

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO



ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES ACATLÁN



TESIS PARA OBTENER EL TÍTULO DE ACTUARIO

**“EL USO DEL COMPUTADOR EN LA APLICACIÓN DE
TÉCNICAS PARA EL PRONÓSTICO DE SERIES DE TIEMPO,
ECONÓMICAS Y FINANCIERAS”**

Ponente: ROBERT HERNÁNDEZ MARTÍNEZ

Contacto:

robert@aei.com.mx

chomchom216@yahoo.com.mx

Asesores:

**ACTUARIO Y MAESTRO EN CIENCIAS LUCIO PÉREZ
RODRÍGUEZ Y LICENCIADA Y MAESTRA EN ECONOMÍA
PATRICIA RODRÍGUEZ LÓPEZ**

enero, 1997

Agradecimientos:

- **A Dios por la vida que me prestó.**
- **A la UNAM y especialmente a la ENEP – ACATLÁN por la formación profesional recibida.**
- **A mis padres Florencio Hernández Valdéz y Victoria Martínez Jiménez, por su preocupación, sus desvelos y su apoyo sin reservas. Este logro es de ellos.**
- **A mis hermanos Víctor Hugo y Guadalupe Hernández Martínez, por su apoyo y motivación permanente en cada paso que doy.**
- **A mi novia Vero por su apoyo y comprensión en todas las situaciones de mi vida. Muchas tardes estuviste a mi lado ayudándome a terminar este trabajo. Te amo.**
- **A mi asesor, el Act. M.C. Lucio Pérez Rodríguez por ayudarme a terminar mi carrera y motivarme a continuar hasta obtener el título, guiándome en el desarrollo de mi trabajo.**
- **A mi asesora la Lic. Patricia Rodríguez López, quien me ha apoyado desinteresadamente hasta culminar con mi tesis, (y por su apoyo para terminar el Diplomado en la Universidad Autónoma Metropolitana).**
- **A los profesores, Act. María del Carmen Videgaray, M.C. Eduardo Godoy, M.A. Leticia Rivas y Act. Yolanda Zepeda, por su dedicación a la revisión de mi trabajo.**
- **A mis amigos, con los que siempre he contado y que de alguna manera u otra colaboraron en este esfuerzo: Ángel Valeriano, Luis Bravo, Mario Flores, Verónica Gutiérrez, Noe Hernández, Salvador Pérez y los que me faltaron.**

ÍNDICE DE CONTENIDO

<i>INTRODUCCIÓN</i>	<i>iii</i>
---------------------	------------

CAPÍTULO I. CONCEPTOS BÁSICOS

Repaso de conceptos estadísticos básicos	2
Estadística descriptiva	2
Distribuciones de probabilidad	4
Distribuciones muestrales	6
Estimación	7
Pruebas de hipótesis	7
Prueba de bondad de ajuste	9
Análisis de correlación	9
Diagramas de dispersión	10
Coeficiente de correlación	11
Necesidad y uso del pronóstico	13
Concepto y tipos de pronósticos	15
Criterios para seleccionar una técnica de pronóstico	16
Técnicas de pronóstico para datos estacionarios	17
Técnicas de pronóstico para datos con tendencia	19
Técnicas de pronóstico para datos con estacionalidad	20
Técnicas de pronóstico para series cíclicas	21
Otros factores a considerar cuando se selecciona una técnica de pronóstico	23

CAPÍTULO II. PRONÓSTICOS A PARTIR DE LA INFORMACIÓN FINANCIERA Y DEL ANÁLISIS NO PARAMÉTRICO

Medias móviles	26
Suavizamiento exponencial simple	29
Suavizamiento exponencial de Holt	33
Suavizamiento exponencial de Winters	35
Descomposición de series de tiempo	38
Desestacionalización de los datos	39
Tendencia	42
Ciclo	42
Pronóstico por medio de descomposición de series de tiempo	43

CAPÍTULO III. MODELOS ECONOMÉTRICOS UNIECUACIONALES

Estimación de un modelo uniecuacional	49
Propiedades de los estimadores	49

Regresión lineal múltiple	50
Modelo econométrico para el pronóstico de la Tasa de Interés Interbancaria Promedio (TIIP)	54
Evaluación de los supuestos de un modelo econométrico	58
Resultados de la regresión	59
Valuación del supuesto de no multicolinealidad	62
Valuación del supuesto de correcta especificación	62
Valuación del supuesto de normalidad	66
Valuación del supuesto de no autocorrelación	67
Valuación del supuesto de no heterocedasticidad	68
Valuación del supuesto de linealidad	72
Valuación del supuesto de permanencia estructural	74
Pronóstico con modelos uniecuacionales	77
Pronóstico de la TIIP para el año de 1996	77
Conclusión	80

CAPÍTULO IV. ANÁLISIS UNIVARIADO DE SERIES DE TIEMPO **82**

Análisis de series de tiempo	83
Modelos estocásticos y análisis de estacionariedad	83
Uso de operadores y polinomios de retraso	84
Procesos estocásticos lineales	87
Procesos estacionarios	88
Diferencias y no estacionariedad homogénea	92
Construcción de modelos para series univariadas	93
Clases de modelos ARIMAS	94
Modelos autoregresivos AR	94
Modelos de promedios móviles MA	100
Modelos ARMA	104
Modelos ARIMA	105
Construcción del modelo	107
Identificación del modelo	107
Estimación de los parámetros	119
Verificación del modelo	121
Pronóstico del Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (IPC), para el período de octubre de 1995 a diciembre 1996	127
Pronóstico del Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (IPC), para el período de octubre de 1995 a marzo 1996	128
Conclusión	131
BIBLIOGRAFÍA	133

INTRODUCCIÓN

Derivado del uso generalizado de la computadora en la mayoría de los ámbitos de trabajo, investigación, negocios, etcétera; así como la creación de programas y paquetería para eficientar y hacer posible realizar tareas que de otra manera serían imposibles o muy laboriosas, se vuelve una realidad la posibilidad de manejar grandes volúmenes de información para su análisis y aún la inferencia sobre su comportamiento, sin necesidad de conocer la "caja negra" de las técnicas de análisis que se apliquen.

Es necesario aclarar, sin embargo, que los datos manejados, los modelos implementados y el análisis de resultados requieren de destreza y capacidad del *analista* para una acertada interpretación.

Una gran variedad de técnicas de análisis estadístico pueden ser implementadas con agilidad y eficiencia, mediante sistemas de cómputo, por lo que actualmente, si bien es necesario que el Actuario cuente con la formación matemática para entender diversos problemas, también es indispensable que esté al día en cuanto a los avances tecnológicos que le permitan explotar su capacidad al máximo en beneficio de su trabajo mediante la actualización profesional.

Por otro lado, la tecnología puede otorgarle mayores elementos de juicio al momento de elaborar predicciones a partir de la simulación de escenarios futuros.

El éxito en los negocios depende de la habilidad del empresario para desarrollar y ejecutar un plan estratégico que vaya a la vanguardia de la organización. Este proceso de planeación involucra un número importante de actividades relacionadas.

En el caso de una organización financiera que administre por ejemplo, portafolios de inversión, tesorerías, mesas de dinero, sociedades de inversión y en general cualquier actividad que implique riesgo, el Actuario debe pronosticar el nivel de oferta y demanda de los valores objeto de su actividad en condiciones de incertidumbre.

Generalmente, los efectos de políticas económicas, actividades promocionales, competencia, expectativas sobre el desarrollo de alguna actividad y en general todas las condiciones económicas y financieras que intervienen, deben ser consideradas para decidir por un plan óptimo de ejecución.

En otras palabras, después de que los datos relevantes han sido recopilados y analizados, el Actuario debe elegir de una gama de escenarios razonables propuestos, el plan que maximice los resultados esperados.

La teoría y práctica del pronóstico merece gran consideración en la administración financiera, ya que hace una importante contribución a la función directiva.

La incertidumbre, siempre presente en el pronóstico, representa uno de los más amplios campos de estudio del Actuario, quien está capacitado para efectuar pronósticos o estimaciones eficientes que den elementos para tomar decisiones de inversión.

El hecho de que la mayoría de las firmas empresariales y consumidores en potencia, basen importantes decisiones en sus expectativas sobre el futuro económico de sus respectivas actividades, crea una sustancial demanda de pronósticos de diversas variables económicas y financieras.

Tan extensa ha sido esta demanda, que los servicios de agencias especializadas en asesoría financiera han crecido notablemente, y aún más; pronósticos económicos de académicos, hombres de negocios y entidades gubernamentales son ampliamente difundidos por los medios de comunicación.

Esta disciplina también considerada un arte por algunos autores, aporta también elementos para la medición del riesgo y sus efectos en la toma de decisiones.

La aportación de técnicas para el pronóstico de variables económicas y financieras por medio del análisis estadístico es de relevante importancia, dada la incertidumbre involucrada. Parece razonable que diversos profesionales otorgarían diferente importancia a una gran variedad de influencias económicas. El juicio de quien pronostica se refleja no sólo en la interpretación que da a los datos generados por los modelos simulados por computadora, sino a los modelos mismos.

Las computadoras pueden generar pronósticos, pero lo hacen sobre la base de suposiciones hechas por el analista. Una modelación y una correcta interpretación de resultados son tarea del Actuario. A lo largo del presente trabajo se utilizarán los paquetes para análisis estadístico, Excel 5.0 y Econometric Views for Windows 1.1B, de uso común en computadoras personales.

No sorprende que el pronóstico y su metodología generen gran controversia. Los usuarios del pronóstico esperan siempre un alto grado de eficiencia en el mismo, olvidando al mismo tiempo que el pronóstico en sí mismo tiene importantes consecuencias.

Cuando el mercado se contrae, en respuesta al pronóstico de una recesión en el corto plazo, por ejemplo, se ha cambiado ya la base inicial de predicción de los analistas que efectuaron el pronóstico. La reacción inesperada del mercado a un pronóstico de esa naturaleza, puede incluso causar una recesión más profunda.

Ese es el dilema del Actuario cuando pronostica; pues el futuro como lo esperamos no existe. De hecho no puede.

Agradezco especialmente al Actuario y Maestro en Ciencias Lucio Pérez Rodríguez (†), a la Lic. y Maestra en Economía Patricia Rodríguez López, y al Lic. Ricardo Zúñiga Alcántara, quienes me asesoraron en el desarrollo de este trabajo con sus sugerencias y comentarios para su mejor culminación.

(†) Este trabajo es también un homenaje al profesor Lucio Pérez, con quien siempre conté y quien me asesoró hasta el momento de su lamentable deceso. Descanse en paz.

CAPÍTULO I. CONCEPTOS BÁSICOS

Repaso de conceptos estadísticos básicos

Muchas de las técnicas de pronóstico están basadas en conceptos estadísticos fundamentales los cuales son generalmente repasados en textos financieros y estadísticos.

La mayoría de los procedimientos hacen inferencias acerca de grupos de interés llamados poblaciones, después de seleccionar y medir un subconjunto de éste llamado muestra. Una cuidadosa selección de una muestra representativa y de suficiente tamaño es un componente importante de un proceso de inferencia estadística que tenga un grado de riesgo aceptable.

Junto con la inferencia estadística, la estadística descriptiva constituyen las dos subdivisiones de la estadística. Los procedimientos de la estadística descriptiva son usados para resumir y describir colecciones de datos con el fin de que sus componentes se vuelvan obvios y puedan ser usados en el proceso de toma de decisiones.

Estadística descriptiva

La manera más común de describir un gran conjunto de valores es promediándolos. Es decir a través del cálculo de la media:

$$\mu = \frac{\Sigma X}{N}$$

donde ΣX representa la suma de todos los valores de la población,
 N representa el tamaño de la población.

Y en el caso de la media muestral \bar{X} se calcula con:

$$\bar{X} = \frac{\Sigma X}{n}$$

donde ΣX representa la suma de todos los valores de la muestra,
 n representa el tamaño de la muestra.

Adicionalmente, para medir la dispersión alrededor de la media se calcula la desviación estándar. Dicha medida puede ser pensada como la diferencia típica entre un grupo de valores y su media. Siguiendo con el caso de la desviación estándar poblacional (σ) y la desviación estándar muestral (s), se tiene:

$$\sigma = \sqrt{\frac{\sum(X - \mu)^2}{N}} = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N}}$$

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{n-1}} = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1}}$$

donde los numeradores representan la suma de los cuadrados de las diferencias entre los valores obtenidos y sus medias.

Finalmente muchos procesos estadísticos utilizan la varianza poblacional o muestral. La varianza de una colección de datos es la desviación estándar al cuadrado y se calcula como sigue:

$$\sigma^2 = \frac{\sum(X - \mu)^2}{N} = \frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N}$$

$$s^2 = \frac{\sum(X - \bar{X})^2}{n-1} = \frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1}$$

El término de grados de libertad es utilizado para indicar el número de datos que son independientes uno de otro en el sentido de que ninguno de ellos puede ser deducido a partir del otro y por lo tanto son piezas únicas de información.

El cálculo de la desviación estándar difiere cuando se trata de muestras, debido a que se utiliza una estimación de la media. Se introduce un sesgo mínimo cuando se calcula $\sum(X - \bar{X})^2$ para una distribución dada y entonces se obtiene una desviación estándar muestral mucho menor a la desviación estándar poblacional. Sin embargo, este sesgo puede ser corregido dividiendo el valor de $\sum(X - \bar{X})^2$ por los grados de libertad apropiados. En general donde quiera que se utilice una muestra estadística como un estimador de un parámetro de una población se pierde un grado de libertad.

Distribuciones de probabilidad

Una variable aleatoria es el nombre de una unidad que puede tomar diferentes valores en cada ensayo de algún experimento, el resultado esperado es un evento aleatorio.

Si sólo ciertos valores específicos son posibles, la variable aleatoria se denomina variable discreta. Si la variable aleatoria puede tomar cualquier valor dentro de un intervalo, se denomina variable continua.

Si se considera una variable discreta, su distribución de probabilidad lista todos los valores posibles que la variable puede tomar junto con la probabilidad de cada uno.

El valor esperado de la variable aleatoria es el valor promedio que la variable asume después de muchos ensayos. El valor esperado para una función de probabilidad discreta puede ser encontrado multiplicando cada valor posible X por su probabilidad y después sumando los productos, es decir:

$$E(X) = \sum [X \cdot P(X)]$$

Para una distribución continua la probabilidad de obtener un valor específico se aproxima a cero. Las distribuciones continuas tratan de encontrar la probabilidad de que un valor caiga dentro de un intervalo cuando se considera aleatoriedad.

Algunas distribuciones teóricas se presentan comúnmente en muchas aplicaciones estadísticas prácticas y por esta razón es importante examinar sus propiedades y aplicaciones. Una de estas distribuciones importantes es la distribución binomial usada para representar una variable discreta. Los requerimientos para un experimento binomial son:

1. Existen n ensayos idénticos, cada uno de ellos tiene dos resultados posibles.
2. La probabilidad de éxito en cada resultado permanece fija en cada ensayo.
3. Los ensayos son independientes.

El interés está en encontrar la probabilidad de X resultados exitosos en n ensayos, donde el resultado exitoso puede ser arbitrariamente definido como cualquiera de los dos posibles resultados, (éxito o fracaso). Los diferentes valores de X junto con sus probabilidades forman la distribución binomial. Estas probabilidades pueden ser encontradas con la siguiente ecuación:

$$P(X) = C_x^n \pi^x (1 - \pi)^{n-x} \quad \text{para cada } X = 0, 1, 2, \dots, n$$

donde:

C_x^n = el número de combinaciones de n elementos tomados X a la vez
 π = la probabilidad de éxito en cada ensayo
X = el número particular de éxitos de interés
n = el número de ensayos

Una distribución continua importante debido a que muchas poblaciones pueden ser aproximadas por ella es la distribución normal, especificada por dos parámetros, la media y la desviación estándar. Gráficamente es una curva simétrica en forma de campana como lo muestra la **figura 1.1**. Esta distribución representa muchas variables reales que son medidas en una escala continua.

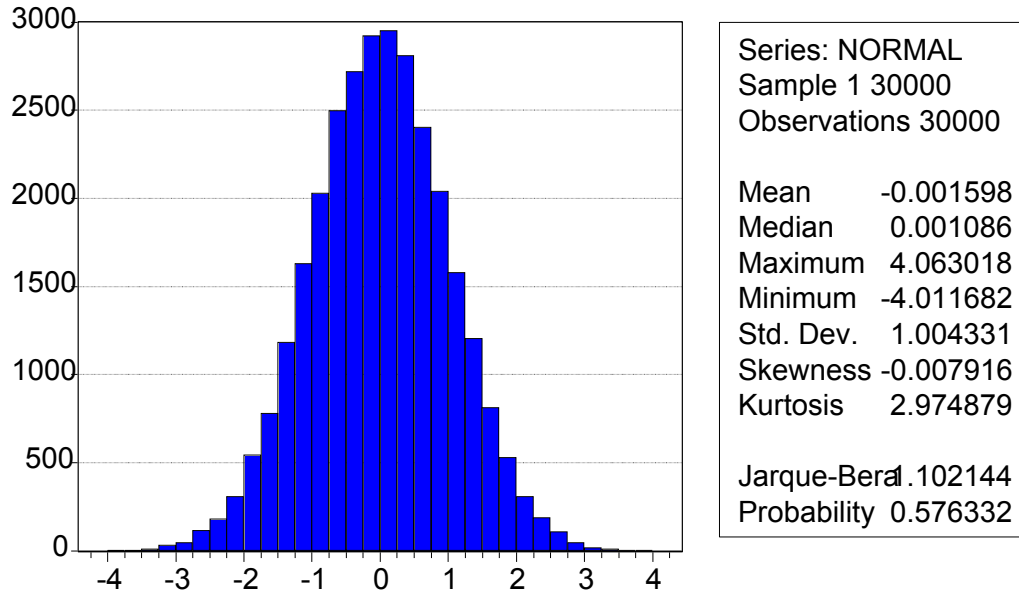


Figura 1.1.: Distribución de números aleatorios con media 0 y varianza 1. Muestra de 30,000 números.

Las probabilidades de una distribución normal en un intervalo pueden ser encontradas estandarizando los intervalos a una Z - escala. La Z - escala de cualquier valor de X es el número de desviaciones estándar del valor central de la curva (μ) a ese valor. La fórmula para Z es:

$$Z = \frac{X - \mu}{\sigma}, \text{ donde: } \begin{array}{ll} X & = \text{el valor particular de interés} \\ \mu & = \text{la media} \\ \sigma & = \text{la desviación estándar} \end{array}$$

Después de que la escala estandarizada Z ha sido calculada, la tabla de la curva normal puede ser consultada para encontrar el área bajo la curva entre el centro de la curva (μ) y el valor de interés (X).

Finalmente la distribución de t es frecuentemente usada en pruebas estadísticas cuando se tienen muestras pequeñas y se puede asumir que la población a ser investigada se distribuye normalmente. Su aplicación fundamental está basada en el siguiente teorema:

Si \bar{X} y s^2 son la media y la varianza de una muestra aleatoria de tamaño n tomada de una población normal con media μ y varianza σ^2 , entonces $t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$ tiene la distribución de t con n-1 grados de libertad.

Existen tablas que muestran la distribución de t y sólo se requiere conocer los grados de libertad para encontrar los valores de t que excluyen los porcentajes deseados de la curva.

Distribuciones muestrales

En muchas aplicaciones estadísticas se toma una muestra aleatoria de la población que se investiga, se calcula un estadístico de los datos de la muestra y se bosquejan conclusiones acerca de la población, sobre la base de la muestra. Una distribución muestral es la distribución de todos los valores posibles de la muestra estadística que pueden ser obtenidos de la población a partir de una muestra de tamaño dado.

El teorema del límite central establece que conforme el tamaño de la muestra es más grande, la distribución muestral de las medias de la muestra tiende a una distribución normal con media μ , la media poblacional, y desviación estándar σ / \sqrt{n} , (éste valor es conocido como el error estándar de la distribución muestral). La distribución muestral tenderá a la normalidad sin importar la forma de la distribución poblacional de donde la muestra fue tomada.

El teorema del límite central es de particular importancia en estadística ya que permite al analista calcular la probabilidad de varios resultados muestrales a través del conocimiento de la curva normal de probabilidades.

Estimación

El estimador puntual de un parámetro de una población es un valor particular calculado a partir de los datos de la muestra que estima el valor desconocido del parámetro de la población. La **figura 1.2** contiene una lista de diversos parámetros poblacionales y el estadístico muestral que proporciona el estimador puntual de ellos.

Parámetro poblacional	Estadístico de la muestra (estimador)
Media (μ)	\bar{X} , la media muestral
Desviación estándar (σ)	s, la desviación estándar muestral
Varianza (σ^2)	s^2 , la varianza muestral
Porcentaje (π)	p, el porcentaje muestral

Figura 1.2.: Valores poblacionales y muestrales.

Un intervalo de confianza es un intervalo dentro del cual posiblemente se encuentra el parámetro de la población de interés. Se construye a través de un intervalo alrededor del estimador puntual y se calcula usando la distribución normal o la distribución t.

Pruebas de hipótesis

En muchas situaciones estadísticas, se tiene interés en probar alguna característica de una población más que estimar alguno de sus parámetros. El procedimiento es conocido como prueba de hipótesis y consta de los siguientes pasos:

- Enunciar la hipótesis a probar (denominada hipótesis nula, H_0) y enunciar también la hipótesis alternativa (la que se acepta si H_0 es rechazada, H_1).
- Tomar una muestra aleatoria de la población, medirla y calcular el estadístico muestral apropiado.

- Asumir la hipótesis nula como verdadera y consultar la distribución muestral del estadístico observado.
- Calcular la probabilidad de que el estadístico muestral no haya sido tomado de esa distribución muestral con un α dado.
- Si la probabilidad es mayor a α , no se rechaza la hipótesis nula, si es menor a α , la hipótesis nula puede ser rechazada con un margen de error mínimo.

Cuando se siguen los pasos anteriores pueden ocurrir dos tipos de errores, como se muestra en la **figura 1.3**. Siempre se espera que después de examinar la evidencia muestral la decisión correcta será la hipótesis nula, sin embargo existe la posibilidad de rechazar H_0 cuando es verdadera y aceptar una H_0 cuando es falsa. Las probabilidades de estos dos eventos se conocen como α y β , respectivamente. Alfa es conocida como el nivel de significancia de la prueba.

Estados posibles	Aceptar H_0	Rechazar H_0
H_0 es verdadera	Decisión correcta	Error tipo I, probabilidad α
H_0 es falsa	Error tipo II, probabilidad: β	Decisión correcta

Figura 1.3.: Resultados de las pruebas de hipótesis.

La **figura 1.4** resume las cuatro hipótesis estadísticas más usadas. En cada una de ellas se calcula el estadístico muestral apropiado. Consultando la distribución muestral que corresponda, después de asumir que la hipótesis nula es verdadera; se obtiene una regla de decisión que especifique el rango dentro del cual es probable que se encuentre el estadístico y que determine rechazar o aceptar la hipótesis nula.

H_0	Estadístico muestral	Distribución muestral	Parámetros de la distribución muestral
$\mu = 50^*$	\bar{X}	Normal	μ y $\frac{\sigma}{\sqrt{n}}$
$\mu = 0.30^*$	p	Normal	π y $\sqrt{\frac{\pi(1-\pi)}{n}}$
$\mu_1 - \mu_2 = 0$	$\bar{X}_1 - \bar{X}_2$	Normal	$\mu_1 - \mu_2$ y $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
$\pi_1 - \pi_2 = 0$	$p_1 - p_2$	Normal	$\pi_1 - \pi_2$ y $\sqrt{\pi(1-\pi)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

Figura 1.4.: Resumen de pruebas de hipótesis.

* Valores para ejemplo

Prueba de bondad de ajuste

A menudo es de interés determinar si una muestra de datos se ajusta o no a alguna distribución teórica conocida. Se toma una muestra de la población de interés y se observan las categorías en las cuales caen los datos. Si las frecuencias observadas son muy próximas a la distribución hipotética esperada, la formulación hipotética es aceptada; si existen grandes diferencias en los valores observados, la formulación es rechazada.

En resumen, las distribuciones observadas pueden ser comparadas a las distribuciones teóricas por medio del uso de la prueba de bondad de ajuste. La distribución de las frecuencias esperadas de la población teórica debe ser calculada y comparada con los valores de la muestra. Si existen grandes diferencias en las categorías, y se calcula un valor grande de una distribución chi - cuadrada, la hipótesis con respecto a la población puede ser rechazada. Si las frecuencias reales están muy cercanas a los valores esperados de las categorías y se calcula un valor pequeño de una chi - cuadrada, entonces la hipótesis con respecto a la población puede ser aceptada.

Muchas distribuciones poblacionales se pueden someter a la prueba de bondad de ajuste. Generalmente la prueba se hace contra la distribución de una población normal debido a que muchas pruebas estadísticas asumen esa distribución. La prueba de bondad de ajuste puede ser usada con los datos actuales contra la distribución del período previo o contra una norma de la industria o contra algunos parámetros estándares que se conozcan sobre el sector de estudio. En todas estas pruebas, no se utilizan datos numéricos. Los datos son medidos en una escala continua, aunque primero son convertidos a categorías antes de efectuar la prueba.

Análisis de correlación

Un objetivo común en muchas aplicaciones estadísticas es examinar la relación que existe entre dos variables numéricas. El énfasis en la correlación y la regresión se justifica por su gran uso en toda clase de aplicaciones, incluyendo el pronóstico. Adicionalmente, extensiones de los conceptos de regresión y correlación en áreas más complejas se utilizan para muchos procedimientos de pronóstico.

Se asume que las variables bajo investigación son mediciones numéricas, esto es, son medidas de un proceso que genera números reales más que categorías.

Diagramas de dispersión

Un estudio de la relación entre variables empieza con el caso más simple, el de la relación entre dos variables. Suponiendo que se han tomado dos medidas de varios objetos. Un analista desea determinar si una de estas variables, por ejemplo Y, tiende a incrementarse o decrementarse conforme la otra variable, X, cambia. Desde luego es peligroso alcanzar conclusiones sobre la base de un tamaño de muestra inadecuado.

Un diagrama de dispersión mapea en dos dimensiones los valores de X a lo largo del eje horizontal y los valores de Y en el eje vertical. El diagrama de dispersión ayuda a ilustrar lo que la intuición sugiere: la relación entre X e Y.

El diagrama (a) de la **figura 1.5** sugiere lo que se denomina relación perfecta lineal positiva. Conforme X se incrementa, Y también, de una manera predecible. Es decir, X e Y parecen comportarse en torno a una línea recta. El diagrama (b) sugiere una relación perfecta lineal negativa. Es decir, conforme X se incrementa, Y decrece de una manera predecible.

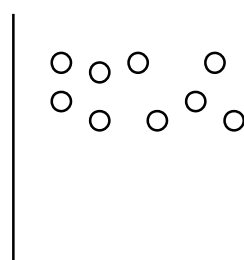
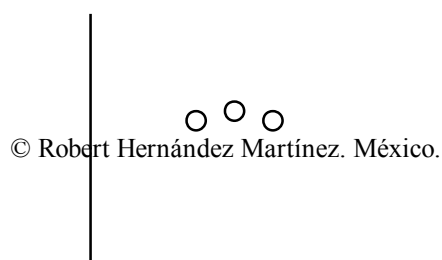
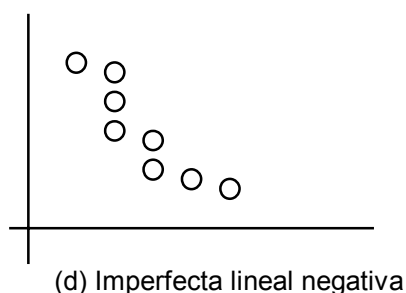
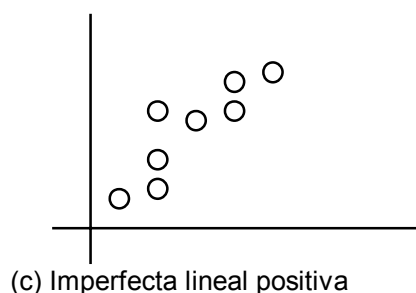
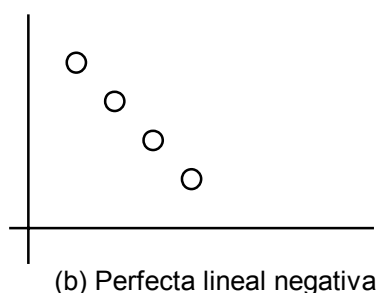
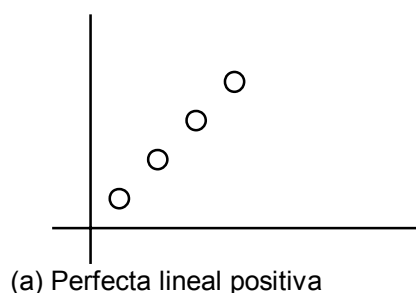




Figura 1.5.: Diagramas de dispersión

Las figuras (c) y (d) ilustran una relación imperfecta lineal positiva y negativa, respectivamente, pero no de una manera predecible. Así, Y puede ser ligeramente superior o inferior a lo esperado. Es decir, los puntos $X - Y$ no se comportan como una línea recta.

Los diagramas de dispersión de las figuras (c) y (d) muestran lo que se conoce como relaciones lineales. La relación $X - Y$, sea perfecta o imperfecta, puede ser resumida por una línea recta. En comparación, una relación de curva aparece en el diagrama (e).

Finalmente, el diagrama de la figura (f), sugiere que no existe relación de ningún tipo entre las variables X e Y . Si X se incrementa, Y no parece incrementarse o decrementarse de una manera predecible. Sobre la base de la evidencia muestral que aparece en el diagrama, se puede concluir que en la población de todos los puntos $X - Y$, no existe relación lineal o de otro tipo entre las variables X, Y .

Ahora considérense los dos diagramas de dispersión de la **figura 1.6**. Ambos diagramas sugieren relaciones imperfectas lineales positivas entre X e Y . La diferencia es que esta relación parece más fuerte en la figura (a) debido a que los datos están más cerca a la línea recta que pasa a través de ellos. En la figura (b) se sugiere una relación más débil. Los puntos de los datos están más lejos de la línea recta que pasa por ellos, sugiriendo que la relación entre X e Y es menos lineal.

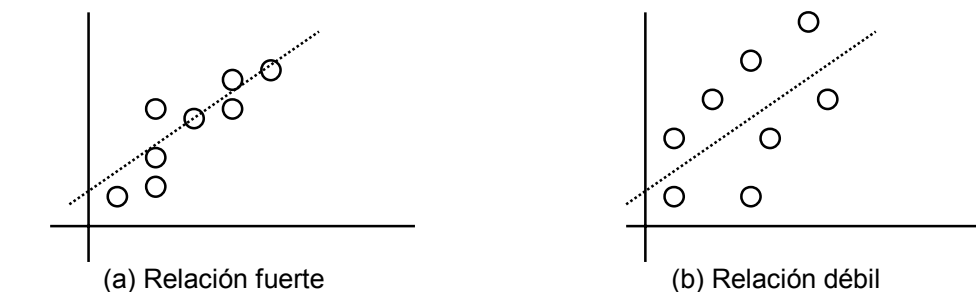


Figura 1.6.: Relación fuerte y débil entre X e Y

Coeficiente de correlación

Como consecuencia de lo anterior, se necesita una manera de medir el grado de relación lineal que existe entre dos variables. En otras palabras, una medida de correlación. La medida más comúnmente utilizada para esta relación es el coeficiente de correlación. Dos variables con una relación perfecta lineal negativa tienen un coeficiente de correlación de igual a -1. En el otro extremo, dos variables con una relación perfecta positiva tienen un coeficiente de correlación de +1. Por lo tanto, el coeficiente de correlación varía entre -1 y +1 incluyéndolos, dependiendo del grado de correlación entre las dos variables a ser medidas.

El diagrama de dispersión (a) de la **figura 1.5** ilustra una situación que produciría un coeficiente de correlación de +1. El diagrama de dispersión (b) de esa misma figura tiene un coeficiente de correlación de -1. Los diagramas de (e) y (f) muestran dos variables que no están linealmente relacionadas. El coeficiente de correlación para esta relación es igual a 0, es decir, no hay relación lineal presente.

Es importante también distinguir entre dos grupos de datos en los cuales esté interesado el analista. En la población que contiene todos los datos X - Y de interés, existe un coeficiente de correlación cuyo símbolo es ρ . Si se toma una muestra aleatoria de los datos X - Y el coeficiente de correlación de esta muestra es r . Un resumen de estas importantes características de ρ y r se da en la **figura 1.7**.

Población	Muestra
Coeficiente de correlación: $\rho(-1 \leq \rho \leq 1)$	Coeficiente de correlación: $r(-1 \leq r \leq 1)$
$\rho = -1$: correlación perfecta negativa	$r = -1$ correlación perfecta negativa
$\rho = 0$: no hay correlación	$r = 0$ no hay correlación
$\rho = 1$: correlación perfecta positiva	$r = 1$ correlación perfecta positiva

Figura 1.7.: Correlación poblacional y muestral

Frecuentemente, X e Y se miden en diferentes unidades, tales como kilos, pesos, unidades vendidas, tasas de interés. A pesar de estas formas diferentes de medir X e Y, es importante el grado en que X e Y están relacionadas. Esta medición se hace primero convirtiendo las medidas de los datos muestrales en unidades estándar.

Las medidas X e Y se convierten en unidades de Z y son multiplicadas, obteniendo productos cruzados para cada dato. Estos productos cruzados son de interés debido a que la media de estos valores es el coeficiente de correlación. Lo anterior se calcula con la siguiente fórmula:

$$\rho = \frac{\sum Z_y Z_x}{N}$$

El cálculo del coeficiente de correlación ρ por medio de los productos cruzados de Z , siempre produce el valor correcto, pero en la mayoría de los casos es más fácil hacer los cálculos. La siguiente fórmula es equivalente para encontrar ρ calculando la media de los productos cruzados de Z .

$$\rho = \frac{\sum Z_y Z_x}{N} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{N \sum X^2 - (\sum X)^2} \sqrt{N \sum Y^2 - (\sum Y)^2}}$$

Cuando se examina una colección de datos, se debe utilizar el juicio para examinar cuando el valor de ρ es demasiado grande para considerar las dos variables altamente correlacionadas linealmente. En la mayoría de los casos la ecuación anterior puede ser usada para encontrar en que dirección se mueven las variables.

Cuando se toma una muestra aleatoria de datos de una población bajo investigación se reemplaza la población N en la ecuación por la muestra n , del tamaño de la muestra. La fórmula para el coeficiente de correlación muestral se convierte en:

$$r = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{n \sum X^2 - (\sum X)^2} \sqrt{n \sum Y^2 - (\sum Y)^2}}$$

Dos aspectos importantes deben ser puntualizados:

Primero: Siempre tener en cuenta que lo que se está midiendo es la correlación, no la causalidad. Es válido decir que dos variables están correlacionadas a partir de un coeficiente de correlación alto. Puede ser válido o no decir que una variable está causando el movimiento de la otra, dependerá del juicio del analista. El error de asumir causalidad sobre la base de correlación es comúnmente hecho por políticos y publicistas.

Segundo: Nótese que el coeficiente de correlación mide la relación lineal entre dos variables. En el caso en que el coeficiente de correlación es bajo, se puede concluir que dos variables no están cercanamente relacionadas de una forma lineal. Pudiera ser que estén muy relacionadas de una manera curva o no lineal. Por lo tanto un coeficiente de correlación bajo no significa que las variables no estén relacionadas, sino solamente que no parece existir una relación lineal directa.

Necesidad y uso del pronóstico

Muchos de los conceptos básicos en estadística pueden ser considerados la base para el entendimiento de las técnicas de pronóstico. Sin embargo, dichos conceptos valen por sí mismo en muchas aplicaciones.

Aún cuando algunas de ellas posiblemente no llevarían el título de "pronóstico", estaría involucrado el uso de datos recopilados para responder a dudas en condiciones de incertidumbre en la operación de muchos negocios, especialmente sobre su desarrollo futuro.

La estadística descriptiva es ampliamente usada cuando se cuenta con grandes volúmenes de datos cuya descripción es importante en un proceso de decisión. Sería prácticamente imposible pensar en un área específica que desarrolle datos numéricos donde los datos observados no sean rutinariamente resumidos por medio de la estadística descriptiva.

Este hecho se aplica particularmente a la media, comúnmente conocida como "promedio" y, - de algún modo menos extendido -, a la desviación estándar. Asimismo, las distribuciones normal y binomial son buenos ejemplos de distribuciones teóricas aplicables a muchas situaciones reales, de ahí que su uso sea muy variado en el pronóstico. La predicción del porcentaje de defectos en un lote de partes, por ejemplo, puede ser encontrada mediante el uso de una distribución binomial.

La estimación y las pruebas de hipótesis son dos de las aplicaciones estadísticas básicas principales. El pronóstico y la estimación de valores de interés de una población a través de la medición de una muestra aleatoria, son de uso corriente donde las restricciones de tiempo y dinero obligan a confiar en resultados de muestras para tomar decisiones. El muestreo se ha vuelto especialmente útil en el proceso de control de calidad. Las pruebas de hipótesis para comparar los valores de un período pasado de una población con sus valores presentes, comparar los valores de los parámetros de dos diferentes sucursales o puntos de negocio y para detectar los cambios en los estándares de un proceso de producción. Esta última aplicación es conocida como control de producción.

Las pruebas de bondad de ajuste son comúnmente usadas para checar la validez de las suposiciones con respecto a las distribuciones poblacionales. Este proceso es especialmente importante cuando se requiere cierta simulación de modelos, por ejemplo; la suposición de un proceso de Poisson para modelar resultados válidos.

Tales simulaciones son usadas para modelar situaciones reales y después validarlas para pronosticar los resultados que se obtendrían cuando se cambian algunas variables del sistema.

La correlación se utiliza para determinar relaciones entre variables numéricas. Estas relaciones son de gran importancia en el pronóstico, ya que pronosticar una variable generalmente implica encontrar relaciones entre variables. El análisis de regresión y la regresión múltiple confían en los coeficientes de correlación para el proceso de pronóstico.

Concepto y tipos de pronósticos

Dos factores básicos deben ser considerados a fin de obtener un pronóstico útil y eficiente:

El primero es que los datos disponibles deben ser relevantes para obtener la mayor eficiencia en el pronóstico. El segundo es escoger una técnica adecuada de pronóstico que utilice al máximo la información contenida en los datos, así como sus patrones de comportamiento.

Después de recopilar los datos el analista debe revisar los trabajos o intentos anteriores para pronosticar la variable de interés. Una investigación bibliográfica y alguna plática con colegas pueden mostrar los intentos exitosos o fallidos al tratar acercamientos diversos al fenómeno. Después de esta introducción inicial el analista está en condiciones de depurar la información con que cuenta.

Con buenos datos a la mano el analista puede comenzar a explorar sus patrones de comportamiento. Este paso incluye la observación de los datos, el tratar de entender lo que sugieren y su visualización gráfica.

Pueden ser empleados dos métodos básicos para pronosticar:

CUALITATIVOS: Son técnicas de pronóstico que confían en el juicio y la intuición humana más que en la manipulación de los datos históricos. Las técnicas cualitativas comunes incluyen el crecimiento de curvas, escenarios de escritorio, investigación de mercado y enfoque de grupos.

CUANTITATIVOS: Son técnicas para el pronóstico cuando se tienen disponibles suficientes datos históricos, y cuando se cree que son representativos para considerar el futuro. Todas las técnicas cuantitativas descansan en la premisa de que el pasado puede ser de alguna manera extendido al futuro para

proveer pronósticos eficientes. Las técnicas cuantitativas se clasifican en dos categorías: estadísticas y determinísticas.

Estadísticas: Las técnicas estadísticas se enfocan completamente en patrones generales, patrones de cambio y alteraciones causadas por influencias aleatorias. Algunas de éstas son las medias móviles, el suavizamiento exponencial, la descomposición de series de tiempo y proyecciones de tendencia.

