de confianza bastante exactos. Una aproximación del intervalo al $(100-\alpha)\%$ de confianza para $y_{n+\tau}$, es:

$$\hat{y}_{n+\tau}(n) \pm z_{\alpha/2} t_\tau \Delta_w(n)$$

donde:

$$\hat{y}_{n+\tau}(n) = [\alpha_0(n) + \beta_1(n)\tau] \hat{S}_{n+\tau}(n+\tau-L)$$

aquí $z_{\alpha/2}$ es el punto en la escala de la curva normal con media 0 y varianza 1, tal que hay un área de $(100-\alpha)/100$ bajo esta curva normal entre $-z_{\alpha/2}$ y $z_{\alpha/2}$, t_τ es una constante que depende solo de τ y no de t , la cual está dada por la fórmula:

$$a_t = 1,25 \left[\frac{1 + \frac{\alpha}{(1+\beta)^t} [(1+4t+5t^2) + 20(1+3t)t + 20^2 t^2]}{1 + \frac{\alpha}{(1+\beta)^t} [(1+4t+5t^2) + 20(1+3t)t + 20^2 t^2]} \right]^{1-\gamma}$$

donde:

$0 = \max\{\alpha, \beta, \gamma\}$ y $t = 1 - 0$, además

$$\Delta_{L_t}(n) = \frac{\sum_{t=1}^n \left| \frac{y_t}{S_t(t-L)} - [a_0(t-1) + b_1(t-1)] \right|}{n}$$

donde: $S_t(t-L)$ es el último factor estacional apropiado para el periodo t .

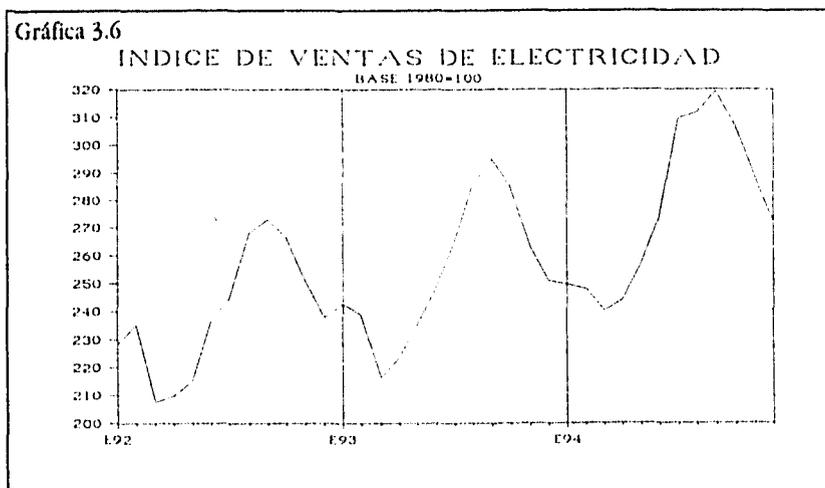
3-4.5 EJEMPLO.

Ahora considérese el método Winters como una herramienta para pronosticar el índice de volumen ventas de electricidad, servicio doméstico de la tabla 3.5. Como se observa en la gráfica 3.6, está serie posee variación estacional, ya que las ventas de electricidad tienden a aumentar en épocas de calor, por el uso excesivo de sistemas de aire acondicionado.

TABLA 3.5 Índice de volumen de ventas de electricidad, servicio doméstico.
Base 1980=100.

Mes	t	γ_t	Mes	t	γ_t	Mes	t	γ_t
E92	1	228.60	E93	13	242.70	E94	25	249.70
F	2	235.10	F	14	239.00	F	26	248.00
M	3	207.60	M	15	216.40	M	27	240.40
A	4	210.00	A	16	223.70	A	28	244.10
M	5	215.40	M	17	235.30	M	29	257.30
J	6	236.30	J	18	248.60	J	30	275.60
J	7	244.80	J	19	265.40	J	31	309.70
A	8	268.10	A	20	285.60	A	32	311.80
S	9	273.10	S	21	294.70	S	33	319.40
O	10	266.50	O	22	284.90	O	34	306.70
N	11	251.20	N	23	263.20	N	35	288.40
D	12	237.90	D	24	250.90	D	36	272.30

Fuente: Banco de México.



Para iniciar este proceso, primero se deben obtener los estimadores $\hat{\alpha}_0(0)$ y $\hat{\beta}_1(0)$ de la siguiente forma:

$$\hat{\beta}_1(0) = \frac{\bar{y}_m - \bar{y}_1}{(m-1)t} = \frac{256.84 - 239.55}{(3-1)(12)}$$

$$= 0.720601$$

$$\hat{\alpha}_0(0) = \bar{y}_1 - \frac{t}{2}(\hat{\beta}_1(0)) = 239.55 - \frac{12}{2}(0.720621)$$

$$= 235.2263$$

y $\hat{S}_t(0)$ para $t = 1, 2, \dots, 12$ los estimadores iniciales de los factores estacionales. (en este caso un factor estacional para cada uno de los 12 meses).

Solo se ilustrarán los cálculos para obtener los factores estacionales del mes de abril, ya que para los meses restantes el procedimiento es similar.

$$\hat{S}_4 = \frac{y_4}{\bar{y}_1 - \left[\frac{(t-1)}{2} - j \right] \hat{\beta}_1(0)}$$

$$= \frac{210.00}{239.55 - \left[\frac{(12-1)}{2} - 4 \right] 0.720601}$$

$$= 0.8832$$

$$\hat{S}_{16} = \frac{y_{16}}{\bar{y}_2 - \left[\frac{(t-1)}{2} - j \right] \hat{\beta}_1(0)}$$

$$= \frac{223.70}{254.20 - \left[\frac{(12-1)}{2} - 4 \right] 0.720601}$$

$$= 0.8262$$

$$\hat{S}_{28} = \frac{y_{28}}{\bar{y}_3 - \left[\frac{(t-1)}{2} - j \right] \hat{\beta}_1(0)}$$

$$= \frac{244.10}{276.78 - \left[\frac{(12-1)}{2} - 4 \right] 0.720601}$$

$$= 0.8876$$

Ahora se promedian estos valores para obtener:

$$\bar{S}_1 = \frac{S_1 + S_{15} + S_{26}}{3} = \frac{0.8832 + 0.8862 + 0.8876}{3} \\ = 0.8857$$

Similarmente se pueden obtener los otros once meses y así encontrar que:

$$\sum_{t=1}^{12} \bar{S}_t = 11.9937$$

Entonces, el estimador inicial $\hat{S}_t(0)$ esta dado por:

$$\hat{S}_t(0) = \bar{S}_t \left[\frac{L}{\sum_{t=1}^{12} \bar{S}_t} \right] = \bar{S}_t \left(\frac{12}{11.9937} \right) = (1.000525) \bar{S}_t$$

Por ejemplo:

$$\hat{S}_1 = 1.000525 \bar{S}_1 = (1.000525)(0.885759) = 0.886224$$

Por lo tanto los estimadores Iniciales usando este método son:

$$\begin{array}{ll} \alpha_0(0) = 235.2263 & \beta_1(0) = 0.720601 \\ \hat{S}_1(0) = 0.952 & \hat{S}_7(0) = 1.061 \\ \hat{S}_2(0) = 0.952 & \hat{S}_8(0) = 1.119 \\ \hat{S}_3(0) = 0.871 & \hat{S}_9(0) = 1.144 \\ \hat{S}_4(0) = 0.886 & \hat{S}_{10}(0) = 1.103 \\ \hat{S}_5(0) = 0.923 & \hat{S}_{11}(0) = 1.030 \\ \hat{S}_6(0) = 0.986 & \hat{S}_{12}(0) = 0.973 \end{array}$$

Con estos estimadores iniciales podemos determinar la combinación óptima de α , β y γ para después hacer pronósticos. Para mostrar una fase del proceso de simulación considérese la combinación $\alpha = 0.42$, $\beta = 0.001$ y $\gamma = 0.001$.

La estimación para enero de 1992 calculada un periodo antes es:

$$\begin{aligned}\hat{y}_1(0) &= [\alpha_0(0) + \beta_1(0)(1)]\hat{S}_1(0) \\ &= [235.226 + 0.720(1)]0.952 \\ &= 224.6986\end{aligned}$$

Ahora se puede obtener la estimación para $\alpha_0(1)$, $\hat{\beta}_1(1)$ y $\hat{S}_1(1)$ como se sigue:

$$\begin{aligned}\alpha_0(1) &= \alpha \frac{y_1}{\hat{S}_1(0)} + (1 - \alpha)[\alpha_0(0) + \hat{\beta}_1(0)] \\ &= 0.42 \frac{228.60}{0.952} + (1 - 0.42)(235.2263 + 0.720601) \\ &= 237.6676\end{aligned}$$

$$\begin{aligned}\hat{\beta}_1(1) &= \beta[\alpha_0(1) - \alpha_0(0)] + (1 - \beta)\hat{\beta}_1(0) \\ &= 0.001(237.6676 - 235.2263) + (1 - 0.001)(0.720601) \\ &= 0.722322\end{aligned}$$

$$\begin{aligned}\hat{S}_1(1) &= \gamma \frac{y_1}{\alpha_0(1)} + (1 - \gamma)\hat{S}_1(0) \\ &= 0.001 \frac{228.60}{237.6676} + (1 - 0.001)0.952 \\ &= 0.952336\end{aligned}$$

Entonces, la estimación para febrero de 1992, calculada un periodo antes es:

$$\begin{aligned}\hat{y}_2(1) &= [\alpha_0(1) + \hat{\beta}_1(1)(1)]\hat{S}_2(0) \\ &= (237.6676 + 0.722322)(0.952) \\ &= 226.8973\end{aligned}$$

Usando los estimadores anteriores se puede calcular:

$$\begin{aligned}
a_0(2) &= \alpha \frac{I_2}{S_2(0)} + (1 - \alpha) \{a_0(1) + \beta_1(1)\} \\
&= 0.42 \frac{235.10}{0.952} + (1 - 0.42)(237.6676 + 0.722322) \\
&= 242.0095
\end{aligned}$$

$$\begin{aligned}
\beta_1(2) &= \beta [a_0(2) - a_0(1)] + (1 - \beta)\beta_1(1) \\
&= 0.001(242.0095 - 237.6676) + (1 - 0.001)(0.722322) \\
&= 0.725942
\end{aligned}$$

$$\begin{aligned}
\hat{S}_2(2) &= \gamma \frac{I_2}{a_0(2)} + (1 - \gamma)\hat{S}_2(0) \\
&= 0.001 \frac{235.10}{242.0095} + (1 - 0.001)0.952 \\
&= 0.951810
\end{aligned}$$

Este proceso se continúa para todas las observaciones, los resultados se muestran en la tabla 3.7 y usando estos resultados se encuentra que para $\alpha = 0.42$, $\beta = 0.001$ y $\gamma = 0.001$ la suma de errores al cuadrado es igual a 1014.407

Para encontrar la combinación óptima de α , β y γ se evalúa la suma de errores al cuadrado para cientos de combinaciones de α , β y γ . Después de hacer todo este proceso de simulación, se encuentra que la combinación óptima fué $\alpha = 0.42$, $\beta = 0.001$ y $\gamma = 0.001$, que es precisamente con la cual se obtuvo la estimación de los estadísticos de la tabla 3.7 y por lo tanto se usará para hacer pronósticos de la serie.

TABLA 3.7 Proceso de simulación con $\alpha = 0.42$, $\beta = 0.001$ y $\gamma = 0.001$.

Mes	t	y_t	$\hat{S}_t(t)$	$\alpha_0(t)$	$\hat{\beta}_1(t)$	$\hat{y}_t(t-1)$	$(y_t - \hat{y}_t(t-1))^2$
E92	1	228.60	0.952336	237.6676	0.722322	224.6986	15.22085
F	2	245.10	0.951810	242.0095	0.725942	226.8973	67.28378
M	3	207.60	0.871185	240.8698	0.724076	211.4698	14.97564
A	4	210.00	0.886214	239.6477	0.722130	214.1061	16.86302
M	5	215.40	0.922507	237.4804	0.719240	221.7467	40.28124
J	6	236.30	0.986209	238.7899	0.719831	234.9138	1.921405
J	7	244.80	1.060678	235.8477	0.716169	254.0484	85.53323
A	8	268.10	1.118946	237.8399	0.717445	264.7004	11.55715
S	9	273.10	1.143673	238.6559	0.717543	272.8317	0.071973
O	10	266.50	1.103502	240.2687	0.718438	264.1478	5.532437
N	11	251.20	1.029512	242.2528	0.719704	248.0975	9.625037
D	12	257.90	0.973423	243.5704	0.720302	236.5143	1.919975
E93	13	242.70	0.952359	248.7243	0.724736	232.6468	101.0651
F	14	239.00	0.951814	250.1426	0.725429	237.4282	2.470438
M	15	216.40	0.871180	249.8303	0.724391	218.5525	4.633534
A	16	223.70	0.886218	251.3389	0.725176	222.0352	2.738242
M	17	255.30	0.922513	253.3247	0.726436	232.5310	7.667302
J	18	248.60	0.986204	253.2217	0.725607	250.5476	3.932006
J	19	265.40	1.060669	252.3806	0.724040	269.3566	15.65475
A	20	285.60	1.118951	254.0015	0.724937	283.2106	5.709052
S	21	294.70	1.143681	255.9663	0.726177	291.3240	11.39715
O	22	284.90	1.103506	257.3163	0.726801	283.2607	2.687044
N	23	263.20	1.029507	257.0401	0.725798	265.6587	6.045393
D	24	250.90	0.973423	257.7592	0.725791	250.9153	0.000236
E94	25	249.70	0.952367	260.0415	0.727347	246.1707	12.45565
F	26	248.00	0.951813	260.6790	0.727258	248.2034	0.041398
M	27	240.40	0.871207	267.5136	0.733365	227.7320	160.4773
A	28	244.10	0.886232	271.2680	0.736386	237.7254	40.63540
M	29	257.30	0.922527	274.9035	0.739287	250.9278	40.60462
J	30	273.60	0.986208	276.3934	0.740036	271.8421	3.089864
J	31	309.70	1.060701	283.3712	0.746273	293.9170	248.1547
A	32	311.80	1.118939	281.8226	0.743979	317.9138	37.37928
S	33	312.48	1.143673	281.1835	0.742596	323.1662	14.18459
O	34	306.70	1.103497	280.2487	0.740918	311.1072	19.42394
N	35	288.40	1.029505	280.6302	0.740559	289.2808	0.775944
D	36	272.30	0.973420	280.6835	0.739871	273.8929	2.537546

De ahí, el pronóstico para Y_{37} es:

$$\begin{aligned} \hat{y}_{37}(36) &= [\alpha_0(36) + \hat{\beta}_1(36)(1)] \hat{S}_{37}(25) \\ &= (280.6835 + 0.73987(1))(0.952) \\ &= 268.0185 \end{aligned}$$

Ahora se obtienen los estimadores $\alpha_0(37)$, $\hat{\beta}_1(37)$ y $\hat{S}_{37}(37)$

$$\begin{aligned} \alpha_0(37) &= \alpha \frac{y_{37}}{\hat{S}_{37}(25)} + (1 - \alpha)[\alpha_0(36) + \hat{\beta}_1(36)] \\ &= 0.42 \frac{235.90}{0.952} + (1 - 0.42)(280.6835 + 0.73987) \\ &= 282.6941 \end{aligned}$$

$$\begin{aligned}\beta_1(37) &= \beta[\alpha_0(37) - \alpha_0(36)] + (1 - \beta)\beta_1(36) \\ &= 0.001(282.6941 - 280.6835) + (1 - 0.001)(0.73887) \\ &= 0.741142\end{aligned}$$

$$\begin{aligned}\hat{S}_{37}(37) &= \gamma \frac{Y_{37}}{\alpha_0(37)} + (1 - \gamma)\hat{S}_{37}(29) \\ &= 0.001 \frac{270.90}{282.6941} + (1 - 0.001)0.952 \\ &= 0.951810\end{aligned}$$

Este proceso se puede continuar tanto como se desee. La tabla 3.8 y la gráfica 3.7 presentan los pronósticos para los siguientes seis meses de la serie.

