



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

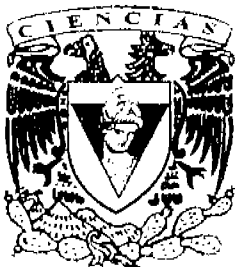
"MODELOS ECONÓMICOS APLICADOS
A LA AUDIENCIA DE LA TELEVISIÓN"

T E S I S

QUE PARA OBTENER EL TÍTULO DE
ACTUARIO.

P R E S E N T A :

DAVID HERRERA FERNÁNDEZ



FACULTAD DE CIENCIAS
UNAM

DIRECTOR DE TESIS: DR. LUIS ANTONIO RINCÓN SOLÍS

2005

17339865

ENCUEN
B
L



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



Autorizo a la Disposición General de Escuelas de la UCR a otorgar el título de licenciado en suceso al contenido de la tesis profesional.

NOMBRE: David Herrera Fernández

FEDATARIO: Enrique Rojas

FIRMA: [Firma]

ACT. MAURICIO AGUILAR GONZÁLEZ
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

"Modelos Econométricos Aplicados a la Audiencia de la Televisión"

realizado por David Herrera Fernández

con número de cuenta 9756936-1, quien cubrió los créditos de la carrera de:
Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis
Propietario

Dr. Luis Antonio Rincón Solís

Propietario

M. en C. Michael Rojas Romero

Propietario

M. en C. Eric Manuel Rodríguez Herrera

Suplente

Act. María Aurora Valdéz Michell

Suplente

M. en A.P. María del Pilar Alonso Reyes

Consejo Departamental de Matemáticas



Act. Jaime Vázquez

Facultad de Ciencias
Consejo Departamental
de
Matemáticas

Dedico este trabajo a mis padres, Franciso y Margarita, con amor y gratitud.

Agradezco al Dr. Luis Antonio Rincón por su orientación en el desarrollo de esta tesis; a los maestros Michael Rojas, Eric Manuel Rodríguez, María Aurora Valdéz y María del Pilar Alonso, por sus comentarios y sugerencias; a mi papá por su enorme ayuda en la gestión de trámites; a Sandra Cornejo por su apoyo; a Azbel García por el software facilitado; a Jorge Sánchez por su consejo y experiencia compartida.

Índice

1. Introducción	4
2. Encendido y <i>rating</i>	6
3. Análisis clásico de la serie encendido	12
3.1. Análisis de la tendencia	17
3.1.1. Método del ajuste analítico	20
3.1.2. Método de la media móvil	21
3.2. Análisis de la estacionalidad	25
3.2.1. Método de la razón a la media móvil	26
3.2.2. Método de la regresión	29
3.3. Análisis del ciclo	35
3.3.1. Método de los residuos	35
3.4. Pronóstico de la serie "PUT"	36
4. Técnicas extrapolativas de pronóstico	39
4.1. Conceptos fundamentales	40
4.1.1. Proceso estocástico	40

4.1.2.	Autocorrelación y correlograma	40
4.1.3.	Ruido blanco	41
4.1.4.	Diferenciación	42
4.2.	Suavizamiento exponencial de Holt-Winters multiplicativo . .	43
4.2.1.	Aplicación a la serie encendido	46
4.3.	Mínimos cuadrados con variables artificiales para estacionali- dad en $\log(W_t)$	52
4.3.1.	Aplicación a la serie encendido	54
4.4.	Mínimos cuadrados y funciones senoidales en $\log(W_t)$	58
4.4.1.	Aplicación a la serie encendido	60
4.5.	Promedios móviles simples en $\nabla \nabla_S \log(W_t)$	64
4.5.1.	Aplicación a la serie encendido	65
4.6.	La elección del modelo	67
5.	Procesos ARMA	70
5.1.	Conceptos fundamentales	70
5.1.1.	Proceso estacionario	70

5.1.2. Rezago	71
5.2. Procesos estacionarios de promedios móviles	72
5.2.1. Proceso de promedios móviles de primer orden	72
5.2.2. Proceso de promedios móviles de orden q	74
5.2.3. Proceso de promedios móviles de orden infinito	75
5.3. Procesos estacionarios autoregresivos	77
5.3.1. Proceso autoregresivo de primer orden	77
5.3.2. Proceso autoregresivo de segundo orden	79
5.3.3. Proceso autoregresivo de orden p	81
5.4. Procesos estacionarios mixtos autoregresivos y de promedios móviles	83
5.5. Pronóstico	85
5.5.1. Conceptos fundamentales	85
5.5.2. Pronóstico de un proceso ARMA(p,q)	87
5.5.3. Pronóstico con base en un número finito de observaciones	88
5.5.4. Procedimiento de Box-Jenkins	91
5.6. Aplicación la metodología ARMA a la serie encendido de TV .	94

1. Introducción

Esta tesis tiene sus antecedentes en una serie de experiencias académicas y profesionales, que tuvieron que ver con el descubrimiento de la aplicabilidad de la econometría en el campo de los medios de comunicación. Una aplicación en particular, que es precisamente el objetivo de esta tesis es la descripción y pronóstico de la serie Encendido de Televisión mediante el uso de ciertos modelos desarrollados para el estudio de las series de tiempo. A la fecha no existen antecedentes en México de la utilización de modelos econométricos para estudiar la serie Encendido de Televisión. La razón de esto es que la medición electrónica de ratings en México es relativamente reciente. Baste mencionar que existen datos a partir de octubre de 1996 y que hay un cambio estructural en los datos a partir de 1999, esto es, un cambio de dirección en la tendencia. Esto se revisará con mayor detalle en el desarrollo de este trabajo.

La hipótesis que se analiza en este trabajo es que el encendido de televisión puede ser descrito en función de su propia dinámica en el tiempo. Esto es relevante desde el punto de vista comercial, ya que la inversión publicitaria en México según cifras de la AMAP (2002), es de 2.5 miles de millones de dólares, de los cuales el 70 % va destinada a la Televisión, y la dinámica de las audiencias de los medios podría eventualmente cambiar este escenario a

largo plazo.

Este trabajo se ha dividido en cuatro partes: la primera trata de la definición de Encendido y de manera más general la de *Rating*, el universo poblacional dentro del cual se circunscribe esta variable. Además se describe el proceso de su estimación, sin ser el objetivo profundizar en la técnica de muestreo utilizada por IBOPE-AGB, empresa dedicada a la producción del *rating* en México. La segunda parte describe las técnicas determinísticas que pueden llegar a ser suficientes para describir, explicar y pronosticar una serie, o bien, servir como parte de un análisis previo al enfoque estocástico. Las partes tercera y cuarta tratan de la aplicación de modelos de series de tiempo univariadas al Encendido de TV. Para ello se utiliza primero un abanico de técnicas de pronóstico que son referidas como "extrapolativas" y finalmente la clase de modelos ARMA.

A manera de conclusión se ha comparado el desempeño de todos estos modelos para dar una recomendación en cuanto al modelo que parece ser adecuado para el análisis del *rating*. Este trabajo ha sido elaborado utilizando los programas Scientific Work Place 4.0, Excel 97 y RATS 4.2.

2. Encendido y *rating*

El *rating* es una medida elemental de audiencia de televisión. Es interesante observarlo, desde el punto de vista social, ya que permite descubrir las preferencias del auditorio y al ser la televisión el medio nacional masivo por excelencia, con una penetración de 84.18 % en viviendas particulares habitadas[3], da cuenta del perfil cultural de nuestra sociedad. Por otro lado está el punto de vista comercial. El *rating* es tanto una “moneda de cambio” en las negociaciones entre anunciantes o agencias de medios y televisoras, como una variable de evaluación - entre otras - de las campañas publicitarias. Como ejemplo, baste recordar que el esquema comercial de TV Azteca tiene como base el costo por punto de *rating*.

Definición 1 *Rating para tiempos discretos*: sea $T = \{t_1, t_2, \dots, t_m\}$ un período de tiempo que consta de m unidades de tiempo; $P = \{p_1, p_2, \dots, p_n\}$ un conjunto de n individuos u hogares y $C = \cup_{i \in I} C_i$ un conjunto de canales o sistemas de video en televisión, entonces

$$Rating(C, T, P) = \frac{1}{nm} \sum_{i=1}^m \sum_{j=1}^n R_{C,i,j},$$

donde

$$R_{C,i,j} = \begin{cases} 1 & \text{Si } p_j \text{ está expuesto a } C \text{ en el momento } t_i, \\ 0 & \text{En otro caso.} \end{cases}$$

Definición 2 *Rating para tiempo continuo:* sea T una unión de intervalos de tiempo; $P = \{p_1, p_2, \dots, p_n\}$ un conjunto de n individuos u hogares y $C = \cup_{i \in I} C_i$ un conjunto de canales o sistemas de video en televisión, entonces

$$\text{Rating}(C, T, P) = \frac{1}{n \|T\|} \sum_{j=1}^n \int_{i \in T} R_{C,i,j} di,$$

donde $R_{C,i,j} = \begin{cases} 1 & \text{Si } p_j \text{ está expuesto a } C \text{ en el momento } i \in T, \\ 0 & \text{En otro caso.} \end{cases}$

Por convención, el rating se expresa como porcentaje y al conjunto P se le llama Grupo Objetivo.