Las técnicas estadísticas de pronóstico utilizan básicamente dos acercamientos. El primero está basado en la premisa de que los datos pueden ser descompuestos en componentes individuales tales como: tendencia, ciclo, estacionalidad e irregularidad. El pronóstico se hace combinando las proyecciones de éstos componentes individuales.

Un segundo acercamiento está asociado con los modelos econométricos de series de tiempo (cuya definición se presenta en el capítulo III). Sus fundamentos teóricos se basan principalmente en conceptos estadísticos (fundamentalmente regresiones que involucren un término de error aleatorio) y no asumen que los datos son representados por componentes individuales.

Determinísticas: Las técnicas determinísticas o causales involucran la identificación y determinación de relaciones entre la variable a ser pronosticada y otras variables que la influyan.

Estas técnicas incluyen indicadores líderes, modelos econométricos, encuestas e investigaciones de anticipación o sondeo y modelos de entrada y salida.

Criterios para seleccionar una técnica de pronóstico

Algunas de las preguntas que deben ser consideradas antes de decidir la técnica de pronóstico más apropiada para un problema en particular son las siguientes:

- ◆ ¿Por qué se necesita un pronóstico?
- ◆ ¿Quién usará el pronóstico?
- ◆ ¿Cuáles son las características de los datos disponibles?
- ◆ ¿Qué período de tiempo va a ser pronosticado?

- ♦ ¿Cuál es la información mínima requerida?
- ♦ ¿Cuáles serán los costos del pronóstico?

Para seleccionar una técnica apropiada de pronóstico el analista debe ser capaz de:

- ♦ Definir la naturaleza del problema a pronosticar
- ♦ Explicar la naturaleza de los datos en estudio
- ♦ Describir las capacidades y limitaciones de las técnicas de pronóstico que le sean útiles
- ♦ Desarrollar algún criterio determinado a partir del cual se tome una decisión.

El mejor factor de selección de una técnica de pronóstico es la identificación y entendimiento de los patrones históricos de los datos. Si se pueden reconocer patrones de tendencia, ciclo o estacionalidad entonces se pueden seleccionar técnicas que modelen efectivamente estas conductas.

Técnicas de pronóstico para datos estacionarios

Una serie estacionaria se define como aquella cuya media y varianza respectivas no cambian a través del tiempo, es decir, dichos parámetros son independientes del tiempo t . Asimismo, la covarianza entre dos observaciones Y_t y Y_{t-j} , $cov(Y_t, Y_{t-j})$, depende solamente de la distancia j que haya entre ellas, en consecuencia la función de autocorrelación tiende a cero rápidamente conforme la distancia en el tiempo es mayor.

Las estimaciones que se obtengan deberán ser actualizadas conforme se tengan más datos históricos, con el fin de captar los cambios en la estructura interna de la serie, **figura 1.8**.

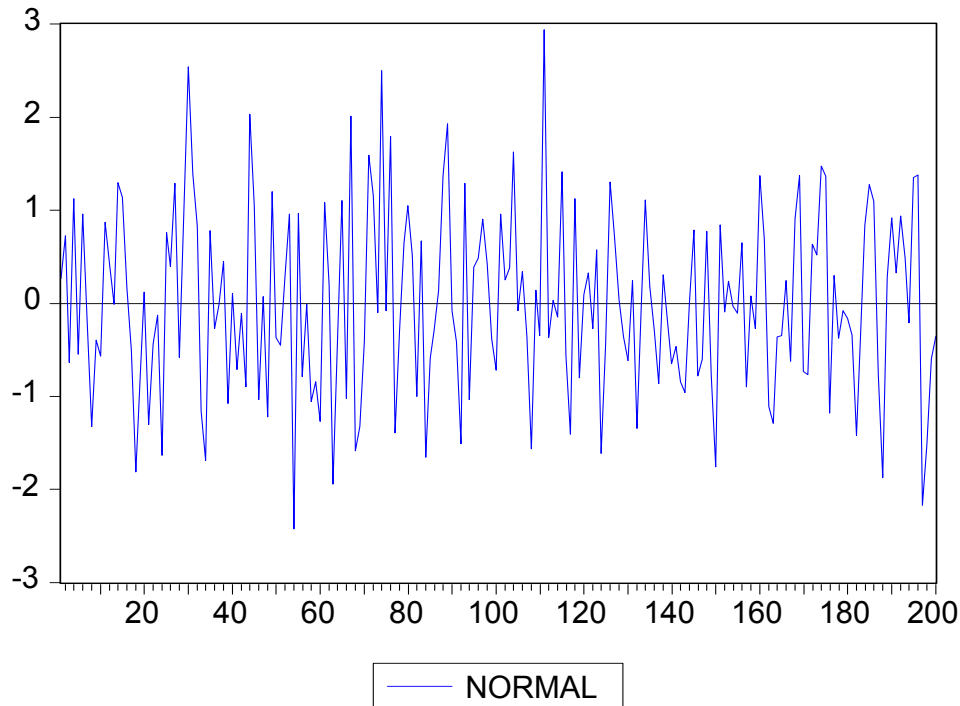


Figura 1.8.: Gráfica de una serie de tiempo estacionaria, con media y varianza constantes. Serie normal con media cero y varianza 1.

Las técnicas de pronóstico estacionarias son usadas en las siguientes situaciones:

- ◆ Donde ocurre que las fuerzas que generan una serie se han estabilizado y el medio en el cual existen permanece relativamente sin cambio. Algunos ejemplos son el número de apagones por semana en una línea de ensamble cuya tasa de producción es uniforme; las unidades vendidas de un producto o servicio en la temporada máxima de su ciclo de ventas y el número de ventas que resultan de un nivel de esfuerzo constante.
- ◆ Donde se necesita un modelo muy simple debido a la escasez de datos o para una fácil explicación o implementación. Un ejemplo es cuando un negocio o una organización es nueva y se tienen muy pocos datos históricos disponibles.
- ◆ Donde la estabilidad puede ser obtenida haciendo simples correcciones por factores tales como el crecimiento de la población o la inflación. Algunos ejemplos son cambiar el ingreso nominal a ingreso per capita o cambiar las ventas en pesos nominales a montos en pesos constantes.

- ♦ Donde las series pueden ser transformadas a series estables. Por ejemplo, transformar una serie tomando logaritmos, raíces o diferencias.
- ♦ Donde la serie es un conjunto de errores de pronóstico producto de la aplicación de una técnica de pronóstico.

Técnicas de pronóstico para datos con tendencia

Una serie tendencial se define como una serie de tiempo que contiene una componente de largo plazo que representa el crecimiento o decrecimiento en las series en un período extenso de tiempo. En otras palabras, se dice que una serie de tiempo es tendencial si su valor esperado cambia a través del tiempo, de tal manera que se incremente o decremente durante el período de pronóstico. Es común que las series de tiempo económicas tengan tendencia, **figura 1.9**.

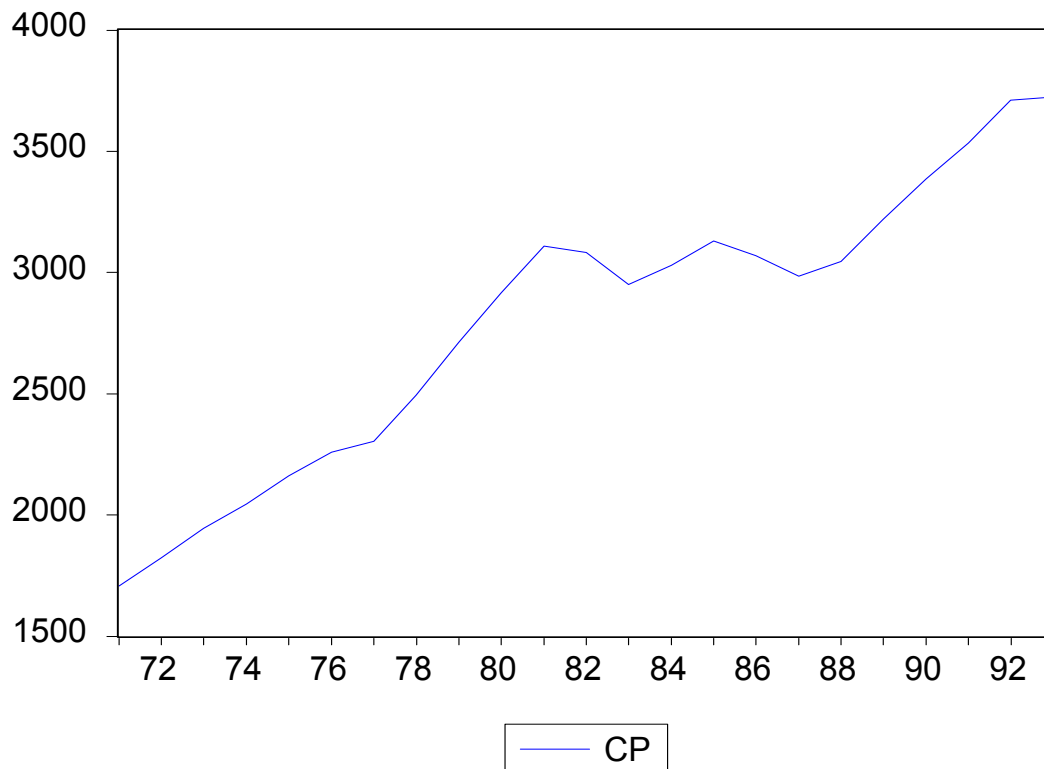


Figura 1.9.: Gráfica de una serie de tiempo con tendencia. Serie del consumo privado en México.

Las técnicas de pronóstico para datos con tendencia son usadas en las siguientes situaciones:

- ♦ Donde el incremento de la productividad y la tecnología llevan a nuevos cambios de estilos de vida. Algunos ejemplos son la demanda de componentes electrónicos, la cual se incrementó con la llegada de las computadoras y el uso del ferrocarril, el cual disminuyó con la aparición de los aviones.
- ♦ Donde el incremento de la población ocasiona incremento en la demanda de artículos, consumo de energía y uso de materias primas.
- ♦ Donde el poder adquisitivo de la moneda afecta las variables económicas debido a la inflación. Algunos ejemplos son los salarios, los costos de producción y los precios.
- ♦ Donde la aceptación del mercado se incrementa. Un ejemplo es el período de crecimiento en el ciclo de vida de un nuevo producto.

Las técnicas que pueden ser consideradas cuando se pronostican series con tendencia incluyen las medias móviles, el suavizamiento exponencial de Holt, regresión lineal, curvas de crecimiento y los modelos exponenciales.

Técnicas de pronóstico para datos con estacionalidad

Una serie estacional se define como una serie de tiempo con patrones de cambio que se repiten cada cierto período específico de tiempo (generalmente año tras año). Desarrollar una técnica de pronóstico estacional usualmente requiere seleccionar un método aditivo o multiplicativo y después estimar los índices de estacionalidad de la serie histórica.

Estos índices son usados para incluir la estacionalidad en los pronósticos o para remover este efecto en los valores observados. Este último proceso es conocido como desestacionalización de los datos, **figura 1.10**.

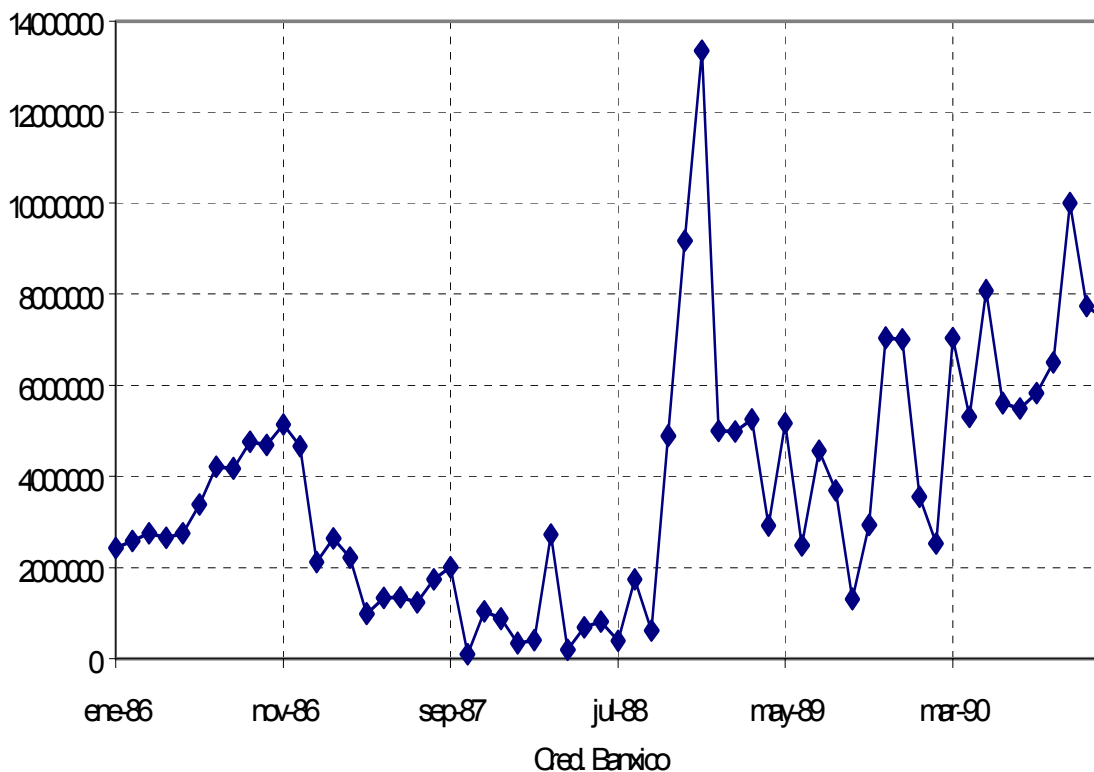


Figura 1.10.: Gráfica de una serie de tiempo con estacionalidad. Serie del crédito vía Banxico.

Las técnicas de pronóstico para datos con estacionalidad son usadas en las siguientes situaciones:

- ♦ Donde el clima afecta la variable de interés. Algunos ejemplos son el consumo de energía eléctrica, actividades de verano e invierno (deportes como el ski), moda y estaciones de cultivo agrícola.
- ♦ Donde el calendario anual tiene influencia en la variable de interés. Algunos ejemplos son las ventas al menudeo en períodos vacacionales, puentes y calendarios escolares.

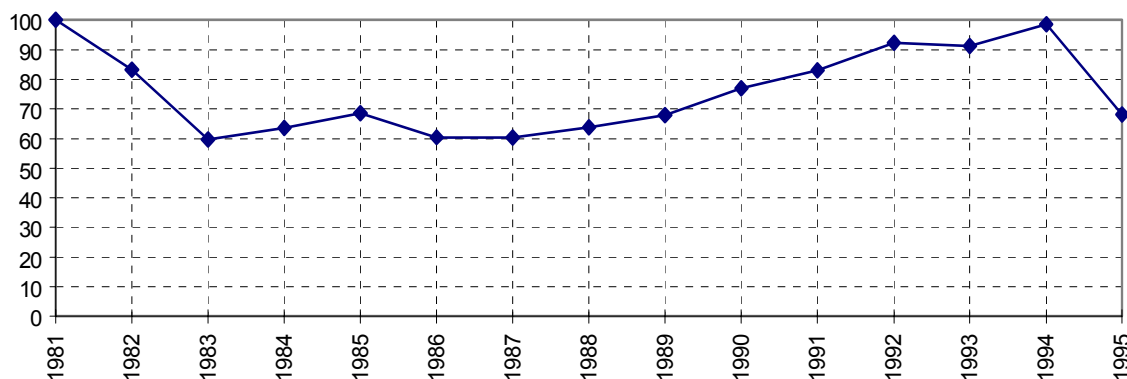
Las técnicas que deben ser consideradas cuando se pronostican series estacionales incluyen la descomposición clásica, Census II, el suavizamiento exponencial de Winter y la regresión múltiple de series de tiempo.

Técnicas de pronóstico para series cíclicas

Una serie cíclica se define como la fluctuación en onda alrededor de la tendencia. Los patrones cíclicos tienden a repetirse en los datos por períodos mayores a un año, (por ejemplo cada dos, tres o más años). Los patrones cíclicos son difíciles de modelar debido a que no son estables. Las fluctuaciones hacia arriba o hacia abajo de la tendencia rara vez se repiten en intervalos fijos de tiempo y la magnitud de las fluctuaciones usualmente varía. Los métodos de descomposición pueden ser extendidos para analizar los datos cíclicos. Sin embargo, debido al comportamiento irregular de los ciclos, analizar la componente cíclica de las series generalmente requiere encontrar indicadores económicos líderes o al menos indicadores sectoriales coincidentes, **figura 1.11**.

La Inversión en México 1982 - 1995

(Índice 1981 = 100)



Fuente: Elaboración Propia. Sistema de Cuentas Nacionales, INEGI

Figura 1.11.: Gráfica de una serie de tiempo cíclica. Serie de la inversión en México.

Las técnicas de pronóstico para datos cíclicos son usadas en las siguientes situaciones:

- ◆ Donde el ciclo de los negocios tiene influencia en la variable de interés. Por ejemplo, los factores económicos, de mercado y de competencia.
- ◆ Donde ocurren alteraciones en el gusto popular. Por ejemplo, la moda, la música y la comida.
- ◆ Donde ocurren cambios en el desarrollo de la población. Por ejemplo, a raíz de guerras, hambrunas, epidemias y desastres naturales.

- ♦ Donde ocurren alteraciones en el ciclo de vida de un producto. Por ejemplo, desde la introducción, crecimiento, consolidación, saturación del mercado y decaimiento del producto.

Las técnicas que pueden ser consideradas cuando se pronostican series cíclicas incluyen la descomposición clásica, los indicadores económicos, modelos econométricos y la regresión múltiple.

Otros factores a considerar cuando se selecciona una técnica de pronóstico

El horizonte de tiempo a ser pronosticado tiene una relación directa con la elección de una técnica de pronóstico. Para el corto y mediano plazo pueden ser aplicadas una gran variedad de técnicas cuantitativas.

Sin embargo, conforme el horizonte del pronóstico se incrementa, algunas de estas técnicas se vuelven menos aplicables. Por ejemplo, las medias móviles y el suavizamiento exponencial son predictores pobres de puntos de intervención; en este caso los modelos econométricos son más útiles. Los modelos de regresión son apropiados para el corto, mediano y largo plazo. Las medias móviles y la descomposición clásica son técnicas cuantitativas apropiadas para horizontes de corto y mediano plazo. Los métodos cualitativos son frecuentemente usados para horizontes de tiempo más largos.

La aplicabilidad de las técnicas de pronóstico es generalmente algo que el analista basa en su experiencia. Los directivos generalmente necesitan pronósticos en muy corto tiempo. El suavizamiento exponencial, los modelos de regresión y los métodos clásicos de descomposición tienen una ventaja ante esta situación.

Finalmente, los pronósticos serán presentados a los directivos para su aprobación y uso en el proceso de toma de decisiones. Por lo tanto, su fácil entendimiento e interpretación de resultados es una consideración importante. Los modelos de regresión, la descomposición clásica y las técnicas de suavizamiento exponencial, manejan altamente este criterio.

Los costos de cálculo ya no significan un obstáculo para seleccionar una técnica en especial. Las nuevas PCs y el software estadístico se han vuelto herramientas comunes en los lugares de trabajo. Debido a este desarrollo, en el futuro algunas de las restricciones a los criterios mencionados, probablemente ya no serán consideradas como costos de cálculo.

Como parte de la selección final, cada técnica debe ser evaluada por el analista en términos de confiabilidad y aplicabilidad al problema, su valor en términos de efectividad comparada con otras técnicas apropiadas, su nivel de eficiencia, su costo y su aceptación por la alta dirección.

CAPÍTULO II. PRONÓSTICOS A PARTIR DE LA INFORMACIÓN FINANCIERA Y DEL ANÁLISIS NO PARAMÉTRICO

Examinaremos cuatro técnicas básicas de suavizamiento, todas ellas tienen la característica de que sólo es necesario el pasado histórico de la serie de tiempo para efectuar un pronóstico, además están basadas en la suposición de que existe un patrón de comportamiento en los datos. Esto es, se asume que todas las series de tiempo a ser pronosticadas tienen algunos ciclos o fluctuaciones que tienden a repetirse. Los cuatro métodos a examinar son:

- ◆ Medias móviles
- ◆ Suavizamiento exponencial simple
- ◆ Suavizamiento exponencial de Holt
- ◆ Suavizamiento exponencial de Winter

Medias móviles

El simple método estadístico de las medias móviles replicará mejor algunos datos que una función matemática compleja. En este capítulo se utilizará la serie del Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (IPC), para ejemplificar las cuatro técnicas no paramétricas. Antes se definirá brevemente el IPC.

El IPC es el indicador que expresa el rendimiento del mercado accionario en función de las variaciones de precios de una muestra balanceada, ponderada y representativa del conjunto de acciones cotizadas en la Bolsa Mexicana de Valores.

Este indicador, aplicado en su actual estructura desde 1978, expresa en forma fidedigna la situación del mercado bursátil y su dinamismo operativo (bursatilidad). Las fluctuaciones en la cotización de cada título responden a la libre concertación de la oferta y la demanda en el salón de remates, relacionadas a su vez con el desarrollo de las empresas emisoras y sus resultados, así como con las condiciones generales de la economía.

La tendencia general de las variaciones de precios de todas las emisoras y series cotizadas en Bolsa, generadas por las operaciones de compra - venta en cada sesión de remates, se refleja automáticamente en el IPC.

EL IPC es un índice ponderado por el valor de capitalización (precio de mercado de las acciones inscritas). Esto significa que el cambio del precio de una acción integrante del IPC, influye en el comportamiento del índice en forma relativa al peso que dicha acción tiene en la muestra.

Así, el impacto en el índice de un movimiento de precios del 5% de una emisora pequeña será menor que el causado por una emisora grande con la misma variación porcentual en precio.

Por otro lado, la selección de la muestra de acciones para calcular el IPC se hace en base al criterio de bursatilidad, el cual considera un período de observación diaria de seis meses de las siguientes variables:

- Importe negociado en el período
- Volumen de acciones negociadas
- Rotación (Importe / valor de capitalización)
- Cantidad de operaciones efectuadas
- Total de días de negociación

La muestra es revisada cada dos meses observando los criterios mencionados, pudiéndose revisar anticipadamente si el caso lo amerita. Actualmente el tamaño de la muestra ha fluctuado entre 35 y 50 en los últimos años.

Una vez definido el tamaño de la muestra, el cálculo del IPC es efectuado en tiempo real, registrando automáticamente cada cambio de precio en las acciones que componen la muestra en el salón de remates de la Bolsa Mexicana de Valores.

La fórmula del cálculo de IPC es la siguiente:

$$I_t = I_{t-1} \left(\frac{\sum_i^n P_{i_t} Q_{i_t}}{\sum_i^n P_{i_{t-1}} Q_{i_{t-1}} F_{i_{t-1}}} \right)$$

Donde:

I_t	=	IPC el día t
P_{i_t}	=	Precio de la acción i el día t
Q_{i_t}	=	Cantidad de acciones inscritas de la acción i el día t
F_{i_t}	=	Factor de ajuste por derechos de la acción i el día t
$t-1$	=	Día hábil inmediato anterior
n	=	Número total de emisoras de la muestra.

Esta fórmula indica que la sumatoria del valor de capitalización de todas las emisoras incluidas en la muestra, dividida entre la sumatoria del valor de capitalización de dicha muestra del día hábil anterior, ajustada en su caso, determina el factor de variación del IPC respecto al día hábil anterior.

El factor de ajuste siempre es igual a uno excepto cuando en la emisora i se aplica un derecho o una reestructuración de capital.

La **figura 2.1** muestra los datos diarios del Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores por el período de 1994 a 1995, obtenidos de la base de datos de INFOSEL (Información Selectiva, S.A. de C.V.). La gráfica no muestra una tendencia lineal, exponencial o cuadrática. La serie parece mostrar cierta aleatoriedad, lo cual se puede eliminar con una técnica que promedie los más recientes valores.

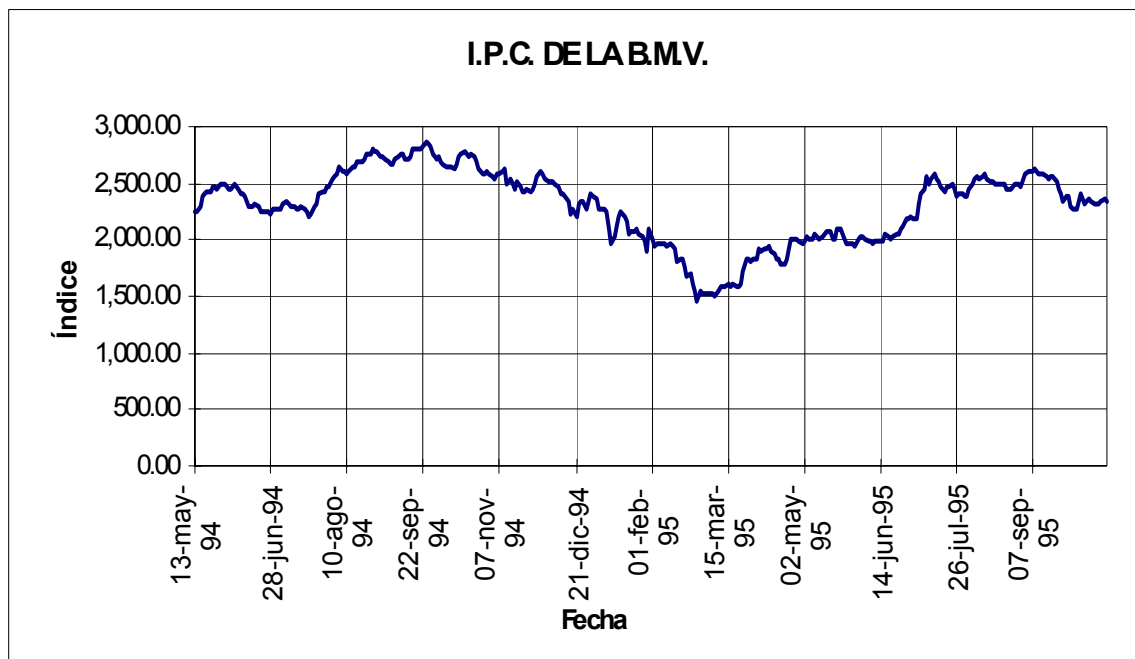


Figura 2.1.: Índice diario de la Bolsa Mexicana de Valores, mayo 94 - octubre 95.

Para calcular la media móvil de treinta días, primero tenemos que sumar las treinta primeras observaciones. Esta suma total se divide por treinta para obtener la “Media móvil de treinta días”. El número suavizado, se convierte en el pronóstico para el período siguiente.

El mismo procedimiento se seguiría si se quisiera obtener una media móvil de sesenta días. Obviamente, una media móvil de sesenta o setenta días no son los únicos tipos de medias. Podemos calcular medias de cien o de ciento veinte días, o de cualesquier unidad, si quisiéramos. La selección del intervalo de la media móvil depende de la longitud del ciclo o el patrón original de los datos.

Si creemos que los datos muestran un ciclo que se repite cada cuatro períodos, escogeríamos una media móvil de cuatro para simular mejor esa fluctuación.

Para conocer si un pronóstico es mejor con un período de treinta o un período de sesenta, es útil calcular la **raíz del error cuadrático medio** (ECM). La forma de cálculo para este valor está dada por la fórmula:

$$ECM = \sqrt{\frac{(X_t - \hat{X}_t)^2}{n}}, \text{ donde } X_t \text{ es el valor real observado y } \hat{X}_t \text{ es el valor estimado.}$$

En esta fórmula cada error o residuo es elevado al cuadrado, después los valores son sumados y divididos por n , lo cual refleja el efecto de residuos de pronóstico grandes. Esto es importante ya que una técnica que proporciona residuos moderados pudiera ser preferible a una que proporciona residuos pequeños pero ocasionalmente arroja residuos muy grandes.

En el caso del IPC el ECM es de 368.65. El pronóstico se presenta en la figura 2.2.

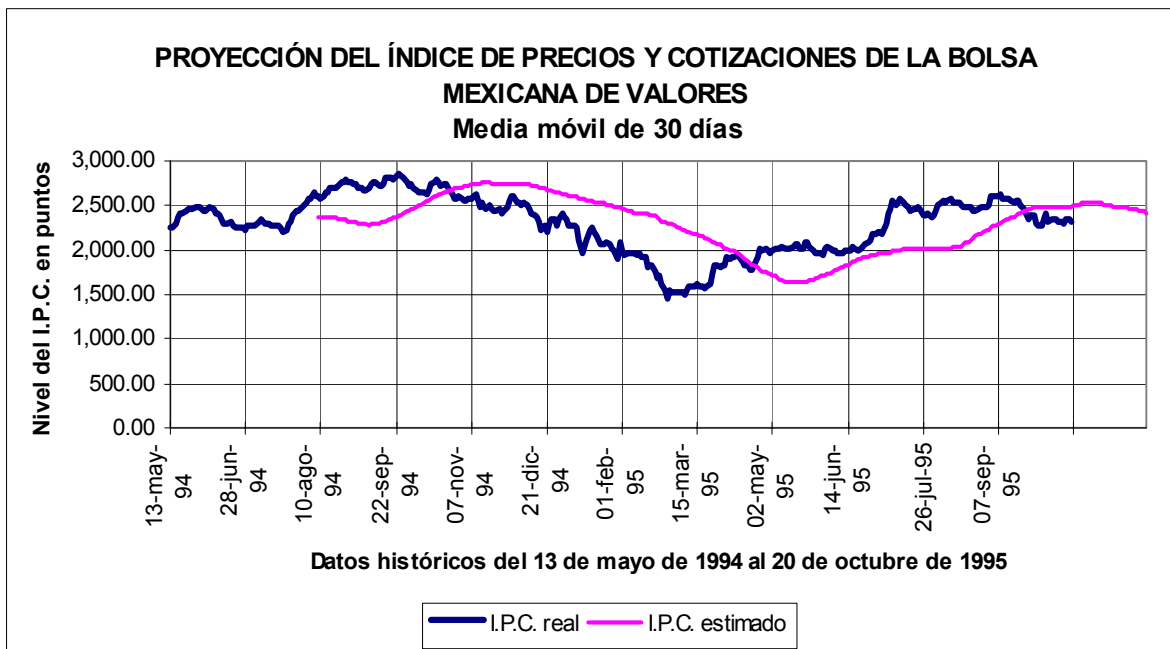


Figura 2.2.: Pronóstico del Índice diario de la Bolsa Mexicana de Valores, noviembre 95.

Una observación final importante, el método de pronóstico por medio de medias móviles, ha engañado a más de un analista cuando parece identificar un ciclo en la serie, cuando en realidad no lo hay. Tal identificación puede ser entendida si se pensara en una serie como una simple sucesión de números aleatorios. Como ninguna media móvil está correlacionada serialmente, debido a que el número de períodos contiguos ha sido promediado, ninguna secuencia de números aleatorios puede tener fluctuaciones cíclicas.

Suavizamiento exponencial simple

Continuando con el mismo ejemplo del IPC, el suavizamiento exponencial simple, al igual que las medias móviles, utiliza sólo valores pasados de la serie de tiempo para pronosticar los valores de la misma serie y es apropiada cuando no existe tendencia o estacionalidad en los datos.

Con el suavizamiento exponencial, el valor del pronóstico es un promedio ponderado de los valores previos disponibles, la ponderación disminuye conforme se retrocede en el tiempo.

El pronóstico de la media móvil proporciona la misma ponderación a los valores incluidos en cada promedio; el suavizamiento exponencial da mayor peso a las observaciones más recientes y menor a las observaciones más remotas. La ponderación declina geométricamente conforme se retrocede en el tiempo, asumiendo que las observaciones más recientes contienen información más relevante, así que deberán tener mayor influencia que las más atrasadas.

El suavizamiento exponencial funciona como las medias móviles al suavizar los valores pasados de la serie; los cálculos para obtener pronósticos suavizados exponencialmente pueden ser expresados en una ecuación. El peso de las observaciones más recientes es asignado multiplicando el valor observado por α , la siguiente observación más reciente por $(1-\alpha)\alpha$, la siguiente por $(1-\alpha)^2\alpha$, y así sucesivamente.

El número que se escoge para α se conoce como constante de suavizamiento, y el modelo puede ser expresado como sigue:

$$F_{t+1} = \alpha X_t + (1 - \alpha)F_t$$

donde:

F_{t+1}	=	Valor del pronóstico para el período t+1
α	=	Constante de suavizamiento para la serie ($0 < \alpha < 1$)
X_t	=	Valor real en el período t
F_t	=	Pronóstico suavizado para el período t, el cual es también el valor suavizado para el período t-1

Para utilizar esta ecuación el analista no necesita trabajar con cada valor pasado en cada paso, sólo son necesarios los valores suavizados exponencialmente para el último período y el valor real de este período. Una forma alternativa de escribir la ecuación anterior resulta del siguiente desarrollo.

$$F_{t+1} = \alpha X_t + (1 - \alpha)F_t$$

$$= \alpha X_t + F_t - \alpha F_t$$

$$= F_t + \alpha(X_t - F_t)$$

A partir de esta forma podemos ver que el suavizamiento exponencial “aprende” de sus errores. El pronóstico para el período $t+1$ se incrementa si el valor real para el período t es más grande que el que se había pronosticado y decrementa si X_t es menor que F_t .

Pronosticar el valor para el próximo período (F_{t+1}) requiere conocer el valor real de éste período (X_t) y el valor pronosticado para éste período (F_t). Sin embargo, todas las observaciones históricas están incluidas como se muestra a continuación:

$$F_{t+1} = \alpha X_t + (1 - \alpha)F_t$$

$$F_t = \alpha X_{t-1} + (1 - \alpha)F_{t-1}$$

$$F_{t+1} = \alpha X_t + (1 - \alpha)\alpha X_{t-1} + (1 - \alpha)^2 F_{t-1}$$

$$F_{t-1} = \alpha X_{t-2} + (1 - \alpha)F_{t-2}$$

$$F_{t+1} = \alpha X_t + (1 - \alpha)\alpha X_{t-1} + (1 - \alpha)^2 \alpha X_{t-2} + (1 - \alpha)^3 F_{t-2}$$

Podríamos continuar con esta expansión para incluir X términos tanto como datos tengamos disponibles con el fin de observar como las ponderaciones para períodos previos de tiempo son cada vez más pequeñas a una tasa que depende del valor de α , como se muestra para dos valores alternativos de α .

**ERROR CUADRÁTICO MEDIO OBTENIDO PARA
DISTINTAS ALFAS PROPUESTAS
ALFA ÓPTIMO: 0.999**

α	ERROR CUADRÁTICO MEDIO
0.1	116.63
0.2	81.46
0.3	66.89
0.4	58.61
0.5	53.28
0.6	49.61

0.7	46.99
0.8	45.10
0.9	43.76
0.999	42.88

El valor de la constante de suavizamiento α debe ser entre 0 y 1. No puede ser igual a 0 o 1; si esto ocurriera la idea de suavizamiento exponencial sería negada. Si se escoge un valor cercano a 1, los valores recientes de la serie de tiempo son fuertemente ponderados en comparación a aquellos de un pasado distante. Por otro lado, si el valor de α es cercano a 0, entonces los valores distantes de la serie tienen un peso comparable al dado a los valores recientes.

Sin importar el valor de la constante de suavización escogida, la suma de las ponderaciones tenderá a 1. El hecho de que la suma de las ponderaciones tienda a 1 rápida o lentamente depende de la constante elegida. Si por ejemplo, escogemos una constante de 0.9, la suma de las ponderaciones se aproximará a 1 mucho más rápido que cuando se escoge 0.01.

Como guía para escoger una α adecuada se consideran los siguientes criterios:

- Se seleccionan valores cercanos a 0 si la serie tiene una gran variación de carácter aleatorio.
- Si se desea que el pronóstico dependa fuertemente de los cambios recientes en los valores reales se escoge un valor cercano a 1.

La raíz cuadrada del error cuadrático medio, ECM, es comúnmente usada como un criterio para asignar una constante de suavizamiento apropiada, en virtud de que proporcionará el error más pequeño al generar un pronóstico.

En la práctica, los valores de α entre 0.05 y 0.30 son los más óptimos cuando se trabaja con modelos de suavizamiento simple.

En el caso del pronóstico para el IPC de la BMV se tiene que el mínimo valor para el error cuadrático medio es de 42.88 con un α de 0.999. El pronóstico se muestra en la **figura 2.3**.

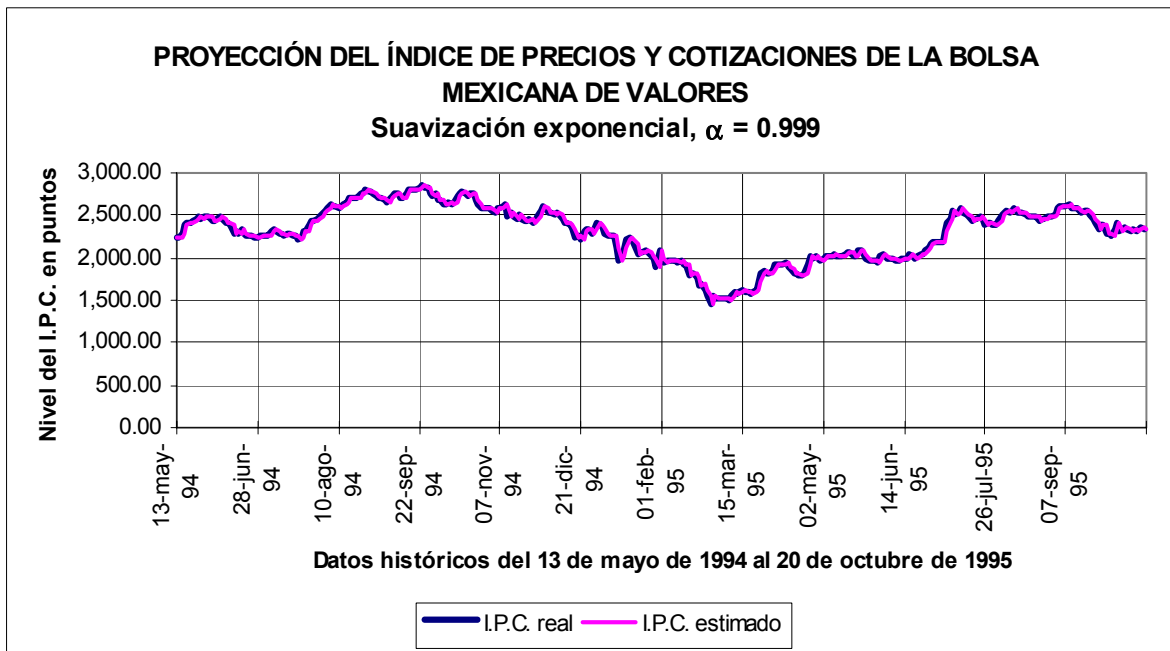


Figura 2.3.: Pronóstico del Índice diario de la Bolsa Mexicana de Valores, noviembre 95.

La razón principal para usar el suavizamiento exponencial simple es un número limitado de datos y porque es más simple que cualquier otro de los métodos de pronóstico. Sin embargo, sus limitaciones son que rezagan los datos reales, además de no tener la habilidad de ajustarse a tendencias o estacionalidad que presenten los datos, por lo que éste modelo tenderá a cometer graves errores en sentido positivo, negativo y viceversa.

Suavizamiento exponencial de Holt

Con efectos de tendencia presente en los datos, la técnica de pronóstico tiene que ser mejorada a través del suavizamiento exponencial de Holt, desarrollado por C. C. Holt. Este método consta de dos parámetros y es una extensión del suavizamiento exponencial simple, incluye un factor de crecimiento o de tendencia en la ecuación de suavizamiento como una manera de ajustarse a la tendencia.