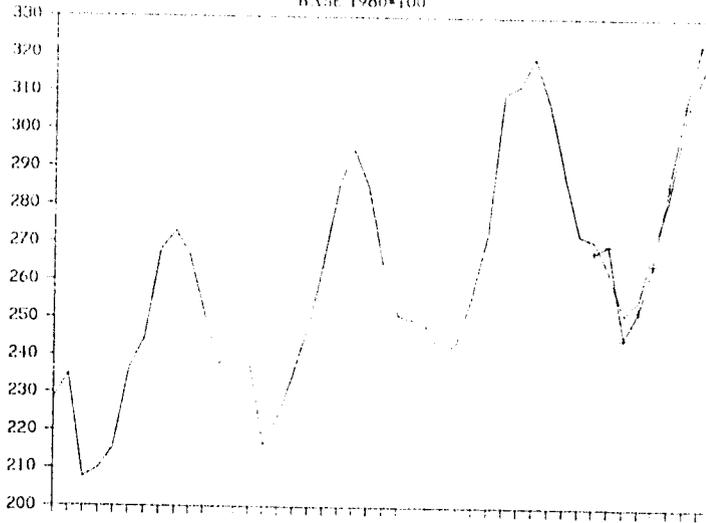
TABLA 3.8 Pronósticos usando el método Winters.

Mes	t	γ_t	$\hat{S}_t(t-12)$	$\alpha_0(t)$	$\beta_1(t)$	$\hat{y}_t(t-1)$
e9s	37	270.9	0.952373	282.6941	0.741142	268.0185
f	38	263.2	0.951800	280.5328	0.738240	269.7775
m	39	251.4	0.871220	284.3345	0.741303	245.0455
a	40	255.3	0.886237	286.3348	0.742562	252.6433
m	41	269.3	0.922536	289.1094	0.744594	264.8367
j	42	282.3	0.986201	288.3394	0.743080	283.8564
j	43	303.4	1.060695	287.8034	0.741800	306.6303
a	44	313.8	1.118920	285.1427	0.738398	322.8645
s	45	322.6	1.143664	284.2819	0.736799	326.9547

Gráfica 3.7

INDICE DE VENTAS DE ELECTRICIDAD

BASE 1980=100



— Observada

• Pronósticos

CAPITULO 4.

METODOLOGIA BOX-JENKINS.

Todos los métodos para pronosticar series de tiempo que se han visto hasta aquí, suponen que los errores en el modelo:

$$y_t = f(\beta_0, \beta_1, \dots, \beta_p; t) + \epsilon_t$$

son estadísticamente independientes uno de otro. En tal caso, si los ϵ_t , son estadísticamente independientes entonces también lo son las y_t . Sin embargo, si los ϵ_t , son estadísticamente dependientes, entonces también lo son las y_t , y se dice que las y_t , están autocorrelacionadas. De aquí se sigue que probablemente se pueden lograr mejores pronósticos empleando un modelo que exprese a y_t como una función del error actual y los errores pasados:

$$y_t = \mu + \psi_0 \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots$$

donde: $\epsilon_t, \epsilon_{t-1}, \dots$ son el error actual y los errores pasados y $\mu, \psi_0, \psi_1, \psi_2, \dots$ son los parámetros del modelo. En este modelo se supone que los $\epsilon_t, \epsilon_{t-1}, \dots$ son independientes y normalmente distribuidos, con media 0 y la misma varianza. Es claro que los valores de la serie de tiempo son dependientes porque están en función de algún error común. Box y Jenkins han desarrollado un proceso sistemático para analizar y pronosticar series de tiempo usando este modelo.

La metodología Box-Jenkins primero desarrolla un modelo apropiado de series de tiempo y lo utiliza para hacer pronósticos. Este desarrollo consiste de tres procesos iterativos: la identificación del modelo, la estimación de los parámetros y la adecuación del modelo. Una vez que el modelo ha sido desarrollado un cuarto paso, llamado obtención de pronósticos, genera predicciones de valores futuros de las series de tiempo. Estos cuatro pasos: identificación, estimación, adecuación y pronóstico son examinados en este capítulo.

Ya que la metodología Box-Jenkins utiliza la autocorrelación de las observaciones más efectivamente que el análisis de regresión y el suavizamiento exponencial, es probablemente el método que produce pronósticos más exactos. Más aun, la metodología Box-Jenkins ofrece una aproximación más sistemática para construir, analizar y pronosticar con modelos de series de tiempo. Sin embargo la metodología Box-Jenkins tiene varios inconvenientes.

Primero, al menos 50 o preferiblemente 100 observaciones son necesarias para construir un buen modelo Box-Jenkins, por lo tanto los modelos Box-Jenkins suelen analizar mejor las series de tiempo en los que el intervalo de muestreo es pequeño, pero que tienen bastante información acumulada. No obstante, cuando se tiene información mensual o anual el intervalo de muestreo debe ser bastante grande y posiblemente no haya suficiente información (50 o 100 observaciones) disponible para desarrollar un buen modelo

Box-Jenkins. Para series de tiempo que muestran variación estacional el problema es más serio, porque cada año es esencialmente una observación, por ejemplo, en tres años de datos mensuales solo se tienen tres observaciones por cada mes.

Otra desventaja de la metodología Box-Jenkins, es que no tiene disponible procesos automáticos para actualizar los estimadores de los parámetros del modelo cuando se ha observado un nuevo dato. Una desventaja final es que, ya que el proceso es más bien complicado y que se necesita una gran cantidad de información (50 o 100 observaciones), por esto su uso puede ser algunas veces muy tardado y costoso.

4-1 PASOS BASICOS DE LA METODOLOGIA BOX-JENKINS.

El método Box-Jenkins [2] consiste de varios pasos. El primero utiliza las observaciones de la serie de tiempo para identificar un modelo alternativo y así poder hacer estimaciones de la serie, para llevar a cabo este paso se deben comprender los conceptos de: autocorrelación, autocorrelación parcial, estacionaridad y no estacionaridad. El segundo paso se refiere a la estimación de los parámetros desconocidos del modelo identificado. El tercer paso involucra las pruebas de adecuación del modelo alternativo y si es necesario sugerir caminos para mejorarlo, a este paso se le llamará adecuación del modelo. Al final cuando ya se ha desarrollado un modelo de series de tiempo es utilizado para hacer pronósticos sobre la serie de tiempo.

4-1.1 IDENTIFICACION DEL MODELO.

La identificación de un modelo alternativo para posteriormente pronosticar valores futuros de una serie de tiempo se hace a través del análisis de toda la información. Se requiere de 50 o preferentemente las últimas 100 observaciones para identificar un buen modelo. Para todo esto es necesario conocer el concepto de estacionaridad.

4-1.1.1 SERIES DE TIEMPO ESTACIONARIAS Y NO ESTACIONARIAS.

Sea la serie de tiempo descrita por el modelo

$$y_t = \mu + \psi_0 \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots$$

Este modelo puede representar tanto series estacionarias como no estacionarias. Si la serie de tiempo es estacionaria estos valores fluctúan al rededor de un nivel medio constante, si la serie de tiempo es no estacionaria esta no tiene media constante. Ahora se verá como determinar si una serie de tiempo es estacionaria o no, esta identificación se lleva a cabo a través del análisis de los n valores de la serie y_1, y_2, \dots, y_n . Si la identificación determina que la serie de tiempo es no estacionaria, es importante transformar ésta serie por otra que sí sea estacionaria. Si la serie de tiempo no posee variación estacional, frecuentemente se podrá transformar en una serie estacionaria tomando las primeras diferencias de los valores originales de la serie, esto es:

$$z_t = \nabla y_t = y_t - y_{t-1} \quad \text{para } t = 2, 3, \dots, n$$

Es decir:

<i>Valores originales</i>	<i>Primeras diferencias</i>
y_1	$z_2 = \nabla y_2 = y_2 - y_1$
y_2	$z_3 = \nabla y_3 = y_3 - y_2$
.	.
.	.
.	.
y_n	$z_n = \nabla y_n = y_n - y_{n-1}$

Cuando se analizan series de tiempo que no poseen variación estacional y además son no estacionarias los valores $\nabla y_2, \nabla y_3, \dots, \nabla y_n$ generalmente son estacionarios. Sin embargo, si estos son todavía no estacionarios, se toman las segundas diferencias de los valores originales y_1, y_2, \dots, y_n generalmente éstas producirán series estacionarias, esto es:

$$z_t = \nabla^2 y_t = y_t - 2y_{t-1} + y_{t-2} \quad \text{para } t = 3, 4, \dots, n$$

Es decir:

<i>Valores originales</i>	<i>Segundas diferencias</i>
y_1	
y_2	
y_3	$z_3 = \nabla^2 y_3 = y_3 - 2y_2 + y_1$
y_4	$z_4 = \nabla^2 y_4 = y_4 - 2y_3 + y_2$
.	.
.	.
.	.
y_n	$z_n = \nabla^2 y_n = y_n - 2y_{n-1} + y_{n-2}$

En ocasiones, para transformar una serie en estacionaria, será necesario tomar ya sea el logaritmo natural, o las primeras o segundas diferencias de los logaritmos naturales de la serie original, sin embargo es poco usual.

4-1.1.2 AUTOCORRELACION Y AUTOCORRELACION PARCIAL.

Sea la serie z_a, z_{a+1}, \dots, z_n (la cual se puede suponer ha sido generada a partir de una serie estacionaria). Cabe notar que el primer valor no se escribe como z_1 , porque los valores z_a, z_{a+1}, \dots, z_n pueden ser primeras o segundas diferencias de una serie no estacionaria y_1, y_2, \dots, y_n .

Si los valores z_a, z_{a+1}, \dots, z_n son primeras diferencias, entonces:

$$z_a = z_2 = y_2 - y_1$$

en este caso $\alpha = 2$; mientras que si los valores z_a, z_{a+1}, \dots, z_n son segundas diferencias, entonces:

$$z_a = z_3 = y_3 - 2y_2 + y_1$$

en este caso $\alpha = 3$. Por supuesto, si y_1, y_2, \dots, y_n es estacionaria, entonces $z_a = y_1$, en tal caso $\alpha = 1$.

Una implicación importante de la suposición que la serie de tiempo z_a, z_{a+1}, \dots, z_n es estacionaria es que las propiedades estadísticas de la serie no se ven afectadas por cambios en el origen. Esto significa, por ejemplo, que las relaciones estadísticas entre n observaciones con origen en t , supóngase $z_t, z_{t+1}, \dots, z_{t+n-1}$, son las mismas relaciones estadísticas entre n observaciones con origen en $t+j$, supóngase $z_{t+j}, z_{t+j+1}, \dots, z_{t+j+n-1}$. una de esas importantes relaciones es medida por ρ_k , el cual es la autocorrelación entre dos observaciones de la serie de tiempo separadas (retrasadas) por k unidades de tiempo. Se puede mostrar que $-1 \leq \rho_k \leq 1$ y que $\rho_k = \rho_{-k}$ lo cual implica que basta considerar solo retrasos positivos [2]. Un valor de ρ_k cercano a 1 indica que observaciones retrasadas k unidades de tiempo tienen una fuerte tendencia a seguir una función lineal con pendiente positiva, mientras un valor de ρ_k cercano a -1 indica que observaciones retrasadas k unidades de tiempo tienen una fuerte tendencia a seguir una función lineal con pendiente negativa. Aunque ρ_k es un parámetro que no se puede conocer con certeza, se puede estimar usando las observaciones z_a, z_{a+1}, \dots, z_n . Al estimador de ρ_k se le llama autocorrelación muestral, se denota por el símbolo r_k y esta definido por la fórmula:

$$r_k = \frac{\sum_{t=a}^{n-k} (z_t - \bar{z})(z_{t-k} - \bar{z})}{\sum_{t=a}^{n-k} (z_t - \bar{z})^2}$$

donde \bar{z} es la media de z_a, z_{a+1}, \dots, z_n y esta dada por:

$$\bar{z} = \frac{\sum_{t=a}^n z_t}{n-a+1}$$

Se define: función de autocorrelación teórica (FACT) al registro o a la gráfica de ρ_k para $k = 1, 2, \dots, n$. Se define: función de autocorrelación muestral (FACM) al registro o a la gráfica de r_k para $k = 1, 2, \dots, n$.