Definición 3 *Encendido:* Utilizando la misma notación de las definiciones anteriores, si C es la unión de todos los canales o sistemas de video en TV, entonces $\text{Rating}(C, T, P)$ se conoce como encendido, y se denotará como $HUT(T, P)$ cuando P sea un conjunto de hogares y $PUT(T, P)$ cuando P sea un conjunto de personas¹.

¹La notación HUT y PUT se deriva de los anglicismos "Households Using Television" y "People Using Television" respectivamente.

Ejemplo 4 *Caso con tiempo discreto: supóngase que la unidad mínima de tiempo es un minuto y que se desea calcular el rating del canal X donde $T=\{9:01, 9:02, \dots, 9:10\}$, para el grupo objetivo $P=\{p_1, p_2, p_3, p_4, p_5\}$ a partir de la siguiente Matriz R en cuyas celdas $r_{i,j}$ se encuentran los valores $R_{X,i,j}$*

$$R = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Entonces, $Rating(X, T, P) = \frac{1}{(5)(10)} \sum_{i=1}^{10} \sum_{j=1}^5 R_{X,i,j} = \frac{5}{50} = 10\%$.

Actualmente el rating es estimado por la empresa IBOPE AGB México. El universo de estudio o Universo IBOPE, está conformado por los hogares y sus miembros de tres o más años de edad, en un conjunto de 27 ciudades categorizadas por zona [4]:

- Area Metropolitana de la Ciudad de México.
- Zona noroeste: Tijuana, Mexicali, Culiacán y Hermosillo.
- Zona centro: Puebla, Toluca, Querétaro y Cuernavaca.
- Zona oeste: Guadalajara, León, Aguascalientes y Morelia.

- Zona noreste: Monterrey, Ciudad Juárez, Torreón, San Luis Potosí, Chihuahua y Saltillo.
- Zona sur: Mérida, Acapulco, Tuxtla Gutiérrez y Oaxaca.
- Zona Golfo: Tampico, Veracruz, Coatzacoalcos y Villahermosa.

Cabe mencionar que, de acuerdo al Establishment Survey 2000[4], este conjunto de ciudades contienen a 45.4 % de la población, 46.3 % de los hogares y 53 % de los hogares con TV a nivel nacional.

El proceso de recolección de datos, estimación de *ratings* y entrega de información a los usuarios consta de los siguientes pasos:

- *Establishment Survey*: consiste en un estudio para estimar los parámetros poblacionales en lo concerniente a la posesión de bienes en los hogares, y especialmente, toda aquella información relacionada con la televisión, como número de televisores en el hogar y sistemas de Televisión de paga contratados. A través de este estudio que contempla las 27 ciudades citadas anteriormente, se obtiene una descripción socioeconómica que refleja tanto los datos demográficos de la población, la vivienda, la familia y los integrantes del hogar, así como los aspectos

más relevantes de la posesión y uso de los televisores, la tecnología de estos aparatos, número de televisores por hogar, su distribución con base en diferentes criterios, los sistemas alternativos de TV - si los hay -, y el lugar físico donde se encuentran los televisores.

- Reclutamiento: Con base en los parámetros poblacionales obtenidos del *Establishment Survey* se levanta la muestra de hogares que constituye el panel.
- Instalación del *People Meter*, microcomputadora conectada a la televisión que registra el rating y tiene un sistema electrónico inalterable que almacena la información, y que se recolecta vía telefónica, ya sea conmutada o celular. Está habilitado para medir segundo a segundo el encendido de la televisión, la selección del canal y distinguir a la persona que está viendo la televisión, ya que cuenta con un control remoto cuya función es registrar la participación de los integrantes de la familia. Dicho instrumento de recolección sólo mide la audiencia de televisión en los hogares.
- Producción de los *ratings* y conversión de la información en bases de datos aptas para determinado software.

- Entrega de la información: diariamente se entregan bases de *ratings* vía Internet.

3. Análisis clásico de la serie encendido

El análisis clásico² sirve para estudiar series univariadas y determinísticas. Se considera que una serie se puede descomponer en: tendencia, ciclo, estacionalidad e irregularidad. La tendencia se refiere a un movimiento suave que indica la orientación de la variable a largo plazo. El ciclo es un movimiento oscilatorio ocasionado por la dinámica del sistema en estudio y está compuesto por las fases de prosperidad, recesión, depresión y recuperación. La estacionalidad se refiere a un patrón o movimiento regular de la serie dentro de un período de tiempo. Finalmente, la irregularidad corresponde a movimientos que carecen de un patrón sistemático o regular, debido a lo cual no es posible pronosticar su comportamiento. Los movimientos irregulares se clasifican como erráticos cuando su causa es identificable - huelgas, terremotos, eventos especiales, etc. - o aleatorios cuando no son atribuibles a algún fenómeno observable sino al azar.

En este capítulo se aplica el análisis clásico a la serie encendido (mensual, total de personas en el universo IBOPE) (figura 1) descrita en la sección

²Los conceptos, métodos y criterios expuestos en este capítulo, han sido recopilados del libro "Análisis de Series Temporales" de Carmen Rodríguez Morilla.

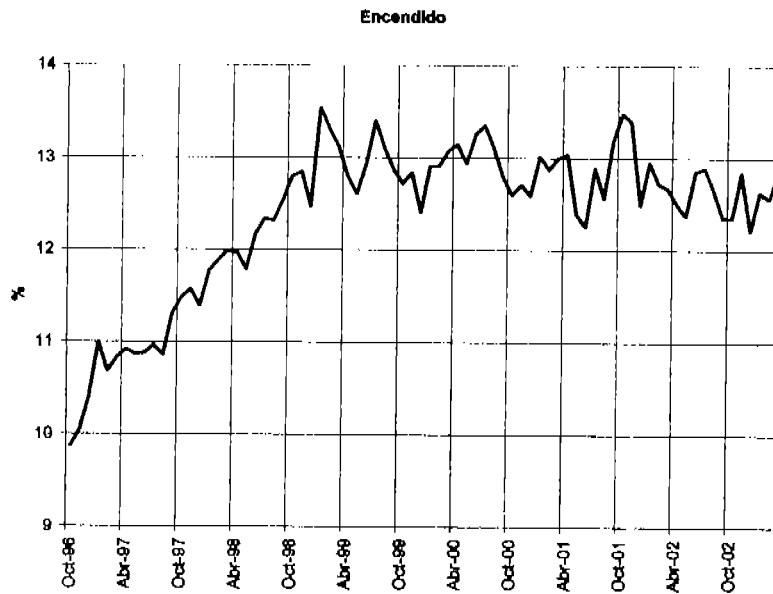


Figura 1: Encendido mensual

anterior, así como su interacción y las técnicas más utilizadas para calcularlos a partir de la serie original. Cabe mencionar que esta forma de descomponer las series sirve como herramienta para analizar, pronosticar y entender su comportamiento, y que no se ha llegado a una convención en cuanto a la definición formal de cada componente en el ámbito teórico de las series de tiempo.

Desde el punto de vista clásico, una serie de tiempo $Y = \{Y_t\}_{t=0}^T$ puede ser expresada como $Y_t = f(T_t, C_t, E_t, I_t)$, donde T_t, C_t, E_t y I_t simbolizan los

componentes tendencia, ciclo, estacionalidad e irregularidad respectivamente en el tiempo t . Conviene considerar a f como una función sencilla y operativa, por lo cual generalmente se trabaja con alguno de los siguientes esquemas:

Esquema aditivo:
$$Y_t = T_t + C_t + E_t + I_t$$

Esquema multiplicativo:
$$Y_t = T_t \times C_t \times E_t \times I_t$$

Conviene señalar que estando en un esquema multiplicativo puede aplicarse logaritmo natural en ambos miembros de la ecuación, para obtener una serie transformada bajo un esquema aditivo y de manera inversa usando la función exponencial:

$$Y_t = T_t \times C_t \times E_t \times I_t \rightarrow \ln Y_t = \ln T_t + \ln C_t + \ln E_t + \ln I_t.$$

Para elegir el modelo aditivo o multiplicativo, los criterios o métodos más utilizados son: a) la representación gráfica de la serie, b) el análisis gráfico de la desviación estándar contra la media y c) el método de diferencias y cocientes estacionales.

La representación gráfica de la serie sirve como un análisis previo para detectar si las componentes se asocian de manera aditiva o multiplicativa.

Es un indicio de aditividad cuando la amplitud de las oscilaciones es más o menos constante a lo largo de la serie. Cuando la amplitud crece en los niveles altos y decrece en los niveles bajos de la serie, da mejores resultados utilizar el esquema multiplicativo(ver figura 1). Para justificar este criterio, considérese que el nivel de la serie está dado por los componentes tendencia y ciclo. En este caso, el cuadrado de las oscilaciones alrededor de la serie bajo un esquema aditivo se pueden expresar como:

$$Oscilación_t^2 = (Y_t - (T_t + C_t))^2 = (E_t + I_t)^2$$

en este caso la oscilación no depende del nivel, lo cual implica que la amplitud de ésta debe ser similar a lo largo de la serie independientemente del nivel local de la misma. Por otra parte, en el esquema multiplicativo:

$$Oscilación_t^2 = (Y_t - T_t C_t)^2 = (T_t C_t)^2 (E_t I_t - 1)^2$$

lo cual indica que la oscilación crece conforme crece el nivel en valor absoluto, siempre y cuando $E_t I_t \neq 1$.