El modelo está representado por tres ecuaciones y dos constantes de suavizamiento.

$$F_{t+1} = \alpha X_t + (1 - \alpha)(F_t + T_t)$$

$$T_{t+1} = \beta(F_{t+1} - F_t) + (1 - \beta)T_t$$

$$H_{t+m} = F_{t+1} + mT_{t+1}$$

donde:

F_{t+1}	=	Valor suavizado para el período $t+1$
α	=	Constante de suavizamiento para la serie ($0 < \alpha < 1$)
X_t	=	Valor real en el período t
F_t	=	Pronóstico suavizado para el período t , el cual es también el valor suavizado para el período $t-1$
T_{t+1}	=	Tendencia estimada
β	=	Constante de suavizamiento para la tendencia estimada ($0 < \beta < 1$)
m	=	Número de períodos futuros a pronosticar
H_{t+m}	=	Pronóstico de Holt para el período $t+m$

La primera ecuación ajusta F_{t+1} al crecimiento del período anterior, T_t , sumando T_t al valor suavizado del período anterior, F_t . La tendencia estimada está representada en la ecuación de T_{t+1} donde se calcula la diferencia de los dos últimos valores suavizados. Como estos dos últimos valores ya están suavizados, se asume que su diferencia es un estimador de la tendencia de los datos. La segunda constante de suavizamiento, β en la ecuación de T_{t+1} se obtiene del mismo principio utilizado en el método de suavizamiento exponencial simple. La tendencia más reciente ($F_{t+1} - F_t$), es ponderada por β y la última tendencia suavizada, T_t , es ponderada por $(1-\beta)$. La suma de los valores ponderados es el nuevo valor de la tendencia suavizada, T_{t+1} .

La ecuación de H_{t+m} es utilizada para pronosticar m períodos futuros sumando el producto de la componente de tendencia T_{t+1} , y el número de períodos a pronosticar, m , al valor real de los datos suavizados F_{t+1} .

Este método es eficiente para cualquier serie con tendencia lineal. Para iniciar se requieren dos valores: uno para el primer valor suavizado y otro para el primer valor de la tendencia. Generalmente el valor inicial suavizado es el último valor real disponible; el valor inicial de la tendencia es 0.00 si no se tienen datos anteriores disponibles.

El modelo de Holt puede ser evaluado examinando los errores y el ECM. Para nuestro ejemplo del IPC los resultados que se obtienen para las constantes propuestas es el siguiente:

ERROR	433.14
CUADRÁTICO MEDIO	

$\alpha =$	0.4
$\beta =$	0.1
$m =$	30

La **figura 2.4** muestra un mapeo de los valores reales y los valores generados por el modelo, el ajuste de la tendencia mejoró el pronóstico y quizá algún cambio en las constantes de suavizamiento incrementaría la eficiencia.

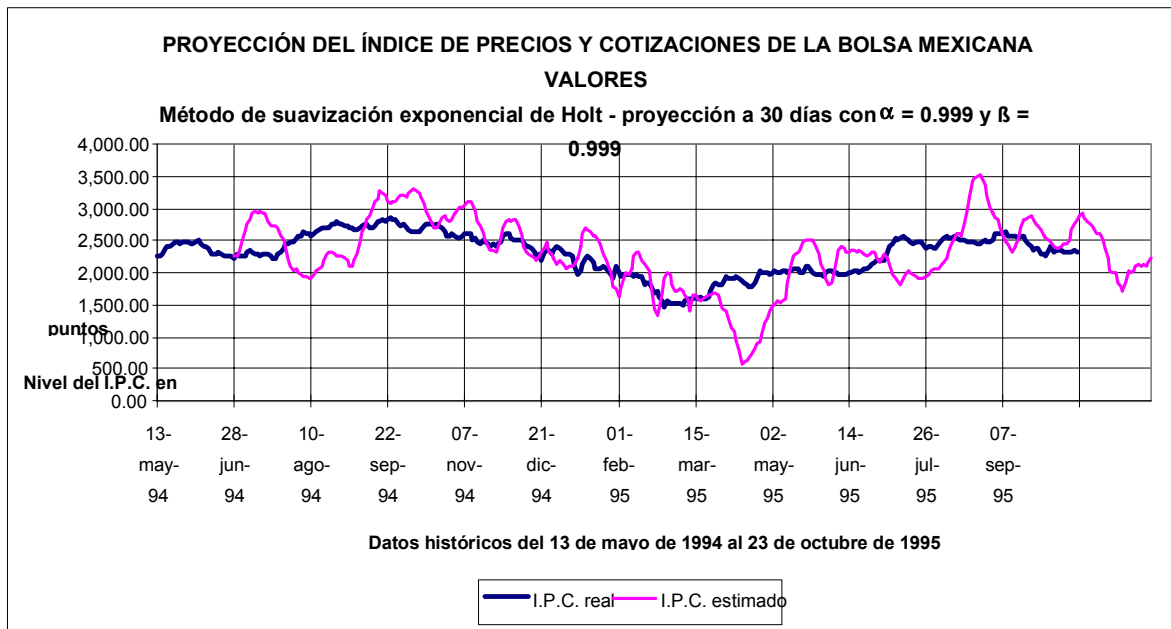


Figura 2.4.: Pronóstico del Índice diario de la Bolsa Mexicana de Valores, noviembre 95.

El método de suavizamiento exponencial de Holt es el apropiado cuando los datos muestran tendencia lineal y poca o nula estacionalidad. Un nombre descriptivo para el suavizamiento de Holt sería suavizamiento de tendencia lineal.

Suavizamiento exponencial de Winters

El suavizamiento exponencial de Winters es la segunda extensión del modelo básico de suavizamiento; es usado cuando los datos muestran tendencia y estacionalidad, además de ser un modelo de tres parámetros que es extensión del modelo de Holt. Una ecuación adicional ajusta el modelo para la componente estacional. Las cuatro ecuaciones necesarias para el modelo de Winters son:

$$F_t = \alpha X_t / S_{t-m} + (1 - \alpha)(F_{t-1} + T_{t-1})$$

$$S_t = \beta X_t / F_t + (1 - \beta) S_{t-m}$$

$$T_t = \gamma (F_t - F_{t-1}) + (1 - \gamma) T_{t-1}$$

$$W_{t+m} = (F_t + m T_t) S_t$$

donde:

F_t	=	Valor suavizado para el período t
α	=	Constante de suavizamiento para la serie ($0 < \alpha < 1$)
X_t	=	Valor real en el período t
F_{t-1}	=	Pronóstico suavizado para el período t, el cual es también el valor suavizado para el período t-1
T_{t+1}	=	Tendencia estimada
S_{t+1}	=	Estacionalidad estimada
β	=	Constante de suavizamiento para la estacionalidad estimada ($0 < \beta < 1$)
γ	=	Constante de suavizamiento para la tendencia estimada ($0 < \gamma < 1$)
m	=	Número de períodos al año (12 para datos mensuales, 4 para datos trimestrales)
W_{t+m}	=	Pronóstico de Winters para m períodos futuros

La ecuación de F_t actualiza la serie suavizada con la tendencia y la estacionalidad; nótese que la ecuación es ligeramente diferente de la ecuación para F_t del modelo de Holt. En la ecuación para S_t , X_t se divide por S_{t-m} para ajustar la estacionalidad; esta operación desestacionaliza los datos o remueve cualquier efecto estacional dejado en ellos. Para observarlo hay que considerar que pasa cuando S_{t-m} es mayor que 1, en tanto como sería cuando el valor en el período t-m es mayor que el promedio en su estacionalidad.

Dividiendo X_t por S_{t-m} se reduce el valor original por un porcentaje igual al porcentaje de estacionalidad del período que fue arriba del promedio. Un ajuste opuesto tendría lugar si el período estuviera por debajo del promedio en términos de estacionalidad.

La estacionalidad estimada a sí misma es suavizada en la ecuación S_{t+1} ; cada uno de estos procesos es exactamente el mismo de un suavizamiento exponencial simple. La última ecuación para W_{t+m} es usada para calcular el pronóstico para m períodos futuros, el procedimiento es idéntico al del modelo de Holt en la ecuación H_{t+m} .

El modelo de Winters puede ser también evaluado examinando los resultados del ECM. Para el caso del IPC de la BMV, que se ha venido trabajando se tienen los siguientes resultados:

ERROR CUADRÁTICO MEDIO	436.17
$\alpha =$	0.8
$\beta =$	0.2
$\gamma =$	0.1
$m =$	30

La RECM de 436.17 obtenida en el modelo de Winters resultó mayor y en consecuencia menos eficiente que los correspondientes al modelo exponencial simple o el modelo de Holt.

ERROR CUADRÁTICO MEDIO OBTENIDO PARA DISTINTAS α , β y γ PROPUESTAS ALEATORIAMENTE, **ALFA ÓPTIMO: 0.8, BETA ÓPTIMO: 0.2, GAMA ÓPTIMO: 0.1**

RESULTADOS: 10³	α	β	γ	ERROR CUADRÁTICO MEDIO
1	0.9	0.3	0.4	745.46
2	0.7	0.9	0.7	1,072.13
3	0.5	0.6	0.7	1,046.06
4	0.8	0.9	0.4	759.52
5	0.3	0.9	0.9	8,258.15
6	0.7	0.7	0.3	660.62
7	0.1	0.4	0.6	1,705.77
8	0.4	0.2	0.5	806.44
9	0.4	0.2	0.1	436.22
10	0.6	0.2	0.9	1,098.07
11	0.6	0.9	0.3	683.66
12	0.4	0.7	0.9	1,911.27
13	0.3	0.4	0.5	892.11
14	0.4	0.2	0.5	806.44
15	0.5	0.1	0.8	994.65
16	0.6	0.8	0.3	675.32
17	0.8	0.2	0.1	436.17
18	0.8	0.4	0.6	912.56
19	0.9	0.1	0.8	1,079.46
20	0.4	0.1	0.4	726.33
21	0.8	0.5	0.7	1,003.43
22	0.1	0.1	0.8	969.91
23	0.1	0.9	0.5	28,053.49
24	0.4	0.4	0.1	440.39
25	0.1	0.8	0.8	20,518.13
26	0.5	0.2	0.3	644.65
27	0.9	0.5	0.1	437.41
28	0.9	0.9	0.5	845.27

La misma conclusión puede ser obtenida examinando la evidencia visual; la **figura 2.5.** muestra el mapeo de los valores reales y pronosticados para este modelo. El modelo exponencial de Winters es el más apropiado cuando los datos

presentan tendencia lineal y estacionalidad, aunque en el caso del IPC no se ajustó como se esperaba debido a la alta volatilidad del mercado.

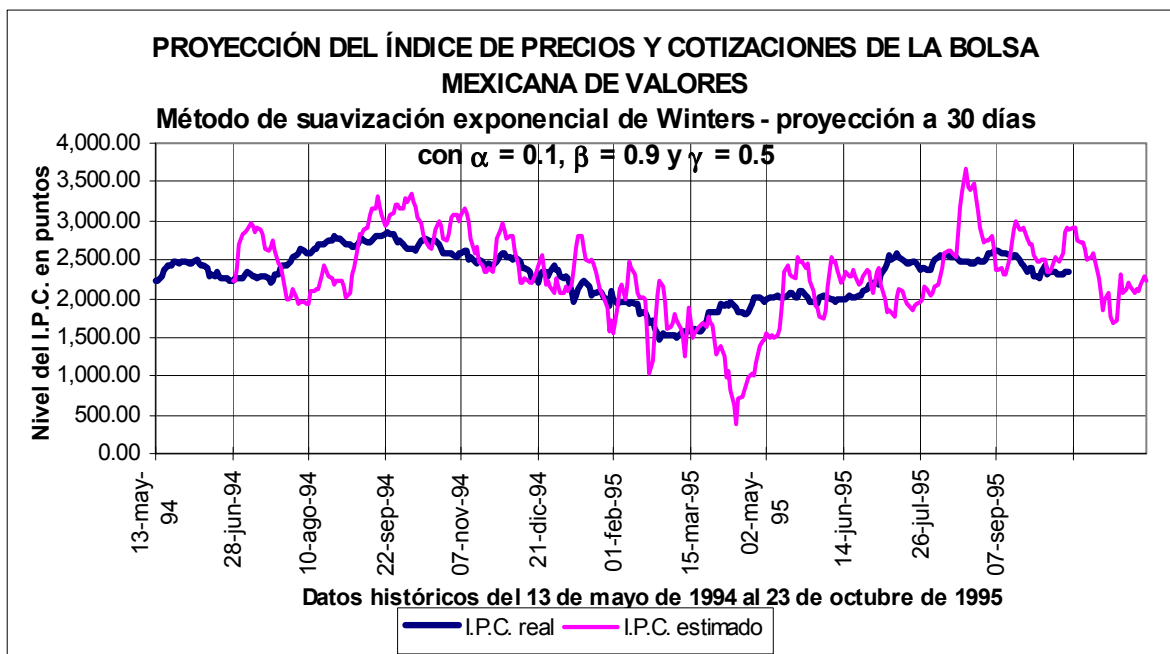


Figura 2.5.: Pronóstico del Índice diario de la Bolsa Mexicana de Valores, noviembre 95.

El Anexo I muestra los cálculos numéricos para cada uno de los métodos descritos en este capítulo y que por razones de espacio sólo se han presentado sus resultados.

Descomposición de series de tiempo

Las series de tiempo están sujetas a tendencias, ciclos, estacionalidad y variaciones irregulares o aleatorias, las cuales cuando se examinan individualmente pueden ayudar a entender mejor los movimientos de los datos y por lo tanto elaborar un mejor pronóstico.

El modelo de descomposición de series de tiempo puede ser usado para identificar tales componentes, reensamblando las partes para construir un pronóstico. Estos modelos se encuentran entre las más antiguas técnicas de pronóstico aunque continúan vigentes. Su popularidad se debe a tres factores principales:

- ♦ En muchas situaciones reales proporciona excelentes pronósticos

- ♦ Son fáciles de entender y de explicar al usuario del pronóstico, esto aumenta la probabilidad de que el pronóstico será correctamente interpretado y usado apropiadamente.
- ♦ La información de series de tiempo proporcionada por la descomposición de series de tiempo es consistente con la manera en que la alta dirección tiende a observar los datos y ayuda a obtener factores de medición que de otra manera no podrían ser cuantificados.

El modelo clásico de descomposición de series de tiempo puede ser representado por la siguiente ecuación algebraica:

$$Y = T \cdot S \cdot C \cdot I$$

Y	=	La variable a ser pronosticada
T	=	Es el término que define la tendencia de los datos
S	=	Es el factor de estacionalidad
C	=	Es el factor de ajuste cíclico
I	=	Variaciones aleatorias

El ejemplo a analizar con la técnica de descomposición de series de tiempo es también la serie del IPC de la BMV, con datos diarios del 13 de mayo de 1994 al 29 de septiembre de 1995.

Desestacionalización de los datos

El primer paso es remover las fluctuaciones de corto plazo con el fin de que la tendencia de largo plazo y la componente del ciclo puedan ser más fácilmente identificadas. Las fluctuaciones de corto plazo incluyen tanto variaciones estacionales como variaciones irregulares. Estas pueden ser removidas calculando la media móvil apropiada para la serie. La media móvil debe contener el mismo número de períodos que existen en la estacionalidad que se pretende identificar. Por ejemplo, si se tienen datos trimestrales y se detecta que la estacionalidad es de ese orden entonces se aplica una media móvil de cuatro períodos y así sucesivamente. La media móvil para el período t se calcula como sigue:

Para datos trimestrales:

$$MA_t = (Y_{t-2} + Y_{t-1} + Y_t + Y_{t+1}) / 4$$

donde MA son las siglas en inglés de la media móvil (moving average).

Para datos mensuales:

$$MA_t = (Y_{t-6} + Y_{t-5} + \dots + Y_t + Y_{t+1} + \dots + Y_{t+5}) / 12$$

La media móvil para cada período contiene un elemento de cada una de las estaciones. Por ejemplo, en el caso de datos trimestrales cada media móvil contendría observaciones del primero, segundo, tercero y cuarto trimestre. El promedio de estas observaciones ya no tendría ninguna estacionalidad. Por lo tanto, la media móvil representa un nivel típico de Y para el año en el que se centra la media móvil. Sin embargo, cuando se utiliza un número par de períodos al calcular la media móvil, realmente no está centrada en el año. Cuando el número de períodos usado es impar la media móvil se centra automáticamente y no se necesita ningún ajuste adicional.

Para centrar la media móvil se calcula una segunda media móvil de la media móvil. Está última es conocida como una media móvil central (central moving average, CMA) y se calcula como sigue:

$$CMA_t = (MA_t + MA_{t+1}) / 2$$

Esta segunda media móvil ayuda a suavizar fluctuaciones aleatorias o irregulares en los datos. El proceso de suavizamiento tiene un costo en términos de la pérdida de datos. Si se utiliza una media móvil de n períodos, al calcular la media móvil central se perderán $n/2$ observaciones.

La media móvil representa los datos desestacionalizados, es decir, las variaciones estacionales han sido removidas a través de un proceso de promediar. Para comparar el valor real de la serie en cualquier período Y_t con el valor desestacionalizado CMA_t se puede obtener la medida del grado de estacionalidad, esto se hace encontrando el ratio del valor real y el del valor desestacionalizado. El resultado se conoce como factor estacional (seasonal factor SF_t), y se denota por:

$$SF_t = \frac{Y_t}{CMA_t}$$

Un SF mayor a uno indica un período en el cual Y es mayor que el promedio anual, mientras que la afirmación contraria es cierta si SF es menor a uno.

Cuando observamos los factores estacionales para un período extenso de tiempo, generalmente se ve consistencia en los valores de cada estación. No se esperaría que todos los factores fueran exactamente iguales. Para establecer un índice de estacionalidad se promedian los factores de estacionalidad para cada estación. La **figura 2.6.** muestra el mapeo de la serie original del I.P.C. junto a la serie desestacionalizada, representada por la media móvil central. Nótese que es mucho más suave, una vez que las variaciones estacionales y las fluctuaciones aleatorias han sido removidas. Recuérdese que los factores de estacionalidad miden que tanto está por arriba o por abajo el valor observado del valor desestacionalizado, ($SF > 1$ y $SF < 1$, respectivamente).

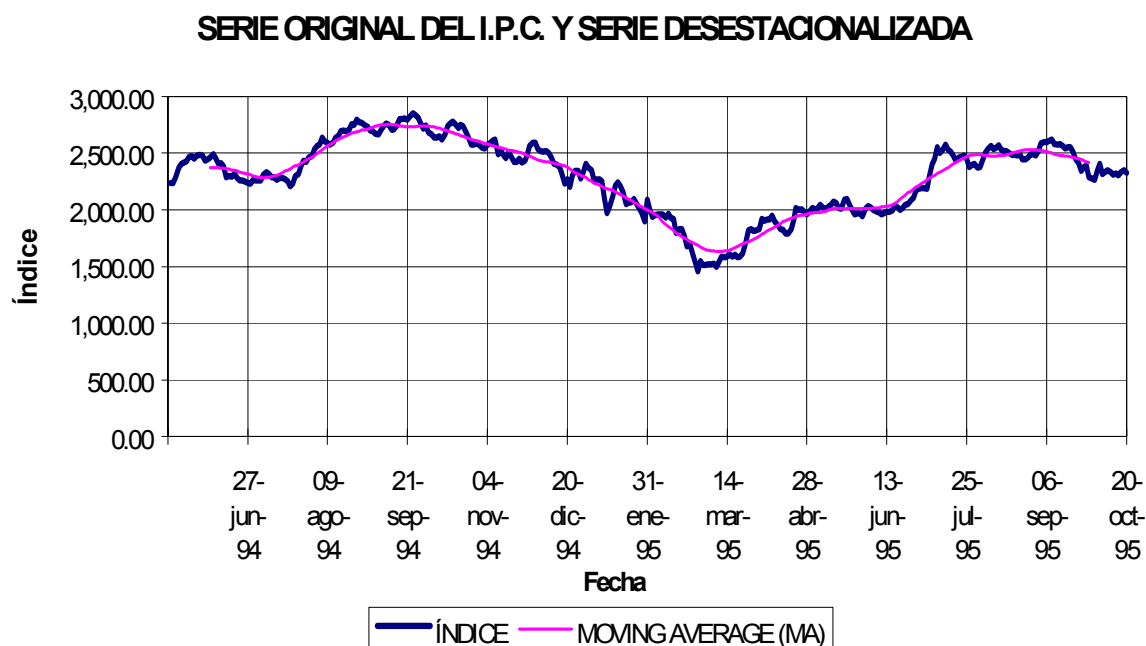


Figura 2.6.: Serie original y desestacionalizada del Índice diario de la Bolsa Mexicana de Valores, mayo 94 - octubre 95.

La determinación de los índices de estacionalidad se muestra en la **tabla 2.7.** La suma del promedio de los factores de estacionalidad debe ser igual al número de períodos (4 para trimestres, 12 para meses). Si no lo son deberán ser normalizados multiplicando cada factor promedio de estacionalidad por el radio del número de períodos (4 ó 12) a la suma del promedio de factores estacionales. Este proceso se muestra al final de la **tabla 2.7.**

Los datos desestacionalizados nos permiten ver mejor el patrón de comportamiento de los datos.

Provee medidas de la magnitud de la estacionalidad en la forma de índices de estacionalidad.

Otra aplicación de los índices de estacionalidad es en la proyección de lo que un período específico puede aportar en el comportamiento de un año entero.

Tendencia

La tendencia de largo plazo se estima a partir de los datos desestacionalizados para la variable a ser pronosticada. Recuérdese que la media móvil central (CMA), es la serie que permanece después que han sido suavizadas la estacionalidad y los componentes aleatorios por medio de la media móvil. Así, para estimar la tendencia de largo plazo, se estima una ecuación lineal de la forma:

$$\begin{aligned} \text{CMA} &= f(\text{tiempo}) \\ &= a + b(\text{tiempo}) \end{aligned}$$

donde tiempo = 1, para el primer período en la serie y se incrementa en los trimestres sucesivos. Los valores a y b son estimados normalmente por medio de una regresión.

Una vez que la ecuación de tendencia ha sido determinada, se utiliza para generar y estimar el valor de la tendencia de la media móvil central para los períodos pasados y futuros. Esta nueva serie es la media móvil central tendencial (central moving average trend, CMAT).

Ciclo

La componente cíclica de las series de tiempo es el siguiente paso después de la tendencia de largo plazo. Se mide con un factor cíclico (cycle factor, CF), el cual es el radio de la media móvil central (CMA), a la media móvil central tendencial, (CMAT). Es decir:

$$\text{CMA} = \text{CMAT} / \text{CF}$$

Un factor cíclico mayor a uno indica que el valor desestacionalizado para ese período es superior a la tendencia de largo plazo.

Si el CF es menor que uno entonces ocurre lo contrario. El factor cíclico es la componente más difícil de analizar y pronosticar de una serie de tiempo. Si se analiza cuidadosamente puede ser la componente que más puede aportar al entendimiento de por qué un sector económico crece. Esta es la ventaja más grande que tiene la técnica de descomposición de series de tiempo.

En la **tabla 2.7.** se muestran los factores cíclicos (CF); se observa que se mueven alrededor de uno, de la misma forma en que la media móvil central lo hace alrededor de la línea de tendencia. Aislando el factor cíclico en la **figura 2.8** se pueden analizar mejor los movimientos a través del tiempo.

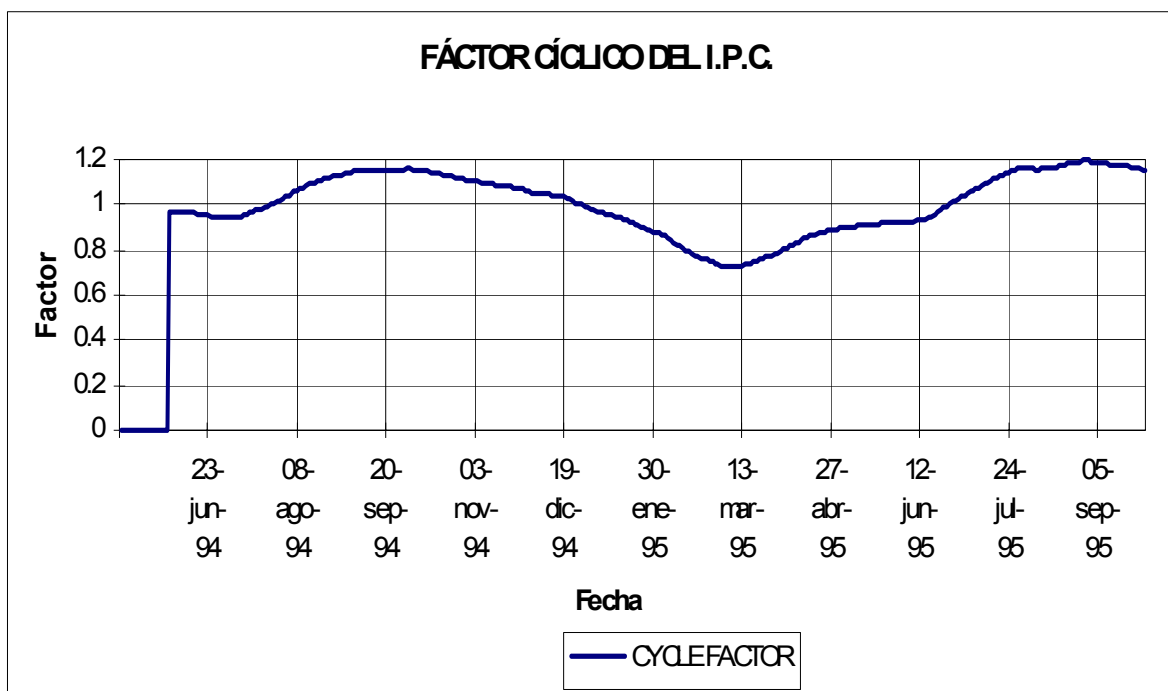


Figura 2.8.: Factor cíclico del Índice diario de la Bolsa Mexicana de Valores, mayo 94 - octubre 95.

La determinación de donde debe estar el factor cíclico en el horizonte del pronóstico es una tarea difícil. Una aproximación será usar otro método de pronóstico para estimar valores del CF. El suavizamiento exponencial de Holt puede ser utilizado para esta tarea.

Pronóstico por medio de descomposición de series de tiempo

Se ha visto que las series de tiempo pueden ser descompuestas en cuatro componentes básicos.

$$Y = T \cdot S \cdot C \cdot I$$

Donde Y es la serie a ser pronosticada. Los cuatro componentes son:

T = La tendencia de largo plazo basada en la desestacionalización de los datos. Generalmente llamada la media móvil central tendencial (CMAT), debido a que los datos desestacionalizados son las medias móviles centrales de los valores Y originales.

S = Los índices de estacionalidad (SI). Estos promedios normalizados de los factores de estacionalidad que son determinados como el radio de los valores reales de cada período (Y) a los valores desestacionalizados (CMA) para ese período.

C = El componente cíclico. El factor cíclico (CF) es el radio de la CMA a la CMAT y representa los movimientos graduales alrededor de la serie.

I = La componente irregular. Se asume que esta es igual a uno a menos que el analista tenga razones para suponer que pueda ocurrir algún choque, en cuyo caso la I puede ser diferente de uno para todo o parte del período de pronóstico.

Para obtener un pronóstico simplemente se reensamblan los componentes. En términos generales el pronóstico para Y, (FY) es:

$$FY = (CMAT)(SI)(CF)(I)$$

En el caso del IPC, se asume que el factor irregular o aleatorio (I) es uno ya que no hay razón para esperar que sea mayor o menor a uno dada su naturaleza aleatoria. Los cálculos se muestran en la **tabla 2.7**. Nótese que este método toma en cuenta la tendencia y hace dos ajustes: el primero por estacionalidad y el segundo por las variaciones cíclicas, **figura 2.9**.

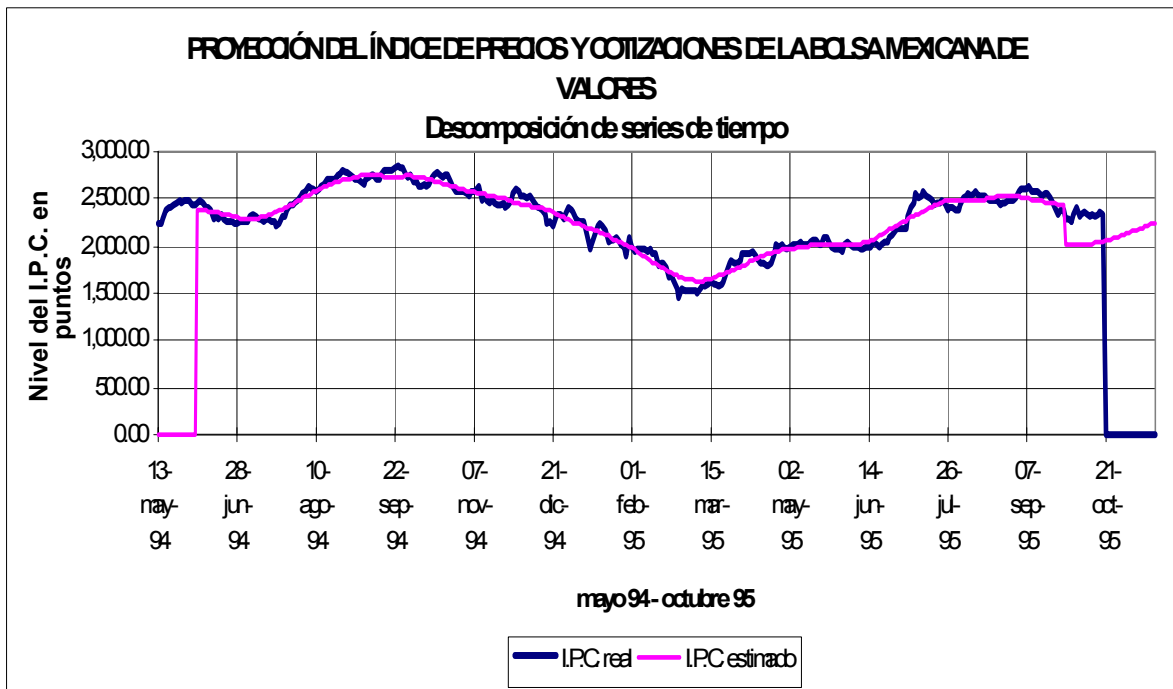


Figura 2.9.: Pronóstico del IPC con el método de descomposición de series de tiempo, nov. 95.

Debido a que el modelo de descomposición de series de tiempo no incluye demasiadas matemáticas o estadística, son relativamente fáciles de explicar al usuario final. Esto tiene la ventaja de poseer una mejor apreciación de como fue hecho el pronóstico y tendrá mayor confianza en el momento de tomar alguna decisión.

PROYECCIÓN DEL ÍNDICE DE PRECIOS Y COTIZACIONES DE LA BOLSA MEXICANA DE VALORES

SUM SEASONAL FACTOR		329.85
SEASONAL INDEX		1
REGRESIÓN		Coefficientes
Intercepción		2,466.23
Variable X 1		-1.07

NÚMERO	FECHA	ÍNDICE	MOVING AVERAGE (MA)	CENTRAL MOVING AVERAGE (CMA)	SEASONAL FACTOR (SF)	CENTERED MOVING AVERAGE TREND (CMAT)	CYCLE FACTOR (CF)	PRONÓSTICO Y = (CMA)(SI)(CF)(I)
1	13-may-94	2,240.17				2,465.16	0.00	-
2	16-may-94	2,239.84				2,464.10	0.00	-
3	17-may-94	2,288.59				2,463.03	0.00	-
.
.
.
343	27-sep-95	2,343.06	2,445.66	2,441.37	0.96	2,100.67	1.16	2,441.37
344	28-sep-95	2,381.49	2,437.08	2,432.97	0.98	2,099.61	1.16	2,432.97
345	29-sep-95	2,392.26	2,428.85	2,424.33	0.99	2,098.54	1.16	2,424.33
346	02-oct-95	2,291.96	2,419.80			2,097.48	0.96	2,008.13
347	03-oct-95	2,277.08				2,096.41	0.96	2,007.89
348	04-oct-95	2,267.13				2,095.34	0.96	2,007.96
349	05-oct-95	2,347.65				2,094.28	0.96	2,008.36
350	06-oct-95	2,409.19				2,093.21	0.96	2,009.19
351	09-oct-95	2,316.77				2,092.15	0.96	2,010.55
352	10-oct-95	2,337.78				2,091.08	0.96	2,012.41
353	11-oct-95	2,353.68				2,090.02	0.96	2,014.73
354	12-oct-95	2,335.49				2,088.95	0.97	2,017.54
355	13-oct-95	2,310.60				2,087.88	0.97	2,020.92
356	16-oct-95	2,325.29				2,086.82	0.97	2,024.91
357	17-oct-95	2,305.19				2,085.75	0.97	2,029.56
358	18-oct-95	2,332.44				2,084.69	0.98	2,034.90
359	19-oct-95	2,351.57				2,083.62	0.98	2,040.91

360	20-oct-95	2,329.42				2,082.56	0.98	2,047.54
361	21-oct-95				PRONÓSTICOS	2,081.49	0.99	2,054.80
362	22-oct-95				PRONÓSTICOS	2,080.42	0.99	2,062.73
363	23-oct-95				PRONÓSTICOS	2,079.36	1.00	2,071.36
364	24-oct-95				PRONÓSTICOS	2,078.29	1.00	2,080.59
365	25-oct-95				PRONÓSTICOS	2,077.23	1.01	2,090.28
366	26-oct-95				PRONÓSTICOS	2,076.16	1.01	2,100.33
367	27-oct-95				PRONÓSTICOS	2,075.10	1.02	2,110.67
368	28-oct-95				PRONÓSTICOS	2,074.03	1.02	2,121.21
369	29-oct-95				PRONÓSTICOS	2,072.96	1.03	2,131.91
370	30-oct-95				PRONÓSTICOS	2,071.90	1.03	2,142.70
371	31-oct-95				PRONÓSTICOS	2,070.83	1.04	2,153.49
372	01-nov-95				PRONÓSTICOS	2,069.77	1.05	2,164.21
373	02-nov-95				PRONÓSTICOS	2,068.70	1.05	2,174.84
374	03-nov-95				PRONÓSTICOS	2,067.64	1.06	2,185.31
375	04-nov-95				PRONÓSTICOS	2,066.57	1.06	2,195.60
376	05-nov-95				PRONÓSTICOS	2,065.50	1.07	2,205.75
377	06-nov-95				PRONÓSTICOS	2,064.44	1.07	2,215.78
378	07-nov-95				PRONÓSTICOS	2,063.37	1.08	2,225.73
379	08-nov-95				PRONÓSTICOS	2,062.31	1.08	2,235.56
380	09-nov-95				PRONÓSTICOS	2,061.24	1.09	2,245.21

Tabla 2.7.: Pronóstico del Índice diario de la Bolsa Mexicana de Valores, may. 94 - nov. 95, a partir de la descomposición de series de tiempo.

CAPÍTULO III. MODELOS ECONOMÉTRICOS UNIECUACIONALES

Estimación de un modelo uniecuacional

Propiedades de los estimadores

Cuando se extrae una muestra de una población de la que desconocemos alguno de sus momentos, es de interés cuestionarse como utilizar la información muestral para obtener una idea del valor del parámetro desconocido. La función utilizada para resumir la información muestral es un número, que se asociará al parámetro desconocido, se llama estimador, y su valor en una muestra determinada se denomina estimación. Un estimador siendo función de la muestra, es una variable aleatoria y tiene su propia distribución de probabilidad.

Como no puede haber un estimador perfecto que siempre dé la respuesta correcta, parecería razonable que un estimador deba serlo cuando menos en promedio. Dicho de otra manera, es deseable que el valor esperado de un estimador sea igual al parámetro que se supone estima. Si éste es el caso, se dice que el estimador es **insesgado**; de lo contrario, se dice que es sesgado. Formalmente se tiene:

Definición: Un estadístico $\hat{\theta}$ es un estimador insesgado del parámetro θ , si y sólo si $E(\hat{\theta}) = \theta$.

Un estimador insesgado de mínima varianza, es decir, un estimador **eficiente**, cumple con el siguiente teorema:

Si $\hat{\theta}$ es un estimador insesgado de θ y

$$\text{Var}(\hat{\theta}) = \frac{1}{n \cdot E \left[\left(\frac{\partial \ln f(x)}{\partial \theta} \right)^2 \right]}$$

entonces $\hat{\theta}$ es un estimador insesgado de varianza mínima de θ .

Cuando la eficiencia del estimador aumenta conforme se incrementa el tamaño de la muestra, y la probabilidad de obtener el parámetro poblacional se acerca a uno, se dice que el estimador es **consistente**, es decir:

Definición: Un estadístico $\hat{\theta}$ es un estimador consistente del parámetro θ , si y sólo si para cada constante positiva c ,

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| \geq c) = 0.$$

o, en forma equivalente, si y sólo si

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| < c) = 1$$

En consecuencia un estadístico $\hat{\theta}$ es un estimador consistente del parámetro θ si,

- $\hat{\theta}$ es insesgado
- $\text{Var}(\hat{\theta}) \rightarrow 0$ cuando $n \rightarrow \infty$.

Regresión lineal múltiple

Es de suma importancia el hecho de que pueden obtenerse propiedades interesantes de los estimadores de mínimos cuadrados, aún sin conocer la distribución de las variables que intervienen en alguna regresión.

El modelo clásico para representar los datos observados está dado por la ecuación:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad i = 1, 2, 3, \dots, n \quad (3.1)$$

Los parámetros, que son estimados mediante el método de mínimos cuadrados, tienen propiedades de linealidad, son insesgados y de mínima varianza.

En la ecuación **(3.1)**, Y_i es la variable aleatoria cuyos valores deseamos predecir en términos de valores dados de $X_{i1}, X_{i2}, \dots, X_{ik}$ y $\beta_0, \beta_1, \beta_2, \dots, \beta_k$, los coeficientes de regresión múltiple son constantes numéricas que deben determinarse a partir de datos observados.

Para n puntos de datos: $\{(x_{i1}, x_{i2}, \dots, x_{ik}, y_i); i = 1, 2, \dots, n\}$

las estimaciones de mínimos cuadrados de las β 's son los valores $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ para los cuales la cantidad

$$q = \sum_{i=1}^n \left[y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik}) \right]^2$$

es un mínimo. En esta notación, x_{i1} es el i -ésimo valor de la variable x_1 , x_{i2} es el i -ésimo valor de la variable x_2 , etcétera.

De este modo al diferenciar parcialmente con respecto a las $\hat{\beta}$'s y al igualar estas derivadas parciales a cero, se obtiene

$$\frac{\partial q}{\partial \hat{\beta}_0} = \sum_{i=1}^n (-2) \left[y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik}) \right] = 0$$

$$\frac{\partial q}{\partial \hat{\beta}_1} = \sum_{i=1}^n (-2) x_{i1} \left[y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik}) \right] = 0$$

$$\frac{\partial q}{\partial \hat{\beta}_2} = \sum_{i=1}^n (-2) x_{i2} \left[y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik}) \right] = 0$$

$$\begin{matrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{matrix}$$

$$\frac{\partial q}{\partial \hat{\beta}_k} = \sum_{i=1}^n (-2) x_{ik} \left[y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik}) \right] = 0$$

y por último las $k+1$ ecuaciones normales

$$\sum y = \hat{\beta}_0 \cdot n + \hat{\beta}_1 \cdot \sum x_1 + \hat{\beta}_2 \cdot \sum x_2 + \dots + \hat{\beta}_k \cdot \sum x_k$$

$$\sum x_1 y = \hat{\beta}_0 \cdot \sum x_1 + \hat{\beta}_1 \cdot \sum x_1^2 + \hat{\beta}_2 \cdot \sum x_1 x_2 + \dots + \hat{\beta}_k \cdot \sum x_1 x_k$$

$$\sum x_2 y = \hat{\beta}_0 \cdot \sum x_2 + \hat{\beta}_1 \cdot \sum x_2 x_1 + \hat{\beta}_2 \cdot \sum x_2^2 + \dots + \hat{\beta}_k \cdot \sum x_2 x_k$$

$$\begin{matrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{matrix}$$

$$\sum x_k y = \hat{\beta}_0 \cdot \sum x_k + \hat{\beta}_1 \cdot \sum x_k x_1 + \hat{\beta}_2 \cdot \sum x_k x_2 + \dots + \hat{\beta}_k \cdot \sum x_k^2$$

Aquí, abreviamos nuestra notación escribiendo $\sum_{i=1}^n x_{i1}$ como $\sum x_1$, $\sum_{i=1}^n x_{i1} x_{i2}$ como $\sum x_1 x_2$, etcétera.