La FACT de una serie de tiempo estacionaria tiende ya sea a disminuir al incrementar k , o, tiende a desaparecer después de un valor particular $k = q$. Cuando se dice que la FACT tiende a desaparecer después de un valor particular $k = q$ significa que:

$$\rho_k = 0 \quad \text{para} \quad k > q$$

Se puede determinar cuando la FACT tiende a desaparecer usando la FACM, sin embargo, aun cuando $\rho_k = 0$ para $k > q$, r_k será probablemente pequeña pero no igual a cero, para $k > q$, debido a variaciones de muestreo. Entonces la pregunta que se origina es: ¿Que tan pequeña debe ser r_k para concluir que $r_k = 0$? Una regla usada frecuentemente para concluir esto es:

$$\rho_k = 0 \quad \text{para} \quad k > q \quad \text{si}$$

$$|r_k| \leq 2 \frac{1}{(n-a+1)^{\frac{1}{2}}} \left(1 + 2 \sum_{j=1}^q r_j^2 \right)^{\frac{1}{2}} \quad \text{para} \quad k > q$$

Otra prueba usada con frecuencia para concluir si o no $\rho_k = 0$ es calculando el estadístico

$$t_{r_k} = \frac{r_k}{s_{r_k}}$$

donde:

$$s_{r_k} = \frac{1}{(n-a+1)^{\frac{1}{2}}} \left(1 + 2 \sum_{j=1}^{k-1} r_j^2 \right)^{\frac{1}{2}}$$

Aunque este estadístico no tiene una distribución t , es asintóticamente normal y a menudo se usa como una t estadística. Se puede concluir que:

$$\rho_k = 0 \quad \text{si} \quad |t_{r_k}| = \left| \frac{r_k}{s_{r_k}} \right| \leq 2$$

Inversamente, si:

$$|t_{r_k}| > 2$$

es razonable concluir que $\rho_k \neq 0$.

Se puede mostrar que si una serie es no estacionaria, entonces la FACM ni tiende a disminuir, ni tiende a desaparecer rápidamente, sino más bien disminuirá muy lentamente [2]. Ya que si se calcula directamente la FACM a y_1, y_2, \dots, y_n y si sucede que tiende a disminuir o tiende a desaparecer rápidamente, se puede suponer que la serie original es estacionaria. Sin embargo, si disminuye "lentamente" entonces se supone que la serie es no estacionaria. El significado de los términos lentamente y rápidamente son arbitrarios y se podrá determinar mejor a través de la experiencia. Si la serie de tiempo no posee variación estacional es muy raro que necesite tomar las segundas diferencias para lograr estacionaridad.

Otro parámetro útil para determinar estacionaridad es ρ_{kk} , el cual es la autocorrelación parcial entre dos observaciones de la serie de tiempo separadas (retrasadas) por k unidades de tiempo, se puede pensar como la autocorrelación de dos observaciones z_t y z_{t+k} , eliminando los efectos de las observaciones $z_{t+1}, z_{t+2}, \dots, z_{t+k-1}$.

Se puede mostrar [2] que $\rho_{11} = \rho_1$. Más aun ρ_{kk} es un parámetro que no se puede conocer con certeza, este puede estimarse a partir de z_a, z_{a+1}, \dots, z_n . El estimador de ρ_{kk} se llama autocorrelación parcial muestral, se denota por el símbolo r_{kk} y está dado por la fórmula:

$$r_{kk} = \begin{cases} r_1 & \text{si } k = 1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & \text{si } k = 2, 3, \dots \end{cases}$$

donde: $r_{kk} = r_{k-1, k} - r_{kk} r_{k-1, k-1}$ para $j = 1, 2, \dots, k-1$

Cabe notar que r_{kk} es la FACM con k retrasos.

Se define: función de autocorrelación parcial teórica (FACPT) al registro o a la gráfica de ρ_{kk} para $k = 1, 2, \dots, n$. Se define: función de autocorrelación parcial muestral (FACPM) al registro o a la gráfica de r_{kk} para $k = 1, 2, \dots, n$.

Igual que con la FACT se puede determinar cuando la FACPT tiende a desaparecer, usando la FACPM. Una regla usada frecuentemente para concluir que $\rho_{kk} = 0$ es la siguiente:

$$\rho_{kk} = 0 \quad \text{para} \quad k > q \quad \text{si}$$

$$|r_{kk}| \leq 2 \frac{1}{(n-a+1)^{1/2}} \quad \text{para} \quad k > q$$

Otra prueba usada a menudo para concluir si o no $\rho_{kk} = 0$, es calculando el estadístico:

$$t_{r_{kk}} = \frac{r_{kk}}{\frac{1}{(n-a+1)^{1/2}}}$$

Aunque este estadístico no tiene una distribución t , es asintóticamente normal y a menudo se usa como una t estadística. Se puede concluir que:

$$\rho_{kk} = 0 \quad \text{si} \quad |t_{r_{kk}}| = \left| \frac{r_{kk}}{\frac{1}{(n-a+1)^{1/2}}} \right| \leq 2$$

Inversamente, si

$$|t_{r_{kk}}| > 2$$

es razonable concluir que $\rho_{kk} \neq 0$.

4-1.1.3 IDENTIFICACION DE UN MODELO PARTICULAR DE ESTACIONARIDAD.

Ya que se tienen conclusiones concernientes a la función de autocorrelación y a la función de autocorrelación parcial de una serie de tiempo, estas se pueden usar para identificar el modelo que se supone generó la serie de tiempo bajo consideración. La metodología Box-Jenkins escoge un modelo de serie de tiempo

estacionario particular de entre una clasificación de modelos de series de tiempo estacionarios y ese modelo se usa para pronosticar valores futuros de la serie de tiempo original. La clasificación antes mencionada se refiere a los modelos $AR(p)$, $MA(q)$ y $ARMA(p, q)$, el significado de estos conceptos se dará a conocer más adelante. Sin embargo, el punto principal, es que cada modelo particular en la clasificación de modelos está caracterizado por el comportamiento de su función de autocorrelación y su función de autocorrelación parcial. Si se concluye que el comportamiento de la función de autocorrelación y la función de autocorrelación parcial de una serie de tiempo estacionaria observada es idéntico al comportamiento de la FACT y la FACPT de un modelo particular, entonces es razonable suponer tentativamente que este modelo particular ha generado la serie de tiempo observada y se puede usar para hacer pronósticos de dicha serie.

Al igual que la propiedad de estacionaridad, existe otra propiedad llamada invertibilidad, la cual se examinará más adelante. La metodología Box-Jenkins requiere que el modelo usado para describir y pronosticar una serie sea estacionario e invertible. Las condiciones de estacionaridad e invertibilidad de un modelo estacionario pueden ser expresadas en términos de los parámetros de los modelos.

Para ilustrar estas ideas, sea el modelo:

$$z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

el cual es llamado promedio móvil de orden uno $MA(1)$.

Se puede mostrar que para algún valor del parámetro θ_1 , este modelo describe una serie de tiempo estacionaria, por lo tanto no se deben imponer condiciones al parámetro θ_1 para hacer el modelo anterior estacionario. Sin embargo, existe una condición para que este modelo sea Invertible [2]. Esta condición es que:

$$|\theta_1| < 1$$

Se verá más adelante que el conocimiento de las condiciones de estacionaridad e Invertibilidad en los parámetros de un modelo de series de tiempo particular es de mucha importancia en la estimación de los parámetros del modelo.

Se puede mostrar que la media del modelo

$$z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

es μ , que su función de autocorrelación parcial tiende a disminuir en forma de una curva exponencial decreciente y que su FACT tiende a desaparecer después de un retraso [2]. En particular

$$\rho_k = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2} & \text{si } k = 1 \\ 0 & \text{si } k > 1 \end{cases}$$

4-1.2 ESTIMACION DE LOS PARAMETROS DEL MODELO.

Después de identificar un buen modelo, se obtienen los estimadores de los parámetros del modelo, usando un criterio similar al de mínimos cuadrados. Se pueden usar programas de computo que encuentren la estimación mínimo cuadrada, algunos de ellos aplican un proceso iterativo que requiere que el usuario especifique un estimador preliminar de los parámetros desconocidos como entrada. Estos estimadores se pueden obtener a través de las relaciones entre la FACT y los parámetros del modelo, dichas relaciones se usan para determinar los parámetros del modelo en términos de las autocorrelaciones teóricas. Algunas veces este proceso produce más de un conjunto de estimaciones preliminares, en tal caso, exactamente un conjunto de estimadores satisficará las condiciones de estacionaridad e invertibilidad en los parámetros del modelo. Este es el que se escogerá.

Por ejemplo, como se vió en la sección anterior, la relación entre ρ_k y θ_1 en el modelo $MA(1)$

$$z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

es:

$$\rho_k = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2} & \text{si } k = 1 \\ 0 & \text{si } k > 1 \end{cases}$$

Por lo tanto:

$$\rho_1 = \frac{-\theta_1}{1 + \theta_1^2}$$

Resolviendo para θ_1 en términos de ρ_1 se obtiene

$$\theta_1 = -\frac{1}{2\rho_1} \pm \left[\frac{1}{(2\rho_1)^2} - 1 \right]^{1/2}$$

reemplazando la autocorrelación teórica ρ_1 con la autocorrelación muestral r_1 , se obtiene un estimador inicial de θ_1

$$\hat{\theta}_1 = -\frac{1}{2r_1} \pm \left[\frac{1}{(2r_1)^2} - 1 \right]^{1/2}$$

Cabe notar que hay dos valores para el estimador preliminar $\hat{\theta}_1$. Para determinar cual de esos estimadores escoger, hay que recordar que la condición de invertibilidad para $MA(1)$ es

$$|\theta_1| < 1$$

En este sentido, se escoge el estimador $\hat{\theta}_1$ que satisfaga la condición

$$|\hat{\theta}_1| < 1$$

4-1.3 ADECUACION DEL MODELO.

Después que se ha identificado un modelo apropiado para una serie de tiempo, es importante probar la adecuación del modelo y, si se necesita, sugerir mejoras potenciales. Un camino puede ser el análisis de residuales, esto es, examinar las diferencias entre los datos observados y la estimación dada por el modelo. Específicamente es encontrar el camino para medir la sobreadecuación del modelo alternativo, examinando una cantidad que determine si o no las primeras k autocorrelaciones de los residuales indican adecuación del modelo. Esta cantidad es el estadístico Box-Pierce y se denota por el símbolo Q

$$Q = (n-d) \sum_{l=1}^k r_l^2(\epsilon)$$

donde:

- n es el número de observaciones en la serie de tiempo original
- d es el grado de diferenciación usado para transformar la serie de tiempo original en una serie estacionaria.
- $r_l(\epsilon)$ es la autocorrelación muestral de los residuales con l retrasos.

Q explica la relación entre los datos. Si Q es pequeña, los residuales serán no relacionados y de aquí la autocorrelación de los residuales será pequeña. Para Q grande, la autocorrelación entre los residuales será grande y los residuales están más relacionados. Así, un valor grande de Q indica que el modelo es inadecuado.

En la práctica es común aceptar la adecuación del modelo si:

$$Q < \chi^2_{(5, k - n_p)}$$

el cual se define como el punto en la escala de la distribución Chi cuadrada con $(k - n_p)$ grados de libertad y una significancia de 5%. Donde n_p es el número de parámetros estimados en el modelo. La selección de k , el número de autocorrelaciones de los residuales usados en el cálculo de Q , es arbitrario, pero comúnmente $k = 12, 24$ ó 36 .

Otra forma de checar la adecuación del modelo es considerar la cantidad:

$$s = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n - n_p}}$$

donde: n es el número de observaciones en la serie original y n_p es el número de parámetros estimados en el modelo. Como en el análisis de regresión s indica el grado de ajuste del modelo, por lo tanto la s más pequeña es el mejor ajuste a considerar. En particular un valor pequeño de s produce mejores intervalos de confianza cuando se esta usando un modelo Box-Jenkins.

Si el modelo se considera inadecuado, se deben hacer mejoras al modelo¹

¹ Véase: Nelson, C. "Applied Time Series Analysis for Managerial Forecasting" Holden-Day, Inc., San Francisco 1973

4-1.4 OBTENCION DE PRONOSTICOS.

Ya que se ha adoptado un modelo, este es usado para pronosticar valores futuros de la serie de tiempo. Sea y_1, y_2, \dots, y_n se desea estimar y_{n+j} . Generalmente, la estimación de y_{n+j} debe construirse recursivamente de las estimaciones $y_{t+1}, y_{t+2}, \dots, y_{n+j-1}$. Sea $\hat{y}_{n+j}(n)$ la estimación de y_{n+j} en el tiempo n . Se encontrará que $\hat{y}_{n+j}(n)$ es una función de y_{n+j} , el cual es un valor de la serie de tiempo más un error ϵ_{n+j} .

Primero, considérese y_{n+j} . Si $j \leq 0$, entonces y_{n+j} ya ha sido observado y se usa este valor observado. Si $j > 0$, entonces y_{n+j} no ha sido observado y se usa $\hat{y}_{n+j}(n)$.

Ahora considérese el término ϵ_{n+j} . Si $j \leq 0$, se estima ϵ_{n+j} por:

$$y_{n+j} - \hat{y}_{n+j}$$

Si $j > 0$, entonces: $\epsilon_{n+j} = 0$

4-1.5 EJEMPLO.

Utilizando la metodología Box-Jenkins, se desea determinar si la serie de tiempo del Índice de Producción en el Sector Manufacturero, Base 1980=100 de la tabla 4.1 y gráfica 4.1(a), es estacionaria o no estacionaria (en caso de no ser estacionaria, transformar la serie, para lograr que lo sea), y poder hacer pronósticos.