El segundo método consiste en apreciar el gráfico de dispersión de la desviación estándar contra la media de los datos de cada año y observar si en apariencia los puntos se distribuyen de manera aleatoria, lo cual indica que un esquema aditivo es más adecuado, o bien alrededor de una recta, lo cual

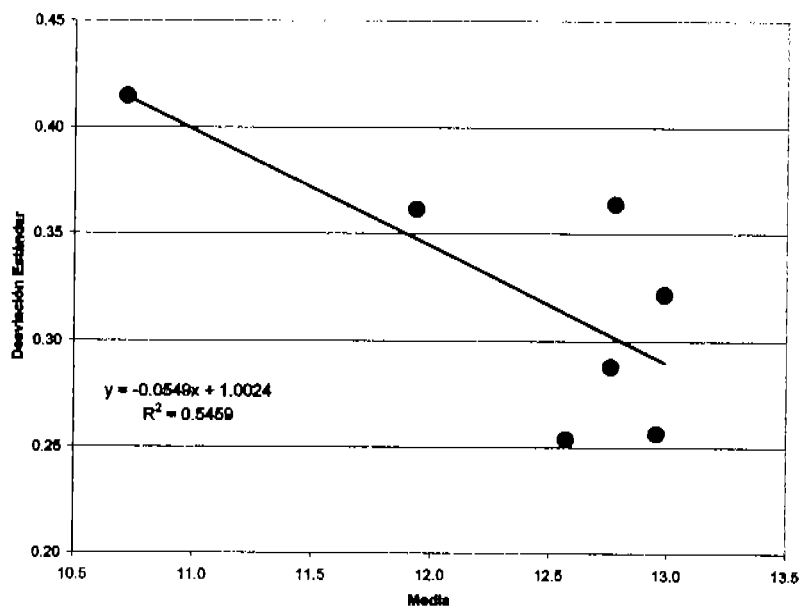


Figura 2: Método del gráfico de dispersión

lleva a un esquema multiplicativo (figura 2).

Finalmente, para exponer el método de diferencias y cocientes estacionales, se denotará $Y_{t,i}$ como la observación de la variable en cuestión en el año t y el mes (o estación) i . El primer paso consiste en calcular $d_{t,i}$ y $c_{t,i}$, las diferencias y cocientes estacionales respectivamente para toda t, i donde:

$$d_{t,i} = Y_{t,i} - Y_{t-1,i} \text{ y } c_{t,i} = \frac{Y_{t,i}}{Y_{t-1,i}},$$

Posteriormente se calculan los coeficientes de variación:

$$CV(d) = \frac{\text{Desviación Estándar (d)}}{\text{Media (d)}} \text{ y } CV(c) = \frac{\text{Desviación Estándar (c)}}{\text{Media (c)}}$$

Si $CV(c) > CV(d)$, entonces se elige el esquema aditivo. Si $CV(c) < CV(d)$, entonces se utiliza el esquema multiplicativo. Este criterio compara la variabilidad de los incrementos de la serie con la de los crecimientos relativos.

Al aplicar este método a la serie PUT (mensual, total personas universo IBOPE) con datos de enero de 1996 a marzo de 2003 (ver figura 3) se obtuvo que $CV(d)=2.11 > CV(c)=0.06$, lo cual indica que para esta serie el esquema multiplicativo resulta más adecuado que el aditivo.

Identificado un esquema aditivo o multiplicativo, el enfoque clásico trata de estimar cada componente de la serie por separado y posteriormente, ya sea por adición o multiplicación, obtener las estimaciones de la variable.

3.1. Análisis de la tendencia

La tendencia es la componente central de una serie e indica la dirección de su movimiento. Debido a que se trata de una componente de movimientos generalmente suaves y persistentes, los métodos utilizados para su estimación consisten en el ajuste de una función matemática dependiente de la variable

Fecha	PUT	Año (l)	Mes (t)	d(t,l)	c(t,l)	Fecha	PUT	Año (l)	Mes (t)	d(t,l)	c(t,l)
Oct-98	8.9	1998	10			Mar-00	13.1	2000	3	-0.1	1.0
Nov-98	10.0	1998	11			Abr-00	13.2	2000	4	0.4	1.0
Dic-98	10.4	1998	12			May-00	12.9	2000	5	0.3	1.0
Ene-97	11.0	1997	1			Jun-00	13.3	2000	6	0.3	1.0
Feb-97	10.7	1997	2			Jul-00	13.4	2000	7	-0.1	1.0
Mar-97	10.8	1997	3			Ago-00	13.1	2000	8	0.0	1.0
Abr-97	10.9	1997	4			Sep-00	12.8	2000	9	-0.1	1.0
May-97	10.9	1997	5			Oct-00	12.6	2000	10	-0.1	1.0
Jun-97	10.9	1997	6			Nov-00	12.7	2000	11	-0.1	1.0
Jul-97	11.0	1997	7			Dic-00	12.8	2000	12	0.2	1.0
Ago-97	10.9	1997	8			Ene-01	13.0	2001	1	0.1	1.0
Sep-97	11.3	1997	9			Feb-01	12.9	2001	2	0.0	1.0
Oct-97	11.5	1997	10	1.8	1.2	Mar-01	13.0	2001	3	-0.1	1.0
Nov-97	11.6	1997	11	1.5	1.2	Abr-01	13.0	2001	4	-0.1	1.0
Dic-97	11.4	1997	12	1.0	1.1	May-01	12.4	2001	5	-0.5	1.0
Ene-98	11.8	1998	1	0.8	1.1	Jun-01	12.3	2001	6	-1.0	0.9
Feb-98	11.9	1998	2	1.2	1.1	Jul-01	12.9	2001	7	-0.5	1.0
Mar-98	12.0	1998	3	1.2	1.1	Ago-01	12.8	2001	8	-0.6	1.0
Abr-98	12.0	1998	4	1.1	1.1	Sep-01	13.2	2001	9	0.4	1.0
May-98	11.6	1998	5	0.9	1.1	Oct-01	13.5	2001	10	0.9	1.1
Jun-98	12.2	1998	6	1.3	1.1	Nov-01	13.4	2001	11	0.7	1.1
Jul-98	12.3	1998	7	1.4	1.1	Dic-01	12.5	2001	12	-0.1	1.0
Ago-98	12.3	1998	8	1.5	1.1	Ene-02	13.0	2002	1	-0.1	1.0
Sep-98	12.6	1998	9	1.2	1.1	Feb-02	12.7	2002	2	-0.1	1.0
Oct-98	12.8	1998	10	1.3	1.1	Mar-02	12.7	2002	3	-0.3	1.0
Nov-98	12.9	1998	11	1.3	1.1	Abr-02	12.5	2002	4	-0.5	1.0
Dic-98	12.5	1998	12	1.1	1.1	May-02	12.4	2002	5	0.0	1.0
Ene-99	13.5	1999	1	1.8	1.2	Jun-02	12.9	2002	6	0.6	1.0
Feb-99	13.3	1999	2	1.4	1.1	Jul-02	12.9	2002	7	0.0	1.0
Mar-99	13.1	1999	3	1.1	1.1	Ago-02	12.7	2002	8	0.1	1.0
Abr-99	12.8	1999	4	0.8	1.1	Sep-02	12.4	2002	9	-0.8	0.9
May-99	12.6	1999	5	0.8	1.1	Oct-02	12.4	2002	10	-1.1	0.9
Jun-99	12.9	1999	6	0.8	1.1	Nov-02	12.8	2002	11	-0.6	1.0
Jul-99	13.4	1999	7	1.1	1.1	Dic-02	12.2	2002	12	-0.3	1.0
Ago-99	13.1	1999	8	0.8	1.1	Ene-03	12.6	2003	1	-0.3	1.0
Sep-99	12.9	1999	9	0.3	1.0	Feb-03	12.6	2003	2	-0.2	1.0
Oct-99	12.7	1999	10	-0.1	1.0	Mar-03	12.8	2003	3	0.2	1.0
Nov-99	12.8	1999	11	0.0	1.0						
Dic-99	12.4	1999	12	-0.1	1.0						
Ene-00	12.9	2000	1	-0.6	1.0						
Feb-00	12.9	2000	2	-0.4	1.0						

Media			0.3395	1.0315
Desviación estándar			0.7156	0.062
Coeficiente de variación			2.108	0.06

Figura 3: Método de las diferencias y cocientes estacionales

tiempo, o bien en la aplicación de medias móviles, con lo cual se busca eliminar las oscilaciones de corto plazo de la serie. Para el análisis de la tendencia puede adoptarse un enfoque global, el cual considera que el modelo de estimación es igual para todo el período de observación, o local para observar los niveles de la tendencia en el corto plazo.