El modelo que utilizamos en la regresión lineal múltiple se presta singularmente a un tratamiento unificado en notación matricial. Esta notación hace posible expresar resultados generales en forma compacta y aprovechar al máximo los resultados de la teoría de matrices.

Las ecuaciones normales anteriores se expresan en notación matricial mediante las siguientes tres matrices:

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, y B = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{pmatrix}$$

La primera, X, es una matriz n x (k+1) que consiste esencialmente en los valores dados de las x, con la columna de unos (1), anexada para dar cabida a los términos constantes. Y es una matriz de n x 1 (o vector columna) consistente en los valores observados de y y B es una matriz (k+1) x 1 (o vector columna) consistente en las estimaciones de mínimos cuadrados de los coeficientes de regresión.

Mediante el uso de estas matrices y la aplicación del siguiente teorema podemos escribir la solución simbólica a las ecuaciones normales obtenidas:

Teorema: Las estimaciones de mínimos cuadrados de los coeficientes de regresión múltiple están dados por:

$$B = (X'X)^{-1}X'Y$$

donde X' es la transpuesta de X y $(X'X)^{-1}$ es la inversa de $X'X$

Demostración:

Primero determinamos $X'X$, $X'XB$ y $X'Y$, con lo que se obtiene

$$X'X = \begin{pmatrix} n & \sum x_1 & \sum x_2 & \cdots & \sum x_k \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 & \cdots & \sum x_1 x_k \\ \sum x_2 & \sum x_2 x_1 & \sum x_2^2 & \cdots & \sum x_2 x_k \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \sum x_k & \sum x_k x_1 & \sum x_k x_2 & \cdots & \sum x_k^2 \end{pmatrix},$$

$$X'XB = \begin{pmatrix} \hat{\beta}_0 \cdot n + \hat{\beta}_1 \cdot \sum x_1 + \hat{\beta}_2 \cdot \sum x_2 + \cdots + \hat{\beta}_k \cdot \sum x_k \\ \hat{\beta}_0 \cdot \sum x_1 + \hat{\beta}_1 \cdot \sum x_1^2 + \hat{\beta}_2 \cdot \sum x_1 x_2 + \cdots + \hat{\beta}_k \cdot \sum x_1 x_k \\ \hat{\beta}_0 \cdot \sum x_2 + \hat{\beta}_1 \cdot \sum x_2 x_1 + \hat{\beta}_2 \cdot \sum x_2^2 + \cdots + \hat{\beta}_k \cdot \sum x_2 x_k \\ \vdots \\ \hat{\beta}_0 \cdot \sum x_k + \hat{\beta}_1 \cdot \sum x_k x_1 + \hat{\beta}_2 \cdot \sum x_k x_2 + \cdots + \hat{\beta}_k \cdot \sum x_k^2 \end{pmatrix} \text{ y } X'Y = \begin{pmatrix} \sum y \\ \sum x_1 y \\ \sum x_2 y \\ \vdots \\ \sum x_k y \end{pmatrix}$$

Identificando los elementos de $X'XB$ como las expresiones del lado derecho de las ecuaciones normales y las de $X'Y$ como las expresiones del lado izquierdo, podemos escribir

$$X'XB = X'Y,$$

Al multiplicar del lado izquierdo por $(X'X)^{-1}$, se obtiene,

$$(X'X)^{-1}X'XB = (X'X)^{-1}X'Y$$

y finalmente $B = (X'X)^{-1}X'Y$, ya que $(X'X)^{-1}X'X$ es igual a la matriz de identidad $(k+1) \times (k+1) = I$ y por la definición $IB=B$. Se ha supuesto aquí que $X'X$ es no singular, de manera que su inversa existe.

Modelo econométrico para el pronóstico de la Tasa de Interés Interbancaria Promedio (TIIP)

El objeto del presente capítulo es el estudio de un modelo que pretenda explicar el comportamiento de una variable financiera, que se denotará como Y_i utilizando la información proporcionada por los valores tomados de un conjunto de variables explicativas $X_{i1}, X_{i2}, \dots, X_{ik}$. Especificando dicha relación por la ecuación **(3.1)** que representa un modelo econométrico, es decir, un conjunto de relaciones matemáticas que expresan una teoría económica.

La variable Y recibe el calificativo de *endógena*, e Y_i denota su valor en el instante $i=1,2,\dots,i$ frente a las variables $X_{i1}, X_{i2}, \dots, X_{ik}$ que se llaman *exógenas*. Los coeficientes $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ denotan la magnitud del efecto que las variables $X_{i1}, X_{i2}, \dots, X_{ik}$ tienen sobre Y_i . El término ε_i es una variable aleatoria de error en la estimación.

El problema fundamental que vamos a abordar es el siguiente: Suponiendo que la relación entre la variable Y y el conjunto de variables $X_{i1}, X_{i2}, \dots, X_{ik}$ es como se ha descrito en **(3.1)**, y que se dispone de un conjunto de i observaciones de las variables endógenas y exógenas, ¿cómo podemos asignar valores numéricos a los parámetros $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ basándonos en la información muestral?. La obtención de dichos valores paramétricos es fundamental para poder hacer un análisis estadístico de la relación **(3.1)**, ya sea mediante la obtención de predicciones acerca del comportamiento futuro de Y_i , o la discusión de determinadas cuestiones de política económica.

A continuación se desarrollará el modelo para pronosticar la Tasa de Interés Interbancaria Promedio (TIIP), cuyo valor es comúnmente una referencia obligada en las operaciones con títulos de crédito que manejan las instituciones bancarias, así como la tasa indicadora para calcular los intereses devengados por la emisión de obligaciones o instrumentos de deuda que cotizan en el Mercado de Valores.

Siendo la TIIP una variable ligada fuertemente al comportamiento económico del país, que en la actualidad se mantiene inestable, su valor se ve afectado sensiblemente, al grado de que antes de la devaluación de 1994 se daba a conocer cada 28 días y después de ésta, cada semana, con el fin de reflejar con oportunidad los efectos que las demás variables económicas le ocasionaban.

En principio la TIIP es una tasa activa que regula el costo del crédito que obtienen las instituciones financieras en el mercado (fondeo en mesas de dinero y del banco central), por lo que es difícil saber si su comportamiento esta en función de las tasas pasivas, además de las variables macroeconómicas, o si ésta es la que influye en el comportamiento de las demás.

Como un primer acercamiento al modelo se exploraran las posibles variables que se supone afectarían el comportamiento de la TIIP mensual, considerándose los resultados de la subasta primaria de Certificados de la Tesorería de la Federación (CETES), a plazo de 28 días; el Índice Nacional de Precios al Consumidor (INPC), mensual; el tipo de cambio promedio del dólar a la venta mensual; el Costo Porcentual Promedio (CPP), mensual; la base monetaria mensual; el monto del crédito otorgado por Banxico a instituciones de crédito e intermediarios públicos y privados mensual; el nivel de las reservas internacionales mensual; el Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (IPC), promedio mensual; las tasas de interés que otorgan los pagarés con rendimiento liquidable al vencimiento (PRLV), a plazo de 28 días; y los depósitos a plazo fijo de 60 días.

La confiabilidad de los datos para un mejor análisis es indispensable, por lo que se tomaron las cifras de las fuentes oficiales que proporciona el Banco de México en su “Resumen de Indicadores Agregados de la Economía” y las que publica la Bolsa Mexicana de Valores en su informe sobre “Indicadores Financieros y Bursátiles”. Todas las cifras son mensuales por el período 1994 - 1995, como se muestra en la **figura 3.1**.

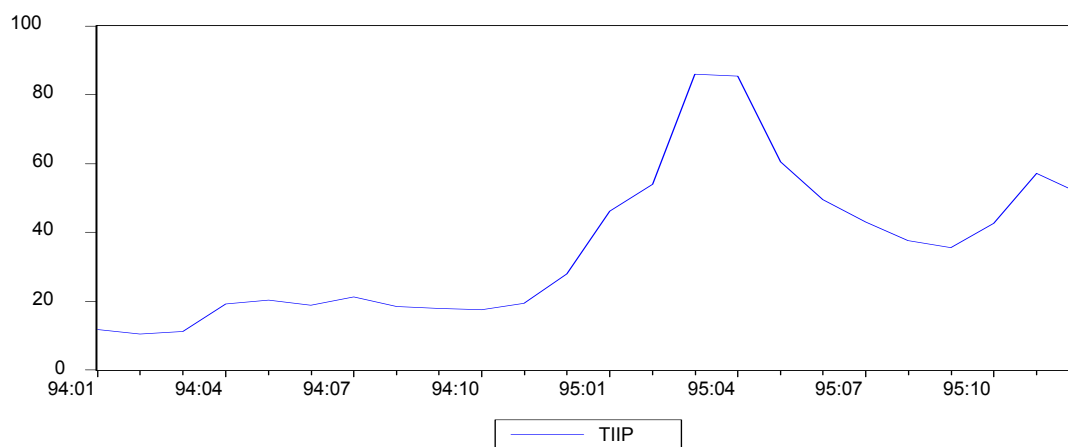


Figura 3.1.: Evolución de la TIIP por el período 1994 1995.

◆ Bases de datos utilizadas

BASE DE DATOS PARA ELABORAR LA REGRESIÓN Y PRONÓSTICO DE LA TIIP
Datos mensuales para el período 1994 - 1995

AÑO	MES	CETES 28 DÍAS ⁽¹⁾ MERCADO PRIMARIO	IP.C. DE LA B.M.V. ⁽²⁾	P.R.L.V. 28 DÍAS ⁽¹⁾	DEPÓSITOS A PLAZO FIJO 60 DÍAS ⁽¹⁾	I.N.P.C. MENSUAL ⁽⁷⁾	TIPO DE CAMBIO DÓLAR PROMEDIO VENTA ⁽⁴⁾	C.P.P. MENSUAL ⁽¹⁾	NIVEL DE RESERVAS INTERNACIONALES ⁽⁶⁾	BASE MONETARIA ⁽³⁾
1994	ENERO	10.53	2,620.24	11.10	10.98	97.20	3.30	13.95	81,613	42,513
	FEBRERO	9.41	2,726.76	10.23	10.25	97.70	3.31	12.59	93,516	41,896
	MARZO	9.63	2,481.96	9.92	10.06	98.21	3.32	11.70	82,816	44,858
	ABRIL	15.74	2,145.06	15.19	15.12	98.69	3.33	12.85	56,514	41,895
	MAYO	16.30	2,316.50	17.66	17.99	99.16	3.34	15.60	56,818	42,586
	JUNIO	16.19	2,339.06	17.30	17.44	99.66	3.35	17.12	54,264	43,382
	JULIO	16.67	2,310.16	17.95	18.35	100.10	3.36	17.50	54,983	45,027
	AGOSTO	14.53	2,665.18	16.70	16.50	100.57	3.37	17.49	55,490	44,235
	SEPTIEMBRE	14.46	2,759.35	15.21	15.26	101.28	3.38	16.99	54,940	43,802
	OCTUBRE	13.86	2,667.56	15.14	15.28	101.81	3.39	16.35	59,141	44,757
	NOVIEMBRE	14.33	2,513.98	15.67	15.94	102.36	3.41	16.11	43,022	49,163
	DICIEMBRE	24.68	2,383.94	16.76	16.64	103.26	4.07	16.65	32,739	56,935
1995	ENERO	37.25	2,127.66	27.77	26.14	107.14	5.64	20.19	19,837	51,200
	FEBRERO	41.69	1,816.37	30.84	29.90	111.68	5.67	31.40	52,406	49,790
	MARZO	69.54	1,629.41	49.58	47.24	118.27	6.76	44.32	46,697	48,805
	ABRIL	74.75	1,907.62	58.36	57.61	127.69	5.18	63.54	50,356	47,553
	MAYO	57.73	2,019.47	50.71	49.17	133.03	5.97	67.16	64,479	46,208
	JUNIO	46.06	2,039.93	42.47	41.14	137.25	6.21	53.27	63,608	47,024
	JULIO	40.94	2,433.24	37.51	37.24	140.05	6.11	45.15	84,824	47,948
	AGOSTO	35.36	2,497.78	33.47	33.30	142.37	6.19	39.69	95,129	47,806
	SEPTIEMBRE	33.46	2,532.85	31.87	31.73	145.32	6.30	35.86	94,358	47,229
	OCTUBRE	40.91	2,307.38	36.17	35.28	148.31	6.73	35.23	96,791	48,480
	NOVIEMBRE	55.85	2,404.20	45.65	43.93	151.96	7.67	41.26	104,014	53,630
	DICIEMBRE	49.74	2,707.93	43.22	41.99	155.00	7.66	47.04	67,103	49,384

(1) Promedio ponderado mensual

(2) Promedio simple mensual

(3) Millones de pesos

(4) Pesos por dólar

(5) Cifras en millones de pesos. Definida según la Ley del Banxico que entró en vigor en abril de 1994. Para su conversión a moneda nacional se utilizó el tipo de cambio para solventar obligaciones en moneda
(6) Crédito interno otorgado a intermediarios financieros: instituciones de crédito y otros intermediarios públicos y privados.

(7) Expresado en puntos, base 1994

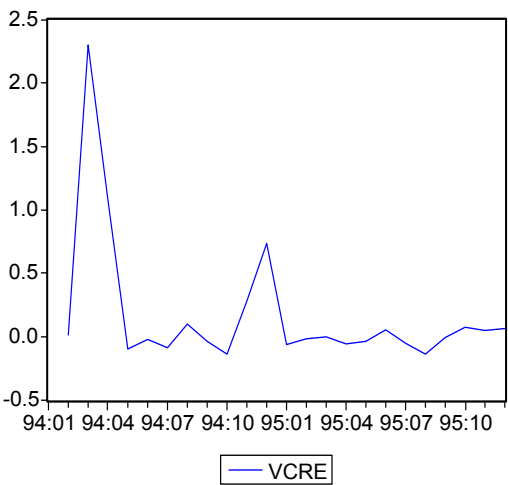
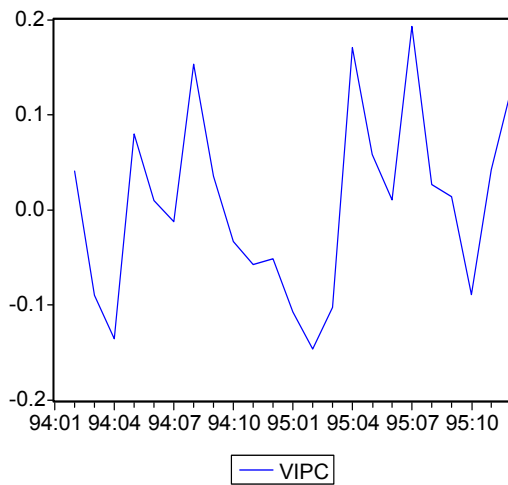
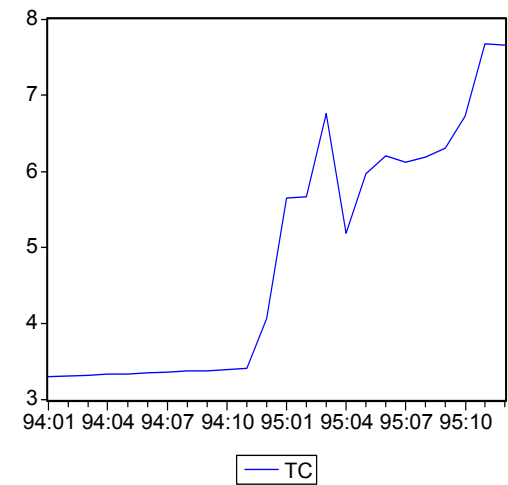
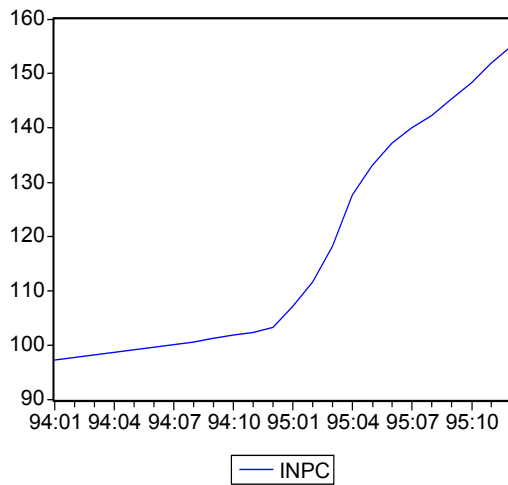
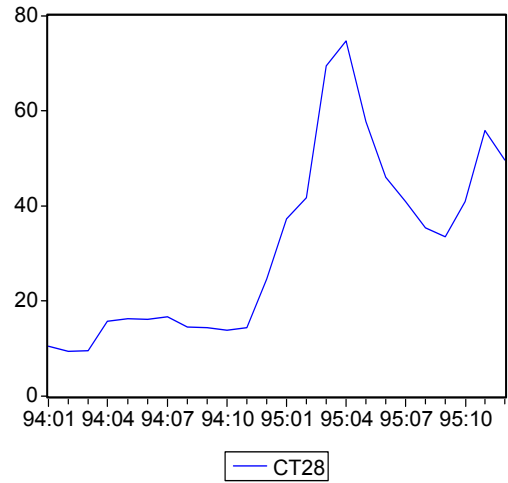
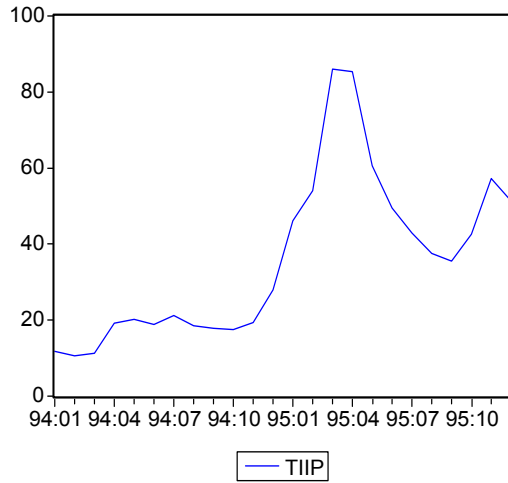


Figura 3.1.: Evolución de la TIIP, CETES, INPC, TC, variación del IPC y del CREDITO por el período 1994 1995.

El modelo a evaluar, de acuerdo a lo expuesto anteriormente es:

TIIP = f(C, CT28, INPC(-1), TC(-1), VIPC, VCRE), que en términos de una ecuación lineal se representa por:

$$\text{TIIP} = \beta_0 + \beta_1\text{CT28} + \beta_2\text{INPC}(-1) + \beta_3\text{TC}(-1) + \beta_4\text{VIPC} + \beta_5\text{VCRE} + \varepsilon_i$$

para $i = 1, 2, 3, \dots n$

Donde cada sigla significa:

C	=	CONSTANTE
CT28	=	RESULTADOS DE LA SUBASTA PRIMARIA DE CERTIFICADOS DE LA TESORERÍA DE LA FEDERACIÓN A 28 DÍAS
INPC(-1)	=	ÍNDICE NACIONAL DE PRECIOS AL CONSUMIDOR DEL MES ANTERIOR
TC(-1)	=	TIPO DE CAMBIO PROMEDIO DEL DÓLAR A LA VENTA DEL MES ANTERIOR
VIPC	=	VARIACIÓN DEL ÍNDICE DE PRECIOS Y COTIZACIONES DE LA BOLSA MEXICANA DE VALORES
VCRE	=	VARIACIÓN EN EL MONTO DEL CRÉDITO OTORGADO POR BANXICO A INSTITUCIONES DE CRÉDITO E INTERMEDIARIOS PÚBLICOS Y PRIVADOS
ε_i	=	ERRORES DE LA REGRESIÓN

Evaluación de los supuestos de un modelo econométrico

Con el objeto de tener cierta uniformidad en la discusión de los métodos apropiados para la solución del problema que acabamos de mencionar, es preciso hacer una serie de supuestos sobre la ecuación de regresión:

SUPUESTO	PRUEBA
Coeficientes significativos	Significancia del estadístico t
No multicolinealidad	Correlaciones
Correcta especificación	Variables omitidas
Normalidad en los errores de la regresión	Normalidad
No autocorrelación	LM correlación serial
No heterocedasticidad	LM ARCH y White
Linealidad	Ramsey's RESET
Permanencia estructural	Chow, CUSUM y CUSUM ²

Los resultados, obtenidos mediante el paquete estadístico Econometric Views for Windows, el cual realiza pruebas de hipótesis sobre los parámetros estimados con un 95% de confianza, se presentan a continuación.

Resultados de la regresión

A partir de los resultados de la regresión mostrados en la **figura 3.2**, se analiza el significado de cada uno de ellos:

Coeficientes de la regresión: Cada coeficiente, multiplicado por su correspondiente variable independiente da la mejor predicción de la variable dependiente. La constante C es el intercepto de la regresión, o sea, la base de predicción cuando los demás coeficientes son cero

Error estándar de los coeficientes: Este valor mide la confiabilidad en los coeficientes de la regresión, un error estándar más grande significa mayor ruido estadístico en los coeficientes. De acuerdo a la teoría de regresión existen dos posibilidades en tres de que los verdaderos coeficientes estén en la vecindad del coeficiente obtenido más (menos) un error estándar y 95 posibilidades en 100 de que se encuentre en la vecindad del coeficiente obtenido más (menos) dos errores estándar.

Estadístico t: Esta es una prueba estadística para la hipótesis de que un coeficiente tiene un valor particular. La hipótesis nula es que alguno de los coeficientes es cero, es decir, que la variable no es significativa para la regresión. El estadístico t es el ratio del coeficiente a su desviación estándar. Si excede el valor de uno, significa que el coeficiente es al menos dos terceras partes más probable que sea diferente de cero y, si el estadístico excede el valor de dos, significa que es diferente de cero con un 95% de confianza.

Probabilidad: La última columna muestra la probabilidad de que el estadístico t sea el que se presenta en la columna previa. Con esta información se puede estar en condiciones de aceptar o rechazar la hipótesis de que el verdadero coeficiente es cero. Normalmente, una probabilidad menor a 0.05 es una fuerte evidencia para rechazar esta hipótesis.

LS // Dependent Variable is TIIP Date: 10/28/96 Time: 13:53 Sample(adjusted): 1994:02 1995:12 Included observations: 23 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	23.303920	2.547510	9.147723	0.000000
CT28	1.099339	0.033074	33.238950	0.000000
INPC(-1)	-0.306154	0.042936	-7.130555	0.000000
TC(-1)	2.881317	0.786445	3.663723	0.001900
VIPC	-13.063410	3.435750	-3.802201	0.001400
VCRE	-1.654870	0.614017	-2.695154	0.015300
R-squared	0.996866	Mean dependent var		36.995880
Adjusted R-squared	0.995945	S.D. dependent var		21.883550
S.E. of regression	1.393576	Akaike info criterion		0.883205
Sum squared resid	33.014920	Schwarz criterion		1.179420
Log likelihood	-36.792440	F-statistic		1,081.592000
Durbin-Watson stat	2.094799	Prob(F-statistic)		0.000000

Figura 3.2.: Resultados de la regresión con la TIIP.

R²: Mide la capacidad de la regresión en la predicción de los valores de la variable dependiente. El R² tiene un valor de uno si la curva de regresión se ajusta perfectamente, y cero si no lo hace más que como un simple promedio de la variable dependiente. El R² es la fracción de la varianza de la variable dependiente explicada por las variables independientes.

R² ajustado por los grados de libertad: Este valor es parecido al R² con una pequeña diferencia ocasionada por la varianza que se utiliza para calcularlo. Es menor que R² debido a que hay más de una variable independiente.

Error estándar de la regresión: Esta es una medida que resume los errores de la predicción. Tiene las mismas unidades que la variable dependiente y mide la magnitud de los residuos. Cerca de dos tercios de los residuos caerán en un rango de más (menos) un error estándar y, el 95% de los residuos caerán dentro del rango de más (menos) dos desviaciones estándar.

Log verosimilitud: Es el valor de la función de Log verosimilitud (asumiendo errores normalmente distribuidos) evaluado sobre los valores estimados de los coeficientes. Las pruebas del radio de verosimilitud pueden ser hechas observando la diferencia entre los valores de verosimilitud de versiones restringidas o no restringidas de una ecuación.

Durbin - Watson: Es el estadístico para la autocorrelación. Si es menor a dos, existe evidencia de autocorrelación positiva.

Criterio de Akaike: Este criterio es una guía para seleccionar el número de términos de una ecuación. Está basado en la suma de los residuos al cuadrado, pero penaliza cuando se tienen coeficientes en exceso. En ciertas situaciones se puede escoger el número de rezagos de una variable a partir del menor valor de Akaike. Los valores más pequeños del criterio de Akaike son mejores.

Criterio de Schwartz: Este criterio es alternativo al de Akaike, con el mismo principio aunque penaliza más el exceso de coeficientes.

Estadístico F: Esta es una prueba para la hipótesis de que todos los coeficientes de una regresión son cero, (excepto la constante). Si el estadístico F excede un nivel crítico, al menos uno de los coeficientes es diferente de cero. Por ejemplo, si hay tres variables independientes y 100 observaciones, un estadístico F superior a 2.7 indica que la probabilidad es de al menos 95% de que uno o los tres coeficientes sean diferentes de cero. La probabilidad que se muestra abajo del estadístico F permite hacer esa prueba con cierto nivel de confianza.

La **figura 3.3** muestra un mapeo de los valores reales de la variable dependiente (Actual), los de la curva de regresión (Fitted), o sea, los valores predichos por la regresión aplicando los coeficientes a las variables independientes, así como los residuos de la regresión (Residual), es decir, las diferencias entre los valores reales y los predichos de la variable dependiente. Esto da una indicación de los errores probables que surgirán cuando se efectué el pronóstico.

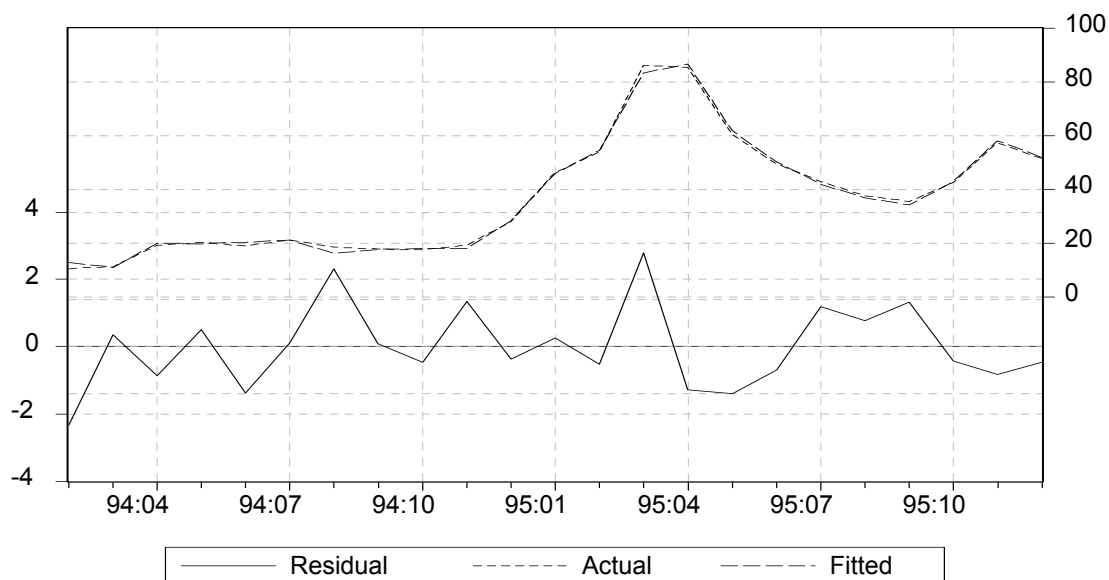


Figura 3.3.: Resultados de la regresión con la TIIP.

Valuación del supuesto de no multicolinealidad

La multicolinealidad es una alta correlación entre las variables explicativas de una regresión. Cuando está presente, es difícil determinar las influencias de cada una de ellas por separado. Los errores estándar elevados, son evidencia de esta dificultad.

En el caso de la TIIP, se observa una alta correlación entre las variables independientes, con valores superiores a 0.8 en la matriz de correlaciones que se muestra en la **figura 3.4**, no obstante esto es factible debido a que se trata de variables económicas y financieras con comportamientos similares.

	CT28	INPC	TC	VIPC	VCRE
CT28	1.000000	0.701021	0.823661	0.192244	-0.333858
INPC	0.701021	1.000000	0.916220	0.325502	-0.305717
TC	0.823661	0.916220	1.000000	0.113516	-0.316310
VIPC	0.192244	0.325502	0.113516	1.000000	-0.374646
VCRE	-0.333858	-0.305717	-0.316310	-0.374646	1.000000

Figura 3.4.: Resultados de la regresión con la TIIP.

La teoría de la regresión asume que las variables independientes son colineales y hace lo mejor posible para separar sus influencias individuales, sin embargo, no es muy recomendable hacer inferencia sobre los resultados de la regresión cuando existe colinealidad.

Valuación del supuesto de correcta especificación

Esta prueba permite agregar u omitir un conjunto de variables a una ecuación existente y verificar si hacen alguna contribución significativa a la explicación de la variable dependiente. El resultado de esta prueba es un estadístico F y un estadístico del radio de verosimilitud (LR), con sus respectivas probabilidades. El estadístico F está basado en la diferencia de la suma de los residuos al cuadrado de la variable que se agrega (o se omite). El estadístico del radio de verosimilitud es aproximado al estadístico F, con la diferencia en que existe un factor por cada variable que se agrega (o se omite) a la regresión.

En el caso de la ecuación de la TIIP, se prueba si la variable Base Monetaria (BM), pudiera ser relevante y no debió ser omitida en el modelo. El resultado es que no es significativa y la probabilidad de la prueba es mayor a 0.05 y por lo tanto se puede rechazar la hipótesis de que es significativa para la regresión, **figura 3.5**.

Omitted Variables: BM				
F-statistic	0.329370	Probability	0.574010	
Log likelihood ratio	0.468662	Probability	0.493603	
Test Equation: LS // Dependent Variable is TIIP Date: 10/28/96 Time: 16:19 Sample: 1994:02 1995:12 Included observations: 23				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	21.158680	4.552857	4.647342	0.000300
CT28	1.093150	0.035427	30.856070	0.000000
INPC(-1)	-0.311944	0.044955	-6.939042	0.000000
TC(-1)	2.939983	0.808917	3.634470	0.002200
VIPC	-12.428760	3.675878	-3.381166	0.003800
VCRE	-1.675191	0.627498	-2.669634	0.016800
BM	0.000058	0.000101	0.573908	0.574000
R-squared	0.996930	Mean dependent var	36.995880	
Adjusted R-squared	0.995778	S.D. dependent var	21.883550	
S.E. of regression	1.421904	Akaike info criterion	0.949784	
Sum squared resid	32.349000	Schwarz criterion	1.295370	
Log likelihood	-36.558110	F-statistic	865.825700	
Durbin-Watson stat	2.163954	Prob(F-statistic)	0.000000	

Figura 3.5.: Valuación de la variable omitida Base Monetaria.

La misma prueba aplicada a la variable Reservas Internacionales (RVAS), arroja un resultado similar y se desecha como una variable que pudiera intervenir en el modelo de la TIIP, **figura 3.6**.

Omitted Variables: RVAS				
F-statistic	0.688787	Probability	0.418789	
Log likelihood ratio	0.969412	Probability	0.324827	
Test Equation: LS// Dependent Variable is TIIP Date: 10/28/96 Time: 16:23 Sample: 1994:02 1995:12 Included observations: 23				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	22.728090	2.663121	8.534384	0.000000
CT28	1.087604	0.036252	30.001000	0.000000
INPC(-1)	-0.290921	0.047061	-6.181824	0.000000
TC(-1)	2.942955	0.797211	3.691563	0.002000
VIPC	-12.673170	3.499372	-3.621555	0.002300
VCRE	-1.533542	0.636725	-2.408484	0.028400
RVAS	-0.000017	0.000021	-0.829932	0.418800
R-squared	0.996996	Mean dependent var	36.995880	
Adjusted R-squared	0.995869	S.D. dependent var	21.883550	
S.E. of regression	1.406510	Akaike info criterion	0.928013	
Sum squared resid	31.652320	Schwarz criterion	1.273598	
Log likelihood	-36.307730	F-statistic	884.941600	
Durbin-Watson stat	2.151700	Prob(F-statistic)	0.000000	

Figura 3.6.: Valuación de la variable omitida Reservas Internacionales.

Al evaluar la variable Costo Porcentual Promedio (CPP), se observa que tampoco es significativa y la probabilidad de aceptar la hipótesis de que no es significativa es mayor a 0.05. Esta variable también se desecha como una variable que pudiera aportar información a la explicación de la TIIP, **figura 3.7.**

Omitted Variables: CPP				
F-statistic	1.254639	Probability	0.279201	
Log likelihood ratio	1.736323	Probability	0.187605	
Test Equation: LS// Dependent Variable is TIIP Date: 10/28/96 Time: 16:25 Sample: 1994:02 1995:12 Included observations: 23				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	22.297960	2.683394	8.309610	0.000000
CT28	1.144098	0.051715	22.122960	0.000000
INPC(-1)	-0.287519	0.045750	-6.284631	0.000000
TQ(-1)	2.706857	0.796007	3.400542	0.003700
VIPC	-10.910570	3.914620	-2.787134	0.013200
VORE	-1.604561	0.611122	-2.625598	0.018400
CPP	-0.059605	0.053213	-1.120107	0.279200
R-squared	0.997094	Mean dependent var		36.995880
Adjusted R-squared	0.996005	S.D. dependent var		21.883550
S.E. of regression	1.383255	Akaike info criterion		0.894669
Sum squared resid	30.614300	Schwarz criterion		1.240254
Log likelihood	-35.924280	F-statistic		915.036900
Durbin-Watson stat	1.970218	Prob(F-statistic)		0.000000

Figura 3.7.: Valuación de la variable omitida Costo Porcentual Promedio.

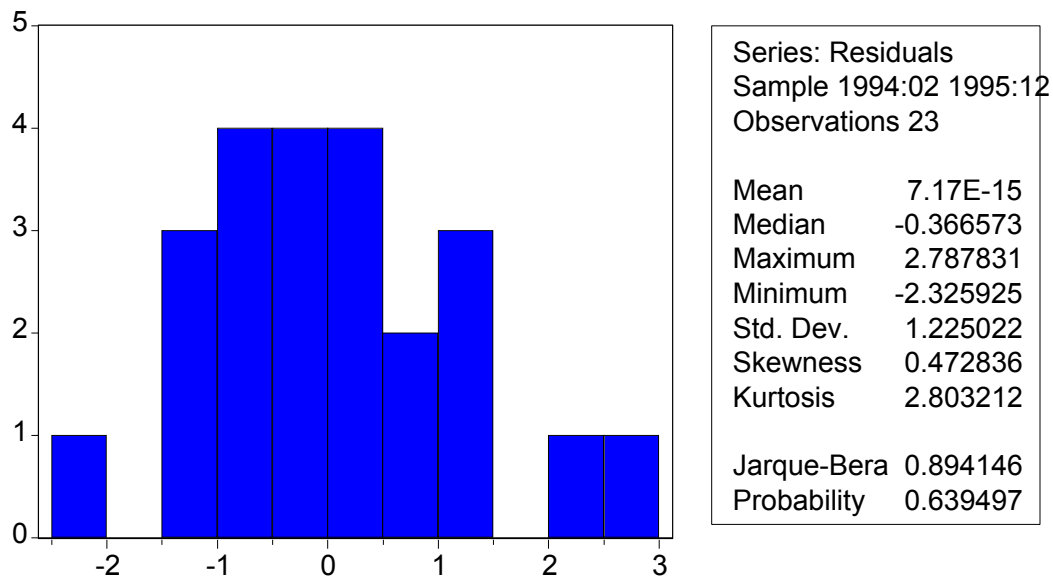
Finalmente, al probar la variable Pagarés con Rendimiento Liquidable al Vencimiento a 28 días (PRLV), se obtiene que no es significativa y la probabilidad de rechazar la hipótesis de que es significativa es mayor a 0.05, por lo que se desecha como una variable que pudiera aportar información al modelo, **figura 3.8**.

Omitted Variables: PRLV				
F-statistic	0.287054	Probability	0.599484	
Log likelihood ratio	0.408982	Probability	0.522487	
Test Equation: LS // Dependent Variable is TIIP Date: 10/28/96 Time: 16:27 Sample: 1994:02 1995:12 Included observations: 23				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	22.632410	2.888730	7.834724	0.000000
CT28	1.165450	0.127937	9.109581	0.000000
INPC(-1)	-0.290389	0.052821	-5.497641	0.000000
TC(-1)	2.763750	0.832900	3.318227	0.004300
VIPC	-11.920510	4.107500	-2.902133	0.010400
VCRE	-1.711251	0.636077	-2.690320	0.016100
PRLV	-0.096335	0.179805	-0.535774	0.599500
R-squared	0.996922	Mean dependent var	36.995880	
Adjusted R-squared	0.995767	S.D. dependent var	21.883550	
S.E. of regression	1.423750	Akaike info criterion	0.952379	
Sum squared resid	32.433050	Schwarz criterion	1.297964	
Log likelihood	-36.587950	F-statistic	863.575000	
Durbin-Watson stat	2.100944	Prob(F-statistic)	0.000000	

Figura 3.8.: Valuación de la variable omitida Pagares con Rendimiento Liquidable al Vencimiento a 28 días.

Valuación del supuesto de normalidad

Esta prueba produce un histograma de los residuos de la regresión y el valor del estadístico de Jarque - Bera para la prueba de normalidad. Bajo la hipótesis nula de normalidad, dicho estadístico se distribuye como una χ^2 con dos grados de libertad.



El estadístico de Jarque - Bera esta dado por $\frac{T-k}{6} \left[S^2 - \frac{1}{4}(K-3)^2 \right]$, donde T es el número de variables independientes de la ecuación de regresión, S es la simetría y K la kurtosis.

Ahora bien, el valor de la simetría de una distribución normal es cero, mientras que el de la kurtosis es tres. Con estos argumentos se puede asumir que los residuos de la regresión de la TIIP se distribuyen normalmente. Dicho de otra manera, se verifica que los errores de la regresión siguen una distribución normal, $\varepsilon_i \approx N[0, \sigma^2]$.

Valuación del supuesto de no autocorrelación

Como se vio al obtener la regresión inicial, el valor del estadístico de Durbin - Watson es superior a dos y se puede presumir que no existe autocorrelación de primer orden, sin embargo, puede existir autocorrelación de segundo, tercero o n-ésimo orden, por lo que la prueba de Breusch - Godfrey prueba la hipótesis nula de que todos los residuos rezagados de la regresión son cero, obteniendo un valor para el estadístico F y un estadístico R^2 que se distribuye como una χ^2 .