De la gráfica 4.1(a) se nota que los valores originales de la serie no parecen fluctuar alrededor de una media constante y por lo tanto parecería que la serie es no estacionaria. Para determinar más precisamente si la serie es estacionaria o no estacionaria, se debe calcular la FACM

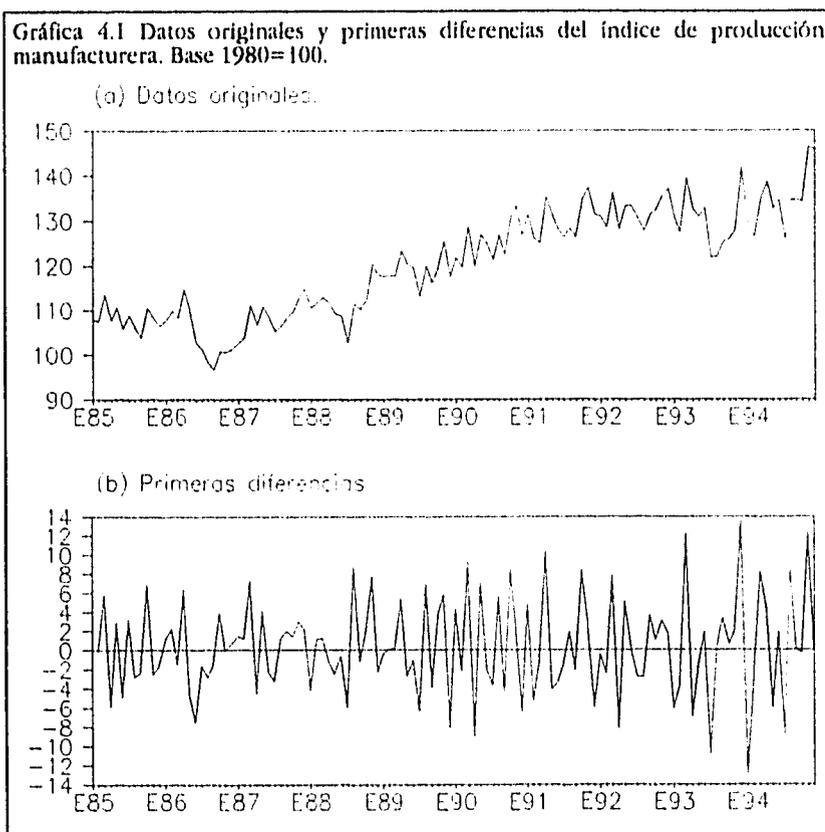


TABLA 4.1 Índice de producción manufacturera. Base 1980=100.

Mes	t	$\%_t$	Mes	t	$\%_t$
E85	1	107.77	E90	61	121.82
F	2	107.66	F	62	119.57
M	3	113.65	M	63	128.82
A	4	107.77	A	64	119.77
M	5	110.67	M	65	126.98
J	6	105.74	J	66	124.85
J	7	108.97	J	67	121.24
A	8	106.22	A	68	126.90
S	9	103.84	S	69	122.43
O	10	110.70	O	70	130.84
N	11	108.15	N	71	133.14
D	12	106.54	D	72	126.61
E86	13	107.69	E91	73	131.41
F	14	109.89	F	74	126.07
M	15	108.35	M	75	125.07
A	16	114.73	A	76	135.43
M	17	110.27	M	77	131.35
J	18	102.76	J	78	127.98
J	19	101.11	J	79	126.40
A	20	98.26	A	80	128.40
S	21	96.70	S	81	126.30
O	22	100.71	O	82	134.70
N	23	100.63	N	83	137.40
D	24	101.23	D	84	131.40
E87	25	102.68	E92	85	131.00
F	26	103.93	F	86	128.50
M	27	111.23	M	87	136.30
A	28	106.69	A	88	128.10
M	29	110.83	M	89	133.20
J	30	108.57	J	90	133.30
J	31	103.29	J	91	130.50
A	32	106.48	A	92	127.70
S	33	108.45	S	93	131.40
O	34	109.81	O	94	132.40
N	35	112.78	N	95	135.50
D	36	114.80	D	96	137.10
E88	37	110.61	E93	97	131.00
F	38	111.74	F	98	127.40
M	39	112.95	M	99	139.60
A	40	111.77	A	100	132.70
M	41	109.27	M	101	130.90
J	42	108.66	J	102	132.80
J	43	102.69	J	103	121.90
A	44	111.42	A	104	122.00
S	45	110.18	S	105	125.40
O	46	112.61	O	106	126.00
N	47	120.34	N	107	128.10
D	48	118.04	D	108	141.80
E89	49	117.67	E94	109	128.90
F	50	117.73	F	110	126.50
M	51	117.81	M	111	134.80
A	52	123.28	A	112	138.70
M	53	120.52	M	113	132.70
J	54	119.48	J	114	134.60
J	55	113.06	J	115	125.80
A	56	119.96	A	116	134.60
S	57	116.03	S	117	134.60
O	58	119.74	O	118	134.30
N	59	123.61	N	119	146.50
D	60	117.50	D	120	145.80

Para esto, primero se calcula:

$$\bar{y} = \frac{\sum_{t=1}^{120} y_t}{120} = \frac{107.77 + 107.66 + \dots + 115.80}{120} = 120.3958$$

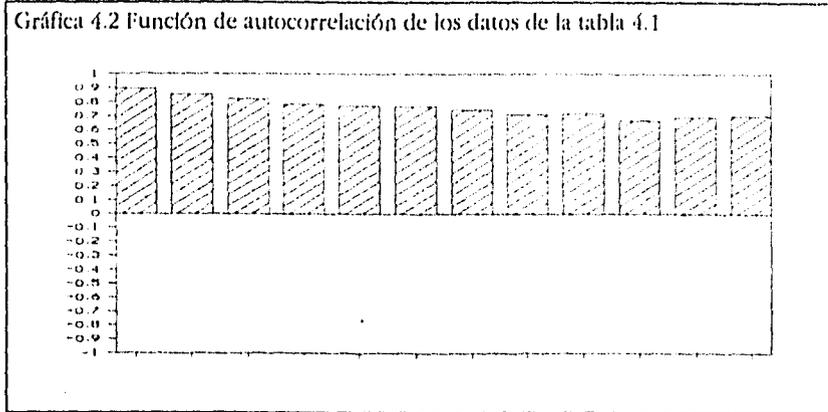
La tabla 4.2 presenta la FACM para los retrasos $k = 1, 2, \dots, 12$ y se dibuja en la gráfica 4.2. Como un ejemplo de los cálculos, se encontrará r_3 , donde: $z_1 = y_1, z_2 = y_2, \dots, z_{120} = y_{120}$ y $\bar{z} = \bar{y}$.

$$\begin{aligned} r_3 &= \frac{\sum_{t=1}^{117} (z_t - \bar{z})(z_{t+3} - \bar{z})}{\sum_{t=1}^{120} (z_t - \bar{z})^2} \\ &= \frac{(z_1 - \bar{z})(z_4 - \bar{z}) + (z_2 - \bar{z})(z_5 - \bar{z}) + \dots + (z_{117} - \bar{z})(z_{120} - \bar{z})}{(z_1 - \bar{z})^2 + (z_2 - \bar{z})^2 + \dots + (z_{120} - \bar{z})^2} \\ &= \frac{(107.77 - 120.39)(107.77 - 120.39) + \dots + (115.80 - 120.39)}{(107.77 - 120.39)^2 + (107.66 - 120.39)^2 + \dots + (115.80 - 120.39)^2} \\ &= 0.8209 \end{aligned}$$

TABLA 4.2 Función de autocorrelación de los datos de la tabla 4.1

k	1	2	3	4	5	6	7	8	9	10	11	12
r_k	0.893	0.853	0.821	0.781	0.772	0.769	0.744	0.711	0.708	0.680	0.691	0.703

Gráfica 4.2 Función de autocorrelación de los datos de la tabla 4.1



Como se nota en la gráfica 4.2, la FACM decrece extremadamente despacio, por lo cual, se supone que la serie es no estacionaria. Ya que se intenta producir estacionaridad, se toma las primeras diferencias de la serie original

$$z_2 = \nabla y_2 = y_2 - y_1 = 107.66 - 107.66 = -0.11$$

$$z_3 = \nabla y_3 = y_3 - y_2 = 113.65 - 107.66 = 5.99$$

.

.

.

$$z_{120} = \nabla y_{120} = y_{120} - y_{119} = 145.80 - 146.50 = -0.70$$

Las primeras diferencias se muestran en la tabla 4.3 y se dibujan en la gráfica 4.1(b). Notar que las primeras diferencias parecen fluctuar alrededor de una media constante muy cercana a cero. Ya que parece que las primeras diferencias son estacionarias, se calculará la FACM de las primeras diferencias para probar si hay estacionaridad o no.

TABLA 4.3 Primeras diferencias de los datos de la tabla 4.1

Mes	t	$Y_t - Y_{t-1}$	Mes	t	$Y_t - Y_{t-1}$
E85	1		E90	61	4.32
F	2	-0.11	F	62	-2.25
M	3	5.99	M	63	9.25
A	4	-5.88	A	64	-9.05
M	5	2.90	M	65	7.21
J	6	-4.93	J	66	-2.13
J	7	3.23	J	67	-3.61
A	8	-2.75	A	68	5.66
S	9	-2.38	S	69	-4.47
O	10	6.86	O	70	8.41
N	11	-2.55	N	71	2.30
D	12	-1.61	D	72	-6.53
E86	13	1.15	E91	73	4.80
F	14	2.20	F	74	-3.34
M	15	-1.54	M	75	-1.00
A	16	6.38	A	76	10.36
M	17	-4.46	M	77	-4.08
J	18	-7.51	J	78	-3.37
J	19	-1.65	J	79	-1.58
A	20	-2.85	A	80	2.00
S	21	-1.56	S	81	-2.10
O	22	-4.01	O	82	8.40
N	23	-0.08	N	83	2.70
D	24	0.60	D	84	-6.00
E87	25	1.45	E92	85	-0.40
F	26	1.25	F	86	-2.50
M	27	7.30	M	87	7.80
A	28	-4.54	A	88	-8.20
M	29	4.14	M	89	5.10
J	30	-2.26	J	90	0.10
J	31	-3.28	J	91	-2.80
A	32	1.19	A	92	-2.80
S	33	1.97	S	93	3.70
O	34	1.36	O	94	1.00
N	35	2.97	N	95	3.10
D	36	2.02	D	96	1.60
E88	37	-4.19	E93	97	-6.10
F	38	1.13	F	98	-3.60
M	39	1.21	M	99	12.20
A	40	-1.18	A	100	-6.90
M	41	-2.50	M	101	-1.80
J	42	-0.61	J	102	1.90
J	43	5.97	J	103	-10.90
A	44	8.73	A	104	0.10
S	45	-1.24	S	105	3.40
O	46	2.43	O	106	0.60
N	47	7.73	N	107	2.10
D	48	-2.30	D	108	13.70
E89	49	-0.37	E94	109	-12.90
F	50	0.06	F	110	-2.40
M	51	0.08	M	111	8.30
A	52	5.47	A	112	3.90
M	53	-2.76	M	113	-6.00
J	54	-1.04	J	114	1.90
J	55	-6.42	J	115	-8.80
A	56	6.90	A	116	8.80
S	57	-3.93	S	117	0.00
O	58	3.71	O	118	-0.30
N	59	5.87	N	119	12.20
D	60	-8.11	D	120	-0.70

Fuente: Banco de México.

La tabla 4.4 muestra la FACM de las primeras diferencias para los retrasos $k = 1, 2, \dots, 12$ y se dibuja en la gráfica 4.3, en la cual se aprecia que r_k es pequeña para $k > 1$. Esto conduce a la hipótesis que la FACM desaparece después de 1 retraso. Es decir: $\rho_k = 0$ para $k > 1$

Sea $q = 1$ y calculando

$$\begin{aligned} 2 \frac{1}{(n-a+1)^{\frac{1}{2}}} \left(1 + 2 \sum_{j=1}^q r_j^2 \right)^{\frac{1}{2}} &= 2 \frac{1}{(120-2+1)^{\frac{1}{2}}} (1 + 2r_1^2)^{\frac{1}{2}} \\ &= 2 \frac{1}{119^{\frac{1}{2}}} (1 + 2(-0.383)^2)^{\frac{1}{2}} \\ &= 0.208 \end{aligned}$$

se ve que:

$$|\rho_k| < 0.208 \text{ para } k > 1$$

entonces se puede concluir que:

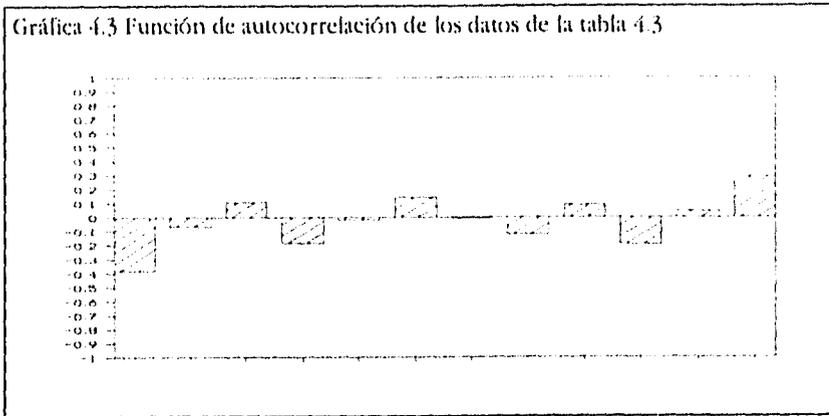
$$\rho_k = 0 \text{ para } k > 1$$

o que la FACM desaparece después de 1 retraso.

TABLA 4.4 FACM de los datos de la tabla 4.3

k	1	2	3	4	5	6	7	8	9	10	11	12
r_k	-0.383	-0.067	0.107	-0.186	-0.019	0.142	-0.009	-0.124	0.088	-0.195	0.043	0.286

Gráfica 4.3 Función de autocorrelación de los datos de la tabla 4.3



Alternativamente se puede calcular:

$$s_{r_k} = \frac{1}{(n - a + 1)^2} \left(1 + 2 \sum_{i=1}^{k-1} r_i^2 \right)^{\frac{1}{2}}$$

$$s_{r_1} = \frac{1}{(120 - 2 + 1)^2} (1 + 2(0))^{\frac{1}{2}}$$

$$= \frac{1}{119^{1/2}}$$

$$= 0.092$$

$$t_{r_1} = \frac{r_1}{s_{r_1}} = \frac{0.383}{0.092}$$

$$= -4.178$$

Ya que $|t_{r_1}| > 2$, se puede concluir que $\rho_1 \neq 0$. También se puede calcular

$$s_{r_k} = \frac{1}{(n - k + 1)^2} \left(1 + 2 \sum_{i=1}^{k-1} r_i^2 \right)^{\frac{1}{2}}$$

$$s_{r_2} = \frac{1}{(120 - 2 + 1)^2} (1 + 2(r_1^2))^{\frac{1}{2}}$$

$$= \frac{1}{119^{1/2}} (1 + 2(-0.067)^2)^{1/2}$$

$$= 0.104$$

Entonces:

$$t_{r_2} = \frac{r_2}{s_{r_2}} = -\frac{0.067}{0.104}$$

$$= -0.642$$

Ya que $|t_{r_2}| < 2$, se concluye que $\beta_2 = 0$. Cálculos similares revelan que $|t_{r_k}| < 2$ para los retrasos $k = 2, \dots, 12$, por lo tanto es razonable concluir que la FACM desaparece después de 1 retraso.