El análisis clásico de tendencia, así como el de estacionalidad, se realiza bajo los siguientes supuestos:

1. Los componentes ciclo y tendencia están asociados formando una sola componente que se denotará $TC_{t,i}$, donde el subíndice t se refiere al año e i a la estación.
2. No existe cambio estructural. En caso de que la tendencia presente cambios bruscos, entonces se procede a dividir el período de observación y estudiar cada parte por separado.
3. Debe tenerse en cuenta todo el período de observación que se encuentra bajo la misma estructura para evitar situaciones tales como un pronóstico decreciente de la tendencia, con base en observaciones correspondientes a una fase de depresión en el ciclo.

4. La influencia de la estacionalidad en el año es nula. Los efectos positivos de la estacionalidad se compensan con los negativos.

A continuación se exponen los principales métodos para la obtención del componente tendencia en una serie de tiempo.

3.1.1. Método del ajuste analítico

Este método consiste en modelar la tendencia mediante una función matemática dependiente del tiempo, $T_t = f(t)$, que ajuste bien y cuyo manejo no sea complicado, donde f es o puede transformarse en una función lineal en parámetros tales como

$$f(t) = \sum_{i=0}^n \beta_i t^i, \quad f(t) = \alpha t^\beta \quad \text{y} \quad f(t) = e^{\sum_{i=0}^n \beta_i t^i},$$

para posteriormente aplicar el análisis de regresión por ejemplo, $f(t) = \sum_{i=0}^n \beta_i t^i + \varepsilon_t$.

Estimando mediante el método de Mínimos Cuadrados Ordinarios³ el

³Nótese que el tiempo no se considera como una variable explicativa, sino como el marco en el cual ocurren los eventos. Debido a esta característica, se dice que el tiempo es "intrínsecamente inerte".

modelo

$$Y_t = \alpha + \beta t + \varepsilon_t$$

Se obtiene el siguiente resultado para el período de octubre de 1996 a diciembre de 1998⁴, los parámetros son:

$$\widehat{\alpha}_1 \approx 10.08$$

$$\widehat{\beta}_1 \approx 0.10$$

Mientras que para enero 1999 a marzo 2003:

$$\widehat{\alpha}_2 \approx 13.11$$

$$\widehat{\beta}_2 \approx -0.01$$

3.1.2. Método de la media móvil

Cuando lo que se desea es observar la tendencia o nivel de la serie localmente, se utiliza lo que se conoce como medias móviles o promedio aritmético móvil (PAM), que consiste en transformar la serie Y_t en

⁴Obsérvese que hay un cambio estructural a partir de enero de 1999.

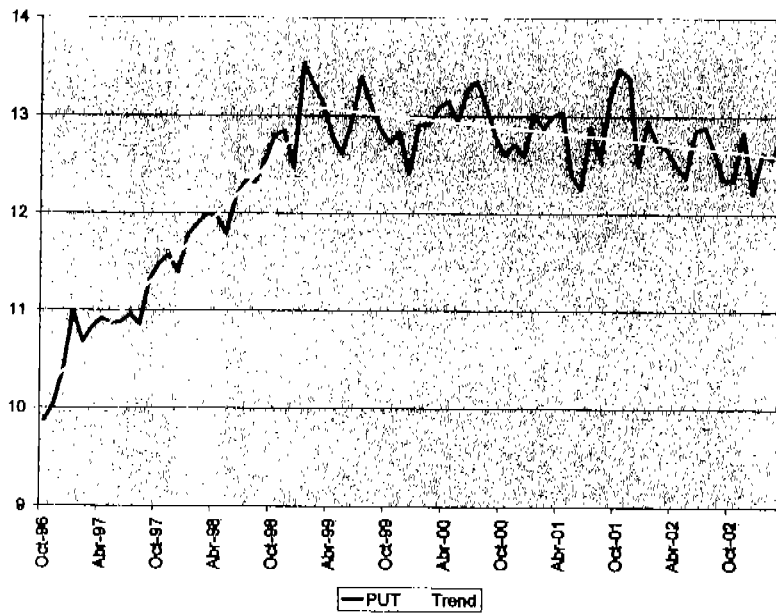


Figura 4: Metodo del ajuste analítico

$$PAM_p(Y_t) = \frac{1}{p} \sum_{i=t-\frac{p-1}{2}}^{t+\frac{p-1}{2}} Y_i$$

con p (longitud del promedio móvil) impar para que $PAM_p(Y_t)$ esté centrada en t . Debe elegirse una longitud que elimine el efecto estacional. La elección de una longitud muy grande suaviza la serie en mayor medida, en tanto que una p pequeña permite observar mejor los movimientos locales, por lo cual la elección de dicho parámetro tiene que ver con las motivaciones del analista. En caso de que p sea par, puede aplicarse

$$PAM_{p,q}(Y_t) = PAM_q(PAM_p(Y_t))$$

con q también par, de forma tal que la nueva variable esté centrada en una estación. La idea de los promedios móviles es suavizar la serie, eliminar tanto la componente irregular como el efecto estacional, ya que como se mencionó en los supuestos de la tendencia, dichos efectos se anulan. De esta forma, lo que se obtiene es la tendencia desde un enfoque local. Por ejemplo, considérese una serie mensual Y_t , a la cual se quiere aplicar promedios móviles con longitud 12. Entonces, para centrar la serie se aplica nuevamente la transfor-

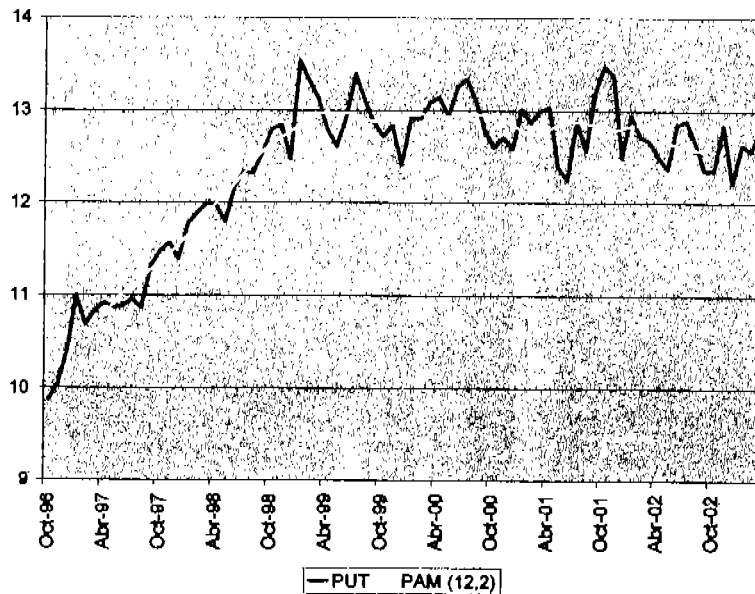


Figura 5: Método de las medias móviles

mación PAM

$$\begin{aligned}
 PAM_{12,2}(Y_t) &= PAM_2 \left(\frac{1}{12} \sum_{i=t-6}^{t+5} Y_i \right) = \frac{1}{2} \left(\frac{1}{12} \sum_{i=t-6}^{t+5} Y_i + \frac{1}{12} \sum_{i=t-5}^{t+6} Y_i \right) \\
 &= \frac{\frac{1}{2}Y_{t-6} + \sum_{i=t-5}^{t+5} Y_i + \frac{1}{2}Y_{t+6}}{12}
 \end{aligned}$$

con lo cual se obtiene un promedio ponderado centrado en el mes t (véase figura 5).

Cuando se lleve a cabo el análisis de tendencia por medio de medias móviles hay que tener presentes algunos inconvenientes del método:

1. Se pierden observaciones.
2. Se dificulta hacer predicciones al no tener una función que permita extrapolar.
3. No se dispone de medidas que evalúen la confiabilidad del resultado obtenido.
4. Los promedios móviles pueden generar ciclos u otro tipo de movimientos que no estaban presentes en la serie original.
5. La aplicación de medias móviles introduce autocorrelación.

3.2. Análisis de la estacionalidad

El análisis de la estacionalidad se realiza con dos propósitos: desestacionalizar y pronosticar. Lo primero significa eliminar el efecto de la estacionalidad sobre la serie, con el objeto de hacer comparables datos de estaciones distintas. Los métodos citados a continuación expresan a la componente estacional como un índice que representa la variación porcentual de la observación en una estación determinada con respecto a la media anual.

3.2.1. Método de la razón a la media móvil

Este método consta de cuatro fases que a continuación se explican bajo un esquema multiplicativo, aunque puede ser fácilmente generalizado al esquema aditivo.