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.358731	Probability	0.704395	
Obs*R-squared	1.049892	Probability	0.591587	
Test Equation: LS // Dependent Variable is RESID Date: 10/28/96 Time: 14:44				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.288084	2.844269	-0.101286	0.920700
CT28	0.007126	0.037555	0.189759	0.852000
INPC(-1)	0.007870	0.050055	0.157229	0.877200
TC(-1)	-0.166403	0.922518	-0.180379	0.859300
VIPC	-0.649354	3.722220	-0.174453	0.863800
VCRE	-0.258016	0.707938	-0.364462	0.720600
RESID(-1)	-0.186095	0.275799	-0.674748	0.510100
RESID(-2)	-0.199176	0.315578	-0.631145	0.537400
R-squared	0.045647	Mean dependent var	0.000000	
Adjusted R-squared	-0.399717	S.D. dependent var	1.225022	
S.E. of regression	1.449319	Akaike info criterion	1.010395	
Sum squared resid	31.507870	Schwarz criterion	1.405350	
Log likelihood	-36.255130	F-statistic	0.102495	
Durbin-Watson stat	1.933086	Prob(F-statistic)	0.997318	

En el caso de la TIIP, se observa que la prueba arroja valores no significativos para los residuos de la regresión con dos rezagos y la probabilidad es mayor a 0.05 por lo que se acepta la hipótesis de que los residuos de la regresión son cero y no están autocorrelacionados de orden dos, es decir, $Cov[\varepsilon_i, \varepsilon_j] = 0$, si $i \neq j$.

Valuación del supuesto de no heterocedasticidad

La heterocedasticidad, al igual que la autocorrelación, invalida las inferencias que se puedan hacer sobre la variable dependiente. Las pruebas para detectarla son: la prueba de ARCH y las de White para términos cruzados y no cruzados.

La prueba de ARCH LM (heterocedasticidad condicional autoregresiva) fue motivada por la observación de que en las variables económicas, la magnitud de los residuos antiguos parecen estar más relacionados con la magnitud de los residuos más recientes, por lo que la prueba está basada en los residuos de la regresión elevados al cuadrado, **figura 3.9**.

ARCH Test:				
F-statistic	0.878182	Probability	0.359879	
Obs*R-squared	0.925368	Probability	0.336069	
Test Equation: LS // Dependent Variable is RESID^2 Date: 10/28/96 Time: 14:52 Sample(adjusted): 1994:03 1995:12 Included observations: 22 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.533831	0.501912	3.055976	0.006200
RESID^2(-1)	-0.187234	0.199798	-0.937114	0.359900
R-squared	0.042062	Mean dependent var	1.254772	
Adjusted R-squared	-0.005835	S.D. dependent var	1.889561	
S.E. of regression	1.895065	Akaike info criterion	1.365014	
Sum squared resid	71.825440	Schwarz criterion	1.464200	
Log likelihood	-44.231810	F-statistic	0.878182	
Durbin-Watson stat	1.956554	Prob(F-statistic)	0.359879	

Figura 3.9.: Prueba de heterocedasticidad del tipo ARCH.

Para el caso de la TIIP se utilizó un rezago al cuadrado encontrándose que no es significativo, además de que la probabilidad asociada al estadístico F y al estadístico R^2 no es significativa, por lo que se asume que no existe heterocedasticidad del tipo ARCH.

Para tener la certeza de que no existe heterocedasticidad de algún otro tipo es conveniente efectuar la prueba de heterocedasticidad de White, la cual está basada en la ecuación de regresión aumentada, es decir, si el modelo original está dado por:

$$Y_i = \beta_0 + \beta_1 X_{i1}$$

entonces la prueba se verifica sobre la ecuación

$$u_t = \beta_0 - \beta_1 X_{i1} - \beta_2 Z_{i2} - \beta_3 X_{i3}^2 - \beta_4 Z_{i4}^2 - \beta_5 X_{i5} Z_{i5} \quad (3.2)$$

El resultado de esta prueba es un estadístico F y un estadístico R^2 que se distribuye como una χ^2 con grados de libertad igual al número de variables independientes en la ecuación de regresión. La prueba es sobre la hipótesis nula de que todos los coeficientes de las variables aumentadas en la ecuación (3.2) son cero. De acuerdo a White la hipótesis nula involucra además que los errores de la regresión son homocedásticos e independientes de las variables y que la especificación lineal del modelo es correcta, **figura 3.10**.

White Heteroskedasticity Test:				
F-statistic	1.586099	Probability	0.222096	
Obs*R-squared	13.09368	Probability	0.218481	
Test Equation: LS // Dependent Variable is RESID^2 Date: 10/28/96 Time: 14:55 Sample: 1994:02 1995:12 Included observations: 23				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	99.562510	59.725550	1.667000	0.121400
CT28	-0.377504	0.180462	-2.091876	0.058400
CT28^2	0.004779	0.001926	2.481277	0.028900
INPC(-1)	-2.295593	1.241718	-1.848723	0.089300
INPC(-1)^2	0.009133	0.004857	1.880429	0.084500
TC(-1)	19.365490	8.444380	2.293300	0.040700
TC(-1)^2	-1.806995	0.744947	-2.425670	0.032000
VIPC	9.992750	9.101248	1.097954	0.293800
VIPC^2	-5.421444	64.837320	-0.083616	0.934700
VCRE	3.957707	3.174162	1.246851	0.236200
VCRE^2	-2.243097	1.372922	-1.633812	0.128200
R-squared	0.569290	Mean dependent var	1.435431	
Adjusted R-squared	0.210365	S.D. dependent var	2.039317	
S.E. of regression	1.812165	Akaike info criterion	1.494979	
Sum squared resid	39.407300	Schwarz criterion	2.038041	
Log likelihood	-38.827840	F-statistic	1.586099	
Durbin-Watson stat	2.716247	Prob(F-statistic)	0.222096	

Figura 3.10.: Prueba de heterocedasticidad del tipo White para términos no cruzados.

Como se observa en el cuadro anterior donde se evalúa la prueba con términos no cruzados, el valor de la probabilidad es mayor a 0.05 por lo que se acepta la hipótesis nula mencionada y se asume que no hay heterocedasticidad en los errores de la regresión de la TIIP con términos no cruzados.

Para evaluar términos cruzados, (se incluyen todas las combinaciones posibles de las variables independientes elevadas al cuadrado), el resultado es similar, por lo que se concluye que no hay heterocedasticidad de White, **figura 3.11.**

White Heteroskedasticity Test:				
F-statistic	5.249102	Probability	0.171979	
Obs*R-squared	22.57002	Probability	0.310397	
Test Equation:				
LS // Dependent Variable is RESID^2				
Date: 10/28/96 Time: 14:58				
Sample: 1994:02 1995:12				
Included observations: 23				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-13.884190	105.502200	-0.131601	0.907300
CT28	0.680829	1.369253	0.497226	0.668300
CT28^2	0.017334	0.007278	2.381770	0.140200
CT28*INPC(-1)	-0.016749	0.022431	-0.746658	0.533100
CT28*TC(-1)	-0.018752	0.188895	-0.099274	0.930000
CT28*VIPC	3.646805	4.446252	0.820198	0.498300
CT28*VCRE	0.459569	0.469708	0.978414	0.431000
INPC(-1)	2.871579	5.106835	0.562301	0.630500
INPC(-1)^2	-0.045203	0.054590	-0.828053	0.494700
INPC(-1)*TC(-1)	1.592824	1.674477	0.951237	0.441900
INPC(-1)*VIPC	2.346591	3.352883	0.699873	0.556500
INPC(-1)*VCRE	0.352371	2.767417	0.127328	0.910300
TC(-1)	-62.668110	88.436670	-0.708621	0.552000
TC(-1)^2	-11.759300	11.086680	-1.060669	0.400000
TC(-1)*VIPC	-70.521780	91.394300	-0.771621	0.521000
TC(-1)*VCRE	-11.783240	41.822310	-0.281745	0.804600
VIPC	-59.200580	110.730500	-0.534637	0.646400
VIPC^2	141.825400	123.502500	1.148361	0.369600
VIPC*VCRE	85.760700	36.758020	2.333115	0.144800
VCRE	5.907042	143.847300	0.041065	0.971000
VCRE^2	-0.928539	3.354135	-0.276834	0.807900
R-squared	0.981305	Mean dependent var	1.435431	
Adjusted R-squared	0.794358	S.D. dependent var	2.039317	
S.E. of regression	0.924785	Akaike info criterion	-0.772649	
Sum squared resid	1.710453	Schwarz criterion	0.264107	
Log likelihood	-2.750122	F-statistic	5.249102	
Durbin-Watson stat	1.937800	Prob(F-statistic)	0.171979	

Figura 3.11.: Prueba de heterocedasticidad del tipo White para términos cruzados.

Valuación del supuesto de linealidad

La prueba de linealidad de Ramsey está basada en la ecuación de regresión aumentada

$$Y_i = X\beta_0 - Z\alpha - \varepsilon$$

donde Z se define como la matriz $Z = [\hat{Y}^2 \quad \hat{Y}^3 \quad \hat{Y}^4 \quad \dots \quad \hat{Y}^n]$ y \hat{Y} es el vector de los valores estimados de Y en X elevados a potencias enteras positivas. La primera potencia no se incluye debido a que es colineal con la matriz X.

El resultado de la prueba es un estadístico F y un estadístico de ratio de verosimilitud para probar la hipótesis nula de que todos los coeficientes de los vectores \hat{Y} son cero.

Ramsey RESET Test:				
F-statistic	0.233704	Probability	0.635340	
Log likelihood ratio	0.333519	Probability	0.563594	
Test Equation: LS // Dependent Variable is TIIP Date: 10/28/96 Time: 15:02 Sample: 1994:02 1995:12 Included observations: 23				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	21.847970	3.983271	5.484933	0.000000
CT28	1.047276	0.112889	9.277078	0.000000
INPC(-1)	-0.278785	0.071663	-3.890228	0.001300
TC(-1)	2.697678	0.889939	3.031306	0.007900
VIPC	-13.420530	3.592672	-3.735529	0.001800
VCRE	-1.713485	0.639933	-2.677599	0.016500
Fitted^2	0.000476	0.000984	0.483429	0.635300
R-squared	0.996911	Mean dependent var	36.995880	
Adjusted R-squared	0.995753	S.D. dependent var	21.883550	
S.E. of regression	1.426088	Akaike info criterion	0.955660	
Sum squared resid	32.539630	Schwarz criterion	1.301245	
Log likelihood	-36.625680	F-statistic	860.737600	
Durbin-Watson stat	2.034975	Prob(F-statistic)	0.000000	

Figura 3.12.: Prueba de Ramsey's RESET para un término \hat{Y}^2 .

Evaluando para un término de \hat{Y} elevado al cuadrado se obtiene que no es significativo y que la probabilidad asociada a los estadísticos F y LR, tampoco lo es, por lo que se acepta la hipótesis nula de linealidad para un término cuadrático de \hat{Y} , **figura 3.12**.

Al hacer una segunda prueba con un término cúbico de \hat{Y} se observa que tampoco es significativo y se concluye que la ecuación de la TIIP es lineal, **figura 3.13**.

Ramsey RESET Test:				
F-statistic	0.686438	Probability	0.518498	
Log likelihood ratio	2.014245	Probability	0.365269	
Test Equation: LS // Dependent Variable is TIIP Date: 10/28/96 Time: 15:15 Sample: 1994:02 1995:12 Included observations: 23				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	25.117570	5.013219	5.010269	0.000200
CT28	1.348973	0.304430	4.431148	0.000500
INPC(-1)	-0.361810	0.105610	-3.425898	0.003800
TC(-1)	3.449434	1.132355	3.046247	0.008200
VIPC	-16.736320	4.739815	-3.531006	0.003000
VCRE	-1.916033	0.664914	-2.881625	0.011400
Fitted^2	-0.005933	0.006089	-0.974368	0.345300
Fitted^3	0.000043	0.000040	1.066381	0.303100
R-squared	0.997129	Mean dependent var	36.995880	
Adjusted R-squared	0.995789	S.D. dependent var	21.883550	
S.E. of regression	1.420014	Akaike info criterion	0.969542	
Sum squared resid	30.246600	Schwarz criterion	1.364496	
Log likelihood	-35.785320	F-statistic	744.262500	
Durbin-Watson stat	2.380888	Prob(F-statistic)	0.000000	

Figura 3.13.: Prueba de Ramsey's RESET para un término \hat{Y}^3 .

Valuación del supuesto de permanencia estructural

La prueba completa de permanencia estructural consta de tres pruebas: la prueba de Chow, la de CUSUM y la de CUSUM².

La prueba de un punto de quiebra de Chow divide el conjunto de datos de cada serie en dos subconjuntos. Cada subconjunto debe contener más observaciones que coeficientes en la ecuación de regresión estimada. El propósito de esta división es probar si el vector de coeficientes puede ser considerado como constante en los subconjuntos. De esta forma se puede probar si la función de la TIIP fue la misma antes y después de un shock económico.

En el caso de la TIIP se probó que existió un impacto entre los meses de abril y mayo de 1995, **figura 3.14**.

Chow Breakpoint Test: 1995:04			
F-statistic	5.133358	Probability	0.009554
Log likelihood	30.70511	Probability	0.000029
Chow Breakpoint Test: 1995:05			
F-statistic	0.995925	Probability	0.473714
Log likelihood	9.979208	Probability	0.125530

Figura 3.14.: Prueba Chow para un punto de shock en los meses de abril - mayo de 1995.

El estadístico F está basado en la suma de los residuos al cuadrado y el estadístico de Log verosimilitud está basado en las restricciones de la función de verosimilitud.

Como se observa, existió un fuerte impacto en las tasas de interés, en particular de la TIIP, la cual repuntó a raíz del fuerte incremento de los CETES, mismas que llegaron a su punto máximo en esos meses; así como a la sobrevaluación de dólar frente al peso.

La prueba CUSUM se realiza sobre la suma acumulada de los residuos de la regresión. En el mapeo producido por esta prueba la suma acumulada se grafica contra el tiempo y se muestran dos líneas críticas. La prueba encuentra el parámetro de inestabilidad si la suma acumulada sale más allá de las líneas críticas.

La prueba CUSUM está basada en el estadístico $W_t = \frac{1}{s} \sum_{\tau=k+1}^t W_\tau$, $t = k+1, \dots, T$; donde s es el error estándar de la regresión sobre la muestra T , y W_t es la suma acumulada en el tiempo t . Si el vector β permanece constante período a período entonces $E(W_t) = 0$, pero si β cambia W_t tenderá a diverger de la media cero.

La significancia de cualquier desviación de la línea igual a cero se ve reflejada por dos líneas rectas que se abren conforme se incrementa t . La prueba CUSUM produce un mapeo de W_t contra t y dos líneas rectas con un valor crítico de 5%. El movimiento de W_t fuera de las líneas críticas sugiere un comportamiento de inestabilidad.

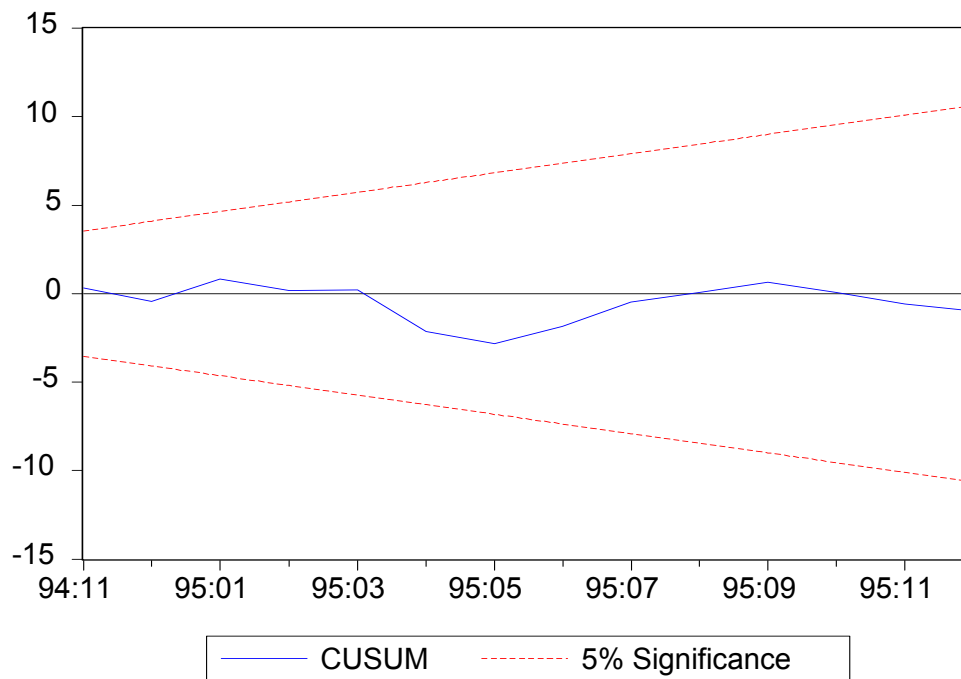


Figura 3.15.: Resultados de la prueba CUSUM.

Como se observa, en el caso de la regresión de la TIIP existe un comportamiento estable de los residuos acumulados de la regresión, **figura 3.15**.

Finalmente, la prueba de CUSUM² está basada en el siguiente estadístico

$$S_t = \frac{\sum_{\tau=k+1}^t W_\tau^2}{\sum_{\tau=k+1}^T W_\tau^2}$$

El valor medio de la línea dado el valor esperado de esta prueba estadística, bajo la hipótesis de constancia en los parámetros es $E(W_t) = \frac{t-k}{T-k}$, con valores que van de cero en $t=K$ a uno en $t=T$. La significancia de la desviación de s_t del valor esperado es mostrado por dos líneas rectas paralelas que arroja la prueba.

La prueba de CUSUM², mapea los valores de s_t contra t en las líneas críticas del 5%. El movimiento fuera de estas líneas sugiere inestabilidad.

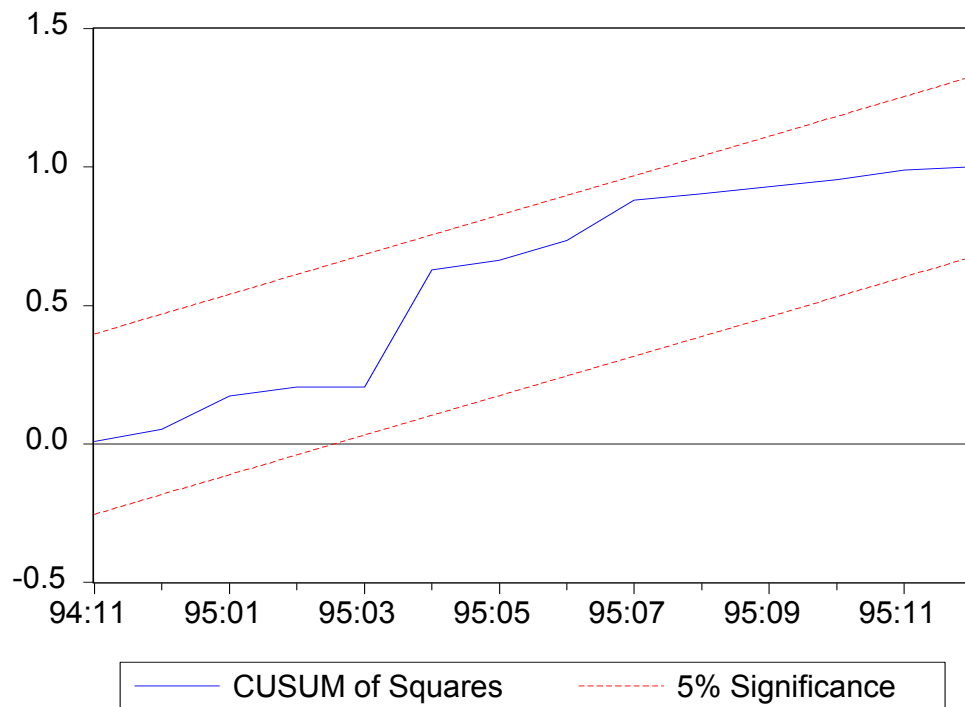


Figura 3.16.: Resultados de la prueba CUSUM².

En el caso de la regresión de la TIIP existe un comportamiento estable de los residuos acumulados al cuadrado de la regresión, **figura 3.16**, por lo que se concluye que el modelo de regresión de la TIIP cumple con el supuesto de permanencia estructural.

Pronóstico con modelos uniecuacionales

Pronóstico de la TIIP para el año de 1996

Para realizar el pronóstico de la TIIP para 1996 se consideraron las principales expectativas de las variables económicas; tal es el caso de la inflación que al primer trimestre de 1996 se incrementó un 8.5%, con un crecimiento mensual promedio del 2%, siendo la inflación de abril la segunda más alta del año, derivada del aumento del 12% a los salarios mínimos que entró en vigor en abril del presente año. Asimismo, se previó una tendencia de estabilidad del entorno económico.

Con estas suposiciones se procedió a efectuar el pronóstico de la TIIP, para el año de 1996, para ello se amplía la muestra original al período 1994 - 1996 asignando los valores supuestos en 1996 a las variables dependientes y el pronóstico a la nueva variable TIIPF.

Los estadísticos asociados al pronóstico son: la raíz del error cuadrático medio, el error medio absoluto, el porcentaje de error absoluto y el coeficiente de Theil y su descomposición, mismos que especifican la capacidad del modelo para pronosticar.

Actual: TIIP Forecast: TIIPF	
Sample: 1994:02 1996:12	
Include observations: 23	
Root Mean Squared Error	1.198095
Mean Absolute Error	0.959484
Mean Absolute Percentage Error	3.741012
Theil Inequality Coefficient	0.014019
Bias Proportion	0.000000
Variance Proportion	0.000785
Covariance Proportion	0.999215

Los valores numéricos para el pronóstico y el mapeo junto con la serie origina se muestran en la siguiente tabla y en la figura **figura 3.17**.

Last updated: 10/30/96 - 13:23 Modified: 1994:02 1996:12 // modtiip.forecast tiipf						
1994:01:00	11.749180	12.847700	10.878840	20.039710	19.721820	20.273970
1994:07:00	21.081290	16.140150	17.724750	17.933270	17.943720	28.386000
1995:01:00	45.860040	54.540560	83.244170	86.613800	61.918270	50.188680
1995:07:00	41.735590	36.805270	34.172880	42.990340	58.050620	51.813840
1996:01:00	46.511550	38.381150	38.799600	26.220000	22.332000	19.961230
1996:07:00	24.245120	16.151300	13.983110	11.643580	9.798152	11.473410

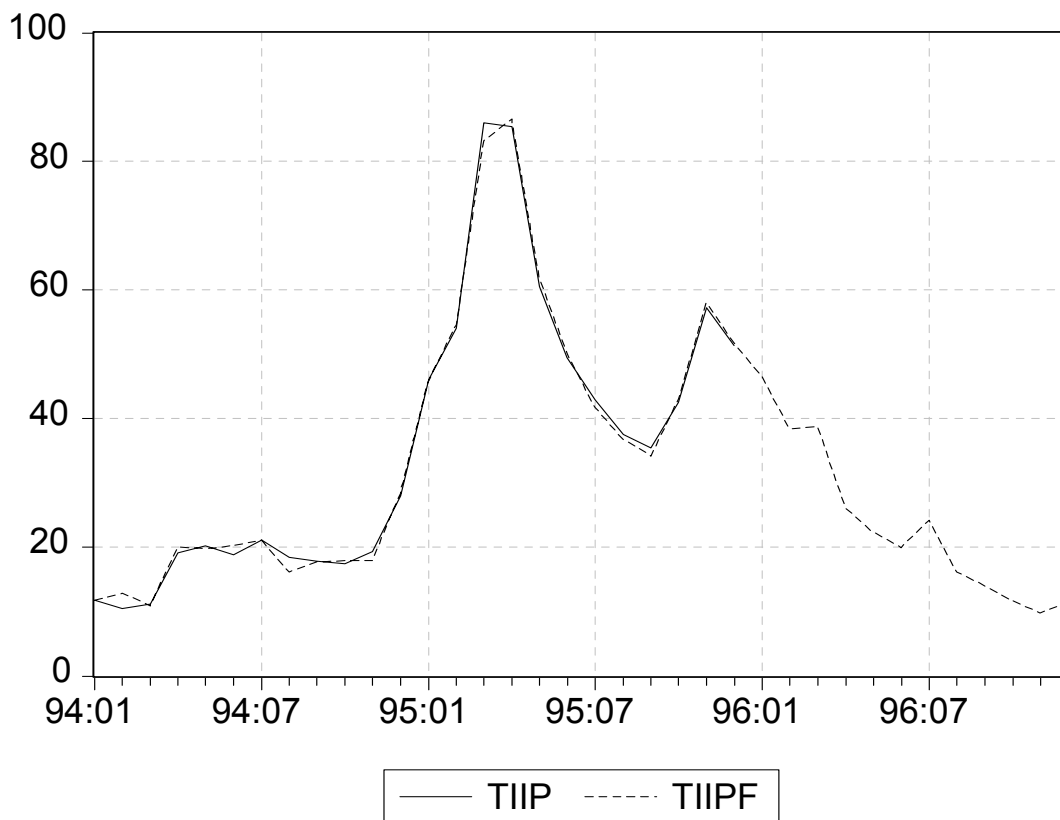
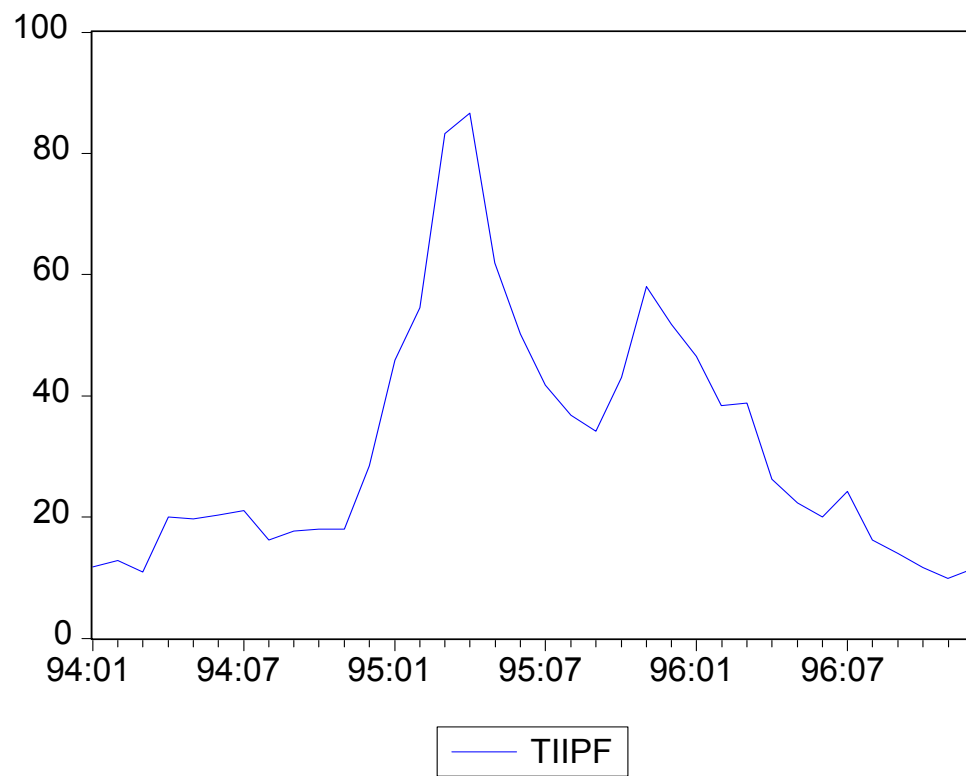
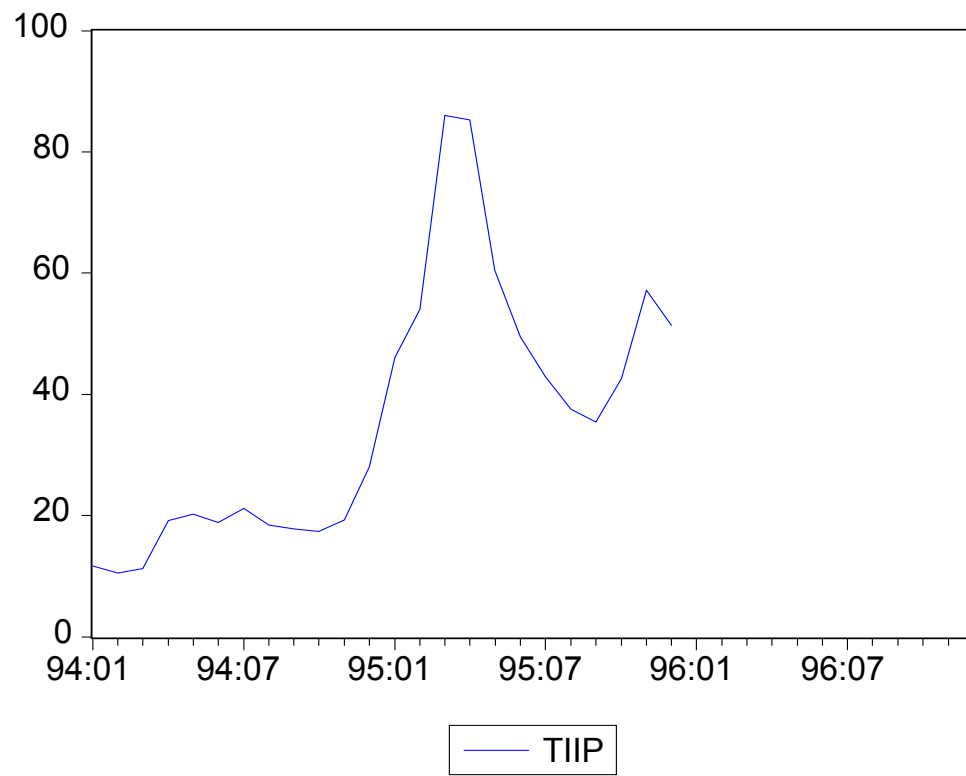


Figura 3.17.: Gráfica del TIIP y la TIIP pronosticada.

El pronóstico de la TIIP, de acuerdo a lo anterior, refleja la tendencia a la baja que se viene registrando actualmente, como signo de recuperación de la economía, aunque a un ritmo muy inestable.



Conclusión

La ecuación final que se utilizó para la regresión solo incluyó cinco de las variables independientes de todas las que se tenían disponibles.

Llama la atención el hecho de que lo que se “suponía” influía en el comportamiento de la TIIP, en realidad no lo hace de manera decisiva; por ejemplo, el caso de la variable NIVEL DE LAS RESERVAS INTERNACIONALES, la cual durante el período 1994 - 1995, era el indicador del grado de liquidez del Gobierno Federal para enfrentar sus compromisos con el exterior y una manera de fijar el deslizamiento del tipo de cambio del dólar. Al incluirla demostró no ser significativa, e incluso la prueba de variables omitidas la descartaba como una variable que aportara información a la regresión de la TIIP.

La explicación puede ser que la variable TIPO DE CAMBIO PROMEDIO DE DÓLAR A LA VENTA, ya contiene esa información en su evolución por lo que una nueva variable resulta redundante.

De igual forma, en el caso de la variable ÍNDICE DE PRECIOS Y COTIZACIONES DE LA BOLSA MEXICANA DE VALORES, cuyo valor acumulado permanece casi constante de un mes a otro, sin embargo, la variación en días particulares es notable cuando es motivada por eventos políticos y/o económicos que vuelven volátil el mercado. En este caso se utilizó una nueva variable que recogiera las variaciones mensuales positivas o negativas para incorporarlas al modelo econométrico.

La misma situación se presentó en la variable MONTO DEL CRÉDITO OTORGADO POR BANXICO A INSTITUCIONES DE CRÉDITO E INTERMEDIARIOS PÚBLICOS Y PRIVADOS, de la cual también se recogieron las variaciones mensuales, debido a que reflejan más representativamente las restricciones en el otorgamiento del crédito, lo que necesariamente influye en el encarecimiento del dinero.

La variable que indudablemente demostró ser significativa para una buena ecuación fueron los RESULTADOS DE LA SUBASTA PRIMARIA DE CERTIFICADOS DE LA TESORERÍA DE LA FEDERACIÓN A 28 DÍAS, que al igual que la TIIP es una tasa indicadora del mercado.

Por ultimo, la influencia del ÍNDICE NACIONAL DE PRECIOS AL CONSUMIDOR, imprimió una tendencia a las tasas de interés, en virtud de la necesidad de los inversionistas por obtener rendimientos reales por sus actividades de ahorro, inversión y crédito.

Asimismo, se muestra una tendencia a la baja en las tasas de los CERTIFICADOS DE LA TESORERÍA DE LA FEDERACIÓN A 28 DÍAS, la cual en la segunda semana de abril registró un rendimiento del 38%. El ÍNDICE DE PRECIOS Y COTIZACIONES DE LA BOLSA MEXICANA DE VALORES, registra un leve repunte en 1996, llegando a rebasar los 3,000 puntos aunque con dificultad para mantenerse, entre otras cosas por la falta de ejercicio del presupuesto para gasto público por parte del Gobierno para el presente año.

Con estos resultados preliminares y la suposición de una leve recuperación de la economía se encontró que, en éste caso en particular, el modelo **no representa eficientemente el fenómeno económico de la TIIP**; por ejemplo, el modelo pronostica para la TIIP en el mes de julio de 1996 un valor de 24.25%, mientras que el resultado oficial fue de 32.82%; para el mes de agosto, el modelo pronosticó un valor de 16.15% y el real fue de 27.33%; por último, en el mes de septiembre el modelo pronosticó un valor de 13.98%, y el valor real a la fecha fue de 26.38%.

La razón se debe principalmente, que no se consideraron otras variables financieras importantes, además de que el modelo necesita ser “cargado” con datos adicionales o expectativas del mercado que no siempre resultan ciertas, dada la alta incertidumbre presente y son tales suposiciones, las que sirven de base para efectuar el pronóstico.

No obstante lo anterior, el modelo puede ser mejorado con la incorporación de nuevos datos y la realización de un nuevo análisis sobre otras variables que se propongan.

La solución más práctica en todo caso es evitar que el modelo dependa de otras variables para funcionar, tal es la propuesta que se manejará en el capítulo IV, donde se analizarán las series de tiempo univariadas y en el cual se supone que la variable misma a pronosticar ya contiene toda la información necesaria de las posibles variables que la influyen en su propio récord histórico, siendo un enfoque distinto al que se ha venido tratando a lo largo de éste trabajo.

CAPÍTULO IV. ANÁLISIS UNIVARIADO DE SERIES DE TIEMPO

Análisis de series de tiempo

Modelos estocásticos y análisis de estacionariedad

Para describir lo que es una serie de tiempo dentro del contexto de procesos estocásticos, es necesario primero definir los procesos estocásticos, esto es: un proceso estocástico es una familia de variables aleatorias asociadas a un conjunto de números reales, de forma tal que a cada elemento del conjunto le corresponda una y sólo una variable aleatoria, esto se escribirá como $\{Z(\tau); \tau \in T\}$, en donde T es el conjunto índice y $Z(\tau)$ es la variable aleatoria correspondiente al elemento τ de T . Si T es un intervalo de números reales ya sea cerrado o abierto, se dirá que el proceso estocástico es continuo, y si T es un conjunto finito o infinito pero numerable, se dirá que es discreto. El hecho de que el proceso estocástico sea continuo o discreto no indica nada acerca de la naturaleza de las variables involucradas, ya que éstas a su vez pueden ser continuas o discretas.

Con base en lo anterior, se concibe entonces a una serie de tiempo como la sucesión de observaciones generadas por un proceso estocástico, cuyo conjunto índice se toma en relación al tiempo. Por lo tanto, la inferencia a realizar será acerca de las características del proceso estocástico generador de la serie observada. En particular, si las observaciones de una serie de tiempo discreta se toman en momentos $\tau_1, \tau_2, \dots, \tau_N$, el proceso estocástico respectivo se denotará por $\{Z(\tau_1), Z(\tau_2), \dots, Z(\tau_N)\}$. Las series económicas y financieras trabajadas hasta ahora han sido discretas, con la característica adicional de que las observaciones son hechas a intervalos con longitud de tiempo fija, (mensual, trimestral, anual, etcétera).

Es importante notar que una serie de tiempo observada no es más que una realización de un proceso estocástico, lo cual significa que bien pudo haberse observado otra realización del mismo proceso, pero cuyo comportamiento fuese distinto del que se observó en la realidad. Con esto se pretende subrayar el elemento probabilístico presente en una serie de tiempo, lo que conducirá a tener en cuenta la función de densidad conjunta de todas las variables aleatorias que constituyen el proceso estocástico.

Como es sabido, el comportamiento de una variable aleatoria Z puede caracterizarse a través de su función de densidad $f(Z)$. Similarmente, dos

variables aleatorias Z_1 y Z_2 quedarán completamente descritas en términos probabilísticos por su función de densidad conjunta $F(Z_1, Z_2)$.

En general N variables aleatorias, y en consecuencia el proceso estocástico podrán describirse mediante la función de densidad conjunta $f(Z_1, Z_2, \dots, Z_N)$.

En prácticamente todo el análisis estadístico, excepto en el análisis de series de tiempo, es costumbre suponer que las observaciones que se tienen provienen de variables aleatorias independientes, de tal forma que con el solo conocimiento de las funciones de densidad individuales es posible obtener fácilmente la función de densidad conjunta. En contraste, en el caso de las series de tiempo se supone que existe toda una estructura de correlación, por consiguiente no es posible obtener la función de densidad conjunta tan directamente y deberá utilizarse alguna otra forma para caracterizar las variables aleatorias que intervienen.

Uso de operadores y polinomios de retraso

El primer operador al que se hará mención es el llamado operador de retraso, el cual se denotará por la letra B (del inglés Backward). Dicho operador se define mediante la relación.

$$BZ_t = Z_{t-1} \text{ para toda } t$$

por aplicación sucesiva del operador B se obtiene:

$$\begin{aligned} B^2 Z_t &= B(BZ_t) = Z_{t-2} \\ B^3 Z_t &= B(B^2 Z_t) = Z_{t-3} \end{aligned}$$

así que, en general, la expresión a la que se llega es:

$$B^k Z_t = Z_{t-k} \text{ para } k = 0, 1, 2, \dots, \text{ y toda } t$$

Adviértase que al “multiplicar” a B^k por Z_t se obtiene la variable retrasada k períodos y, debido a que $B^0 = 1$, se tiene $B^0 Z_t = Z_t$ (si se fuera estricto, debería escribirse $B^0 = I$, en donde I denotaría el operador identidad que deja intacta la variable sobre la cual opera). Debe notarse también que, de hecho, el operador modifica toda la sucesión de valores $\{Z_1, Z_2, \dots, Z_t, \dots, Z_N\}$ y la transforma en la nueva sucesión $\{Z_{1-k}, Z_{2-k}, \dots, Z_{t-k}, \dots, Z_{N-k}\}$. A este respecto conviene apreciar que si únicamente se cuenta con las observaciones Z_1, \dots, Z_N , las observaciones $Z_{1-k},$

Z_{2-k}, \dots, Z_0 no se tendrán y, por lo tanto la serie que originalmente constaba de N observaciones, se reducirá a una serie de solamente $N-k$ observaciones, por el solo hecho de aplicar B^k .