También, es importante analizar la FACP, la cual se muestra en la tabla 4.5 y se dibuja en la gráfica 4.4.

$$r_{kk} = \begin{cases} r_1 & \text{si } k = 1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & \text{si } k = 2, 3, \dots \end{cases}$$

donde: $r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j}$ para $j = 1, 2, \dots, k-1$

$$r_{11} = r_1 = -0.383$$

$$r_{22} = \frac{r_2 - \sum_{j=1}^{2-1} r_{2-1,j} r_{2-j}}{1 - \sum_{j=1}^{2-1} r_{2-1,j} r_j} = \frac{r_2 - r_{11} r_1}{1 - r_{11} r_1} = \frac{(-0.067) - (-0.383)(-0.383)}{1 - (-0.383)(-0.383)} = -0.251$$

$$r_{21} = r_{11} - r_{22} r_{11} = (-0.383) - (-0.251)(-0.383) = -0.479$$

$$r_{33} = \frac{r_3 - \sum_{j=1}^{3-1} r_{3-1,j} r_{3-j}}{1 - \sum_{j=1}^{3-1} r_{3-1,j} r_j} = \frac{r_3 - (r_{21} r_2 + r_{22} r_1)}{1 - r_{21} r_1 + r_{22} r_2} = \frac{(0.107) - [(-0.479)(-0.067) + (-0.251)(-0.383)]}{1 - (-0.479)(-0.383) + (-0.251)(-0.067)} = -0.027$$

$$r_{31} = r_{21} - r_{33} r_{22} = (-0.479) - (-0.027)(-0.251) = -0.485$$

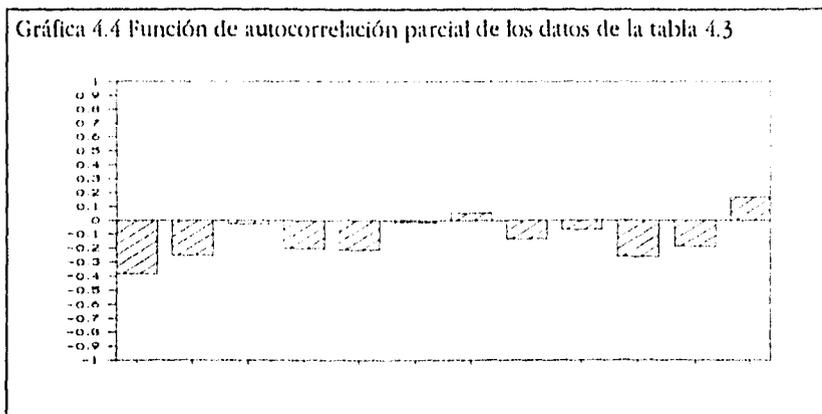
$$r_{32} = r_{22} - r_{33} r_{21} = (-0.027) - (-0.479)(-0.263) = -0.263$$

$$r_{44} = \frac{r_4 - \sum_{j=1}^{4-1} r_{4-1,j} r_{4-j}}{1 - \sum_{j=1}^{4-1} r_{4-1,j} r_j} = \frac{r_4 - (r_{31} r_3 + r_{32} r_2 + r_{33} r_1)}{1 - r_{31} r_1 + r_{32} r_2 + r_{33} r_3} = \frac{(-0.186) - [(-0.485)(0.107) + (-0.263)(-0.067) + (-0.027)(-0.383)]}{1 - (-0.485)(-0.383) + (-0.263)(-0.067) + (-0.027)(0.107)} = -0.202$$

TABLA 4.5 Función de autocorrelación parcial de los datos de la tabla 4.3

k	1	2	3	4	5	6	7	8	9	10	11	12
r_k	-0.383	-0.251	-0.027	-0.202	-0.212	-0.017	0.051	-0.134	-0.068	-0.262	-0.189	0.165

Gráfica 4.4 Función de autocorrelación parcial de los datos de la tabla 4.3



Ahora se puede calcular $t_{r_{11}}$ para los retrasos $k = 1, 2, \dots, 12$

$$t_{r_{11}} = \frac{r_{11}}{\frac{1}{(n-k+1)^2}} = \frac{r_{11}}{0.092} = -4.17$$

como $|t_{r_{11}}| > 2$, se puede concluir que $\rho_{11} \neq 0$

Cálculos similares revelan que $|t_{r_{kk}}| < 2$ para los retrasos $k = 3, 4, 5, 6, 7, 8, 9$

Sin embargo, para los retrasos 2, 10 y 11 $|t_{r_{kk}}|$ es razonablemente cercano a 2, por lo tanto tentativamente se puede concluir que $\rho_{kk} = 0$ para $k = 2, 3, \dots, 12$ y por lo tanto concluir que la FACP desaparece después de 1 retraso o tiende a desaparecer.

Resumiendo, se concluye que la FACM desaparece después de 1 retraso y la FACPM desaparece después de 1 retraso o tiende a desaparecer.

Entonces, se puede suponer que el comportamiento de la FAC y la FACP de las primeras diferencias es idéntico al comportamiento de esas funciones para el modelo MA(1)

$$z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

por lo tanto es razonable concluir que éste modelo ha generado las primeras diferencias de la tabla 4.1. Entonces, sea

$$Z_t = \nabla y_t = y_t - y_{t-1}$$

es razonable concluir que el modelo:

$$z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

$$\forall y_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

$$y_t - y_{t-1} = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

$$y_t = y_{t-1} + \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

ha generado la serie de tiempo de la tabla 4.1

Para hacer la estimación de θ_1 del modelo:

$$z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

y, ya que $\bar{z} = 0.093$ es muy pequeña comparada con las primeras diferencias, se concluye que $\mu = 0$, por lo tanto el modelo:

$$z_t = \epsilon_t - \theta_1 \epsilon_{t-1}$$

describe las primeras diferencias.

Ahora se requiere de un estimador preliminar de θ_1 , como dato inicial para encontrar el estimador mínimo cuadrado de θ_1 .

Recordando que $r_1 = -0.383$ y calculando

$$\frac{-1}{2r_1} + \left[\frac{1}{2r_1^2} - 1 \right]^{1/2} = 2.144 \quad \frac{-1}{2r_1} - \left[\frac{1}{2r_1^2} - 1 \right]^{1/2} = 0.166$$

El estimador $\theta_1 = 2.144$ no satisface la condición de $|\theta_1| < 1$ por lo tanto el estimador $\theta_1 = 0.166$ es utilizado como dato de entrada en algún programa de computo para que después de varias iteraciones, el estimador final sea:

$$\theta_1 = 0.4336$$

Por lo tanto el modelo:

$$\hat{c}_t = c_t - (0.4336)c_{t-1}$$

describe las primeras diferencias de la serie de la tabla 4.1. y por consiguiente el modelo:

$$y_t = \gamma_t + c_t - (0.4336)c_{t-1}$$

describe la serie de la tabla 4.1.

El calculo de los residuales es como sigue:

c_1 no se puede calcular, sin embargo, para $t = 2$

$$\hat{y}_2 = \gamma_1 + c_2 + (0.4336)c_1$$

ya que no se tiene información para c_1 y c_2 el mejor estimador para ambos es cero. Entonces:

$$\hat{y}_2 = y_1 = 107.77$$

De la tabla 4.1 $y_2 = 107.66$ entonces se tiene el residual:

$$y_2 - \hat{y}_2 = 107.66 - 107.77 = -0.11$$

Ahora se puede usar este valor como una estimación de c_2 . Similarmente, para $t = 3$

$$\hat{y}_3 = \gamma_2 + c_3 + (0.4336)c_2$$

ya que el mejor estimador para c_3 es cero, se tiene:

$$\hat{y}_3 = 107.66 + (0.4336)(-0.11) = 107.72$$

De la tabla 4.1 $y_3 = 113.65$ entonces se tiene el residual:

$$y_3 - \hat{y}_3 = 113.65 - 107.72 = 5.93$$

Ahora se puede usar este valor como una estimación de c_3 . El resto de los residuales se pueden estimar de una manera similar.

$$\begin{aligned}
 \hat{y}_{120} &= y_{119} + (0.4336)\epsilon_{119} \\
 &= 133.73 + (0.4336)(12.77) \\
 &= 139.2683
 \end{aligned}$$

Las estimaciones de y_t se muestran en la gráfica 4.5

Ya que $y_{120} = 145.80$ se obtiene el residual:

$$y_{120} - \hat{y}_{120} = 145.80 - 139.26 = 6.53$$

Entonces un pronóstico para y_{121} es:

$$\begin{aligned}
 \hat{y}_{121}(120) &= y_{120} + \epsilon_{121} + (0.4336)\epsilon_{120} \\
 &= 145.80 + 0 + (0.4336)(6.53) \\
 &= 142.10
 \end{aligned}$$

El pronóstico para y_{122} es:

$$\begin{aligned}
 \hat{y}_{122}(120) &= y_{121} + \epsilon_{122} + (0.4336)\epsilon_{121} \\
 &= \hat{y}_{121}(120) + 0 + 0 \\
 &= 142.10
 \end{aligned}$$

Notar que se usó $\hat{y}_{121}(120)$ como la estimación de y_{121} .

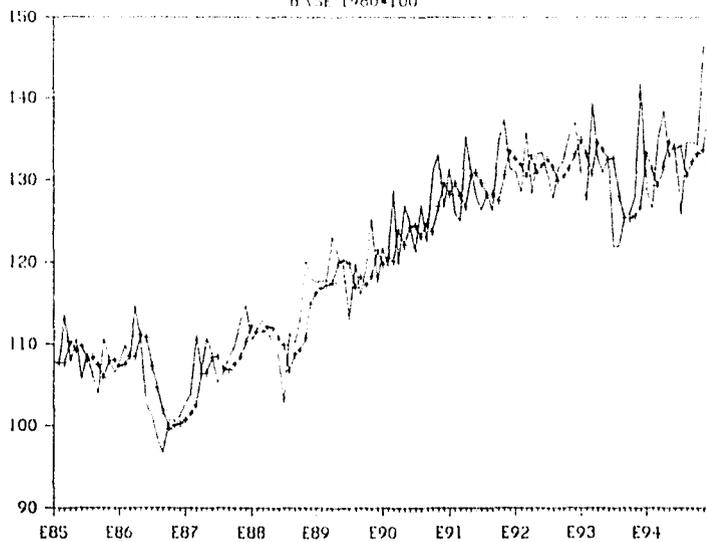
En general, y_{120+t} para $t \geq 2$, es:

$$\begin{aligned}
 \hat{y}_{120+t}(120) &= y_{120+t-1} + \epsilon_{120+t} + (0.4336)\epsilon_{120+t-1} \\
 &= \hat{y}_{120+t-1}(120) \\
 &= 142.10
 \end{aligned}$$

Gráfica 4.5

INDICE DE PRODUCCION MANUFACTURERA

BASE 1980=100



— Observada + Estimada

4-2 MODELOS BOX-JENKINS MAS COMUNES.

La metodología Box-Jenkins elige de entre una clasificación de modelos de series de tiempo estacionarios un modelo particular, el cual utiliza para pronosticar valores futuros de la serie de tiempo. En esta sección se analizarán los modelos más comunes de la metodología Box-Jenkins. Se presentan modelos utilizados en la estimación de series de tiempo no estacionales.

4-2.1 MODELOS DE PROMEDIOS MOVILES.

El modelo $MA(q)$

$$z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$$

es llamado de promedios móviles de orden q , el signo menos se introduce por convención. Se puede mostrar que no se deben imponer condiciones a los parámetros $\theta_1, \theta_2, \dots, \theta_q$ para que este modelo sea estacionario, pero existen condiciones que se deben aplicar a los parámetros para que el modelo sea invertible. También se puede mostrar que la FACPT de este modelo es decreciente y que la FACT tiende a desaparecer después de q retrasos [2].

Un caso particular del modelo de promedios móviles, es el de promedios móviles de orden uno, $MA(1)$

$$z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$$

La condición que debe cumplir el parámetro θ_1 para que el modelo sea invertible es:

$$|\theta_1| < 1$$

Se puede mostrar que la media del modelo es μ , que la FACPT decrece en forma exponencial y que la FACT tiende a desaparecer después de un retraso [2]. En particular se puede mostrar que:

$$\rho_k = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2} & \text{si } k = 1 \\ 0 & \text{si } k > 1 \end{cases}$$

Por lo tanto:

$$\rho_t = \frac{-\theta_t}{1 + \theta_t^2}$$

resolviendo para θ_t en términos de ρ_t se tiene:

$$\theta_t = \frac{-1}{2\rho_t} \pm \left[\frac{1}{(2\rho_t)^2} - 1 \right]^{1/2}$$

reemplazando ρ_t con r_t se obtiene el estimador preliminar de θ_t

$$\hat{\theta}_t = \frac{-1}{2r_t} \pm \left[\frac{1}{(2r_t)^2} - 1 \right]^{1/2}$$

Para determinar cual de esos estimadores preliminares se debe usar, se aplica la condición de invertibilidad, en este sentido, se escoge $\hat{\theta}_t$ tal que satisfaga la condición

$$|\hat{\theta}_t| < 1$$

Por último, un estimador preliminar para μ puede ser:

$$\bar{z} = \frac{\sum_{t=a}^n z_t}{n - a + 1}$$

Otro caso particular del modelo de promedios móviles, es el de promedios móviles de orden dos, $MA(2)$

$$z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}$$

Las condiciones que se deben aplicar a los parámetros θ_1 y θ_2 para que el modelo sea invertible son:

$$\theta_1 + \theta_2 < 1$$

$$\theta_2 - \theta_1 < 1$$

$$|\theta_2| < 1$$

Se puede mostrar que la media del modelo es μ , que la FACPT decrece siguiendo una mezcla de exponenciales y/o senos y que la FACT tiende a desaparecer después de dos retrasos [2]. En particular se puede demostrar que:

$$\rho_k = \begin{cases} \frac{-\theta_1(1-\theta_1)}{1+\theta_1^2+\theta_2^2} & \text{si } k=1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{si } k=2 \\ 0 & \text{si } k>2 \end{cases}$$

Estas relaciones se pueden usar para encontrar los parámetros θ_1 y θ_2 en términos de ρ_1 y ρ_2 , los cuales son estimados por r_1 y r_2 y así producir un conjunto de estimadores preliminares $\hat{\theta}_1$ y $\hat{\theta}_2$ que satisfagan las condiciones de invertibilidad.