1. Estimación de la tendencia-ciclo a través de promedios móviles PAM_p con un orden p que permita la eliminación del efecto estacional. En caso de tener un número par de estaciones por año, se estima la tendencia ciclo mediante $PAM_{p,q}$ donde p y q son pares.
2. Obtención de los índices brutos de estacionalidad (IBE), mediante la razón

$$IBE_{t,i} = \frac{Y_{t,i}}{PAM_p(Y_{t,i})}$$

3. Cálculo de los índices de estacionalidad (IE) por medio de una medida de tendencia central, ya sea la media o la mediana para eliminar el efecto de datos atípicos

$$IE_i = \overline{IBE_{t,i}}$$

o bien

$$IE_i = \text{Mediana}(IBE_{t,i}).$$

Fecha	PUT	PAM (12,2)	IBE
Oct-96	9.9		
Nov-96	10.0		
Dic-96	10.4		
Ene-97	11.0		
Feb-97	10.7		
Mar-97	10.8		
Abr-97	10.9	10.8	1.01
May-97	10.9	10.9	1.00
Jun-97	10.9	11.0	0.99
Jul-97	11.0	11.1	0.99
Ago-97	10.9	11.2	0.97
Sep-97	11.3	11.3	1.00
Oct-97	11.5	11.4	1.01
Nov-97	11.6	11.5	1.01
Dic-97	11.4	11.5	0.99
Abr-02	12.5	12.7	0.98
May-02	12.4	12.7	0.98
Jun-02	12.9	12.6	1.02
Jul-02	12.9	12.6	1.02
Ago-02	12.7	12.6	1.01
Sep-02	12.4	12.6	0.98

Figura 6: Índices brutos de estacionalidad

4. Cálculo de los índices de estacionalidad normalizados (IEN), dados por

$$IEN_i = IE_i \frac{k}{\sum_{j=1}^k IE_j},$$

donde k representa el número de estaciones por año.

La interpretación de los índices de estacionalidad normalizados es la razón que existe entre los valores de cada estación y el promedio anual, y es en sí mismo el componente estacional del modelo. Así, intérpretese por ejemplo $IEN_{mayo} = 0.9748$ como que los valores de una serie en mayo son 2.52% menores que los de la media anual.

	Indices brutos de estacionalidad						Indices de Estacionalidad	Indice de Estacionalidad Normalizado
	1997	1998	1999	2000	2001	2002		
Enero		1.0100	1.0538	0.9958	1.0179	1.0087	1.0100	1.0089
Febrero		1.0101	1.0305	0.9959	1.0095	0.9905	1.0095	1.0084
Marzo		1.0089	1.0121	1.0093	1.0202	0.9890	1.0093	1.0082
Abril	1.0125	0.9983	0.9859	1.0154	1.0191	0.9820	1.0054	1.0043
Mayo	0.9958	0.9744	0.9716	1.0000	0.9634	0.9772	0.9758	0.9748
Junio	0.9872	0.9977	0.9965	1.0245	0.9507	1.0179	0.9971	0.9960
Julio	0.9879	1.0019	1.0350	1.0305	1.0009	1.0231	1.0125	1.0114
Agosto	0.9716	0.9896	1.0160	1.0118	0.9760	1.0056	0.9976	0.9965
Septiembre	1.0030	0.9995	0.9989	0.9875	1.0266	0.9826	0.9992	0.9981
Octubre	1.0099	1.0128	0.9863	0.9734	1.0521		1.0099	1.0088
Noviembre	1.0105	1.0112	0.9934	0.9840	1.0470		1.0105	1.0094
Diciembre	0.9868	0.9763	0.9581	0.9796	0.9748		0.9763	0.9752

Figura 7: Indices de estacionalidad normalizados

Estacionalidad no estable. Puede ocurrir que se observe cierta tendencia en la estacionalidad. En tal caso no es adecuado sintetizar el comportamiento estacional en k índices, donde k es el número de estaciones por año. Dada esta situación procédase a ajustar una función a partir de los índices brutos de estacionalidad

$$IBE_i(t) = IBE_{t,i} = f_i(t) + \varepsilon_{t,i}$$

y estimense los parámetros de f mediante mínimos cuadrados ordinarios.

De esta forma se obtiene un índice distinto para cada trimestre y año. Por ejemplo, en el caso lineal de una serie trimestral, se estiman los parámetros α_i y β_i de las ecuaciones

$$\widehat{IBE}_1(t) = \widehat{\alpha}_1 + \widehat{\beta}_1 t$$

$$\widehat{IBE}_2(t) = \widehat{\alpha}_2 + \widehat{\beta}_2 t$$

$$\widehat{IBE}_3(t) = \widehat{\alpha}_3 + \widehat{\beta}_3 t$$

$$\widehat{IBE}_4(t) = \widehat{\alpha}_4 + \widehat{\beta}_4 t$$

En este caso, $\widehat{IBE}_i(t)$ constituye el componente estacional.

3.2.2. Método de la regresión

El método de la regresión consiste en plantear un modelo econométrico en función del tiempo y de variables dicotómicas que representen la estación.

Considérese, por ejemplo, una serie trimestral con tendencia lineal:

$$Y_t = \beta_0 + \beta_1 t + \lambda_2 \omega_{2t} + \lambda_3 \omega_{3t} + \lambda_4 \omega_{4t}$$

donde

$$\omega_{it} = \begin{cases} 1 & \text{si el tiempo } t \text{ corresponde al trimestre } i \\ 0 & \text{en otro caso} \end{cases}, \quad i \in \{2, 3, 4\}$$

Interprétense los coeficientes λ_i como la variación media, en términos absolutos, que presenta la variable en la estación i con respecto a la primera estación.

Las ventajas principales de este método son:

1. La tendencia y estacionalidad se estiman conjuntamente
2. Se pueden aplicar pruebas estadísticas para cada parámetro estimado
3. Se pueden incorporar en el modelo otras variables explicativas asociadas a la serie en estudio

Los cuadros y tablas 8 y 9 muestran el resultado de las regresiones aplicadas a los datos en estudio antes y a partir de enero de 1999. Obsérvese que en ambos casos los parámetros en su conjunto (estadístico F) son significativos con $\alpha = 0.05$. Sin embargo, las oscilaciones de los residuos siguen estando relacionados positivamente con el nivel de la serie (ver fig.11), esto se debe a que el método asocia los componentes tendencia y estacionalidad de manera aditiva. Este resultado confirma que la serie puede ser descrita mejor con un modelo multiplicativo.

El método más utilizado para el cálculo de la estacionalidad es el de razón a la media móvil, el cual, a pesar de no permitir inferencia estadística como lo hacen los métodos de regresión, en la práctica suele proporcionar mejores

Dependent Variable: PUT
Method: Least Squares
Date: 03/22/03 Time: 12:47
Sample: 1996:10 1998:12
Included observations: 27

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	10.45	0.14	75.56	0.0000
TREND	0.10	0.00	21.66	0.0000
W2	-0.20	0.19	-1.09	0.2922
W3	-0.18	0.19	-0.98	0.3445
W4	-0.25	0.19	-1.34	0.2002
W5	-0.47	0.19	-2.51	0.0250
W6	-0.38	0.19	-2.02	0.0635
W7	-0.36	0.19	-1.89	0.0790
W8	-0.52	0.19	-2.75	0.0156
W9	-0.28	0.19	-1.49	0.1574
W10	-0.31	0.17	-1.83	0.0880
W11	-0.31	0.17	-1.83	0.0886
W12	-0.48	0.17	-2.82	0.0137
R-squared	0.97	Mean dependent var		11.48
Adjusted R-squared	0.95	S.D. dependent var		0.82
S.E. of regression	0.19	Akaike info criterion		-0.22
Sum squared resid	0.48	Schwarz criterion		0.40
Log likelihood	15.98	F-statistic		40.73
Durbin-Watson stat	1.23	Prob(F-statistic)		0.00

Figura 8: Regresión 10/1996 a 12/1998

Dependent Variable: PUT				
Method: Least Squares				
Date: 04/01/03 Time: 22:58				
Sample: 1999:01 2003:03				
Included observations: 51				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	13.25	0.14	97.79	0
TREND	-0.01	0.00	-3.83	0.0005
W2	-0.12	0.17	-0.73	0.4706
W3	-0.05	0.17	-0.28	0.7789
W4	-0.17	0.18	-0.93	0.3605
W5	-0.45	0.18	-2.50	0.0167
W6	-0.20	0.18	-1.09	0.2811
W7	0.12	0.18	0.68	0.5016
W8	-0.14	0.18	-0.79	0.4347
W9	-0.19	0.18	-1.05	0.2991
W10	-0.19	0.18	-1.07	0.2929
W11	-0.03	0.18	-0.14	0.8894
W12	-0.54	0.18	-2.96	0.0053
R-squared	0.489598	Mean dependent va	12.84353	
Adjusted R-squared	0.328418	S.D. dependent var	0.328432	
S.E. of regression	0.26915	Akaike info criterion	0.428468	
Sum squared resid	2.752785	Schwarz criterion	0.920894	
Log likelihood	2.074058	F-statistic	3.037592	
Durbin-Watson stat	0.81132	Prob(F-statistic)	0.004453	