Otro operador de uso muy frecuente y que está íntimamente ligado con B es el operador diferencia ∇ . Este operador se utiliza para expresar relaciones del tipo $Y_t = Z_t - Z_{t-1}$, donde, si Z_t es una variable de saldo, entonces Y_t será la correspondiente variable de flujo; es decir, si se define a ∇ mediante:

$$\nabla Z_t = Z_t - Z_{t-1} \text{ para toda } t$$

se tiene que Y_t puede escribirse como $Y_t = \nabla Z_t$

La relación que liga a ∇ con B es la siguiente:

$$\nabla = 1 - B \text{ o sea } \nabla Z_t = (1 - B)Z_t$$

de esta manera, también podría obtenerse la siguiente forma general para ∇^k :

$$\nabla^k Z_t = \sum_{j=0}^k \frac{k!}{j!(k-j)!} (-1)^j Z_{t-j} \text{ para } k = 0, 1, 2, \dots \text{ y toda } t$$

sin embargo, no se requiere la aplicación sucesiva del operador ∇ para llegar a la ecuación anterior, ya que esta expresión es fácilmente comprobable a través del empleo del Teorema de Binomio, porque ∇^k es en realidad un binomio elevado a la k - ésima potencia, o sea $\nabla^k Z_t = (1 - B)^k Z_t$

Ahora bien, en el análisis de series de tiempo se utilizan operadores de retraso en forma de polinomios, es decir, el polinomio:

$$Z_t - g_1 Z_{t-1} - g_2 Z_{t-2} - \dots - g_k Z_{t-k} = Z_t - \sum_{j=1}^k g_j Z_{t-j}$$

es un polinomio de retraso que puede expresarse como $G(B)Z_t$, en donde:

$$G(B) = 1 - g_1 B - g_2 B^2 - \dots - g_k B^k = 1 - \sum_{j=1}^k g_j B^j$$

y los coeficientes g_1, \dots, g_k son constantes que ponderan la importancia de los retrasos con los cuales están asociados, además k puede ser $1, 2, \dots$

También es frecuente trabajar con polinomios de retraso racionales, los cuales pueden expresarse como cocientes de dos polinomios de retraso, o sea, si a_j y c_j son constantes, $G(B)$ será un polinomio racional si:

$$G(B) = A(B) / C(B) \text{ con } A(B) = 1 - \sum_{j=1}^k a_j B^j \text{ y } C(B) = 1 - \sum_{j=1}^m c_j B^j$$

El hecho de que $G(B)$ tenga tal representación, equivale a restringir el número de coeficientes g_j de este polinomio.

El ejemplo más simple de este tipo de polinomios viene dado por:

$$G(B) = 1 + gB - g^2B^2 + g^3B^3 + \dots \text{ con } |g| < 1 \quad (4.1)$$

el cual por su naturaleza de serie geométrica generada por el término $(gB)^j$ se puede expresar como:

$$G(B) = 1 / (1 - gB), \text{ para } |g| < 1 \quad (4.2)$$

es decir, $G(B) = A(B) / C(B)$ con $A(B) = 1$ y $C(B) = 1 - gB$. La igualdad entre (4.1) y (4.2) puede apreciarse multiplicando ambas relaciones por $1 - gB$. El requerimiento de que $|g| < 1$ surge del siguiente argumento: supóngase que se sugiere operar con el polinomio (4.1) sobre Z_t , así que:

$$G(B)Z_t = (1 - gB - g^2B^2 + g^3B^3 + \dots)Z_t = \sum_{j=0}^{\infty} g^j Z_{t-j}$$

y considérese el caso en que $Z_t = \bar{Z}$, para toda t , entonces se tiene:

$$G(B)Z_t = \bar{Z} \sum_{j=0}^{\infty} g^j$$

en donde se observa que esta última suma será infinita, lo cual no tendría ningún sentido práctico a menos que $|g| < 1$.

El uso de polinomios de retraso es de particular importancia en el estudio de series de tiempo porque permite expresar, de una manera concisa y simple, algunos de los modelos que han probado ser de mayor utilidad en la práctica para representar los fenómenos reales. A reserva de presentar con cierto detenimiento las características más sobresalientes de dichos modelos, cabe por ahora

mencionar simplemente que dentro de estos modelos se encuentran los de promedios móviles, a los cuales se representa mediante la expresión:

$$Z_t - \mu = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

en donde μ denota la media de la serie (o, dicho de otra manera, al nivel del proceso), de tal forma que $Z_t - \mu$ representa la desviación de Z_t respecto a la media, $\{a_t\}$ es una sucesión de variables aleatorias con ciertas características que facilitan su manejo y $\theta_1, \theta_2, \theta_3, \dots, \theta_q$ son parámetros que sirven para relacionar a las sucesiones $\{a_t\}$ y $\{Z_t\}$. Por consiguiente, un modelo de promedios móviles se denota de manera compacta por:

$$Z_t - \mu = \theta(B) a_t. \quad (4.3)$$

Asimismo, los modelos autoregresivos, que se definen como:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(Z_t - \mu) = a_t$$

en donde $\phi_1, \phi_2, \dots, \phi_p$ son parámetros que pueden expresarse mediante la relación

$$\phi(B)(Z_t - \mu) = a_t \quad (4.4)$$

A las combinaciones de modelos anteriores, a las cuales se conoce con el nombre de autoregresivos de promedios móviles, se les puede también representar con la expresión:

$$\phi(B) \nabla^d Z_t = \theta(B) a_t \quad (4.5)$$

que constituyen los llamados modelos autoregresivos integrados y de promedios móviles.

Procesos estocásticos lineales

Los modelos para los procesos estocásticos que serán estudiados en éste capítulo se basan en la idea de que, una serie de tiempo cuyos valores sucesivos pueden ser altamente dependientes, puede considerarse generada a partir de una serie de choques aleatorios independientes $\{a_t\}$. Estos choques aleatorios se supone que son realizaciones independientes de una variable aleatoria cuya media es constante (generalmente se le considera igual a cero) y cuya varianza

es σ_a^2 . A esta sucesión de variables aleatorias $\{a_t\}$ se le conoce como proceso de ruido blanco.

Dicha terminología surge del hecho de que el concepto de ruido blanco se utiliza frecuentemente, en Ingeniería Electrónica y de Comunicaciones, en estudios donde se pretende distinguir a las señales con mensaje de simples ruidos.

La idea expuesta previamente fue concebida por Yule (1927), quien propuso en esencia, expresar al proceso $\{Z_t\}$ en función de $\{a_t\}$ mediante la relación lineal:

$$\begin{aligned} Z_t &= \mu + a_t + \phi_1 a_{t-1} - \phi_2 a_{t-2} - \dots \\ &= \mu + \phi(B)a_t \end{aligned} \quad (4.6)$$

en donde μ es un parámetro que determina el nivel (no necesariamente la media) del proceso y $\phi(B)$ es el polinomio de retraso:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots$$

que convierte el proceso $\{a_t\}$ en el proceso $\{Z_t\}$. De hecho, aquí se tiene en mente un filtro lineal el cual se basa en el operador lineal $\phi(B)$.

Adviértase que los modelos (4.3), (4.4) y (4.5) son casos especiales del modelo (4.6), si se considera respectivamente que $\phi(B)$ está dado por:

$$\phi(B) = \theta(B), \quad \phi(B) = 1/\theta(B) \quad \text{ó} \quad \phi(B) = \theta(B)/\phi(B)$$

Procesos estacionarios

El concepto de estacionariedad es de suma importancia para el análisis de series de tiempo. En general, para poder caracterizar completamente a un proceso estocástico, es necesario conocer la función de densidad conjunta de todas las variables aleatorias involucradas, sin embargo, en la práctica no es muy realista pensar que esto pueda lograrse con una serie de tiempo. No obstante, puesto que los primeros momentos de las variables aleatorias resumen en buena medida a su distribución, se ha sugerido estudiar los momentos de primer y

segundo orden, es decir, las medias, varianzas y covarianzas de las variables involucradas en el proceso estocástico.

Supóngase que la media de Z_t es μ , o sea:

$$E(Z_t) = \mu_t$$

entonces, de acuerdo con la ecuación **(4.6)** se tiene:

$$\mu_t = \mu + E(a_t - \phi_1 a_{t-1} - \phi_2 a_{t-2} - \dots)$$

ya con esta expresión, podría intentarse obtener la esperanza de la suma en paréntesis mediante la suma de las esperanzas de cada uno de los sumandos, pero esto no es válido necesariamente, a menos que se tenga:

$$\phi_0 + \sum_{i=1}^{\infty} |\phi_i| < \infty \quad (4.7)$$

donde $\phi_0 = 1$, esto es, para que la esperanza de la suma exista se requiere que la serie de ponderaciones en valor absoluto converja, si esto ocurre se tendría entonces:

$$E(Z_t) = \mu \quad (4.8)$$

ya que $E(a_t) = 0$ para toda t . Por consiguiente, si **(4.7)** se cumple, la media del proceso no depende del tiempo, lo cual implica en particular que aún cuando durante cierto período el proceso se aleje de la media, éste siempre regresará a una vecindad de la misma.

La varianza del proceso **(4.6)** se obtiene fácilmente a partir de su definición, es decir, si γ_0 denota a la varianza de Z_t :

$$\begin{aligned} \gamma_0 &= E[(Z_t - \mu)^2] \\ &= E[(a_t - \phi_1 a_{t-1} - \phi_2 a_{t-2} - \dots)^2] \\ &= E(a_t - \phi_1 a_{t-1} - \phi_2 a_{t-2} - \dots)^2 - E(\text{productos cruzados de } \phi_i a_{t-i} \text{ y } \phi_j a_{t-j}) \end{aligned}$$

$$= \sigma_a^2 \sum_{i=0}^{\infty} \phi_i^2 \quad (4.9)$$

expresión que tendrá sentido solamente cuando $\sum_{i=0}^{\infty} \phi_i^2$ converja, lo cual ocurre si (4.7) se cumple. La razón por lo cual la esperanza de los productos cruzados desaparece, es por el supuesto sobre $\{a_t\}$ de que son variables aleatorias independientes con media cero, entonces se sigue que:

$$E(\phi_i a_{t-i} \cdot \phi_j a_{t-j}) = \phi_i E(a_{t-i}) \cdot \phi_j E(a_{t-j}) = 0 \text{ para toda } i \neq j$$

Finalmente, al considerar a una serie de tiempo como un proceso estocástico se debe tener en mente a todas las variables aleatorias $Z_1, Z_2, \dots, Z_t, \dots, Z_N$, de forma tal que resulta necesaria estudiar la covarianza entre Z_t y Z_{t+k} .

Dicha covarianza se denota por $\gamma_k = \text{Cov}(Z_t, Z_{t+k})$ y viene dada por:

$$\begin{aligned} \gamma_0 &= E[(Z_t - \mu)(Z_{t+k} - \mu)] \\ &= E[(a_t - \phi_1 a_{t-1} - \phi_2 a_{t-2} - \dots)(a_{t+k} - \phi_1 a_{t+k-1} - \phi_2 a_{t+k-2} - \dots)] \\ &= E(-\phi_k a_t^2 + \phi_1 \phi_{k+1} a_{t-1}^2 + \phi_2 \phi_{k+2} a_{t-2}^2 + \dots) - E(\text{productos cruzados de } \phi_i a_{t-i} \text{ y } \phi_j a_{t-j}) \\ &= \sigma_a^2 (-\phi_k + \phi_1 \phi_{k+1} + \phi_2 \phi_{k+2} + \dots) \\ &= \sigma_a^2 \left(\sum_{i=1}^{\infty} \phi_i \phi_{k+i} - \phi_k \right) \end{aligned} \quad (4.10)$$

Es importante apreciar que, en las expresiones (4.8), (4.9) y (4.10) no aparece t , lo cual implica que ni el nivel de la serie, ni su variabilidad dependen del tiempo, y en lo que respecta a la covarianza, no existe dependencia de tiempo, pero si de la separación (k) que hay entre las variables.

Lo anterior conduce a pensar que la serie mostrará el mismo comportamiento en términos generales sin importar el momento en el que se observe. Esto es, si se graficara un cierto número de observaciones contiguas de una serie, la gráfica que se obtendría sería bastante similar a la que se obtuviera

al graficar el mismo número de observaciones contiguas, pero k períodos hacia adelante o hacia atrás de los considerados inicialmente; así el comportamiento de $(Z_{t_0}, Z_{t_0+1}, \dots, Z_{t_0+m})$ y el de $(Z_{t_0+k}, Z_{t_0+k+1}, \dots, Z_{t_0+k+m})$ será esencialmente el mismo, para cualquier valor k que se considere.

Los procesos que cumplen con que sus momentos de primero y segundo orden no dependen del tiempo, se les denomina estacionarios de segundo orden. Entonces, si las expresiones (4.8), (4.9) y (4.10) son validas para una serie de tiempo (lo cual ocurre siempre que la suma $\sum_{i=1}^{\infty} |\varphi_i|$ es convergente) se dirá que la serie $\{Z_t\}$ es estacionaria de segundo orden.

Asimismo, un proceso estocástico será estrictamente estacionario, si la función de densidad para un conjunto arbitrario de variables $(Z_t, Z_{t+1}, \dots, Z_{t+m})$ es invariante respecto a desplazamientos en el tiempo, es decir, si se cumple que:

$$f(Z_t, Z_{t+1}, \dots, Z_{t+m}) = f(Z_{t+k}, Z_{t+k+1}, \dots, Z_{t+k+m}) \text{ para toda } t, m \text{ y } k$$

De esta expresión, si los primeros momentos de f existen, se sigue que:

$$E(Z_t) = E(Z_{t+m}) = \mu, \text{ que } E[(Z_t - \mu)^2] = E[(Z_{t+m} - \mu)^2] = \gamma_0 \text{ y que}$$

$$E[(Z_t - \mu)(Z_{t+k} - \mu)] = E[(Z_{t+m} - \mu)(Z_{t+m+k} - \mu)] = \gamma_k \text{ para toda } t, m \text{ y } k$$

Por consiguiente, la estacionariedad estricta implica la estacionariedad de segundo orden, y aunque la implicación inversa no siempre se cumple, sí se satisface cuando el proceso tiene como distribución conjunta de todas las variables que lo integran a la normal multivariada.

Debido a que en la práctica es común suponer que la distribución asociada con las series de tiempo es la normal, se sigue que es suficiente conocer la media μ y la función de autocovarianza $\{\gamma_k\}$ para caracterizar completamente a una serie estacionaria (adviértase que se habla de autocovarianza, lo cual sirve para indicar que la covarianza es entre observaciones de la misma variable, aunque retrasada). Sin embargo, para evitar la influencia de las unidades de medida, es preferible trabajar con las autocorrelaciones $\rho_0, \rho_1, \rho_2, \dots$ definidas a través de:

$$\rho_k = \frac{E[(Z_t - \mu)(Z_{t+k} - \mu)]}{E[(Z_t - \mu)^2]} = \frac{\gamma_k}{\gamma_0}, \text{ para } k = 0, \pm 1, \pm 2, \pm 3, \dots$$

las cuales generan la función de autocorrelación $\{\rho_k\}$, la cual sirve para especificar, junto con μ y σ_a^2 , un proceso estocástico estacionario.

Un fenómeno típico de las series de tiempo no estacionarias es que la función de autocorrelación tiende a cero muy lentamente, cabe notar que en la vida real, lo más frecuente es que las series que se observan tengan funciones de autocorrelación que no desaparecen rápidamente, es decir, lo común es que las series de tiempo económicas y financieras que se manejan en la práctica no sean estacionarias.

Diferencias y no estacionariedad homogénea

El hecho de que la gran mayoría de las series de tiempo con las que uno trabaja en la práctica no sean estacionarias, no es tan grave como puede parecer a primera vista, ya que existen formas para solucionar este problema y convertir en estacionarias a series que originalmente no lo son. En muchos casos, la no estacionariedad de la serie es homogénea, lo cual significa que es únicamente el nivel de la serie el que se ve afectado por la no estacionariedad, debido a que existe alguna tendencia polinomial. En estos casos es posible eliminar dicha tendencia, y por lo tanto volver la serie estacionaria, mediante la aplicación del operador ∇ un número apropiado de veces. Para ver esto, considérese el proceso definido por:

$$Z_t = \alpha_0 + \beta_0 t + \gamma_0 t^2 + \dots + \delta_0 t^d + a_t \text{ si } t = 1, \dots, n_1$$

$$Z_t = \alpha_1 + \beta_1 t + \gamma_1 t^2 + \dots + \delta_0 t^d + a_t \text{ si } t = n_1 + 1, \dots, n_2$$

$$Z_t = \alpha_e + \beta_e t + \gamma_e t^2 + \dots + \delta_0 t^d + a_t \text{ si } t = n_e + 1, \dots, N$$

en donde $\{a_t\}$ es ruido blanco, $\alpha_i, \beta_i, \gamma_i, \dots, \delta_0$ ($i = 0, \dots, e$) son parámetros, no todos necesariamente distintos de cero y $d \geq 1$, de tal forma que Z_t sigue una tendencia polinomial de orden d . Entonces, puede verificarse que, si $t \neq n_i + 1, \dots, n_i + d$ con $i = 1, \dots, e$.

$$\nabla^d Z_t = (d!) \delta_0 + \nabla^d a_t \quad (4.11)$$

La fórmula (4.11) es a final de cuentas, la que define básicamente el comportamiento de $\nabla^d Z_t$ y por lo tanto, la nueva serie $\{\nabla^d Z_t\}$ tendrá las mismas

características que $\{\nabla^d a_t\}$ excepto por el nivel, que para $\{\nabla^d a_t\}$ es cero y para $\{\nabla^d Z_t\}$ es $(d!) \delta_0$, pero que igualmente no depende de t .

La conclusión es por consiguiente que la nueva serie, obtenida mediante la aplicación del operador ∇^d es estacionaria en cuanto a su nivel.

Es importante también observar que al aplicar el operador diferencia a una serie de tiempo, en éste caso $\{a_t\}$, se obtiene otra serie también estacionaria, $\{\nabla^d a_t\}$, con varianza $(2d)! \sigma_a^2 / (d!)^2$.

El propósito de tomar diferencias es pues, volver estacionario el nivel de una serie, pero debe recordarse que si se toman diferencias de series que son estacionarias, éstas seguirán siendo estacionarias, lo cual significa que fácilmente podría sobrediferenciarse a una serie al tratar de volverla estacionaria. El hecho de sobrediferenciar una serie acarrea problemas en la identificación de algún modelo para representarla, se incrementa la varianza de la serie y se pierden observaciones innecesariamente, ya que al aplicar d veces el operador ∇ , o sea, al aplicar ∇^d , se pierden automáticamente d observaciones. Es necesario entonces tener cuidado de no sobrediferenciar la serie de estudio, lo cual requiere esencialmente práctica en el manejo de series de tiempo.

La razón por la cual la función de autocorrelación asociada con procesos homogéneos no estacionarios no decae rápidamente a cero, es porque la tendencia polinomial produce que las observaciones de la serie dependan de observaciones distantes en el tiempo, y por consiguiente que las autocorrelaciones sean altas aún para retrasos muy separados. De aquí se sigue que la función de autocorrelación mide básicamente lo marcado de la tendencia polinomial; para evitar que esto ocurra se debe eliminar dicha tendencia, lo cual se logra mediante un “número apropiado” de aplicaciones del operador diferencia. La forma como se determina en la práctica el número de diferencias se basa en que las autocorrelaciones muestrales sean prácticamente cero después de un cierto número de rezagos (comúnmente tres o cuatro retrasos, si es que no existe algún efecto estacional que actúe sobre la serie y que requiera de tratamiento especial para series de tiempo estacionales).

Construcción de modelos para series univariadas

La estrategia de construcción de modelos que se presentará a continuación fue desarrollada por Box y Jenkins (1970). Dicha estrategia consta de cuatro etapas fundamentales:

Identificación de un posible modelo dentro de la clase de modelos ARIMA; es decir, determinación de los valores p , d , y q que especifiquen el modelo ARIMA apropiado para la serie de estudio

Estimación de los parámetros involucrados en el modelo, a través de técnicas de regresión.

Verificación de que el modelo proporciona un ajuste adecuado y de que los supuestos básicos implícitos en el modelo, se satisfacen; de no cumplirse los supuestos, se determinan las modificaciones necesarias y de hecho, se repiten las etapas anteriores hasta que la verificación indique resultados aceptables, y por último:

Uso del modelo para los fines que el analista haya tenido en mente al construirlo; dichos fines son por lo general de **pronóstico**, control, simulación o explicación del fenómeno de estudio.

Clases de modelos ARIMAS

Modelos autoregresivos AR

Sea la ecuación $A(B)Z_t = \text{constante}$, en donde $A(B)$ representa un polinomio de retraso. Una generalización de este tipo consiste en introducir una variable aleatoria en el lado derecho, de tal manera que se tenga:

$$A(B)Z_t = \text{constante} + a_t \quad (4.12)$$

en donde, por simplicidad, se supone que $\{a_t\}$ es un proceso de ruido blanco. La ecuación (4.12) permite representar los procesos autoregresivos mediante la equivalencia:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)Z_t = \text{constante} + a_t \quad (4.13)$$

en donde la constante es igual a $(1 - \phi_1 - \phi_2 - \dots - \phi_p)\mu$, en caso de que el proceso Z_t tenga un nivel medio constante dado por $E(Z_t) = \mu$ para toda t ; de esta forma, la ecuación (4.13) se convierte en:

$$\phi(B)\tilde{Z}_t = a_t, \text{ con } \tilde{Z}_t = Z_t - \mu \quad (4.14)$$

El término autoregresivo (AR) que se le da al proceso representado por (4.14) se refiere al hecho de que también puede expresarse como:

$$Z_t = (1 - \phi_1 - \phi_2 - \dots - \phi_p)\mu + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t$$

la cual es básicamente una ecuación de regresión lineal con la característica especial de que el valor de la variable dependiente Z en el período t depende, no de los valores de un cierto conjunto de variables independientes, como sucede en el modelo de regresión, sino de sus propios valores, observados en períodos anteriores a t y ponderados de acuerdo con los coeficientes autoregresivos ϕ_1, \dots, ϕ_p .

Un proceso AR podrá ser estacionario o no estacionario dependiendo de los valores que tomen las raíces de la ecuación característica $\phi(x) = 0$, la cual dicta el comportamiento del proceso autoregresivo.

Modelo AR(1)

El caso más simple es el de un modelo autoregresivo de orden uno, es decir, un AR(1), que se representa como:

$$\tilde{Z}_t - \phi\tilde{Z}_{t-1} = a_t$$

Para que dicha ecuación sea estacionaria se requiere que la raíz de la ecuación $1 - \phi x = 0$ se encuentre fuera del círculo unitario; o sea, se requiere que $|\phi| < 1$ para asegurar la estacionariedad del proceso AR(1).

Otra posible representación de dicho proceso autoregresivo, en términos de la serie de errores aleatorios exclusivamente, es

$$\tilde{Z}_t = (1 - \phi B)^{-1} a_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots$$

en donde se observa que, si $|\phi| < 1$

$$(\tilde{Z}_t) = E(a_t) + \phi E(a_{t-1}) + \phi^2 E(a_{t-2}) + \dots = 0$$

$$\gamma_0 = \text{Var}(\tilde{Z}_t) = \sigma_a^2(1 + \phi^2 + \phi^4 + \dots)$$

de tal forma que tanto la media como la varianza de la serie son constantes y además se obtiene:

$$\gamma_0 = \frac{\sigma_a^2}{(1 - \phi^2)}$$

Asimismo, las autocovarianzas pueden obtenerse a partir de la expresión **(4.10)** con $\phi_i = -\phi^i$, $i = 1, 2, \dots$, es decir:

$$\begin{aligned} \gamma_k &= \sigma_a^2 \left(\sum_{i=1}^{\infty} \phi_i \phi^{k+i} - \phi^k \right) \\ &= \sigma_a^2 \phi^k \sum_{i=1}^{\infty} \phi^{2i}, \quad k = 1, 2, \dots \end{aligned}$$

por lo cual, ya que se supone $|\phi| < 1$,

$$\gamma_k = \frac{\sigma_a^2 \phi^k}{(1 - \phi^2)} = \phi \gamma_{k-1}, \quad k = 1, 2, \dots$$

y debido a que $\gamma_k = \gamma_{-k}$ se obtiene la fórmula general:

$$\gamma_k = \frac{\sigma_a^2 \phi^{|k|}}{(1 - \phi^2)}, \quad k = 0, \pm 1, \pm 2, \dots$$

de donde se sigue que las autocorrelaciones deben ser de la forma

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi^{|k|}, \quad k = 0, \pm 1, \pm 2, \dots \quad \mathbf{(4.15)}$$

lo cual indica que conforme $k > 0$ crece, la función de autocorrelación (FAC) tiende a cero, con decaimiento del tipo exponencial cuando $0 < \phi < 1$. Esto mismo puede

observarse al expresar la FAC mediante la ecuación en diferencia de primer orden:

$$(1-\phi B)\rho_k = 0, \quad k = 1, 2, \dots \quad (4.16)$$

sujeta a la condición inicial $\rho_0 = 1$, adviértase que en (4.15) el operador B actúa sobre el índice k . Además, debido a (4.16), la condición de estacionariedad del proceso $AR(1)$, $|\phi| < 1$, en términos de las autocorrelaciones, se convierte en:

$$|\rho_1| < 1$$

Modelo $AR(2)$

El siguiente esquema autoregresivo se define por

$$(1 - \phi_1 B - \phi_2 B^2)\tilde{Z}_t = a_t \quad (4.17)$$

que fue introducido por Yule (1927). Para que el proceso (4.17) sea estacionario, debe cumplirse que las raíces de

$$1 - \phi_1 x - \phi_2 x^2 = 0 \quad (4.18)$$

se encuentren fuera del círculo unitario.

El Teorema de Schur, proporciona una manera de encontrar las condiciones estacionariedad, a causa de la equivalencia entre equilibrio de un proceso determinista y estacionariedad de un proceso estocástico; por lo tanto, para que un proceso $AR(2)$ sea estacionario se debe cumplir que:

$$D_1 = \begin{vmatrix} -1 & \phi_2 \\ \phi_2 & -1 \end{vmatrix} > 0 \quad y \quad D_1 = \begin{vmatrix} -1 & 0 & \phi_2 & \phi_1 \\ \phi_1 & -1 & 0 & \phi_2 \\ \phi_2 & 0 & -1 & \phi_1 \\ \phi_1 & \phi_2 & 0 & -1 \end{vmatrix} > 0 \quad (4.19)$$

Las condiciones enunciadas en (4.19) son equivalentes a requerir que:

$$|\phi_2| < 1, \phi_2 + \phi_1 < 1 \text{ y } \phi_2 - \phi_1 < 1 \quad (4.20)$$

Si el proceso AR(2) en estudio resulta ser estacionario, lo único que hace falta para describirlo completamente es obtener su función de autocovarianza, lo cual se logra como sigue: si $E(Z_t) = \mu$ entonces $E(\tilde{Z}_t) = E(Z_t - \mu) = 0$ y se tiene

$$E(\tilde{Z}_t \tilde{Z}_{t+k}) = \phi_1 E(\tilde{Z}_t \tilde{Z}_{t+k-1}) + \phi_2 E(\tilde{Z}_t \tilde{Z}_{t+k-2}) + E(\tilde{Z}_t a_{t+k})$$

de tal forma que

$$\gamma_k = \begin{cases} \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_a^2, & \text{si } k = 0 \\ \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma_a^2, & \text{si } k > 0 \end{cases}$$

De aquí que, si $\gamma_0 < \infty$, se puede dividir las autocovarianzas γ_1 y γ_2 entre γ_0 para obtener las ecuaciones de Yule - Walker (llamadas así en honor de Yule (1927) y Walker (1931))

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \end{aligned} \quad (4.21)$$

que permiten obtener los valores de ρ_1 y ρ_2 en función de los parámetros autoregresivos ϕ_1 y ϕ_2 .

En general, la FAC satisface

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \quad k \geq 3 \quad (4.22)$$

de tal forma que ρ_k sigue un comportamiento autoregresivo similar al de (4.17), pero con la diferencia de que en (4.22) no aparece ningún término aleatorio. Por tanto, las raíces de la ecuación característica (4.18) sirven para indicar el comportamiento dinámico de la función de autocorrelación.

De acuerdo con los dos casos de solución de la ecuación cuadrática (4.18) se tiene que

$$x = \frac{-\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2}$$

en la que si $\phi_1^2 + 4\phi_2 \geq 0$, entonces las raíces son reales y las autocorrelaciones decaerán exponencialmente a cero (todas las autocorrelaciones serán positivas si la primera lo es y tendrán signos alternados si la primera autocorrelación es negativa), y si $\phi_1^2 + 4\phi_2 \leq 0$ las raíces serán complejas, así que la FAC correspondiente seguirá un comportamiento sinusoidal convergente a cero.

De las ecuaciones (4.21) se sigue que

$$\phi_1 = \rho_1 (1 - \rho_2) / (1 - \rho_1^2)$$

$$\phi_2 = (\rho_2 - \rho_1^2) / (1 - \rho_1^2)$$

de manera que al sustituir los valores de las autocorrelaciones ρ_1 y ρ_2 por sus valores estimados, se obtienen estimaciones para los parámetros ϕ_1 y ϕ_2 . Además, al resolver las ecuaciones de Yule - Walker para ρ_1 y ρ_2 se obtiene:

$$\rho_1 = \phi_1 / (1 - \phi_2)$$

$$\rho_2 = \phi_2 + \phi_1^2 / (1 - \phi_2)$$

de tal forma que las condiciones de estacionariedad (4.20) pueden también expresarse en términos de autocorrelaciones, es decir, dichas condiciones son equivalentes a:

$$|\rho_2 - \rho_1^2| < |1 - \rho_1^2|, \rho_2 + \rho_1^2 + \rho_1 - \rho_1\rho_2 < 1 - \rho_1^2 \text{ y } \rho_2 - \rho_1^2 - \rho_1 + \rho_1\rho_2 < 1 - \rho_1^2$$

las cuales a su vez se cumplen si y sólo si $\rho_2 < 1$ y $\rho_1^2 < (\rho_2 + 1)/2$, $\rho_1 < 1$ y $\rho_1 > -1$, así que las condiciones de estacionariedad de segundo orden, en términos de las autocorrelaciones, vienen a ser:

$$\begin{aligned} \rho_2 &< 1 \\ \rho_1^2 &< (\rho_2 + 1)/2 \\ -1 &< \rho_1 < 1 \end{aligned}$$

Modelo AR(p)

Como caso general de un proceso autoregresivo se mencionará el proceso AR(p) que se describe mediante la ecuación

$$Z_t = (1 - \phi_1 - \phi_2 - \dots - \phi_p)\mu + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t$$

la cual es equivalente a

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t$$

en donde $\tilde{Z}_t = Z_t - \mu$

Un proceso AR(p) será estacionario si y sólo si las raíces de la ecuación característica

$$1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0$$

se encuentran fuera del círculo unitario.

Modelos de promedios móviles MA

Los modelos de promedios móviles fueron introducidos por Yule (1926) y Slutsky (1927); la idea básica de estos modelos consiste en presentar a un proceso estocástico $\{Z_t\}$, como una suma finita ponderada de choques aleatorios independientes $\{a_t\}$, o sea

$$\tilde{Z}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t = \theta(B) a_t \quad (4.23)$$

en donde $\{\tilde{Z}_t\}$ representa a las desviaciones de $\{Z_t\}$ respecto a su nivel medio μ y $\theta_1, \theta_2, \dots, \theta_q$ son las ponderaciones (parámetros de promedios móviles) asociados con los choques aleatorios en los períodos $t - 1, t - 2, \dots, t - q$, respectivamente.

El término de promedios móviles parece sugerir que el modelo se obtiene como un promedio de los choques aleatorios que intervienen, pero no es así puesto que los parámetros no tienen que ser necesariamente positivos ni su suma debe ser la unidad, como requeriría un promedio.

Obsérvese que el modelo (4.23) está expresado en forma de proceso lineal general y que la suma $\sum_{q=1}^{\infty} |\theta_i|$, al considerar un número finito de sumandos, es una

constante finita (a menos que algún θ_i fuese infinito, pero esto implicaría que \tilde{Z}_t es infinito), por consiguiente, todo proceso MA es estacionario.

Esto sugiere la siguiente interpretación para los modelos MA: dado un proceso que se encuentre en equilibrio, $\{\tilde{Z}_t\}$, son causadas por choques asociados con eventos inesperados; tales choques no necesariamente se asimilan de manera instantánea, sino que pueden seguir causando efectos aún después de transcurrido un cierto número de períodos y además, la intensidad del choque se refleja en el valor de su ponderación θ_i .

Modelo MA(1)

El proceso de promedios móviles de orden uno es el más simple, dicho proceso MA(1) se expresa mediante

$$\tilde{Z}_t = (1 - \theta B)a_t,$$

por lo cual es inmediato obtener

$$E(\tilde{Z}_t) = 0 \text{ y } \gamma_0 = \text{Var}(\tilde{Z}_t) = \sigma_a^2(1 + \theta^2)$$

además las autocovarianzas vienen dadas por

$$\gamma_k = E[(a_t - \theta a_{t-1})(a_{t-k} - \theta a_{t-k-1})] = \begin{cases} -\theta \sigma_a^2 & \text{si } k = 1 \\ 0 & \text{si } k \geq 2 \end{cases}$$

de donde se sigue que la FAC es

$$\rho_k = \begin{cases} \frac{-\theta}{1 + \theta^2} & \text{si } k = 1 \\ 0 & \text{si } k \geq 2 \end{cases} \quad (4.24)$$

El hecho de que las autocorrelaciones para retrasos mayores que un período sean de cero, indica que el proceso de MA(1) “no recuerda” más allá de lo que ocurrió el período anterior, es decir, tiene una memoria limitada a solo un

período. Es importante notar también que aún cuando solamente la primera autocorrelación es distinta de cero, eso implicaría una fuerte dependencia de la observación actual con la anterior y así sucesivamente, por lo cual sería más adecuado pensar en un modelo autoregresivo para esa situación; por otro lado, dada la ecuación (4.24) se puede demostrar que $|\rho_1| < 0.5$ lo cual corrobora la idea intuitiva.

Por otro lado, este hecho implica que aún cuando una cierta FAC muestre un sólo pico, correspondiente al retraso 1, dicha FAC corresponderá a un proceso MA(1) sólo cuando satisfaga $|\rho_1| < 0.5$.

Modelo MA(2)

El siguiente proceso a considerar será el de medias móviles de orden dos, cuyo modelo es:

$$\tilde{Z}_t = (1 - \theta_1 B - \theta_2 B^2) a_t$$

dicho proceso tiene como media y varianza a

$$E(\tilde{Z}_t) = E(Z_t) - \mu = 0 \text{ y } \gamma_0 = \sigma_a^2(1 + \theta_1^2 + \theta_2^2)$$

y como función de autocovarianza a

$$\begin{aligned} \gamma_k &= E[(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2})(a_{t-k} - \theta_1 a_{t-k-1} - \theta_2 a_{t-k-2})] \\ &= E[a_t a_{t-k} + \theta_1^2 a_{t-1} a_{t-k-1} - \theta_1(a_t a_{t-k-1} + a_{t-1} a_{t-k}) \\ &\quad - \theta_1 \theta_2(a_{t-1} a_{t-k-2} - a_{t-2} a_{t-k-1}) - \theta_2(a_t a_{t-k-2} + a_{t-2} a_{t-k}) + \theta_2^2 a_{t-2} a_{t-k-2}] \\ &= \begin{cases} -(\theta_1 + \theta_1 \theta_2) \sigma_a^2 & \text{si } k = 1 \\ -(\theta_2) \sigma_a^2 & \text{si } k = 2 \\ 0 & \text{si } k \geq 3 \end{cases} \end{aligned}$$

por lo cual, la correspondiente FAC viene a ser:

$$\rho_k = \begin{cases} \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} & \text{si } k = 1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{si } k = 2 \\ 0 & \text{si } k \geq 3 \end{cases}$$

De las fórmulas anteriores resulta evidente que el proceso estacionario; para que sea también invertible se requiere que las raíces de la ecuación:

$$1 - \theta_1 x - \theta_2 x^2 = 0$$

se encuentren fuera del círculo unitario. Las condiciones de invertibilidad pueden expresarse como:

$$|\theta_2| < 1, \theta_2 + \theta_1 < 1 \text{ y } \theta_2 - \theta_1 < 1$$

Así como se mencionó para un proceso MA(1), la primera autocorrelación no puede ser mayor en valor absoluto que 0.5 de igual manera existen restricciones para que los valores de un conjunto de autocorrelaciones pertenezcan a un proceso MA(2) y se puede demostrar que dichas restricciones son:

$$|\rho_1^2| \leq 0.5 \text{ y } |\rho_2| \leq 0.5$$

Modelo MA(q)

En general, se dirá que un proceso estocástico sigue un esquema de promedios móviles de orden $q \geq 1$ si se le puede representar mediante.

$$\tilde{Z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

con $\tilde{Z}_t = Z_t - \mu$, en donde μ es el nivel del proceso, $\theta_1, \theta_2, \dots, \theta_q$ son los parámetros de promedios móviles y $\{a_t\}$ es un proceso de ruido blanco con media cero y varianza constante σ_a^2 . Como ya se mencionó anteriormente, todo proceso MA es estacionario, y en particular se puede observar que ni la media ni la varianza, ni las covarianzas del proceso MA(q), dependen del tiempo.

Modelos ARMA

Una generalización de los modelos AR y MA previamente descritos, consiste en combinar ambas clases de modelos para obtener lo que se conoce como modelos autoregresivos y de promedios móviles (ARMA), los cuales fueron estudiados por Wold (1938) y Bartlett (1946). El proceso ARMA(p,q) se representa mediante:

$$\phi(B)\tilde{Z}_t = \theta(B)a_t$$

en donde \tilde{Z}_t son polinomios de retraso de orden p y q respectivamente, es decir,

$$\tilde{Z}_t - \phi_1\tilde{Z}_{t-1} - \dots - \phi_p\tilde{Z}_{t-p} = a_t - \theta_1a_{t-1} - \dots - \theta_qa_{t-q}$$

donde $\{a_t\}$ es un proceso de ruido blanco y \tilde{Z}_t es la serie de desviaciones de la variable Z_t , respecto a su nivel μ .

Para que este proceso sea estacionario se requiere que las raíces de $\phi(x) = 0$ estén fuera del círculo unitario, y para que sea invertible la condición es que las raíces de la ecuación $\theta(x) = 0$ se encuentren también fuera del círculo unitario, si ambos casos ocurren, entonces el proceso admite las representaciones AR y MA siguientes:

$$\tilde{Z}_t = \frac{\theta(B)}{\phi(B)}a_t = \varphi(B)a_t$$

$$\frac{\theta(B)}{\phi(B)}\tilde{Z}_t = \pi(B)\tilde{Z}_t = a_t$$

Dicha generalización surge del hecho de que las series de tiempo que se observan en la práctica, muchas veces presentan características tanto de procesos AR como de procesos MA. Por otro lado, Granger y Morris (1976) presentan casos particulares de procesos ARMA que surgen al considerar series obtenidas con agregación de componentes (por ejemplo, el PIB, que se obtiene al agregar series de diversos sectores de la economía) o series con errores en los datos observados, lo cual es común en series macroeconómicas.

El proceso autoregresivo y de promedios móviles de orden (1,1), aún siendo el más sencillo de los procesos ARMA, es de gran interés desde el punto de vista práctico, porque proporciona representaciones adecuadas para muchas series de fenómenos reales. El modelo ARMA(1,1) está definido por:

$$(1 - \phi B)\tilde{Z}_t = (1 - \theta B)a_t$$

y, puesto que contiene tanto características autoregresivas como de promedios móviles, las condiciones de invertibilidad y estacionariedad se derivan de las condiciones respectivas para los procesos AR(1) y MA(1).