Por último, un estimador preliminar para μ puede ser:

$$\bar{z} = \frac{\sum_{t=a}^n z_t}{n-a+1}$$

4-2.2 MODELOS AUTORREGRESIVOS.

El modelo $AR(p)$

$$z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + \epsilon_t$$

es llamado proceso autorregresivo de orden p . El término autorregresivo se usa porque z_t es "retrasado" o expresado como una función de $z_{t-1}, z_{t-2}, \dots, z_{t-p}$. Se puede mostrar que no se deben imponer condiciones a los parámetros $\phi_1, \phi_2, \dots, \phi_p$ para que este modelo sea invertible, pero existen condiciones que se deben aplicar a los parámetros para que el modelo sea estacionario. También se puede mostrar que la FACPT de este modelo tiende a desaparecer después de p retrasos y que la FACT es decreciente [2].

Un caso particular del modelo autorregresivo, es el autorregresivo de orden uno, $AR(1)$

$$z_t = \delta + \phi_1 z_{t-1} + \epsilon_t$$

La condición que debe cumplir el parámetro ϕ_1 para que el modelo sea estacionario es:

$$|\phi_1| < 1$$

Se puede mostrar que:

$$\mu = \frac{b}{1 - \phi_1}$$

y que la FACPT tiende a desaparecer después de un retraso y que la FACT decrece en forma exponencial [2]. En particular se puede mostrar que:

$$\rho_k = \phi_1^k \quad \text{para} \quad k \geq 1$$

Por lo tanto:

$$\rho_1 = \phi_1$$

resolviendo para ϕ_1 en términos de ρ_1 , el cual es estimado por r_1 , se puede calcular el estimador preliminar de ϕ_1

$$\hat{\phi}_1 = r_1$$

Este estimador preliminar debe satisfacer la condición de estacionaridad

$$|\hat{\phi}_1| < 1$$

También se puede ver que

$$\delta = \mu(1 - \phi_1)$$

Así un estimador preliminar de δ es

$$\hat{\delta} = \bar{z}(1 - \hat{\phi}_1)$$

donde

$$\bar{z} = \frac{\sum_{t=0}^n z_t}{n - a + 1}$$

Otro caso particular del modelo autorregresivo, es el autorregresivo de orden dos, $AR(2)$

$$z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \epsilon_t$$

Las condiciones que se deben aplicar a los parámetros ϕ_1 y ϕ_2 para que el modelo sea estacionario son:

$$\phi_1 + \phi_2 < 1 \qquad \phi_2 - \phi_1 < 1 \qquad |\phi_2| < 1$$

Se puede mostrar que la media de este modelo es

$$\mu = \frac{\delta}{1 - \phi_1 - \phi_2}$$

y que la FACPT tiende a desaparecer después de dos retraso y que la FACT decrece siguiendo una mezcla de exponenciales y/o senos. En particular se pueden mostrar las siguientes relaciones para la FACT [2].

Primero se resuelven las ecuaciones:

$$\rho_1 = \phi_1 + \phi_2 \rho_1 \quad \gamma \quad \rho_2 = \phi_1 \rho_2 + \phi_2$$

para ρ_1 y ρ_2 en términos de ϕ_1 y ϕ_2 . Esto nos lleva a:

$$\rho_1 = \frac{\phi_1}{1 - \phi_2} \quad \gamma \quad \rho_2 = \frac{\phi_2}{1 - \phi_2} + \phi_2$$

Entonces se puede obtener ρ_k para $k \geq 3$ usando la ecuación

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}$$

Las ecuaciones $\rho_1 = \phi_1 + \phi_2 \rho_1$ y $\rho_2 = \phi_1 \rho_2 + \phi_2$ generalmente llamadas Ecuaciones de Yule-Walker, se pueden resolver para ϕ_1 y ϕ_2 en términos de ρ_1 y ρ_2 , los cuales son estimados por r_1 y r_2 y de esta manera se puede calcular los estimadores preliminares de ϕ_1 y ϕ_2

$$\hat{\phi}_1 = r_1 \left(\frac{1 - r_2}{1 - r_1^2} \right) \quad \gamma \quad \hat{\phi}_2 = \frac{r_2 - r_1^2}{1 - r_1^2}$$

Estos estimadores preliminares deben satisfacer las condiciones de estacionaridad:

$$\hat{\phi}_1 + \hat{\phi}_2 < 1$$

$$\hat{\phi}_2 - \hat{\phi}_1 < 1$$

$$|\hat{\phi}_2| < 1$$

También, la media del modelo $AR(2)$ es

$$\mu = \frac{\delta}{1 - \hat{\phi}_1 - \hat{\phi}_2}$$

Esto implica que $\delta = \mu(1 - \hat{\phi}_1 - \hat{\phi}_2)$ y así, un estimador de δ es

$$\hat{\delta} = \bar{z}(1 - \hat{\phi}_1 - \hat{\phi}_2)$$

donde

$$\bar{z} = \frac{\sum_{t=\alpha}^n z_t}{n - \alpha + 1}$$

4.2.3 MODELOS AUTORREGRESIVOS Y DE PROMEDIOS MOVILES.

El modelo $ARMA(p, q)$

$$z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} + \epsilon_t$$

es llamado autorregresivo y de promedios móviles de orden (p, q) . Las condiciones de estacionaridad del modelo $AR(p)$ y las condiciones de invertibilidad del modelo $MA(q)$ son las mismas condiciones de estacionaridad e invertibilidad respectivamente para el modelo $ARMA(p, q)$. La FACPT y la FACT son decrecientes.

Un modelo particular de gran utilidad es el $ARMA(1, 1)$

$$z_t = \delta + \phi_1 z_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}$$

Este modelo es estacionario si $|\phi_1| < 1$ e invertible si $|\theta_1| < 1$.

Se puede mostrar que la media de este modelo es

$$\mu = \frac{\delta}{1 - \phi_1}$$

y que la FACPT y la FACT decrecen en forma exponencial [2]. En particular

$$\rho_k = \begin{cases} \frac{(1 - \phi_1 \theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\theta_1 \phi_1} & \text{si } k = 1 \\ \phi_1 \rho_1 & \text{si } k = 2 \\ \phi_1 \rho_{k-1} & \text{si } k \geq 3 \end{cases}$$

Estas relaciones se pueden utilizar para determinar ϕ_1 y θ_1 en términos de ρ_1 y ρ_2 los cuales están estimados por r_1 y r_2 , y así producir un conjunto de estimadores preliminares $\hat{\phi}_1$ y $\hat{\theta}_1$ que satisfagan las condiciones de estacionaridad e invertibilidad

$$|\hat{\phi}_1| < 1 \quad \text{y} \quad |\hat{\theta}_1| < 1$$

También, ya que la media de modelo es

$$\mu = \frac{b}{1 - \phi_1}$$

implica que

$$b = \mu(1 - \phi_1)$$

un estimador preliminar de b es

$$\hat{b} = \bar{z}(1 - \hat{\phi}_1)$$

donde

$$\bar{z} = \frac{\sum_{t=1}^n z_t}{n - \alpha + 1}$$

4-2.4 RESUMEN DE MODELOS BOX-JENKINS NO ESTACIONALES.

Ya se habló de algunos modelos usados por la metodología Box-Jenkins para estimar series de tiempo que no poseen variación estacional. Para iniciar la discusión de modelos estacionales, se resumirán los modelos no estacionales y sus propiedades.

TABLA 4.6 Modelos generales.

MODELO	FACPT	FACT	CONDICIONES DE ESTACIONARIDAD	CONDICIONES DE INVERTIBILIDAD
El modelo $MA(q)$ $z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}$	Decrece	Desaparece después de q retrasos	No	Si
El modelo $AR(p)$ $z_t = \delta + \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + \epsilon_t$	Desaparece después de p retrasos	Decrece	Si	No
El modelo $ARMA(p, q)$ $z_t = \delta + \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} + \epsilon_t$	Decrece	Decrece	Si	Si

TABLA 4.7 Modelos específicos.

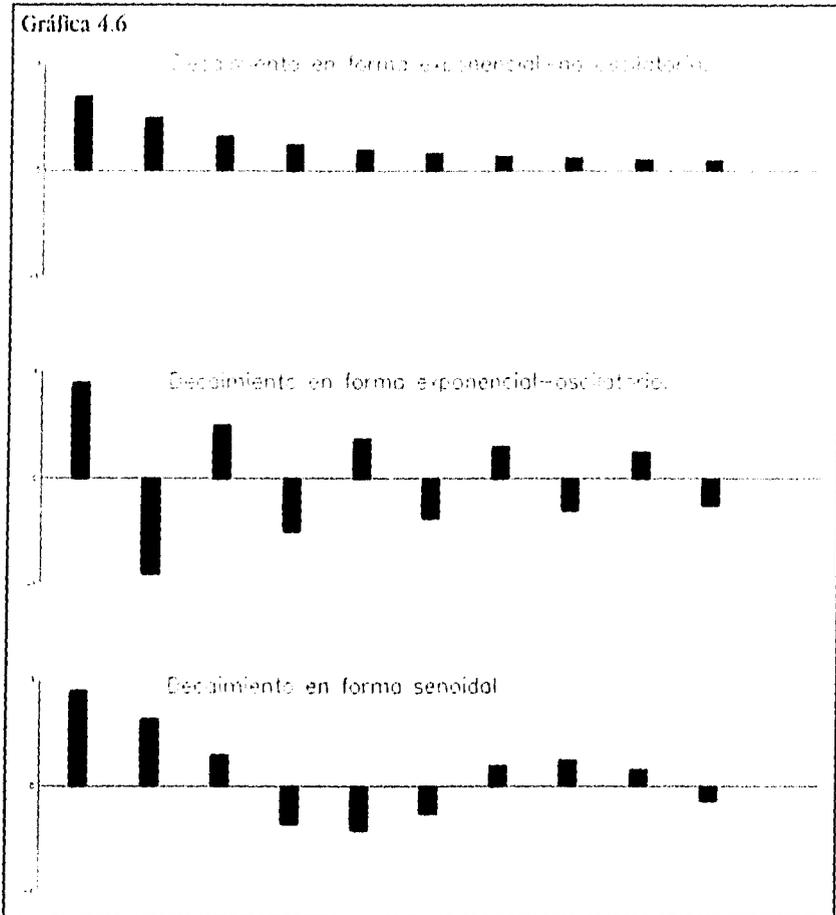
MODELO	FACPT	FACT	CONDICIONES DE ESTACIONARIDAD	CONDICIONES DE INVERTIBILIDAD
El modelo $MA(1)$ $z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1}$	Decrece en forma exponencial	Desaparece después de 1 retraso	Ninguna	$ \theta_1 < 1$
El modelo $MA(2)$ $z_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}$	Decrece de acuerdo a una mezcla de exponenciales y/o senos	Desaparece después de 2 retrasos	Ninguna	$0_1 + 0_2 < 1$ $0_2 - 0_1 < 1$ $ \theta_1 < 1$
El modelo $AR(1)$ $z_t = \delta + \phi_1 z_{t-1} + \epsilon_t$	Desaparece después de 1 retraso	Decrece en forma exponencial	$ \phi_1 < 1$	Ninguna
El modelo $AR(2)$ $z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \epsilon_t$	Desaparece después de 2 retrasos	Decrece de acuerdo a una mezcla de exponenciales y/o senos	$0_1 + 0_2 < 1$ $0_2 - 0_1 < 1$ $ \phi_1 < 1$	Ninguna
El modelo $ARMA(1, 1)$ $z_t = \delta + \phi_1 z_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}$	Decrece en forma exponencial	Decrece en forma exponencial	$ \phi_1 < 1$	$ \theta_1 < 1$

La tabla 4.6 presenta los modelos $AR(p)$, $MA(q)$ y $ARMA(p, q)$ e indica si o no un modelo requiere condiciones de estacionaridad e invertibilidad e incluye una descripción del comportamiento de la FACT y la FACPT de cada modelo. La tabla 4.7

presenta modelos específicos, indica condiciones de estacionaridad e invertibilidad requeridas para cada modelo e incluye una descripción del comportamiento de la FACT y la FACPT de cada modelo.

Examinando una y otra tabla se ve que ninguno de estos modelos tienen una FACT y una FACPT que tiendan ambas a desaparecer. Este hecho es de mucha ayuda en la identificación de modelos Box-Jenkins más apropiados.

El significado intuitivo de las expresiones "modelo exponencial decreciente" y "modelo senoidal decreciente" se ilustra en la gráfica 4.6.



Considérese primero el proceso $AR(p)$. La diferencia entre los modelos en los cuales la FACT de los procesos $AR(1)$ y $AR(2)$ decrecen, pueden usarse algunas veces para explicar, si o no una serie de tiempo fué generada por un proceso $AR(1)$ o $AR(2)$ sin examinar la FACPT. En la tabla 4.7 es claro que la FACT de los procesos $AR(1)$ y $AR(2)$ pueden decrecer de forma similar, por esta razón, es algunas veces muy difícil o imposible explicar las diferencias entre ambos procesos examinando solamente la FACM. Por lo tanto, es importante también examinar la FACPM y determinar si tiende a desaparecer después de uno o dos retrasos. Similarmente, la diferencia entre los modelos en los cuales la FACPT de los procesos $MA(1)$ y $MA(2)$ decrecen, pueden algunas veces usarse para explicar, sin examinar la FACT, si o no una serie de tiempo fué generada por un proceso $MA(1)$ o $MA(2)$. Como se ve en la tabla 4.7 la FACT de los procesos $MA(1)$ y $MA(2)$ pueden decrecer de forma similar, esto es, algunas veces es muy difícil explicar las diferencias entre ambos procesos examinando solamente la FACPM. Por lo tanto, es importante también examinar la FACM y determinar si tiende a desaparecer después de uno o dos retrasos.