Figura 9: Regresión 01/1999 a 03/2003

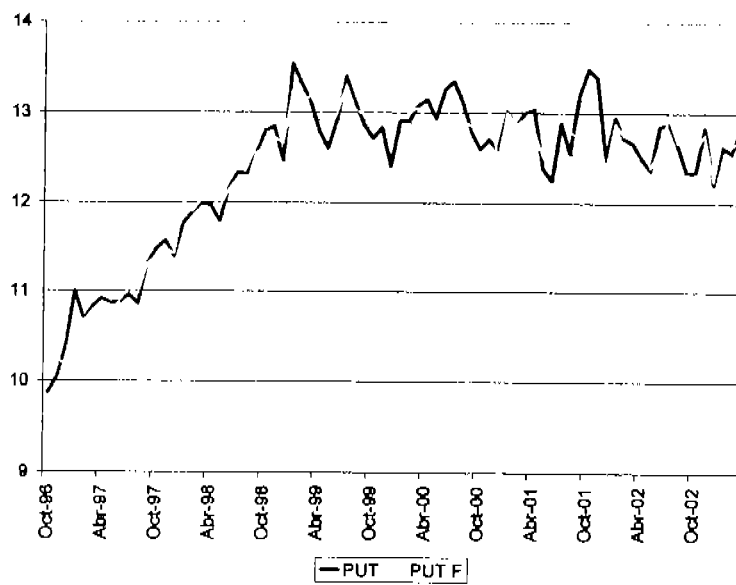


Figura 10: Ajuste tendencia-estacionalidad con método de regresión

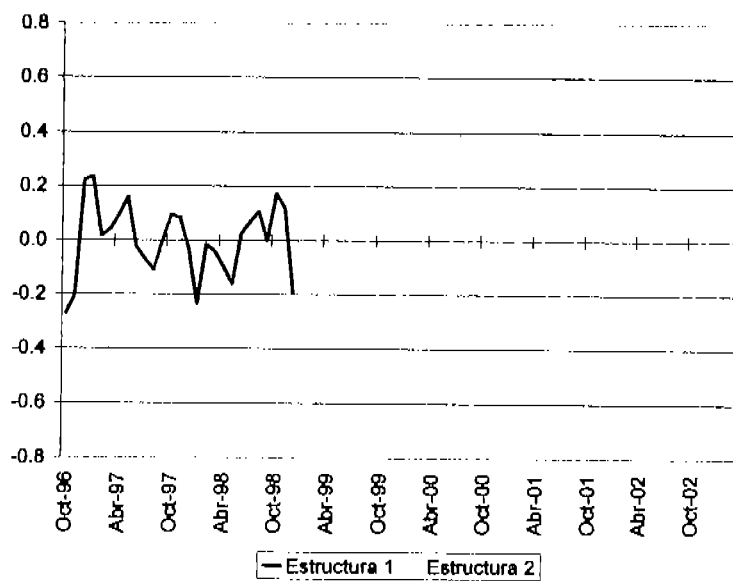


Figura 11: Residuales regresión

resultados en el sentido de que la variable Irregular explica en menor medida el comportamiento de la serie.

3.3. Análisis del ciclo

Ya se mencionó que las componentes ciclo y tendencia suelen analizarse en conjunto. Sin embargo, cuando se requiere aislar el ciclo, dentro del análisis de las componentes que integran una serie, éste es el más difícil de estudiar debido a las grandes variaciones que presentan tanto su duración como su amplitud. A continuación se explica el método de los residuos, que es el más utilizado para la identificación del componente cíclico.

3.3.1. Método de los residuos

Este método consiste en identificar la tendencia, estacionalidad e irregularidad para eliminarlas de la serie ya sea por sustracción o por división. Si $Y_t = T_t \times C_t \times E_t \times I_t$, entonces

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

para eliminar las variaciones irregulares se puede someter la serie $E_t \times T_t$ a un suavizamiento por medias móviles.

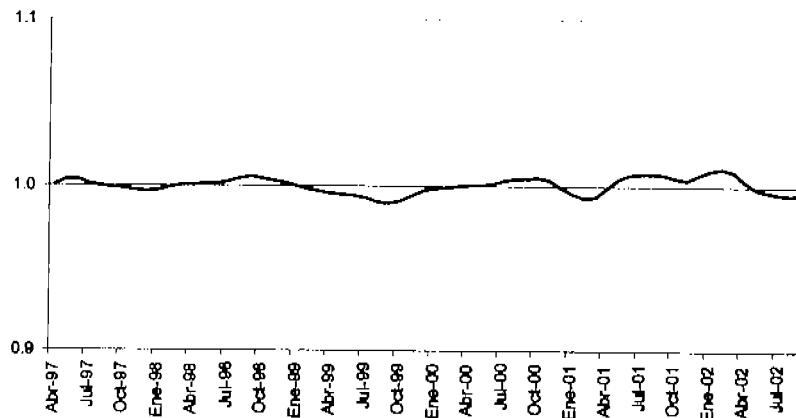


Figura 12: Ciclo

En el caso de la serie PUT, la componente ciclo no tiene gran influencia, tal como se puede apreciar en la figura 12 tras someter la serie $E_t \times T_t$ a un suavizamiento $PAM(12, 2)$, con E_t y T_t obtenidas con los métodos de razón a la media móvil y analítico respectivamente.

3.4. Pronóstico de la serie “PUT”

Con base en los resultados anteriores de este capítulo, se estimó el valor del encendido mensual del grupo objetivo “Total personas en el universo IBOPE” (ver figura 13) y se pronosticaron los valores para enero, febrero y marzo de 2003 comparándose con los datos reales, considerando como valores

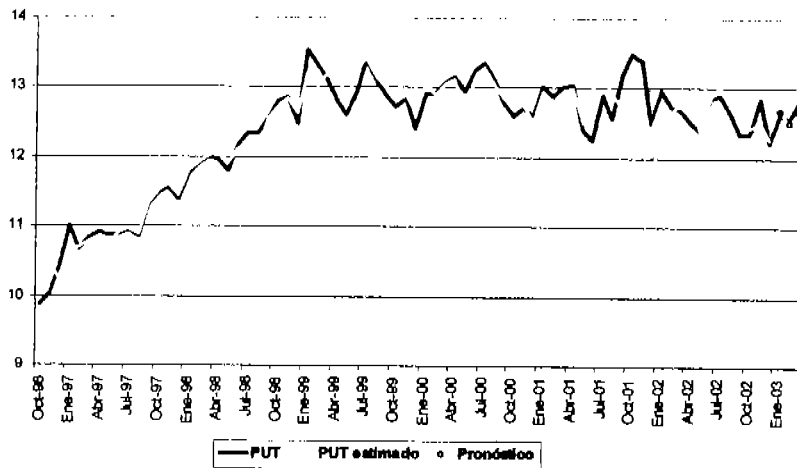


Figura 13: PUT real vs estimado

conocidos únicamente los datos de octubre de 1996 a diciembre de 2002 (ver figura 14). Se utilizó el esquema multiplicativo tal como lo sugiere tanto el método gráfico (ver figur 1), como el método de las diferencias y cocientes estacionales (ver figura 3); el método del ajuste analítico para la estimación del componente tendencia (ver figura 4) y el método de la razón a la media móvil para la estimación de la estacionalidad (ver figura 7). Se estimó cada uno de los componentes antes y a partir de enero de 1999, ya que es el punto a partir del cual se observa un cambio estructural.

Finalmente, se puede apreciar en la gráfica del componente irregular que oscila alrededor de 1, tal como es de esperarse en el esquema aditivo. Hay

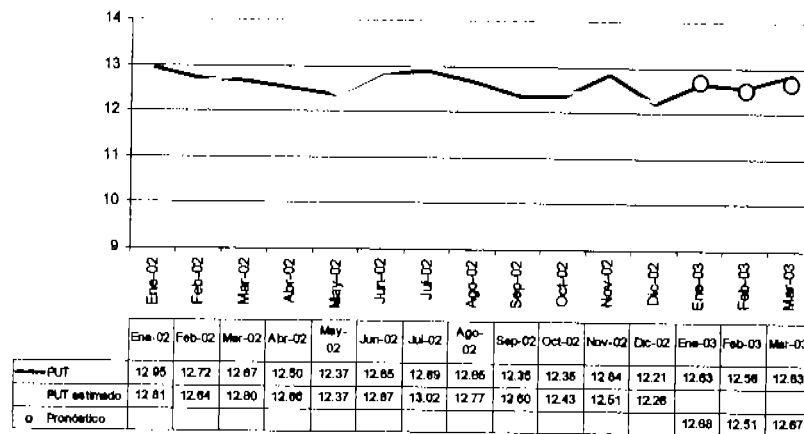


Figura 14: Pronóstico enero a marzo de 2003

un factor errático en los meses de septiembre y octubre de 2001 debido a los atentados de Nueva York y la guerra en Afganistán que estimularon la audiencia en TV (ver figura 15).

Tal como se ha visto hasta ahora, los métodos utilizados en el análisis clásico de series de tiempo son relativamente sencillos desde el punto de vista técnico, en parte debido a su enfoque determinístico, y los resultados son satisfactorios tanto en el ajuste (ver figura 13) - el error cuadrático medio del modelo tomando el período de tiempo completo es 0.056, mientras que si se considera únicamente a partir de enero de 1999 es 0.072 - como en el pronóstico (ver figura 14).