Modelos ARIMA

Los modelos autoregresivos e integrados de promedios móviles (ARIMA) pueden ser vistos como una generalización de los modelos ARMA. Yaglom (1955) sugirió la posibilidad de que un cierto tipo de no estacionariedad mostrado por algunas series de tiempo, podía representarse mediante la simple toma sucesiva de diferencias de la serie original. Esto permite gran flexibilidad a los modelos ARMA, puesto que en realidad lo que se hace al aplicar el operador diferencia ∇^d es eliminar una posible tendencia polinomial de orden d , presente en la serie que se analice. Visto desde otro ángulo, si el proceso original $\{\tilde{Z}_t\}$ adolece de no estacionariedad causada por una tendencia polinomial no determinista (a la cual se denomina no estacionariedad homogénea) es posible construir el proceso estacionario $\{W_t\}$, en donde:

$$W_t = \nabla^d \tilde{Z}_t \quad \text{para toda } t \quad (4.25)$$

para esta nueva serie podría ser ya posible obtener un modelo ARMA; $\phi(B)W_t = \theta(B)a_t$, lo cual sería equivalente a considerar el modelo ARIMA:

$$\phi(B)\nabla^d \tilde{Z}_t = \theta(B)a_t, \quad d \geq 1$$

para $\{\tilde{Z}_t\}$, en donde $\{a_t\}$ es un proceso de ruido blanco. El término “integrado” se refiere a que \tilde{Z}_t se obtiene de la relación (4.25) por inversión del operador ∇^d , dando como resultado una suma infinita (o una integración) de términos de W_t .

El orden del polinomio de retraso $\phi(B)$, del exponente en el operador diferencia y el orden del polinomio de retraso $\theta(B)$, se acostumbra mencionarlos en esa secuencia, de manera que un modelo ARIMA(p,d,q) indica que consta de un polinomio autoregresivo de orden p, de una diferencia de orden d y de un polinomio de promedios móviles de orden q.

A continuación, para ejemplificar lo descrito anteriormente, se presentará un análisis estadístico de una serie de tiempo económica, vista como un proceso estocástico (variable aleatoria como función del tiempo), con el propósito de identificar un modelo que represente lo más fielmente posible el comportamiento de los datos observados y que, adicionalmente, dicho modelo sea sugerido por los datos mismos.

Cabe mencionar que el fin último de éste análisis es la posibilidad de simular un fenómeno económico y de efectuar pronósticos verosímiles sobre el desarrollo futuro de la variable de estudio.

La variable económica escogida para ser analizada es el Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (IPC), base octubre 1978 = 0.78, por el periodo de enero de 1983 a septiembre de 1995; siendo tomados 153 valores al último día de operaciones de cada mes, según datos proporcionados por el BDINEGI, (Banco de Datos del Instituto Nacional de Estadística, Geografía e Informática).

Como se mencionó, el proceso de construcción del modelo se divide en cuatro etapas fundamentales:

- I. Identificación del modelo dentro de las diferentes clases de ARIMAS posibles.
- II. Estimación de los parámetros involucrados en el modelo por medio de técnicas de estimación, (regresiones).
- III. Verificación de que el modelo presenta un ajuste adecuado, determinando las modificaciones necesarias, hasta que la verificación indique resultados aceptables.
- IV. Uso del modelo para la realización de un pronóstico por el período del octubre de 1995 a diciembre 1996 y otro por el período de octubre de 1995 a marzo de 1996.

Para la valoración del modelo y la elaboración del pronóstico, se cuenta con la ayuda del paquete de cómputo Econometric Views for Windows, con el cual se obtendrán todas las pruebas para llegar a un modelo adecuado.

Construcción del modelo

Identificación del modelo

Como un primer acercamiento a la variable I.P.C. de la B.M.V., se presenta la gráfica de los datos históricos, **figura 4.1** y de la cual se desprenden las siguientes observaciones:

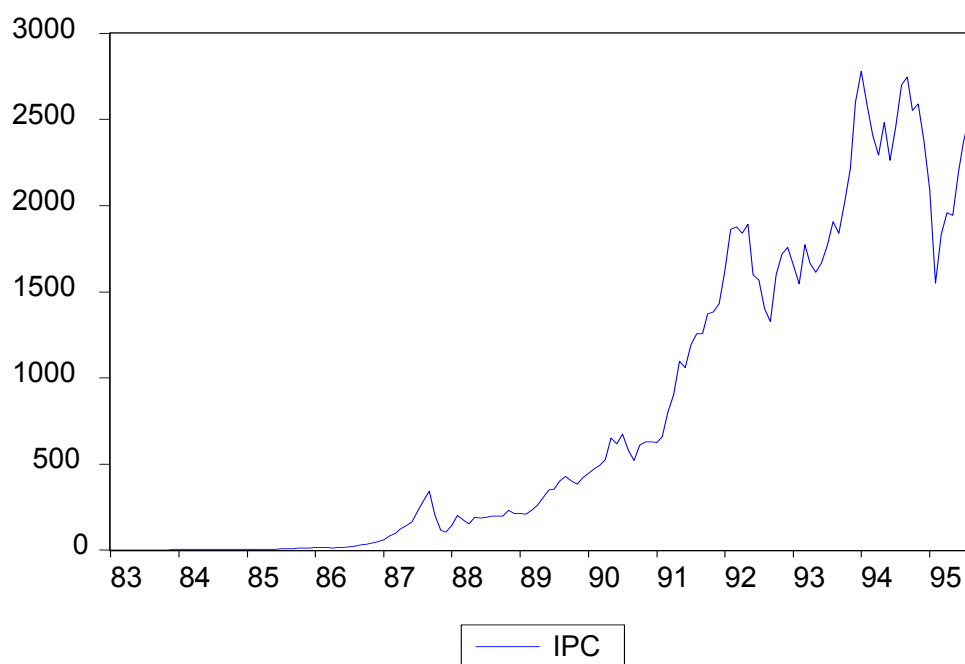
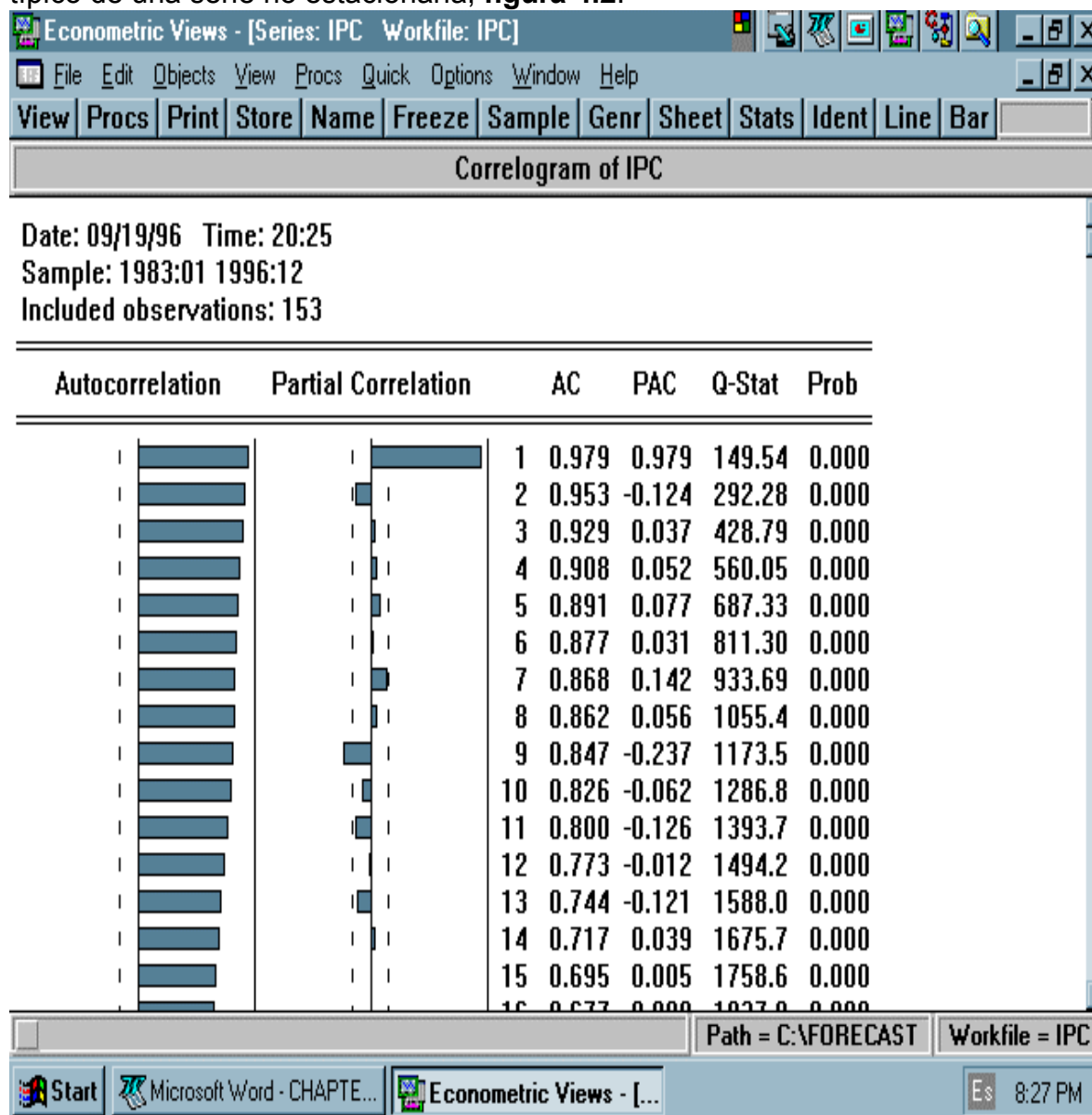


Figura 4.1.: Gráfica del I.P.C. de la B.M.V. Datos mensuales 1983 - 1995.

- a) Los datos reflejan tendencia creciente, aproximadamente cuadrática.
- b) La media de los datos no es constante, y parece estar en función del tiempo.
- c) La varianza no es única y también parece depender del tiempo.
- d) En consecuencia el proceso estocástico representado por esta variable no es estacionario.

Al observar su correlograma se tiene la certeza de una alto grado de **no** estacionariedad, debido a que su Función de Autocorrelación no desaparece rápidamente, es decir, las salientes del FAC tienden hacia el valor de cero muy

lentamente, además de que la probabilidad asignada para los valores del estadístico Q son iguales a cero (menor a 0.05), lo cual implica que dichos valores son significativos, por lo que se tiene la certeza de estar ante un caso típico de una serie no estacionaria, **figura 4.2**.



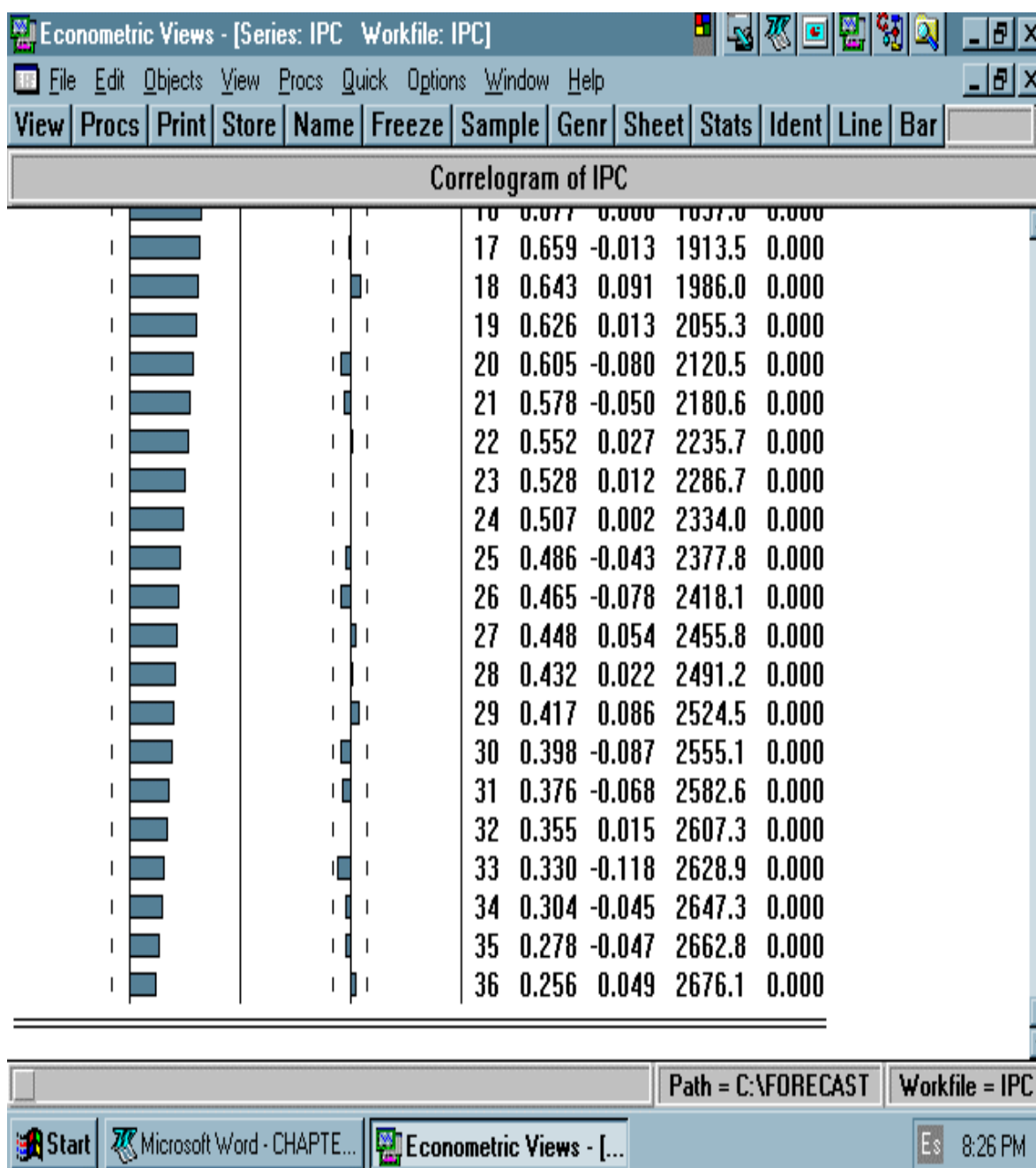


Figura 4.2.: Correlograma del I.P.C. de la B.M.V. con 36 rezagos.

Como una manera de estabilizar la media y la varianza se recurre a obtener una nueva serie a partir del logaritmo natural de la serie original, la cual se interpreta como la tasa de crecimiento de los valores del I.P.C. Dicha serie es denominada LIPC, y muestra una varianza notablemente estabilizada, aún cuando se conserva una media variable y cierta tendencia, la serie LIPC es mucho más “suave” que la original, **figura 4.3**.

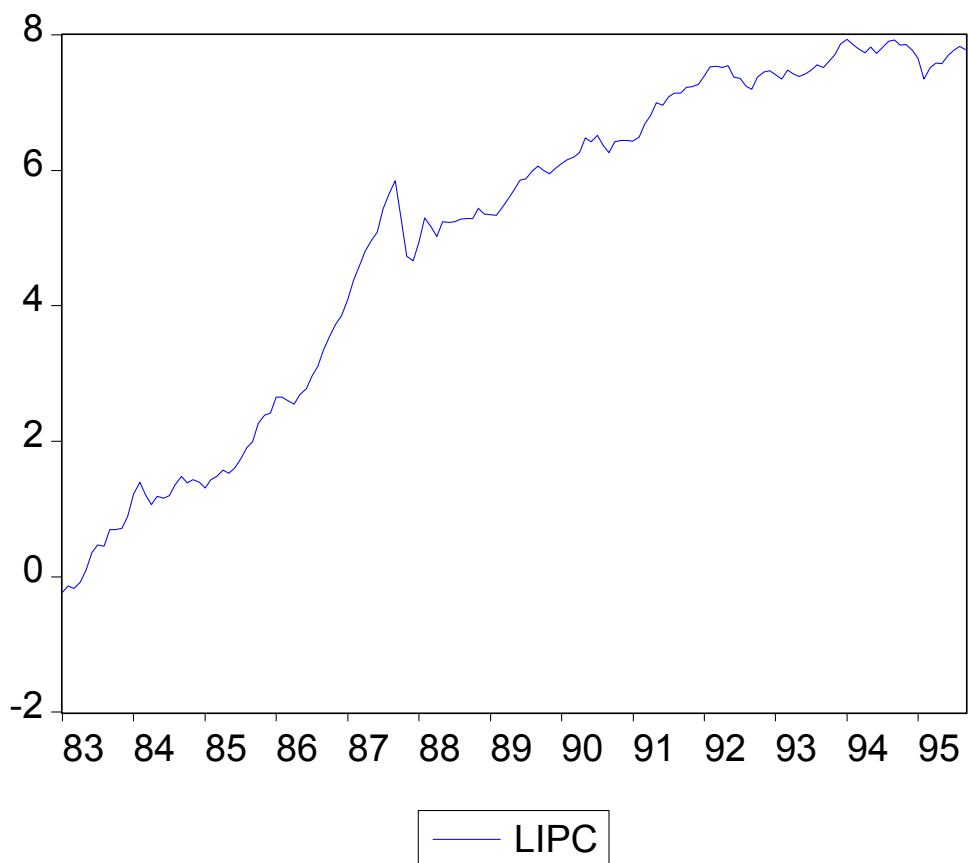


Figura 4.3.: Gráfica del logaritmo natural del I.P.C.

Al obtener su correlograma se observa que aún persiste un comportamiento no estacionario, claramente representado por la FAC que continúa cayendo lentamente al mismo tiempo que el estadístico Q tiene una probabilidad asociada de cero (menor a 0.05), para todos sus valores y por lo tanto significativos; concluyendo en éste caso que no corresponde a “ruidos blancos” y que aún falta trabajar la serie antes de proponer un modelo, **figura 4.4.**

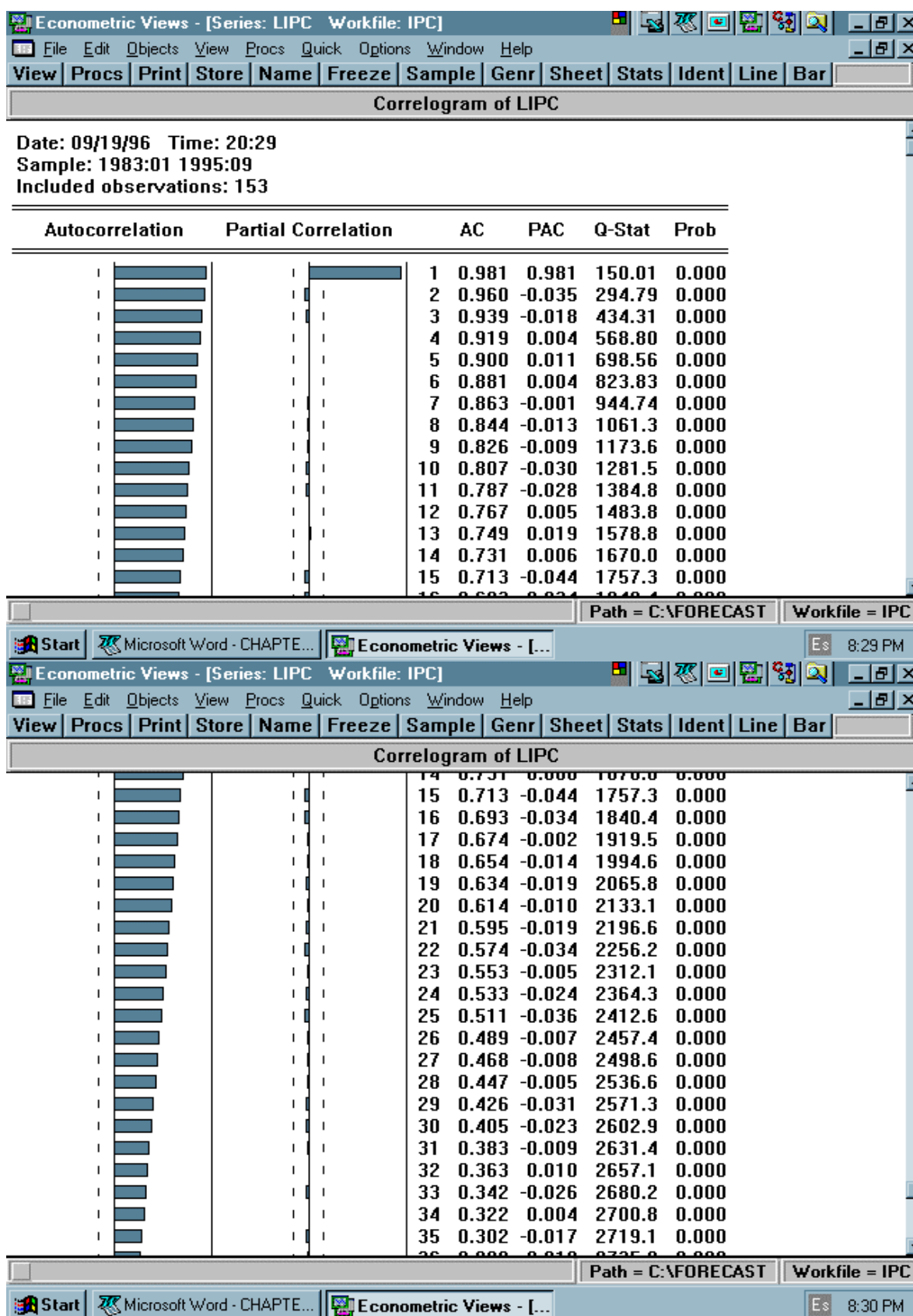


Figura 4.4.: Correlograma de la variable logaritmo natural del I.P.C. (LIPC) con 36 rezagos.

A continuación y con el objeto de eliminar la tendencia de la serie y estabilizar la media se procede a obtener primeras diferencias de la serie LIPC asignándolas a una nueva serie llamada DLIPC que contiene los valores del LIPC menos LIPC rezagada un período. La gráfica de esta nueva serie presenta un comportamiento estacionario con una media tendiente a cero y una varianza “constante”, en términos generales sin importar el tiempo t en el que se observe, es decir, un cierto número de observaciones contiguas pueden alejarse de la media en cierto período, sin embargo, siempre regresan a una vecindad de la misma. Lo mismo ocurre con cualquier número de observaciones contiguas k períodos adelante o hacia atrás del considerado inicialmente, **figura 4.5**.

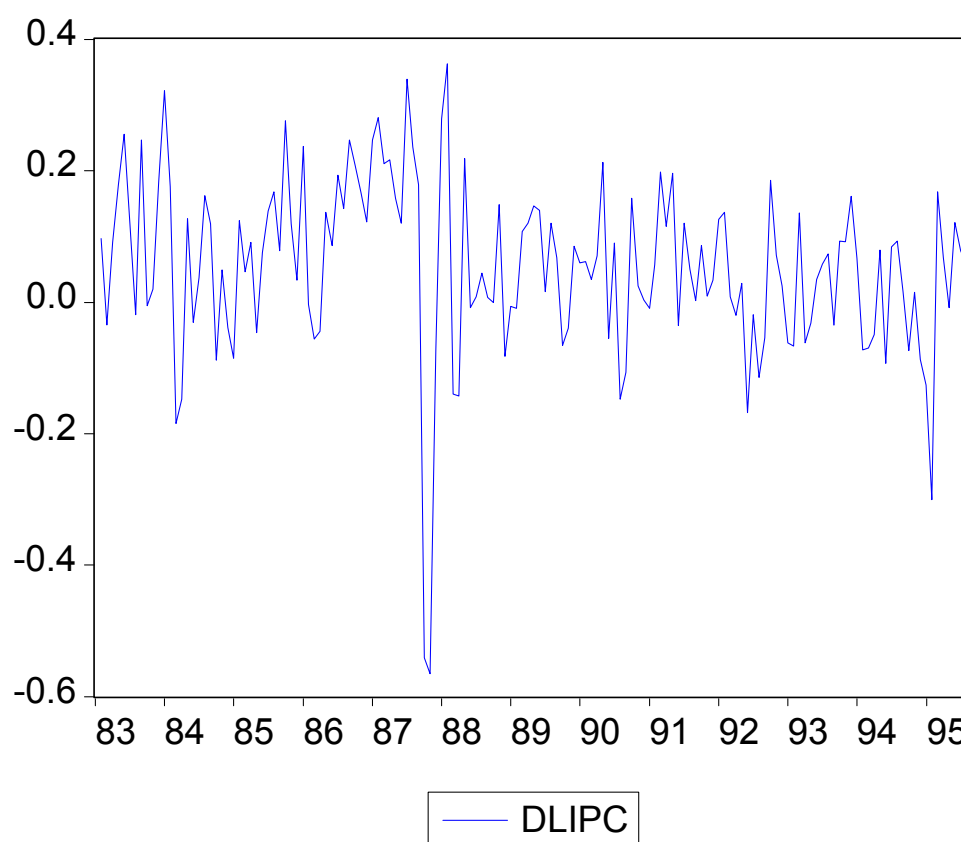


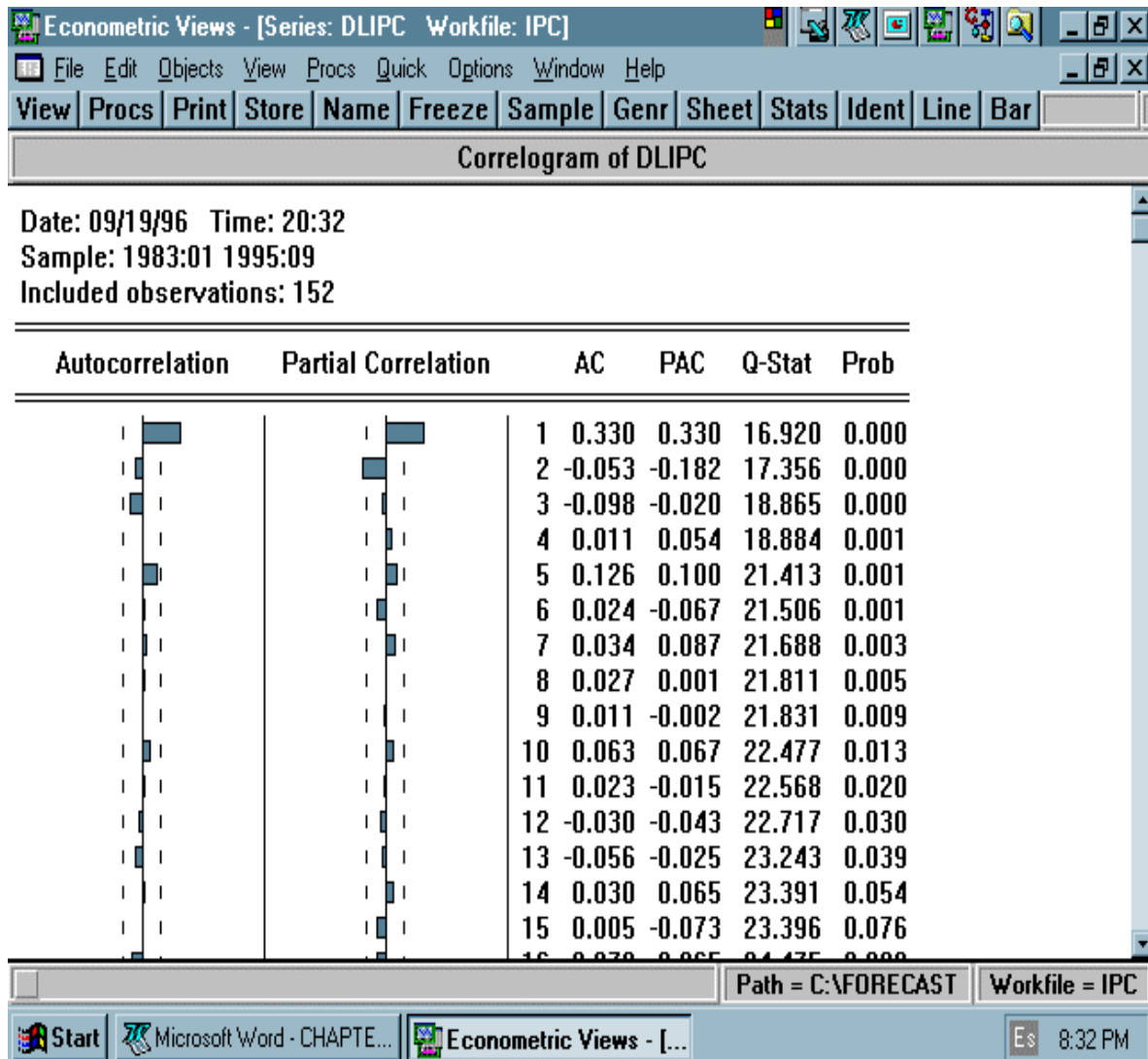
Figura 4.5.: Gráfica de las primeras diferencias de la variable LIPC, (DLIPC).

Un detalle importante en ésta última gráfica es la detección de una única observación que no corresponde al comportamiento descrito anteriormente, y que se encuentra en el mes de septiembre de 1987, dicha observación constituye una intervención o una “aberración “ en los datos, como muestran los datos anteriores y posteriores:

agosto 1987	287.40
septiembre 1987	343.54
octubre 1987	200.02

La explicación a éste dato particularmente anormal se puede atribuir al crack mundial de las Bolsas de Valores Internacionales, y que afectó fuertemente a la Bolsa Mexicana de Valores a final de 1987, debido a que México es un mercado emergente que está subordinado a economías más fuertes.

No obstante esta particularidad en los datos, el correlograma obtenido muestra una FAC con sólo una pequeña saliente fuera de las bandas de aceptación, en el caso de la Función de Autocorrelación Parcial (FACP), se observan dos salientes pequeñas de signos contrarios. En adelante, todas las salientes están contenidas dentro de las bandas de aceptación, **figura 4.6**.



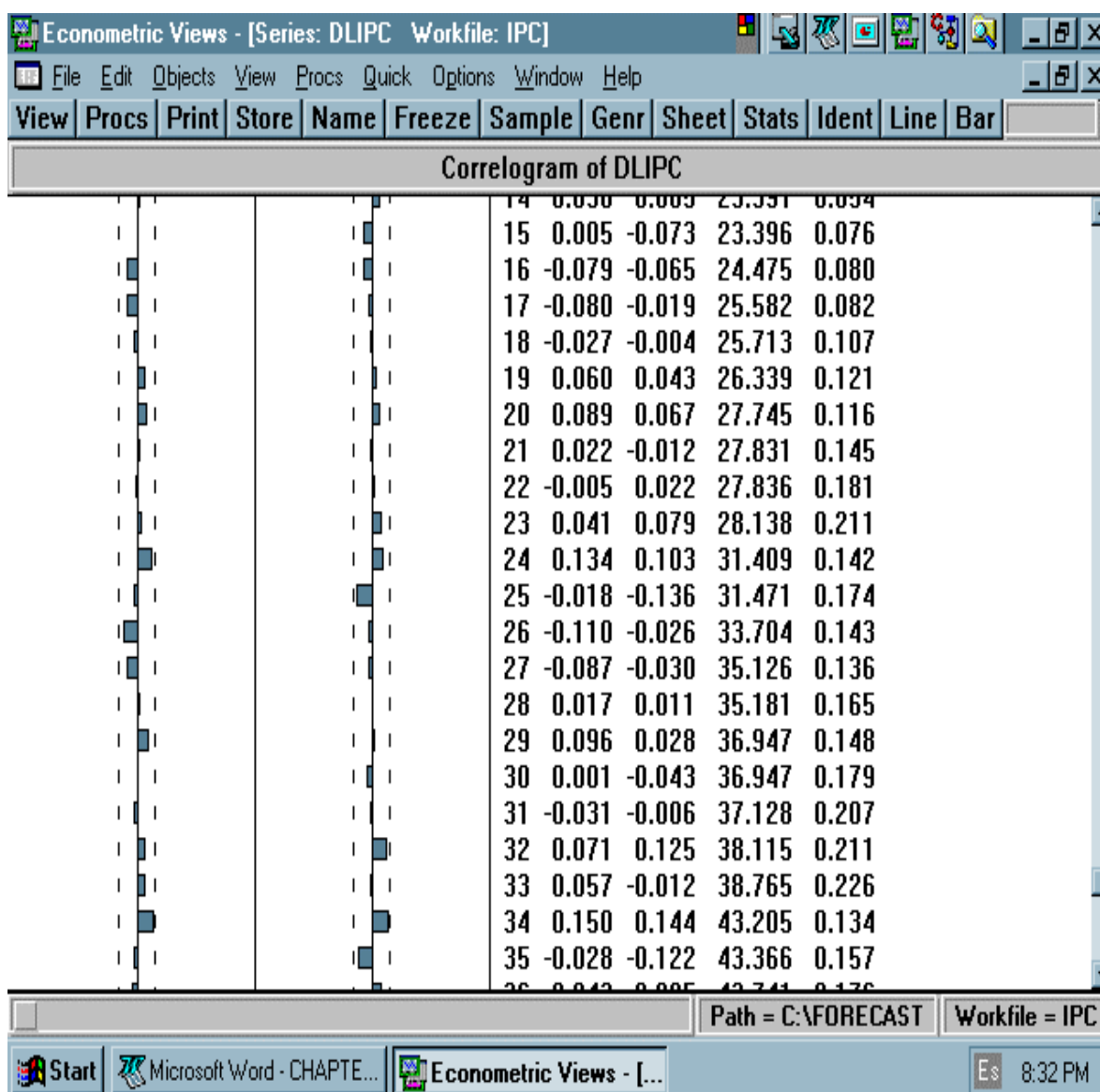


Figura 4.6.: Correlograma de la variable DLIPC con 36 rezagos.

Por otro lado, la probabilidad asociada a los valores de la columna del estadístico Q, es significativa durante los primeros trece rezagos y posteriormente es mayor a 0.05, lo cual puede atribuirse a “ruidos blancos”.

Sin embargo, al haberse advertido al inicio del análisis una tendencia cuadrática, es importante verificar lo que ocurre con la serie aplicando una segunda diferencia, lo cual se observa al obtener los valores de DLIPC menos DLIPC rezagada un período y asignarlos a una serie denominada D2LIPC; ésta última gráficamente tiene una mejor apariencia que la de primeras diferencias (se acentúa el comportamiento de la media cero y la varianza constante y prácticamente se ha eliminado la tendencia), no obstante que ya se han perdido dos observaciones de la serie original en esta segunda diferenciación, **figura 4.7**.

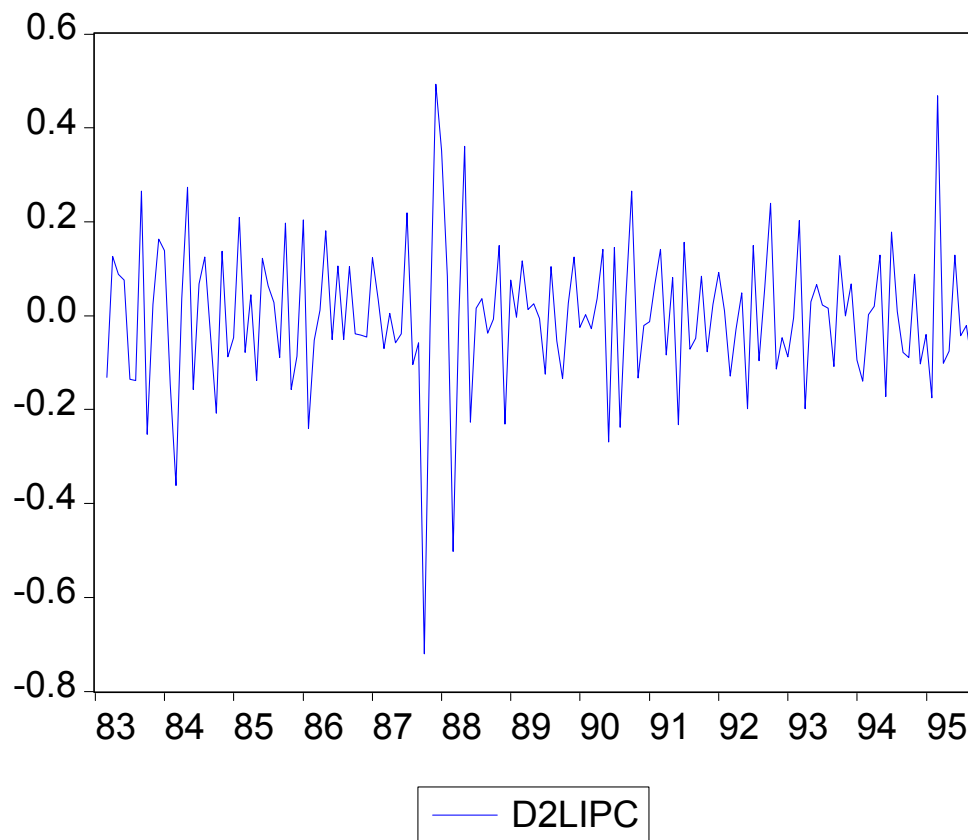


Figura 4.7.: Gráfica de las segundas diferencias de la variable LIPC, (D2LIPC).

Al obtener el correlograma se observa que la FAC se ha deformado presentando cinco salientes importantes y la FACP muestra también cinco salientes donde antes no había. Asimismo, los valores del estadístico Q son significativos durante los primeros dieciocho rezagos y en adelante sólo se presenta el fenómeno para ciertos valores particulares, por lo que no se pueden atribuir a “ruido blanco” y en consecuencia se ha sobrediferenciado la serie, lo que en principio acarrea problemas en la identificación de algún modelo para representarla, **figura 4.8**.

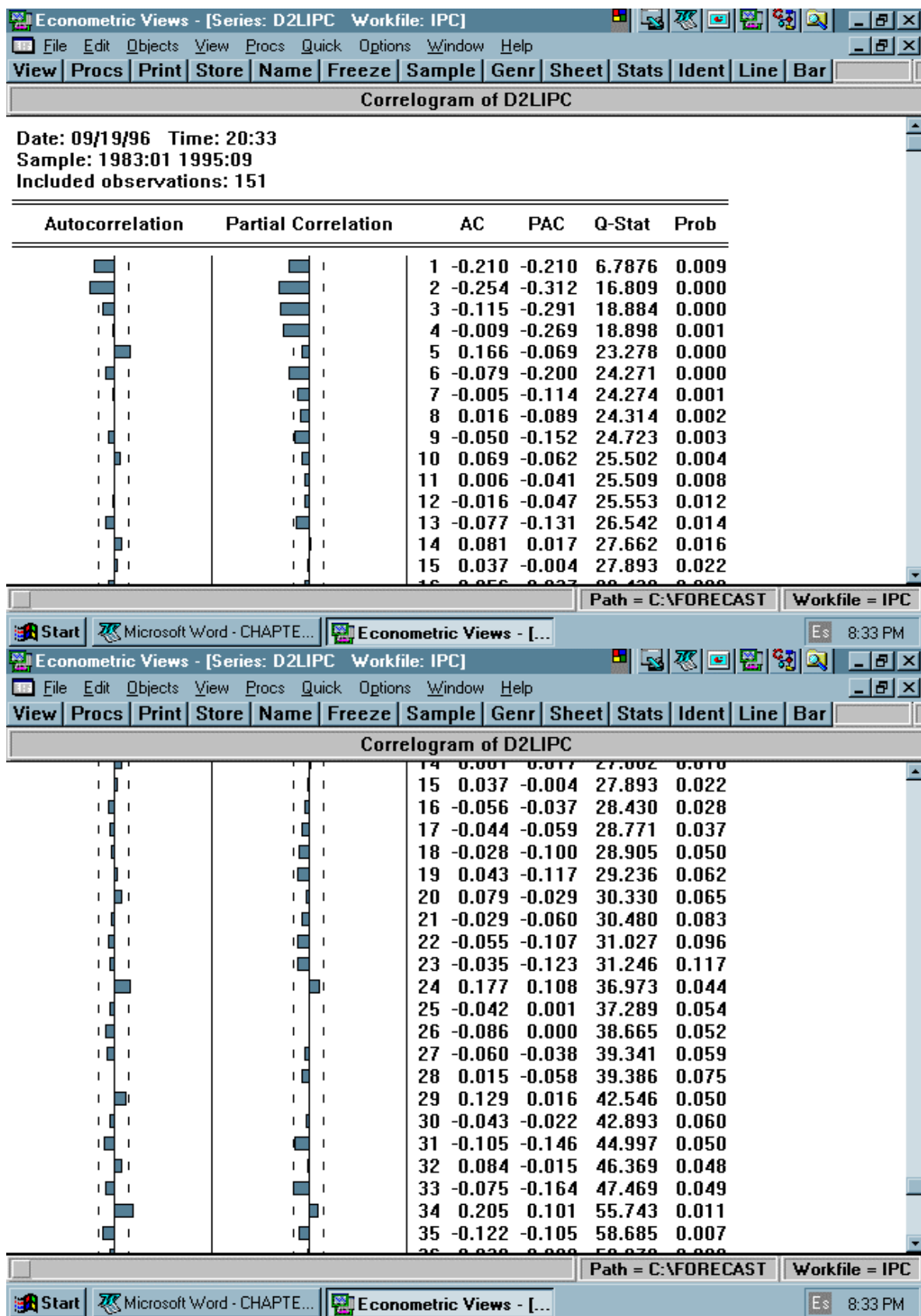


Figura 4.8.: Correlograma de la variable D2LIPC con 36 rezagos.