Hay que notar que los modelos usados por la metodología Box-Jenkins contienen relativamente pocos parámetros. Por ejemplo, ninguno de los modelos presentados en la tabla 4.7 poseen más de tres parámetros, en general, un objetivo importante de la metodología Box-Jenkins es el desarrollo de modelos parsimoniosos, esto es, modelos que describan adecuadamente a las series de tiempo y además utilicen pocos parámetros. En particular, si dos modelos diferentes describen igualmente bien una serie de tiempo dada y además uno de los modelos contiene menos parámetros que el otro, entonces se sugiere escoger el modelo con menos parámetros.

4.2.5 MODELOS ESTACIONALES.

Ahora se discutirán modelos estacionales usando la metodología Box-Jenkins², para simplificar este estudio es necesario familiarizarse con los operadores de retraso: B (Backshift), el cual resta un periodo al subíndice de la observación y B^k el cual resta k periodos al subíndice de la observación. Esto es:

$$By_t = y_{t-1}$$

y

$$B^k y_t = y_{t-k}$$

Como en el análisis de series de tiempo no estacionales, el primer paso es encontrar una serie estacionaria z_0, z_1, \dots, z_n . Supóngase la serie de tiempo y_1, y_2, \dots, y_n recordar que si esta no posee variación estacional y es no estacionaria, entonces se puede transformar en una serie de tiempo estacionaria usando la transformación

² Otra discusión y ejemplos de modelos estacionales se presenta en: Nelson, C. "Applied Time Series Analysis for Managerial Forecasting". Holden-Day, Inc., San Francisco, 1973.

$$z_t = \nabla y_t = y_t - y_{t-1}$$

o la transformación

$$z_t = \nabla^2 y_t = y_t - 2y_{t-1} + y_{t-2}$$

Si la serie y_1, y_2, \dots, y_n posee variación estacional, puede ser no estacionaria y es posible que una de las transformaciones $z_t = \nabla y_t$ o $z_t = \nabla^2 y_t$ produzca estacionaridad. Sin embargo frecuentemente se requiere del "diferenciador estacional" para poder producir series estacionarias. Para estudiar el diferenciador estacional se define el operador ∇ como

$$\nabla = 1 - B$$

Entonces

$$\nabla y_t = (1 - B)y_t = y_t - By_t = y_t - y_{t-1}$$

y

$$\begin{aligned} \nabla^2 y_t &= (1 - B)^2 y_t \\ &= (1 - 2B + B^2)y_t \\ &= y_t - 2By_t + B^2 y_t \\ &= y_t - 2y_{t-1} + y_{t-2} \end{aligned}$$

En general

$$\nabla^d y_t = (1 - B)^d y_t$$

donde d es el grado de diferenciación no estacional requerido para producir series de tiempo estacionarias.

Ahora, sea L el número de periodos en un año, así $L = 12$ para datos mensuales y $L = 4$ para datos trimestrales. Se define el operador estacional ∇_L como

$$\nabla_L = (1 - B^L)$$

Entonces

$$\nabla_L y_t = (1 - B^L)y_t = y_t - B^L y_t = y_t - y_{t-L}$$

Similarmente

$$\begin{aligned}I_1^2 y_t &= (1 - B^L)^2 y_t \\ &= (1 - 2B^L + B^{2L}) y_t \\ &= y_t - 2B^L y_t + B^{2L} y_t \\ &= y_t - 2y_{t-L} + y_{t-2L}\end{aligned}$$

En general

$$I_1^D y_t = (1 - B^L)^D y_t$$

donde D es el grado de diferenciación estacional requerido para producir series de tiempo estacionarias.

En general cuando se usa la metodología Box-Jenkins para analizar una serie de tiempo que posee variación estacional, primero se grafican los valores de la serie y se determina si esta variación crece o decrece conforme avanza el tiempo. Si la variabilidad de una serie de tiempo estacional es constante en el tiempo, entonces la serie posee estacionalidad aditiva. Si la variabilidad de una serie de tiempo estacional cambia con el tiempo, entonces la serie posee estacionalidad multiplicativa.

Si la serie de tiempo y_1, y_2, \dots, y_n parece exhibir variación estacional multiplicativa, se aplica la transformación

$$y_t^* = \ln y_t$$

para producir una serie $y_1^*, y_2^*, \dots, y_n^*$ con variación estacional aditiva.

Ahora se verá la forma de transformar la serie $y_1^*, y_2^*, \dots, y_n^*$ en una serie estacionaria. En general la transformación que produce series de tiempo estacionarias es de la forma:

$$z_t = \nabla_t^D \nabla^d y_t^* = (1 - B^L)^D (1 - B)^d y_t^*$$

Donde se supone que la transformación $y_t^* = \ln y_t$ ya ha sido aplicada, esto es porque la mayoría de las series de tiempo económicas poseen variación estacional multiplicativa. Para series de tiempo con variación estacional aditiva no se aplica dicha transformación.

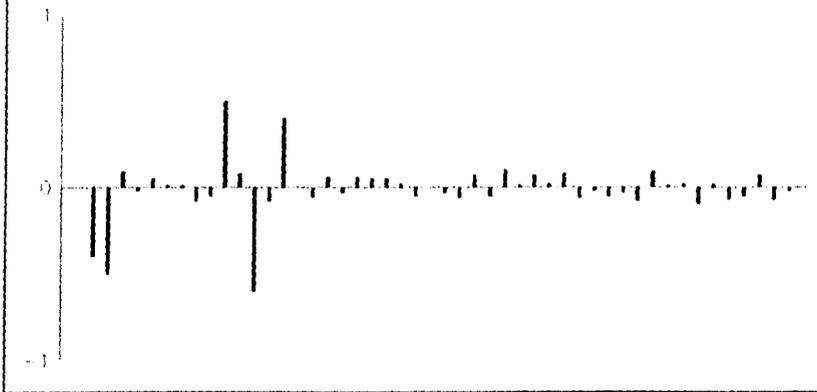
Considerando las particularidades de esta transformación, si hubiera sucedido que la serie z_a, z_{a+1}, \dots, z_n producida por alguna transformación tiene una FACM que decrece extremadamente rápido, entonces la serie se puede considerar estacionaria; en cambio si su FACM decrece extremadamente despacio se considera no estacionaria. El significado de los términos decrece extremadamente rápido y decrece extremadamente despacio es aún más arbitrario para series estacionales que para no estacionales y debe determinarse en base a la experiencia. La FACM y la FACPM se calculan exactamente como se calcularon para series no estacionales y se aplican las mismas reglas empíricas.

Generalmente, cuando se estiman series de tiempo no estacionales, la FACM y la FACPM se estiman con 12 o menos retrasos, una razón de esto es que si la serie no posee variación estacional y si alguna de las funciones de autocorrelación tienden a desaparecer, esto sucederá generalmente después de q retrasos, con $q \leq 2$. Por otro lado, si una serie de tiempo estacionaria posee variación estacional y si alguna de las funciones de autocorrelación tienden a desaparecer, esto sucederá generalmente después de q retrasos, con $q \leq 2L + 2$ y muy frecuentemente $q \leq L + 2$.

Se debe notar que si la FACT o la FACPT tienden a desaparecer después de $L + 2$ (o $2L + 2$) retrasos, algunas de las autocorrelaciones teóricas o algunas de las autocorrelaciones parciales teóricas con retrasos menores que $L + 2$ (o $2L + 2$) serán diferentes de cero, mientras otras serán igual a cero. En este sentido cuando se dice que la FACM (o la FACPM) tienen picos en un conjunto dado de retrasos, significa que las autocorrelaciones muestrales (o que las autocorrelaciones parciales muestrales) en este conjunto de retrasos son mayores en valor absoluto que las autocorrelaciones muestrales (o las autocorrelaciones parciales muestrales) en otros retrasos. Picos en la FACM (o la FACPM) indican autocorrelaciones (o autocorrelaciones parciales) distintas de cero.

Por ejemplo, considérese una serie mensual estacional ($L=12$) y la FACM de un conjunto particular de valores z_a, z_{a+1}, \dots, z_n tiene picos en los retrasos 1, 2, 10, 12 y 14 y tiende a desaparecer después del retraso 14 como ilustra la gráfica 11.1, es razonable creer que la serie z_a, z_{a+1}, \dots, z_n es estacionaria.

Gráfica 4.7



Para determinar la transformación

$$z_t = \nabla_t^D \nabla_t^d y_t^* = (1 - B^L)^D (1 - B)^d y_t^*$$

que se requiere para producir estacionaridad en una serie estacional, la FACM y la FACPM usualmente se calculan para todos los retrasos de 1 a $4L$. Para la gran mayoría de series de tiempo económicas con variación estacional generalmente d es 0 ó 1 y D es 0 ó 1.

Si $d = 0$ y $D = 0$ entonces

$$\begin{aligned} z_t &= \nabla_t^0 \nabla_t^0 y_t^* \\ &= (1 - B^L)^0 (1 - B)^0 y_t^* \\ &= y_t^* \end{aligned}$$

Por lo tanto

$$z_a = z_1 = y_1^*$$

$$z_{a+1} = z_2 = y_2^*$$

.

.

.

$$z_n = y_n^*$$

Si la FACM de la serie $y_1^*, y_2^*, \dots, y_n^*$ tiende a desaparecer o disminuir rápidamente entonces se considera estacionaria.

Si la serie $y_1^*, y_2^*, \dots, y_n^*$ es no estacionaria se debe considerar una de las siguientes transformaciones: $d = 1$ y $D = 0$ o $d = 0$ y $D = 1$, o $d = 1$ y $D = 1$

Si $d = 1$ y $D = 0$, entonces

$$\begin{aligned} z_t &= \nabla_t^0 \nabla^1 y_t^* \\ &= (1 - B^t)^0 (1 - B)^1 y_t^* \\ &= (1 - B) y_t^* \\ &= y_t^* - B y_t^* \\ &= y_t^* - y_{t-1}^* \end{aligned}$$

Por lo tanto

$$z_a = z_2 = y_2^* - y_1^*$$

$$z_{a+1} = z_3 = y_3^* - y_2^*$$

.

.

.

$$z_n = y_n^* - y_{n-1}^*$$

Si $d = 1$ y $D = 1$, entonces

$$\begin{aligned}
 z_t &= \nabla_t^1 \nabla_t^1 y_t^* \\
 &= (1 - B^L)^1 (1 - B)^1 y_t^* \\
 &= (1 - B^L)(y_t^* - y_{t-1}^*) \\
 &= (y_t^* - y_{t-1}^*) - B^L(y_t^* - y_{t-1}^*) \\
 &= y_t^* - y_{t-1}^* - B^L y_t^* + B^L y_{t-1}^* \\
 &= y_t^* - y_{t-1}^* - y_{t-L}^* + y_{t-1-L}^*
 \end{aligned}$$

Por lo tanto

$$\begin{aligned}
 z_n &= z_{t-2} = y_{t-2}^* - y_{t-1}^* - y_2^* + y_1^* \\
 z_{n-1} &= z_{t-3} = y_{t-3}^* - y_{t-2}^* - y_3^* + y_2^* \\
 &\vdots \\
 &\vdots \\
 &\vdots \\
 z_n &= y_n^* - y_{n-1}^* - y_{n-L}^* + y_{n-1-L}^*
 \end{aligned}$$

Si la FACM de estos valores es decreciente o tiende a desaparecer rápidamente, entonces la serie se considera estacionaria.

Si $d = 0$ y $D = 1$, entonces

$$\begin{aligned}
 z_t &= \nabla_t^1 \nabla_t^0 y_t^* \\
 &= (1 - B^L)^1 (1 - B)^0 y_t^* \\
 &= (1 - B^L) y_t^* \\
 &= y_t^* - B^L y_t^* \\
 &= y_t^* - y_{t-L}^*
 \end{aligned}$$

Por lo tanto

$$\begin{aligned}
z_a &= z_{t+1} = Y_{t+1}^* - Y_t^* \\
z_{a+1} &= z_{t+2} = Y_{t+2}^* - Y_{t+1}^* \\
&\vdots \\
&\vdots \\
&\vdots \\
z_n &= Y_n^* - Y_{n-1}^*
\end{aligned}$$

Si la FACM de estos valores es decreciente o tiende a desaparecer rápidamente, entonces la serie se considera estacionaria.

4-2.5.1 DIFERENTES TIPOS DE ESTACIONARIDAD.

Ya que la serie de tiempo estacional ha sido transformada a una serie de tiempo estacionaria z_a, z_{a+1}, \dots, z_n se debe identificar el modelo particular de estacionaridad que se supone generó los valores z_a, z_{a+1}, \dots, z_n .

Para iniciar esta discusión, se debe recordar el modelo $ARMA(p, q)$

$$\begin{aligned}
z_t &= b + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} \\
&\quad - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} + \epsilon_t
\end{aligned}$$

Esto es equivalente a:

$$\begin{aligned}
z_t - \phi_1 z_{t-1} - \phi_2 z_{t-2} - \dots - \phi_p z_{t-p} &= \\
b + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} & \\
\text{O} & \\
z_t - \phi_1 B z_t - \phi_2 B^2 z_t - \dots - \phi_p B^p z_t &= \\
b + \epsilon_t - \theta_1 B \epsilon_t - \theta_2 B^2 \epsilon_t - \dots - \theta_q B^q \epsilon_t & \\
\text{O} & \\
(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) z_t &= \\
b + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \epsilon_t &
\end{aligned}$$

Bajo esta notación, se le llama a

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

operador autorregresivo no estacional de orden p y a

$$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

operador de promedios móviles no estacional de orden q

De igual forma, se le llama a

$$\phi_P(B^L) = (1 - \phi_{1,L} B^L - \phi_{2,L} B^{2L} - \dots - \phi_{P,L} B^{PL})$$

operador autorregresivo estacional de orden P y a

$$\theta_Q(B^L) = (1 - \theta_{1,L} B^L - \theta_{2,L} B^{2L} - \dots - \theta_{Q,L} B^{QL})$$

operador de promedios móviles estacional de orden Q

Combinando estos cuatro operadores se obtiene el modelo general multiplicativo estacional

$$\phi_p(B)\phi_P(B^L)\theta_q(B)\theta_Q(B^L)z_t = b + \epsilon_t$$

o

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \times (1 - \phi_{1,L} B^L - \phi_{2,L} B^{2L} - \dots - \phi_{P,L} B^{PL}) z_t = b + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \times (1 - \theta_{1,L} B^L - \theta_{2,L} B^{2L} - \dots - \theta_{Q,L} B^{QL}) \epsilon_t$$

también llamado modelo mixto autorregresivo y de promedios móviles de orden (p, P, q, Q) .