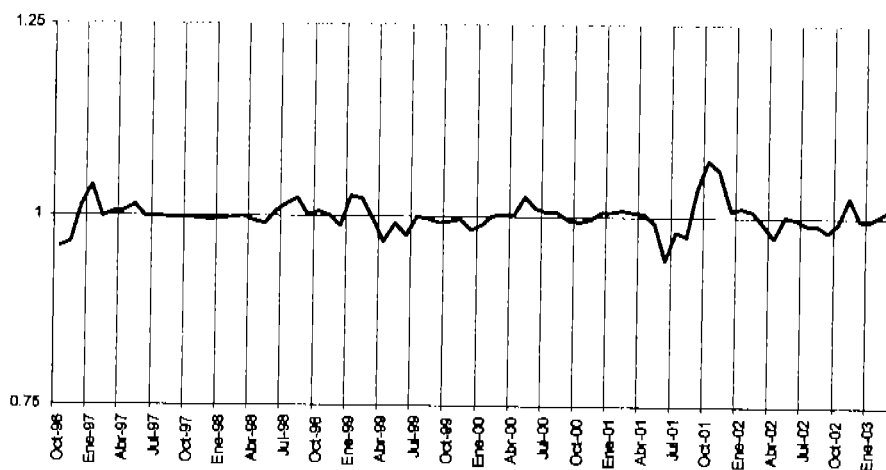


Figura 15: Irregularidad

4. Técnicas extrapolativas de pronóstico

En esta sección se presentan cuatro técnicas de pronóstico que consideran la existencia del efecto estacional multiplicativo con respecto a otros componentes de la serie, con base en los resultados obtenidos en el capítulo anterior. A diferencia del análisis clásico, aquí ya no se aborda la serie desde un enfoque determinístico, sino estocástico. Al inicio de este capítulo se exponen algunos conceptos fundamentales para la aplicación de los métodos descritos⁵.

⁵Los métodos fueron recopilados de la versión preliminar del libro "Modelos Estadísticos de Pronóstico para Series de Tiempo" del Doctor Víctor Guerrero.

4.1. Conceptos fundamentales

4.1.1. Proceso estocástico

Se llama proceso estocástico a un conjunto de variables aleatorias $\{X_t\}$, donde t es un parámetro que corre sobre un conjunto T . Los elementos distintivos de un proceso estocástico se encuentran en la naturaleza del espacio de estados, que es el conjunto de posibles valores de X_t ; el parámetro de índice T , generalmente asociado al tiempo ya sea continuo o discreto; y la relación de dependencia entre las distintas X_t , que está determinado por una función de distribución conjunta.

4.1.2. Autocorrelación y correlograma

El coeficiente de autocorrelación entre observaciones sucesivas es un caso particular del coeficiente de correlación de Pearson y se calcula para una colección de parejas $\{(x_i, x_{i+1})\}_{i=1,2,\dots,n}$ de la siguiente manera:

$$r_1 = \frac{\sum_{i=1}^{n-1} (x_i - \bar{x}_{(1)}) (x_{i+1} - \bar{x}_{(2)})}{\sqrt{\sum_{i=1}^{n-1} (x_i - \bar{x}_{(1)})^2 \sum_{i=1}^{n-1} (x_{i+1} - \bar{x}_{(2)})^2}}$$

donde

$$\bar{x}_{(1)} = \frac{1}{n-1} \sum_{i=1}^{n-1} x_i \quad \text{y} \quad \bar{x}_{(2)} = \frac{1}{n-1} \sum_{i=1}^{n-1} x_{i+1}.$$

Para valores de n suficientemente grandes sucede que $\bar{x}_{(1)} \approx \bar{x}_{(2)}$ y de ahí

$$\sum_{i=1}^{n-1} (x_i - \bar{x}_{(1)})^2 \approx \sum_{i=1}^{n-1} (x_{i+1} - \bar{x}_{(2)})^2 \approx \sum_{i=1}^n (x_i - \bar{x})^2$$

de lo cual se obtiene una expresión más simple para el coeficiente de autocorrelación entre observaciones sucesivas

$$r_1 \approx \frac{\sum_{i=1}^{n-1} (x_i - \bar{x})(x_{i+1} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

De forma similar, para observaciones rezagadas más de un período se extiende esta fórmula al coeficiente de autocorrelación al desplazamiento k :

$$r_k \approx \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Se le llama correlograma a la representación gráfica de las parejas (k, r_k) y es una herramienta fundamental para las técnicas que se revisarán en este capítulo.

4.1.3. Ruido blanco

A un proceso estocástico $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ se le llama ruido blanco si satisface las siguientes condiciones:

- $E(\varepsilon_t) = 0$,

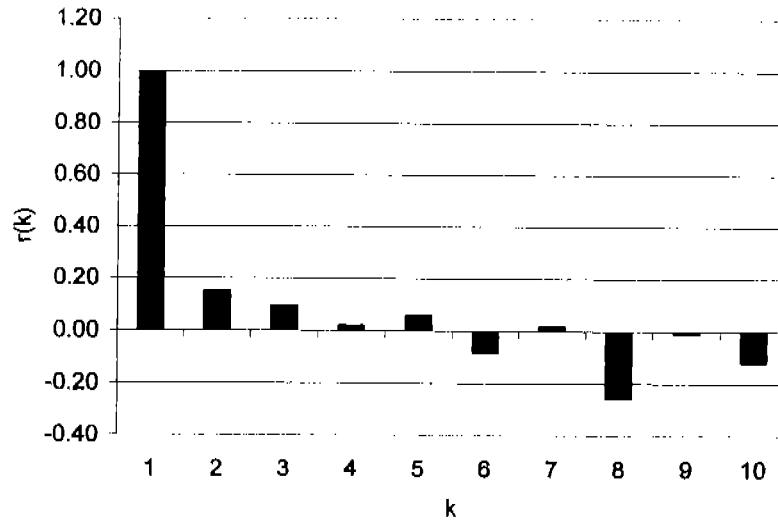


Figura 16: Ejemplo autocorrelograma

- $E(\varepsilon_t^2) = \sigma^2$,
- $E(\varepsilon_t \varepsilon_\tau) = 0$, si $\tau \neq t$.

Si además $\varepsilon_t, \varepsilon_\tau$, son independientes para $\tau \neq t$, entonces se le llama proceso de ruido blanco independiente, y si se agrega la condición $\varepsilon_t \sim N(0, \sigma^2)$ se tiene un proceso de ruido blanco gaussiano[2].

4.1.4. Diferenciación

Si $\{Y_t\}$ es una serie de tiempo, entonces a la resultante de la transformación $\nabla Y_t = Y_t - Y_{t-1}$ se le llama serie de las primeras diferencias de la

serie $\{Y_t\}$. Ahora bien, considérese el operador rezago: $B(Y_t) = BY_t = Y_{t-1}$, entonces se tiene que $\nabla = 1 - B$, es decir $\nabla Y_t = Y_t - Y_{t-1} = Y_t - BY_t = (1 - B) Y_t$.

Una propiedad muy importante de la diferenciación es que si una serie se puede expresar como un polinomio de grado n en t , entonces la k -ésima diferencia es una constante, de lo cual se concluye que con un número suficiente de diferencias se pueden remover los efectos polinomiales en las series de tiempo dadas.

Análogamente, se denotará con el nombre de primera diferencia estacional al operador $\nabla_S Y_t = Y_t - Y_{t-S}$.

4.2. Suavizamiento exponencial de Holt-Winters multiplicativo

Para una serie $\{W_t\}$ con S períodos anuales se parte del modelo siguiente:

$$W_{t+h} = T_t E_{t+h} + a_{t+h} \text{ para } t = 1, \dots, N \text{ y } h = 1, \dots, H,$$

$$T_t = T_{t-1} + \beta_{t-1} + \frac{\alpha_1 a_t}{E_{t-s}},$$

$$\beta_t = \beta_{t-1} + \frac{\alpha_1 \alpha_2 a_t}{E_{t-s}},$$

$$E_{t+h} = E_{t+h-S} + \frac{\alpha_3 (1 - \alpha_1) a_t}{T_t}.$$

Supuestos: $\{a_t\}$ es ruido blanco normal.

Método: suavizamiento exponencial de Holt-Winters multiplicativo.