Con la evidencia presentada hasta ahora se puede afirmar que la serie es **estacionaria** con sólo obtener primeras diferencias, sin embargo, a fin de corroborarlo se efectúa la prueba Aumentada de Dickey - Fuller sobre la existencia de raíces unitarias en la serie.

Prueba aumentada de Dickey - Fuller

La prueba sobre la no estacionariedad de las series, es decir, la existencia de raíces unitarias, puede llevarse a cabo a través de la prueba Aumentada de Dickey - Fuller, la cual consiste en correr una regresión con las primeras diferencias de la serie contra la misma serie rezagada un período y opcionalmente una constante o un valor de tendencia en el tiempo.

Con dos diferencias rezagadas, la regresión es:

$$\Delta y_t = \beta_1 y_{t-1} - \beta_2 \Delta y_{t-1} - \beta_3 \Delta y_{t-2} - \beta_4 - \beta_5 t$$

Existen tres opciones para correr la prueba Aumentada de Dickey - Fuller. La primera es incluir una constante en la regresión. La segunda es incluir una línea de tendencia, y la tercera es un número de diferencias rezagadas que deben ser incluidas en la regresión.

En cada caso la prueba de raíces unitarias es una prueba sobre el coeficiente de y_{t-1} en la regresión. El resultado de la prueba consiste en un estadístico t sobre el coeficiente de la variable rezagada probada y los valores críticos para la hipótesis nula de que el coeficiente es cero. Si el coeficiente es significativamente diferente de cero entonces la hipótesis de que y_t contiene una raíz unitaria es rechazada y se acepta la hipótesis de que la serie es estacionaria.

El valor más negativo del estadístico t rechaza la hipótesis de raíz unitaria y sugiere que la serie es estacionaria. Bajo la hipótesis nula de raíz unitaria, el estadístico t no se distribuye como una t estándar, por lo que se tiene que recurrir a los valores críticos que resultan en la prueba, obtenidos mediante la aproximación de MacKinnon ("Critical Values of Cointegration Test"). Si el estadístico de t de Dickey - Fuller es menor en valor absoluto que el valor reportado en los valores críticos, se acepta la hipótesis de no estacionariedad de la serie y la existencia de una raíz unitaria.

En la **figura 4.9** se presenta el resultado de la prueba de Aumentada de Dickey - Fuller, donde se observa que el valor del estadístico t es mayor en valor absoluto a cualquiera de los valores críticos de MacKinnon para el 1%, 5% y 10% de confianza.

ADF Test Statistic	-7.294075	1% Critical Value*	-2.5793
		5% Critical Value	-1.942
		10% Critical Value	-1.6168
*MacKinnon critical values for rejection of hypothesis of a unit root.			
Augmented Dickey-Fuller Test Equation			
LS // Dependent Variable is D(DLIPC)			
Date: 10/31/96 Time: 20:18			
Sample(adjusted): 1983:04 1995:09			
Included observations: 150 after adjusting endpoints			

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLIPC(-1)	-0.640476	0.087808	-7.294075	0.000000
D(DLIPC(-1))	0.109092	0.081630	1.336424	0.183500
R-squared	0.297107	Mean dependent var	-0.000105	
Adjusted R-squared	0.292358	S.D. dependent var	0.156401	
S.E. of regression	0.131567	Akaike info criterion	-4.043231	
Sum squared resid	2.561870	Schwarz criterion	-4.003089	
Log likelihood	92.401550	F-statistic	62.558400	
Durbin-Watson stat	1.971542	Prob(F-statistic)	0.000000	

Figura 4.9.: Resultados de la prueba Aumentada de Dickey Fuller.

De ésta última prueba se concluye que la serie es estacionaria con primeras diferencias y es la adecuada para identificar el modelo y proceder a la estimación de sus parámetros (elaboración de regresiones); por lo que la serie escogida es la denominada DLIPC.

Conviene aclarar, que existe la posibilidad de incluir una variable dicotómica (Dommy), para simular la intervención ocasionada por el crack bursátil de septiembre de 1987, generando la variable $I1 = 1$ en los meses donde se ubica la intervención y cero en cualquier otro período, efectuando las regresiones correspondientes.

A partir del correlograma de la serie DLIPC se detectan los siguientes modelos:

Modelo 1:	MA(1)
Modelo 2:	AR(1) AR(2)
Modelo 3:	AR(1)
Modelo 4:	MA(1) AR(2)

Éste último corresponde a un ARIMA(2,1,1) donde el 2 denota un AR(2), el 1 representa las primeras diferencias de la serie y el tercer número denota una MA(1).

Estimación de los parámetros

Usando el cálculo de regresiones para estimar los parámetros más adecuados al modelo, se obtienen siempre estimadores insesgados para DLIPC, dado que se basa en el procedimiento de mínimos cuadrados ordinarios, que a su vez es consistente con los estimadores obtenidos por el método de máxima verosimilitud.

En todos los casos la variable dependiente es DLIPC y las variables independientes son las que incluya el modelo de que se trate, además de considerar la constante **C** en todas las regresiones.

Los resultados obtenidos se muestran en los cuadros siguientes:

RESULTADOS PROPORCIONADOS POR LA REGRESIÓN PARA LOS MODELOS PROPUESTOS
VARIABLE DEPENDIENTE: DLIPC

		COEFICIENTES	SUMA ERRORES DE LA REGRESIÓN	DESVIACIÓN ESTÁNDAR DE DLIPC	CRITERIO DE AKAIKE	CRITERIO DE SCHWARTZ
		SIGNIFICATIVOS				
MODELO 1:	MA(1)	SI	0.125914	0.134800	(4.131235)	(4.091447)
MODELO 2:	AR(1) AR(2)	SI	0.126397	0.135463	(4.116860)	(4.056648)
MODELO 3:	AR(1)	SI	0.128000	0.135201	(4.098288)	(4.058324)
MODELO 4:	MA(1) AR(2)	NO	0.126775	0.135463	(4.110880)	(4.050667)

		FUNCIÓN DE AUTOCORRELACIÓN	FUNCIÓN DE AUTOCORRELACIÓN
			PARCIAL
MODELO 1:	MA(1)	Todas las salientes dentro de las bandas de aceptación	Todas las salientes dentro de las bandas de aceptación
MODELO 2:	AR(1) AR(2)	Todas las salientes dentro de las bandas de aceptación	Todas las salientes dentro de las bandas de aceptación
MODELO 3:	AR(1)	Todas las salientes dentro de las bandas de aceptación	Todas las salientes dentro de las bandas de aceptación
MODELO 4:	MA(1) AR(2)	Todas las salientes dentro de las bandas de aceptación	Todas las salientes dentro de las bandas de aceptación

		PROBABILIDAD DEL ESTADÍSTICO Q	HIPÓTESIS DE RUIDO BLANCO
MODELO 1:	MA(1)	Mayor que 0.05 en todos los rezagos	Se presume alteraciones de FAC y FACP ocasionadas por "ruido blanco"
MODELO 2:	AR(1) AR(2)	Mayor que 0.05 en todos los rezagos	Se presume alteraciones de FAC y FACP ocasionadas por "ruido blanco"
MODELO 3:	AR(1)	Mayor que 0.05 a partir del cuarto rezago	Se presume alteraciones de FAC y FACP ocasionadas por "ruido blanco"
MODELO 4:	MA(1) AR(2)	Mayor que 0.05 en todos los rezagos	Se presume alteraciones de FAC y FACP ocasionadas por "ruido blanco"

Verificación del modelo

Los criterios para determinar la optimalidad de los modelos propuestos están ya expresados de alguna forma en los cuadros anteriores. Baste mencionar que uno de los más importantes criterios para elegir un modelo es comparar la suma de los errores de la regresión y decidirse por el que proporcione el menor valor posible.

En caso de mantenerse la duda se deben considerar los siguientes criterios:

- a) Los coeficientes de la regresión para las variables independientes deben ser significativos.
- b) La desviación estándar de la variable dependiente debe ser la menor posible.
- c) El valor obtenido para el criterio de Akaike debe ser el menor posible.
- d) El valor obtenido para el criterio de Schwartz debe ser el menor posible.
- e) La Función de Autocorrelación de la regresión debe mostrar todas las salientes dentro de las bandas de aceptación.
- f) La Función de Autocorrelación Parcial de la regresión debe mostrar todas las salientes dentro de las bandas de aceptación.
- g) La probabilidad asignada al estadístico Q debe ser mayor a 0.05 (para un 95% de confianza), en el mayor número de rezagos posibles.
- h) Del punto anterior se infiere que se debe estar en posibilidad de suponer que las alteraciones observadas en el correlograma son resultado de “ruidos blancos”.

Después de comparar los modelos propuestos se concluye que el que mejor se ajusta a los criterios señalados es el Modelo 1 que sólo incluye C MA(1), **figura 4.10**.

LS // Dependent Variable is DLIPC				
Date: 06/14/96 Time: 18:36				
Sample: 1983:02 1995:09				
Included observations: 152 after adjusting endpoints				
Convergence achieved after 3 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.052553	0.014235	3.691725	0.000300
MA(1)	0.394711	0.075027	5.260941	0.000000
R-squared	0.133269	Mean dependent var		0.052735
Adjusted R-squ	0.127491	S.D. dependent var		0.134800
S.E. of regressi	0.125914	Akaike info criterion		-4.131235
Sum squared re	2.378167	Schwarz criterion		-4.091447
Log likelihood	100.295200	F-statistic		23.064060
Durbin-Watson	1.999100	Prob(F-statistic)		0.000004
Inverted MA Roots		-0.390000		

Figura 4.10.: Resultados de la regresión con C MA(1).

Finalmente, las gráficas de la serie real, ajustada y el mapeo de los residuos, presentan un comportamiento consistente con los argumentos presentados anteriormente, **figura 4.11**.

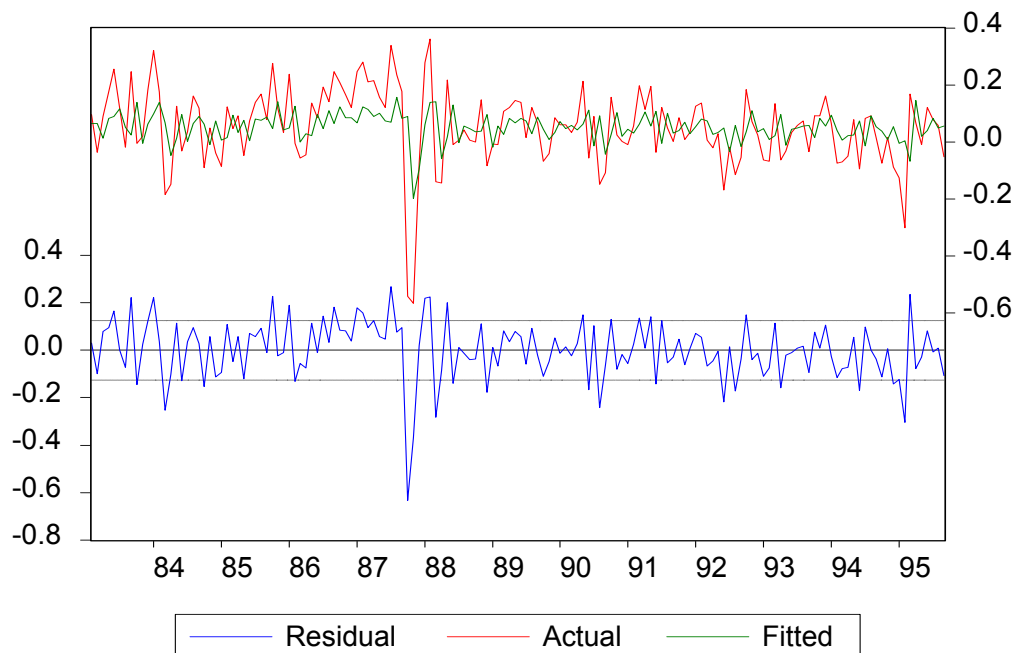


Figura 4.11.: Mapeo de los valores reales, estimados y los residuos de la regresión.

Una vez que ha sido detectado el modelo se procede a mejorarlo por medio del análisis de intervención, cuyo efecto se simula mediante la variable Dommy denominada I1, que contiene los valores iguales a 1 en los meses de agosto, septiembre y octubre de 1987 y cero en cualquier otro período. Recuérdese que en esta fecha fue cuando se dio el crack bursátil y se observó un shock en la serie. Al correr la regresión, el modelo final mejorado está representado por C MA(1) I1, cuyos parámetros son aceptables y el correlograma muestra un comportamiento de “ruido blanco”, sin embargo, la variable I1 resulta no significativa en la regresión por lo que finalmente se desecha ésta variable, **figura 4.12.**

LS // Dependent Variable is DLIPC				
Date: 06/19/96 Time: 18:11				
Sample: 1983:02 1995:09				
Included observations: 152 after adjusting endpoints				
Convergence achieved after 4 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.050230	0.014112	3.559263	0.000500
I1	0.177528	0.104722	1.695239	0.092100
MA(1)	0.385622	0.075713	5.093227	0.000000
R-squared	0.149608	Mean dependent var		0.052735
Adjusted R-squared	0.138194	S.D. dependent var		0.134800
S.E. of regression	0.125140	Akaike info criterion		-4.137109
Sum squared resid	2.333334	Schwarz criterion		-4.077427
Log likelihood	101.741600	F-statistic		13.106700
Durbin-Watson stat	1.976250	Prob(F-statistic)		0.000006
Inverted MA Roots	-0.39			

Figura 4.12.: Resultados de la regresión con C MA(1) I1.

El resultado anterior obliga a replantear la variable Dommy y ésta se modifica a I2 asumiendo el valor de 1 únicamente en septiembre de 1987 y cero en cualquier otro período. Se propone entonces un nuevo modelo C MA(1) I2, **figura 4.13.**

LS // Dependent Variable is DLIPC				
Date: 06/19/96 Time: 18:24				
Sample: 1983:02 1995:09				
Included observations: 152 after adjusting endpoints				
Convergence achieved after 4 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.050883	0.014403	3.532687	0.000500
I2	0.250268	0.112603	2.222557	0.027800
MA(1)	0.427984	0.074127	5.773654	0.000000
R-squared	0.159994	Mean dependent var		0.052735
Adjusted R-squared	0.148719	S.D. dependent var		0.134800
S.E. of regression	0.124373	Akaike info criterion		-4.149397
Sum squared resid	2.304837	Schwarz criterion		-4.089715
Log likelihood	102.675500	F-statistic		14.189860
Durbin-Watson stat	1.977488	Prob(F-statistic)		0.000002
Inverted MA Roots	-0.43			

Figura 4.13.: Resultados de la regresión con C MA(1) I2.

Al estimar los parámetros de éste nuevo modelo se encuentran los siguientes resultados:

- Los coeficientes estimados por la regresión para las variables independientes son significativos.
- La suma de los errores de la regresión es 0.124373.
- La desviación estándar de DLIPC es 0.1348.
- El criterio de Akaike es de -4.149397.
- El criterio de Schwartz es de -4.089715.
- La probabilidad del estadístico Q es mayor que 0.05 en todos los rezagos.
- La FAC presenta todas las salientes dentro de las bandas de aceptación.
- La FACP presenta igualmente todas las salientes dentro de las bandas de aceptación.

- i) En consecuencia se acepta la hipótesis de que las alteraciones de la FAC y de la FACP son ocasionadas por “ruido blanco”.
- j) La gráfica de la regresión denota un buen ajuste, aunque permanece la inestabilidad en la varianza al final de 1987.
- k) Una rápida observación al mapeo de los residuos de la regresión muestra que, la alteración se mantiene en los últimos seis meses de 1987.
- l) Las gráficas de los residuos de la regresión utilizando I1 e I2 muestra gran semejanza entre ellas, aún cuando la primera no es una variable significativa, **figura 4.14**.

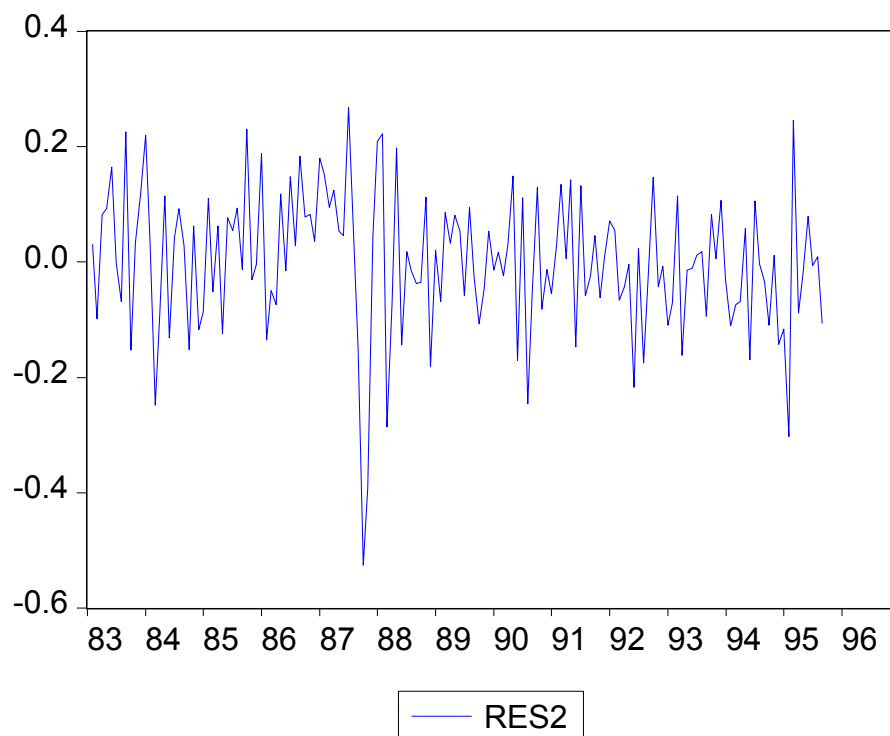
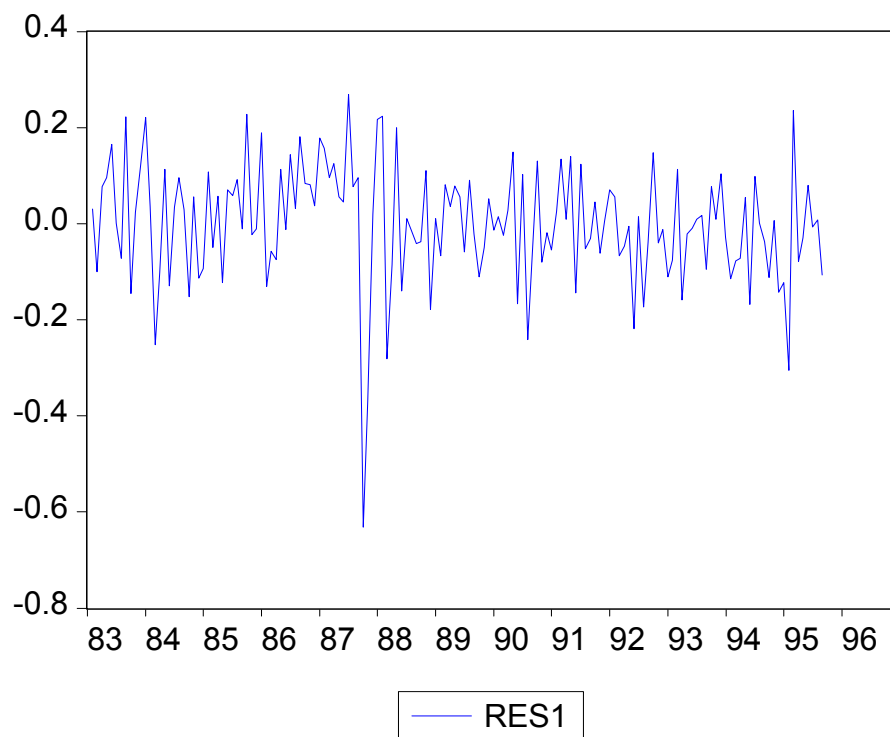


Figura 4.14.: Gráfica de los residuos de las regresiones de los modelos C MA(1) I1 y C MA(1) I2; RES1 y RES2, respectivamente.

En virtud de que el modelo MA(1) mejora cuando se le incluye la variable Dommy I2 entonces, siguiendo los criterios de aceptación de los modelos, se concluye que el modelo óptimo para simular el comportamiento del IPC de la BMV es un MA(1) I2.

Pronóstico del Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (IPC), para el período de octubre de 1995 a diciembre 1996

Una vez conocido el modelo que mejor representa el fenómeno se procede a elaborar un pronóstico por el período señalado asignándolo a la variable IPCF; encontrándose que para finales de 1995 estima 4,904.45 puntos para el IPC de la BMV, lo cual evidentemente no corresponde a la realidad y condiciones actuales, por lo que la decisión inmediata es recortar el período de pronóstico pues conforme se aleja, su eficiencia se ve notablemente disminuida, **figura 4.15**.

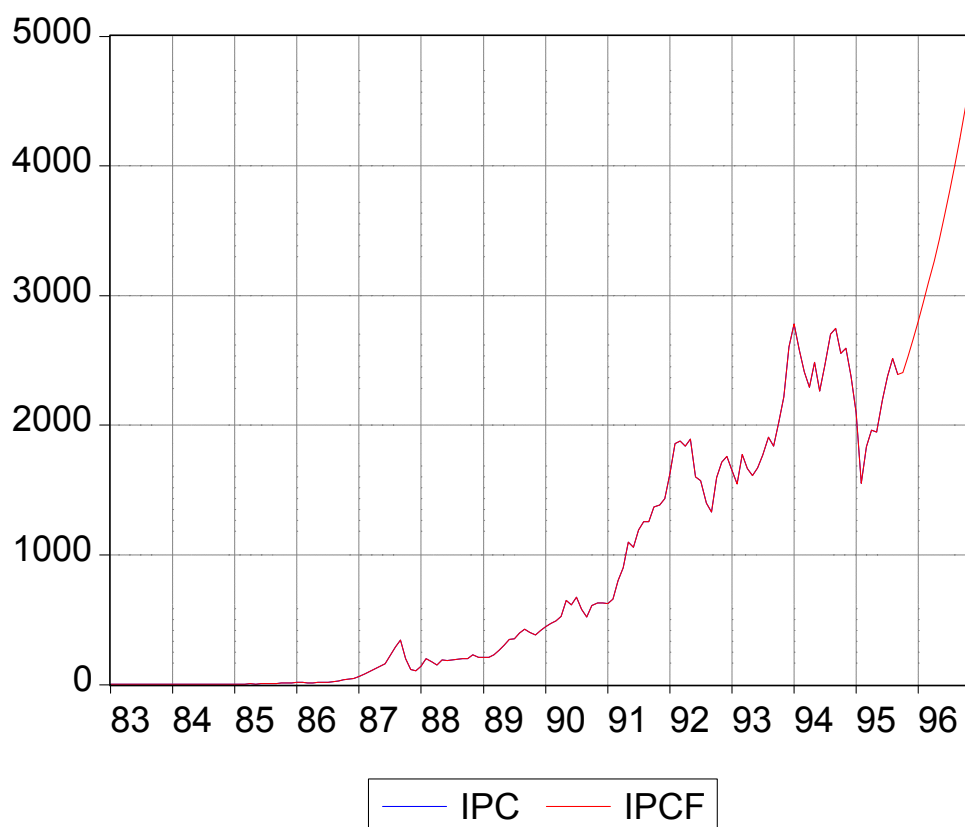


Figura 4.15.: Pronóstico del I.P.C. por el período de octubre 1995 - diciembre 1996.

Pronóstico del Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (IPC), para el período de octubre de 1995 a marzo 1996

Recortando el período a pronosticar se obtiene una buena estimación de 3,102.47 puntos al final de marzo lo cual se muestra en la variable IPCF1, lo que en efecto ocurrió al IPC, por lo que se deduce que el modelo propuesto es eficiente sólo en el pronóstico de períodos cortos y es necesario enriquecerlo con información nueva cada mes para volver a efectuar un pronóstico, **figura 4.16**.

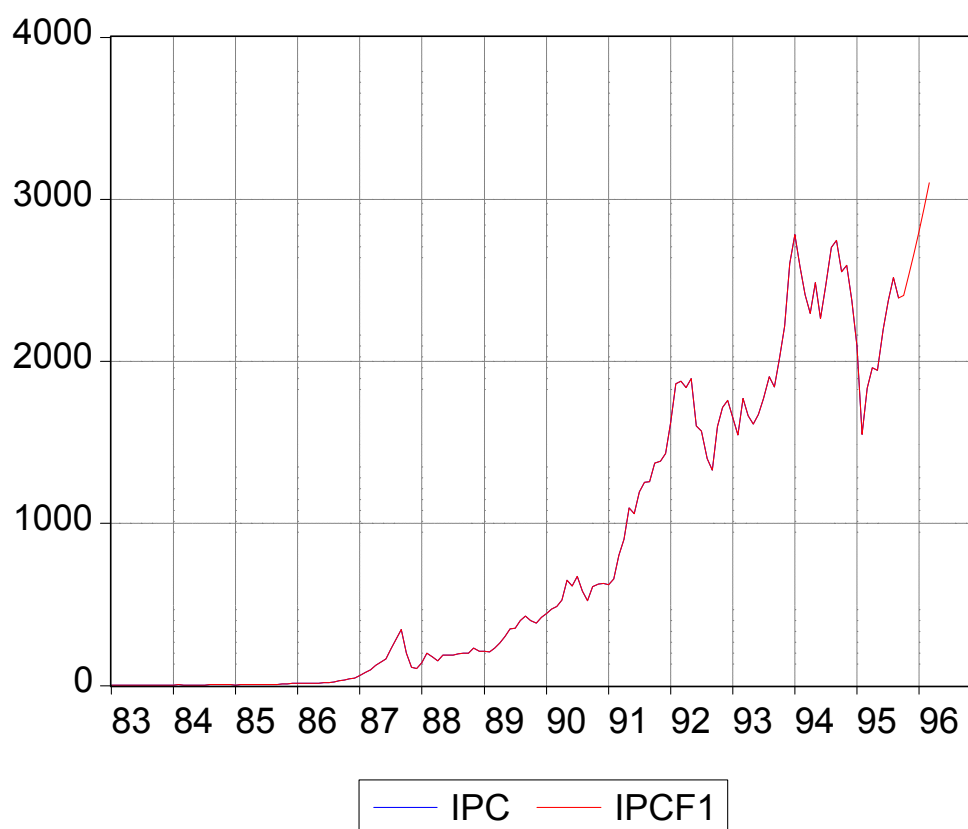


Figura 4.16.: Pronóstico del I.P.C. por el período de octubre 1995 - marzo 1996.

Esta afirmación se ve soportada por la comparación de los valores de la serie original, el pronóstico hasta diciembre de 1996, y el pronóstico hasta marzo de 1996 que son prácticamente iguales en el corto plazo pero divergen conforme el pronóstico se aleja en el tiempo, **figura 4.17**.

Last updated: 06/19/96 - 18:43

Modified: 1983:01 1996:12 // ipcf=exp(lipcf)

1983:01:00	0.79	0.87	0.84	0.92	1.10	1.42
1983:07:00	1.60	1.57	2.01	2.00	2.04	2.45
1984:01:00	3.38	4.03	3.35	2.89	3.28	3.18
1984:07:00	3.30	3.88	4.37	4.00	4.20	4.04
1985:01:00	3.71	4.20	4.40	4.82	4.60	4.96
1985:07:00	5.70	6.74	7.29	9.61	10.83	11.20
1986:01:00	14.20	14.15	13.38	12.80	14.68	16.00
1986:07:00	19.40	22.37	28.63	35.25	41.67	47.10
1987:01:00	60.28	79.82	98.53	122.30	143.31	161.67
1987:07:00	226.99	287.40	343.54	200.02	113.63	105.67
1988:01:00	139.62	200.59	174.40	151.16	188.07	186.57
1988:07:00	188.08	196.52	197.87	197.82	229.58	211.53
1989:01:00	210.21	208.31	232.02	261.73	302.94	348.49
1989:07:00	354.25	399.66	427.67	400.36	384.75	418.93
1990:01:00	444.75	473.02	489.62	525.61	650.29	615.33
1990:07:00	673.14	580.98	522.08	611.38	626.71	628.79
1991:01:00	622.99	659.17	803.35	901.14	1,096.17	1,058.02
1991:07:00	1,193.65	1,254.60	1,257.27	1,371.01	1,384.18	1,431.46
1992:01:00	1,623.47	1,860.63	1,875.73	1,838.30	1,892.33	1,599.26
1992:07:00	1,569.73	1,400.37	1,327.07	1,597.33	1,715.69	1,759.44
1993:01:00	1,653.22	1,546.68	1,771.71	1,665.41	1,612.99	1,670.29
1993:07:00	1,769.71	1,905.59	1,840.72	2,020.26	2,215.69	2,602.63
1994:01:00	2,781.37	2,585.44	2,410.38	2,294.10	2,483.73	2,262.58
1994:07:00	2,462.27	2,702.73	2,746.11	2,552.08	2,591.34	2,375.66
1995:01:00	2,093.98	1,549.84	1,832.83	1,960.54	1,945.13	2,196.08
1995:07:00	2,375.17	2,516.99	2,392.26	2,405.56	2,531.13	2,663.26
1996:01:00	2,802.28	2,948.55	3,102.47	3,264.42	NA	NA
1996:07:00	NA	NA	NA	NA	NA	NA

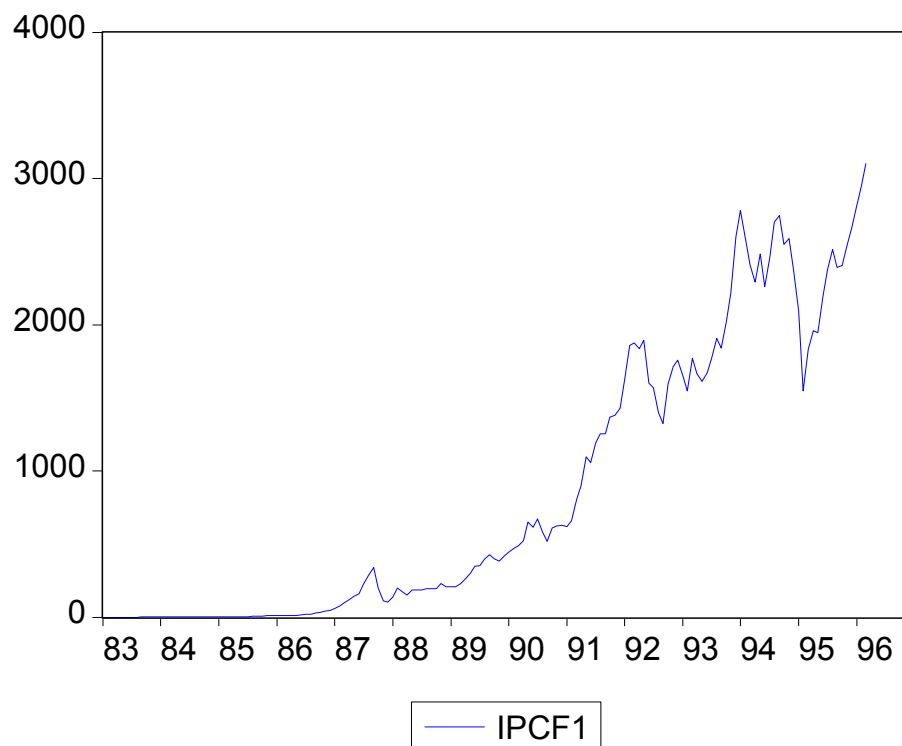
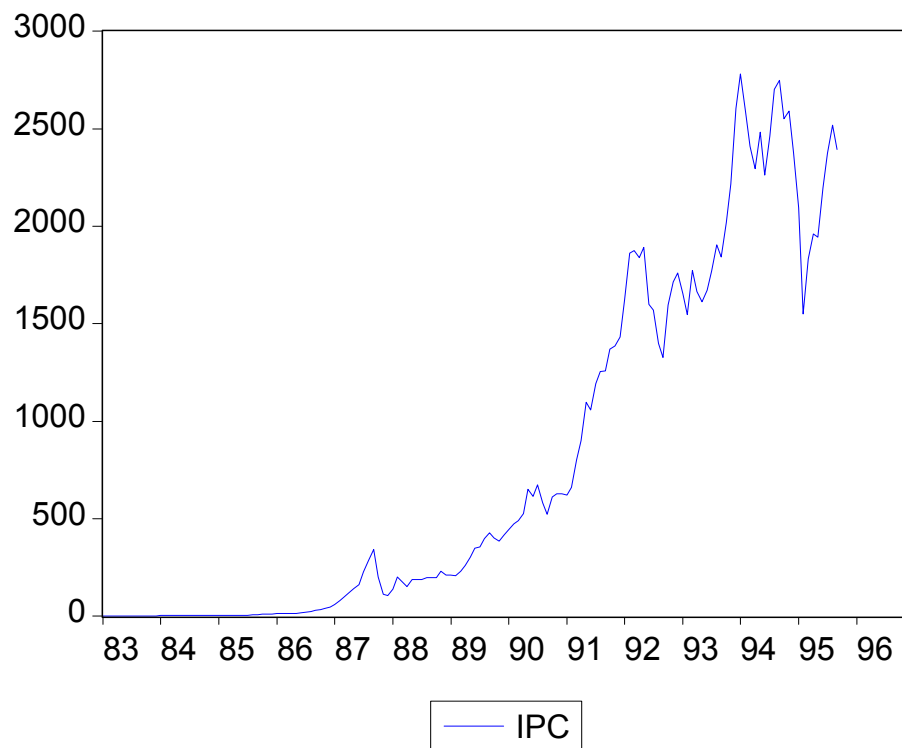


Figura 4.17.: Pronóstico del I.P.C. por el período de octubre 1995 - marzo 1996.

Conclusión

El modelo encontrado resultó ser un buen estimador para el corto plazo ya que el valor esperado de 3,102.47 puntos es muy aproximado al que realmente ocurrió al final de marzo de 1996 que fue de 3,072.40 puntos. Llama la atención que basta agregar un nuevo dato a la serie para estimar con mayor eficiencia una fecha futura cercana con una gran precisión, lo cual demuestra que el modelo de series de tiempo “recoge” la experiencia pasada y la incorpora a sí mismo para efectuar un pronóstico certero, siempre que se trate del corto plazo, aunque es posible que las nuevas observaciones sugieran replantear el modelo o elaborar uno totalmente nuevo.

Comparado con los modelos tradicionales de regresión para efectuar pronósticos, el modelo de series de tiempo no necesita ser “cargado” con valores supuestos de otras variables para obtener un pronóstico, lo que lo hace superior a esas herramientas generalmente usadas.

Cabe señalar que el uso del software estadístico facilita el proceso de análisis del modelo, ya que de otra manera los costos de obtener una simple regresión o las pruebas de hipótesis sobre los estimadores serían bastante considerables, sin embargo, a través del computador se pueden efectuar corridas una y otra vez en cuestión de segundos por lo que un modelo de series de tiempo como el visto en el capítulo IV puede ser rápidamente implementado y de muy alta calidad en el pronóstico en situaciones donde la alta dirección requiere opciones para tomar decisiones inmediatas.

La dificultad básica, - que el pasado nunca puede ser una perfecta guía para el futuro - nos advierte que el pronosticar no debe considerarse como una aplicación rutinaria de ciertas técnicas o ideas teóricas de una lista de variables no cambiantes. Un buen pronóstico requiere de una mezcla experta de la teoría económica, conocimientos estadísticos importantes y perfecta familiarización con los datos estadísticos relevantes.

Un aspecto adicional de la mayoría de las técnicas aplicadas a lo largo de este trabajo es el juicio personal del analista apoyado en el análisis estadístico, pues si se encuentra involucrado directamente en la actividad que va a pronosticar, es indudable que tendrá mayores elementos para considerar todos los factores que la influyen, lo cual es importante cuando se busca un buen pronóstico.

De vez en cuando es necesario hacer alguna pausa en el proceso de pronosticar y reconsiderar los procedimientos utilizados cuando se actualizan los datos históricos con nuevas observaciones, verificando si el modelo incrementa su eficiencia. Si es así, podrá seguir usándose hasta el siguiente período de actualización, de lo contrario conviene ver la posibilidad de aplicar una técnica diferente.

Lo anterior puede lograrse mediante el seguimiento de los errores que arroja el pronóstico hasta obtener criterios de comparación entre los errores obtenidos y los errores esperados para suponer si cierta técnica es o no adecuada.

Aún cuando sólo las grandes firmas empresariales o entidades gubernamentales cuentan con los recursos técnicos y humanos para elaborar pronósticos, quienes los efectúan se han convertido en equipos de staff de la alta dirección y asesores en el área de estadística con recursos de software cada vez más sofisticado que incrementa su capacidad, haciendo posible procesos que hubieran sido muy costosos hace diez o quince años.

Derivado de este progreso, el uso del pronóstico continúa ganando importancia en las organizaciones modernas debido al incremento en la complejidad de los negocios y al dinamismo de las actividades económicas, por lo que nuevas técnicas son desarrolladas para pronosticar series específicas.

Es claro que una constante actualización de los profesionales dedicados al pronóstico es indispensable para una mejor implementación de técnicas innovadoras, ya que de hecho no existe una técnica definitiva, sino que la mejor para los propósitos del analista es resultado de la aplicación una y otra vez de diversas técnicas y de la acumulación de experiencia.

BIBLIOGRAFÍA

John Hanke, (1992), **Bussines Forecasting**, Ed. Allyn and Bacon, E.U.

J. Holton Wilson, (1990), **Bussines Forecasting**, Ed. Irwin, E.U.

Damodar N. Gujarati, (1990), **Econometría**, Ed. Mc Graw Hill, México.

Terence C. Mills, (1994), **The Econometric Modelling of Financial Time Series**, Ed. Cambridge University Press, E.U.

Guerrero Víctor M., (1992), **Análisis Estadístico de Series de Tiempo Económicas**, Ed. Universidad Autónoma Metropolitana, México.

Dale G. Vails, (1990), **Bussines Fluctuation**, Ed. Prentice Hall, E.U.

Makridakis Whelwihtrs, (1990), **Forecasting Methods for Managers**, Ed. Willy, E.U.

Barry Keatins J., Halton Will, (1991), **Managerial Economics**, Ed. H.B.S., E.U.

Jae A. Shim and Joel G. Siegel, (1992), **Estrategic Bussines Forecasting**, Ed. Probus, E.U.

John Hanke, Arthur G. Rich, (1991), **Bussines Forecasting**, Ed. Prentice Hall, E.U.

Pindik and Robin Field, (1993), **Modelos Econométricos de Series de Tiempo**, Ed. Prentice Hall, México.

John E. Freud & Ronald E. Walpole, (1990), **Estadística Matemática con Aplicaciones**, Ed. Prentice Hall, México.

A. Novales, (1988), **Econometría**, Ed. Mc. Graw Hill, México.

Quantitative Micro Software QMS, (1995), **Econometric Views for Windows and Macintosh**, Ed. QMS, E.U..

Ya - Lun Chou, (1992), **Análisis Estadístico**, Ed. Mc. Graw Hill, México.