La identificación de la forma particular que debe tomar este modelo para describir adecuadamente una serie de tiempo estacionaria con variación estacional, involucra la determinación de los operadores que deberían estar en el modelo y el orden de cada operador. Esta identificación se puede hacer teóricamente comparando la FACM y la FACMP de una serie de tiempo estacionaria con la FACT y la FACTP de varias formas particulares del modelo general estacional multiplicativo.

Box-Jenkins presenta un resumen de las funciones de autocorrelación teóricas de varias formas particulares y se utiliza para identificar modelos. Sin embargo, en la práctica, esta comparación puede ser difícil, por esta razón el método que se usa para desarrollar un modelo apropiado, es algunas veces intuitivo y generalmente efectúa varias iteraciones del proceso. En cada etapa se usará la FACM y la FACMP de la serie observada para identificar los componentes más convenientes para un buen modelo; entonces se utiliza el análisis de residuales con el modelo adoptado para desarrollar un modelo más apropiado.

A continuación se enumeran algunos lineamientos generales para identificar cuales componentes $\phi_p(B^L)$, $\phi_p(B)$, $\theta_q(B^L)$, y $\theta_q(B)$ se usarán en el modelo general para representar adecuadamente a una serie de tiempo estacionaria.

Primero, si la FACM de la serie observada decrece rápidamente y la FACMP tiende a desaparecer después de algún retraso que es sustancialmente menor que L , entonces alguna forma del operador

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

puede ser adecuada. El retraso en el cual la FACMP tiende a desaparecer, sugiere el orden p del operador $\phi_p(B)$ a utilizar.

Segundo, si la FACM de la serie observada decrece rápidamente y la FACMP tiende a desaparecer en un retraso que es casi igual a L , $L+1$, $L+2$, $2L$, $2L+1$ o $2L+2$, entonces alguna forma del operador

$$\phi_p(B^L) = (1 - \phi_{1,L} B^L - \phi_{2,L} B^{2L} - \dots - \phi_{p,L} B^{pL})$$

puede ser adecuada si no existieran picos en retrasos sustancialmente menores que L .

Sin embargo, si tales picos existieran, entonces alguna forma del operador

$$\phi_p(B)\phi_p(B^L) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \times (1 - \phi_{1,L} B^L - \phi_{2,L} B^{2L} - \dots - \phi_{p,L} B^{pL})$$

puede ser adecuada. Los retrasos en los cuales existen picos en la FACMP ayudan a sugerir el orden P del operador autorregresivo estacional o los ordenes de los operadores a usarse en el producto.

Tercero, si la FACMP de la serie observada decrece rápidamente y la FACMP tiende a desaparecer después de algún retraso que es sustancialmente menor que L , entonces alguna forma del operador

$$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$$

puede ser adecuada. El retraso en el cual la FACM tiende a desaparecer sugiere el orden q del operador $\theta_q(b)$ a utilizar.

Cuarto, si la FACMP de la serie observada decrece rápidamente y la FACMP tiende a desaparecer en un retraso que es casi igual a $L, L+1, L+2, 2L, 2L+1$ o $2L+2$, entonces alguna forma del operador

$$\theta_q(B^L) = (1 - 0_{1,L}B^L - 0_{2,L}B^{2L} - \dots - 0_{q,L}B^{qL})$$

puede ser adecuada si no existieran picos en retrasos sustancialmente menores que L .

Sin embargo, si tales picos existieran, entonces alguna forma del operador

$$\theta_q(B)\theta_q(B^L) = (1 - 0_1B - 0_2B^2 - \dots - 0_qB^q) \times (1 - 0_{1,L}B^L - 0_{2,L}B^{2L} - \dots - 0_{q,L}B^{qL})$$

puede ser adecuada. Los retrasos en los cuales existen picos en la FACM ayudan a sugerir el orden Q del operador de promedios móviles o los órdenes de los operadores a usarse en el producto.

Quinto, si la FACM y la FACMP decrecen rápidamente, entonces alguna forma de los operadores $\theta_q(B)\theta_q(B^L)$ y $\phi_p(B)\phi_p(B^L)$ puede ser adecuada, aunque la forma probablemente sea muy simple.

Se comprenderá que los lineamientos generales dados, no representan la totalidad del método para identificar los "mejores" componentes del modelo general multiplicativo estacional; algunas veces las modificaciones a los operadores y/o modelos vistos anteriormente podrían ser apropiados en un sistema dado, mientras que en otros no. Por ejemplo, considérese una situación en la que la FACM de una serie de tiempo estacional se puede interpretar como que tiende a desaparecer después de algún retraso casi igual a $L, L+1, L+2, 2L, 2L+1$ o $2L+2$, lo cual sugiere el uso de alguna forma del operador de promedios móviles estacionales y puede también interpretarse como que tiende a desaparecer bastante rápido, lo cual sugiere el uso de alguna forma del operador autorregresivo no estacional. Entonces como ambos patrones parecen estar presentes, se usa alguna forma del modelo general multiplicativo.

Ya que la serie de tiempo estacional y_1, y_2, \dots, y_n se ha transformado en una serie estacional z_a, z_{a+1}, \dots, z_n y ya que se ha identificado una forma particular del modelo

$$\phi_p(B)\phi_p(B^L)z_t = b + \theta_q(B)\theta_q(B^L)\epsilon_t$$

como el que ha generado a la serie z_a, z_{a+1}, \dots, z_n se deben estimar los parámetros en un intento por identificar un modelo.

Las condiciones de estacionaridad e invertibilidad existen para cada forma especial del modelo general multiplicativo estacional. Se puede mostrar que usar 0.1 como estimador preliminar de todos y cada uno de los parámetros contenidos en los operadores $\phi_p(B^L)$, $\phi_p(B)$, $\theta_q(B^L)$, y $\theta_q(B)$ generalmente satisface las condiciones de estacionaridad e invertibilidad de un modelo particular bajo consideración. También se puede mostrar que si μ es la media del modelo general entonces:

$$\delta = (1 - \phi_1 - \phi_2 - \dots - \phi_p)(1 - \phi_{1,L} - \phi_{2,L} - \dots - \phi_{p,L})\mu$$

El estimador preliminar de μ es:

$$\bar{z} = \frac{\sum_{t=a}^n z_t}{n-a+1}$$

Así el estimador preliminar de δ es:

$$\hat{\delta} = (1 - \hat{\phi}_1 - \hat{\phi}_2 - \dots - \hat{\phi}_p)(1 - \hat{\phi}_{1,L} - \hat{\phi}_{2,L} - \dots - \hat{\phi}_{p,L})\bar{z}$$

donde: $\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p, \hat{\phi}_{1,L}, \hat{\phi}_{2,L}, \dots, \hat{\phi}_{p,L}$

son los estimadores preliminares de $\phi_1, \phi_2, \dots, \phi_p, \phi_{1,L}, \phi_{2,L}, \dots, \phi_{p,L}$

Ya que se han obtenido los estimadores preliminares de los parámetros desconocidos, estos son usados como entrada en los programas de computo para producir estimadores finales por mínimos cuadrados de los parámetros desconocidos.

Después que un modelo tentativo se ha ajustado a los datos, es importante revisar la suficiencia del modelo y si es necesario sugerir mejoras potenciales. Como ya se explico en la sección 1.3, esto se puede lograr a través del calculo de autocorrelaciones de los residuales y la cantidad Q . Para series de tiempo estacionales es habitual escoger k (el número de autocorrelaciones en el calculo de Q) igual a $3L$.

Después que se ha adoptado un modelo apropiado este se usa para pronosticar valores futuros de la serie de tiempo utilizando los principios analizados en la sección 4-1.4.

CONCLUSIONES.

La mayoría de las decisiones tomadas por distintas empresas se fundamentan en pronósticos de eventos futuros. Estos pronósticos a menudo se basan en información histórica en forma de series de tiempo, la cual se usa para identificar el comportamiento de la serie. Este patrón es entonces extrapolado a futuro, suponiendo que el patrón observado continuará a través del tiempo. Para identificar este patrón, es conveniente pensar que la serie de tiempo consiste de cuatro componentes principales: tendencia, ciclo, variación estacional y fluctuación irregular. Ya que esas componentes pueden ocurrir al mismo tiempo, no existe un método óptimo para pronosticar. Sin duda, es importante escoger la técnica de estimación según el patrón que caracteriza los datos.

Al hacer pronósticos de series de tiempo se producen errores, cada error es causado por la componente irregular y nuestra inhabilidad para predecir perfectamente la tendencia, el ciclo y la estacionalidad de la serie de tiempo. Frecuentemente se requiere de una estimación de que tan malo puede ser un pronóstico y los intervalos de confianza proporcionan tal información.

Existe una multitud de métodos de estimación. La selección de que método usar, depende de varios factores, como son: el pronóstico deseado, el tiempo de realización, el costo, la precisión deseada, la disponibilidad de los datos, la facilidad de operación, etc. Se puede decir que el mejor método de estimación no necesariamente es el más exacto, más bien, el método que deberá usarse es el que resuelve las necesidades del problema en función de exactitud al menor costo.

Por otra parte, para analizar y pronosticar series de tiempo descritas por la tendencia y el factor irregular, se puede usar la aproximación por regresión o por suavizamiento exponencial. La estimación por mínimos cuadrados en un modelo de regresión se puede usar para producir pronósticos puntuales e intervalos de confianza. La construcción de

un modelo de regresión múltiple involucra la especificación de un conjunto apropiado de variables independientes, para esto, existen varias herramientas estadísticas que son de gran ayuda en la construcción de un buen modelo, por ejemplo los estadísticos:

$$t_{\beta_j}, r^2 \text{ y } s^2$$

El método de suavizamiento exponencial, involucra varios pasos, entre los cuales la selección de la constante de suavizamiento es la más importante ya que de esto depende un mejor resultado en la obtención de los pronósticos y se hace a través de un proceso de simulación.

Hablando de la componente estacional, es importante reconocer la diferencia entre variación estacional aditiva y variación estacional multiplicativa para escoger un modelo apropiado para hacer pronósticos. Los números llamados factores estacionales, son factores de corrección que consideran la estacionalidad de las series de tiempo y que es desconocida para el observador, por lo tanto, se deben estimar a lo largo de la tendencia y así poder hacer mejores pronósticos de la serie de tiempo.

La metodología Box-Jenkins usualmente se emplea cuando las observaciones de una serie de tiempo están relacionadas entre sí o autocorrelacionadas. El primer paso en esta metodología es la identificación, para esto es importante producir una serie de tiempo estacionaria por medio de la diferenciación de los valores originales o aplicando alguna otra transformación. Aquí la función de autocorrelación muestral y la función de autocorrelación parcial muestral son muy importantes por que ya que se ha encontrado una serie de tiempo estacionaria, el comportamiento de esas funciones pueden usarse para identificar el modelo particular de series de tiempo que se supone a generado a la serie de tiempo estacionarla. El segundo paso en esta metodología involucra la estimación de los parámetros en el modelo. El tercer paso es llamado adecuación del modelo, el estadístico de prueba es Q , el cual determina si o no las autocorrelaciones de los residuales indican adecuación. El último paso es la estimación de pronósticos.

Los modelos Box-Jenkins no estacionales más comúnmente usados son: "promedios móviles" (MA), "autorregresivos" (AR) y "autorregresivos y de promedios móviles" (ARMA). Para analizar los modelos Box-Jenkins estacionales es necesario familiarizarse con el operador de retraso B , el operador no estacional ∇ y el operador estacional ∇_L . Cuando se está tratando con series de tiempo estacionales a menudo es necesario usar la transformación logaritmo junto con diferentes órdenes de diferenciación estacional y no estacional para producir una serie de tiempo estacionaria. Y ya que se ha obtenido una serie de tiempo estacionaria, se debe identificar un modelo Box-Jenkins estacional. La identificación del modelo será una forma especial del modelo general estacional multiplicativo, este modelo emplea el operador no estacional autorregresivo de orden p , el operador no estacional de promedios móviles de orden q , el operador estacional autorregresivo de orden P y el operador estacional de promedios móviles de orden Q . Existen varios lineamientos generales que se pueden usar para identificar el modelo general estacional multiplicativo que se debe usar para representar una serie de tiempo estacionaria. Finalmente la estimación, el diagnóstico y el cálculo de pronósticos para modelos Box-Jenkins estacionales son esencialmente los mismos que los descritos para los modelos Box-Jenkins no estacionales.

Motivo de otro trabajo, sería la técnica llamada análisis de intervención, que sirve para medir la influencia de algún evento anormal o extraño sobre la serie en estudio.

BIBLIOGRAFIA.

- [1] Bowerman, B. y O'Connell R. "An Applied Approach, Time Series and Forecasting". Duxbury Press. Delmont, Calif. 1979.
- [2] Box, G. E. P., y Jenkins, G. "Time Series Analysis, Forecasting and Control". Holden-Day, Inc. San Francisco. 1970.
- [3] Draper, N. R. y Smith, H. "Applied Regression Analysis". John Wiley & Sons, Inc. New York. 1976.
- [4] Guerrero, V. M. "Desestacionalización de Series de Tiempo Económicas: Parte I" 1983.
- [5] Guerrero, V. M. "Desestacionalización de Series de Tiempo Económicas: Parte II" 1983.
- [6] Hoel, P. "Introduction to Mathematical Statistics". John Wiley, New York. 1962.
- [7] Hoel, P. "Introduction to Probability Theory". Houghton Mifflin, Co. 1971.
- [8] Johnson, L. y Montgomery, D. "Forecasting and Time Series Analysis". McGraw-Hill Book Company. New York. 1976.
- [9] Makridakis, S. y Wheelwright S. "Forecasting: Methods and Applications". John Wiley & Sons, Inc. New York. 1978.
- [10] Shiskin, J. y Young, A. "The X-11 Variant of the Census II Method Seasonal Adjustment Program". Bureau of the Census, Technical Paper No. 15.
- [11] Winters, P. "Forecasting Sales by Exponentially Weighted Moving Averages". Management Science. Vol 6. 1960.