En realidad el modelo no es totalmente multiplicativo, sino un híbrido, ya que aunque la estacionalidad E y la tendencia T están asociadas de manera multiplicativa, el error aleatorio a lo está aditivamente. Las estimaciones de cada uno de estos componentes están dadas por:

$$\begin{aligned}\hat{T}_t &= \alpha_1 \frac{W_t}{\hat{E}_{t-s}} + (1 - \alpha_1) (\hat{T}_{t-1} + \hat{\beta}_{t-1}), \\ \hat{\beta}_t &= \alpha_2 (\hat{T}_t - \hat{T}_{t-1}) + (1 - \alpha_2) \hat{\beta}_{t-1}, \\ \hat{E}_t &= \alpha_3 \frac{W_t}{\hat{T}_t} + (1 - \alpha_3) \hat{E}_{t-s}.\end{aligned}$$

Esto lleva a un proceso de recursiones que requiere de los valores iniciales \hat{T}_0 , $\hat{\beta}_0$ y $\{\hat{E}_0, \dots, \hat{E}_{s-1}\}$ para poder realizarse, al igual que la determinación de α_1 , α_2 y α_3 , que son las respectivas constantes de suavizamiento para la tendencia, pendiente y estacionalidad. Se puede hacer un ajuste por mínimos cuadrados tomando las primeras observaciones, utilizar los resultados del análisis clásico o bien el método sugerido por Montgomery y Johnson [8] que consiste en utilizar en principio $\bar{W}_1, \bar{W}_2, \dots, \bar{W}_m$ que denotan la media de las observaciones para los años 1, 2, ..., m. Los valores iniciales de la pendiente,

tendencia y estacionalidad están dados por:

$$\begin{aligned}\widehat{\beta}_0 &= \frac{\overline{W}_m - \overline{W}_1}{(m-1)S}, \\ \widehat{T}_0 &= \overline{W}_1 - \frac{\widehat{\beta}_0 S}{2}, \\ \widehat{E}_{q-S} &= E'_q \left(\frac{S}{\sum_{q=1}^S E'_q} \right) \text{ para } q = 1, \dots, S, \\ E'_q &= \frac{1}{m} \sum_{i=1}^m W_{q+(1-i)}.\end{aligned}$$

En cuanto a α_1 , α_2 y α_3 , se acostumbra elegirlos arbitrariamente dentro del intervalo (0.01,0.3), pero una selección más adecuada sería aquella que disminuya la suma de los cuadrados de los residuales vista como una función de estos tres parámetros.

El pronóstico para el período $N+h$, donde N es el último período observado, se realiza mediante las siguientes ecuaciones:

$$\widehat{W}_N(h) = \left(\widehat{T}_N + \widehat{\beta}_N h \right) \widehat{E}_{N-h-mS}, \text{ para } h = (m-1)S+1, (m-1)S+2, \dots, mS.$$

El intervalo de predicción del 100(1- η) % de probabilidad para W_{N+h} es:

$$\widehat{W}_N(h) \pm t_{\eta/2} \sqrt{\widehat{Var}[e_N(h)]},$$

con

$$\widehat{Var}[e_N(h)] \approx \left(\sum_{i=0}^{h-1} v_i^2 \right) \widehat{E}_{N+h-S}^2 \widehat{\sigma}_E^2, \text{ para } h = 1, \dots, S,$$

donde $v_0 = 1$ y $v_i = \alpha_1 + \alpha_1 \alpha_2 i$ para $i = 1, \dots, h - 1$ y además

$$\hat{\sigma}_E^2 = \frac{S(SCR)}{(N - S - 1) \sum_{i=0}^{S-1} \hat{E}_{N-i}},$$

donde (SCR) significa la suma de los cuadrados de los residuos.

4.2.1. Aplicación a la serie encendido

Con los datos del encendido de la televisión de enero de 1999 a diciembre de 2002 se aplicó el método de Holt-Winters multiplicativo para estimar el encendido de enero a marzo de 2003, con los parámetros $(\alpha_1, \alpha_2, \alpha_3) = (0.3, 0.01, 0.3)$ habiendo obtenido un $SCR = 1.06$ (ver fig 17). Vale la pena mencionar que en estos modelos los parámetros que den un mejor ajuste no necesariamente proporcionan un mejor pronóstico, ya que cuando $\alpha_1 \rightarrow 1$ entonces $SCR \rightarrow 0$, sin embargo, en este caso el modelo se vuelve insensible a los patrones observados y se obtiene un pronóstico plano con un valor muy cercano a la última observación, es por esto que se recomienda que los parámetros se busquen en el intervalo $[0.01, 0.3]$.

Los valores reales del encendido durante el período enero a marzo de 2003 se encontraron dentro de los intervalos de confianza del 95 % del modelo (asumiendo ruido blanco). La estimación de enero y febrero fue especialmente

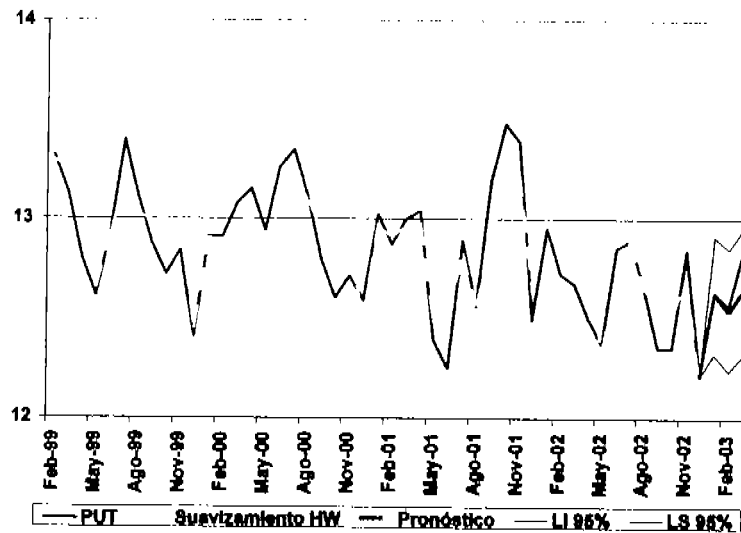


Figura 17: Suavizamiento y pronóstico HW multiplicativo

certera (ver figura 18), desde este punto de vista el resultado fue satisfactorio.

Pruebas de hipótesis

Normalidad de los residuos. Se aplicó la prueba de normalidad de Lilliefors ([9]) para la hipótesis nula:

H_0 : La variable aleatoria viene de una población con distribución normal con media y desviación estándar desconocida; vs.

H_1 : La función de distribución de la variable aleatoria no es normal.

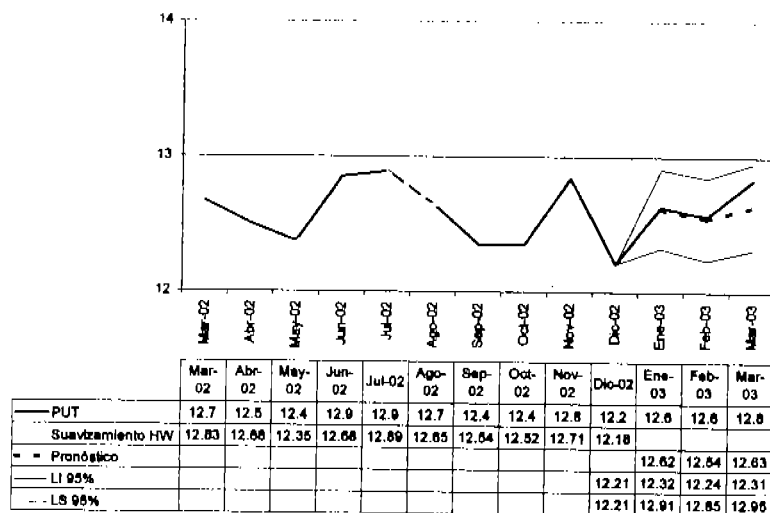


Figura 18: Pronóstico HW multiplicativo

Para aplicar esta prueba se tiene que estandarizar previamente la serie original a_t en

$$x_t = \frac{a_t - \bar{a}}{s}$$

El estadístico de esta prueba

$$T_1 = \sup_x |F(x) - S(x)|,$$

donde F representa la función de distribución normal estándar y S la función de distribución empírica, se compara con el cuantil $q_{1-\alpha}$ de la tabla de Lilliefors, y si éste resulta menor que T_1 , entonces se rechaza la hipótesis nula. En este ejercicio, para $n \geq 31$, se tiene que $q_{0.95} = 0.13 > 0.07 = T_1$; por

lo tanto no se rechaza la hipótesis de normalidad con $\alpha = 0.05$.

Ruido Blanco. Para la serie de residuos a_t obtenidos del modelo de Holt-Winters multiplicativo, se probó la hipótesis $H_0 : \mu = 0$ vs. $H_1 : \mu \neq 0$ mediante el estadístico:

$$T = \frac{\bar{a}}{S/\sqrt{n}}$$

que tiene una distribución t de *Student* con $n - 1$ grados de libertad ([10]). Se rechaza H_0 cuando $t \leq t_{\alpha/2, n-1}$ o $t \geq t_{1-\alpha/2, n-1}$. En este ejercicio, se obtuvo $t = -0.27 \in [t_{0.025, 47}, t_{0.975, 47}] = [-2.01, 2.01]$, por lo tanto no se rechaza hipótesis nula con $\alpha = 5\%$ y así la media es cero.

En cuanto a la homocedasticidad, procédase a realizar una regresión tomando como variable independiente los valores estimados de la serie y como variable dependiente a^2 . Rechazamos la hipótesis de heteroscedasticidad si el coeficiente de la variable independiente no es significativamente distinto de cero. En el caso aquí presentado, el coeficiente de a^2 es 0.04 con $p = 6.4\%$, por lo tanto podemos rechazar la hipótesis de heteroscedasticidad en la serie a con $\alpha = 5\%$ (ver fig19).

La característica restante del ruido blanco es la ausencia de autocorrelación. Al apreciar el correlograma de la serie a es evidente que no se

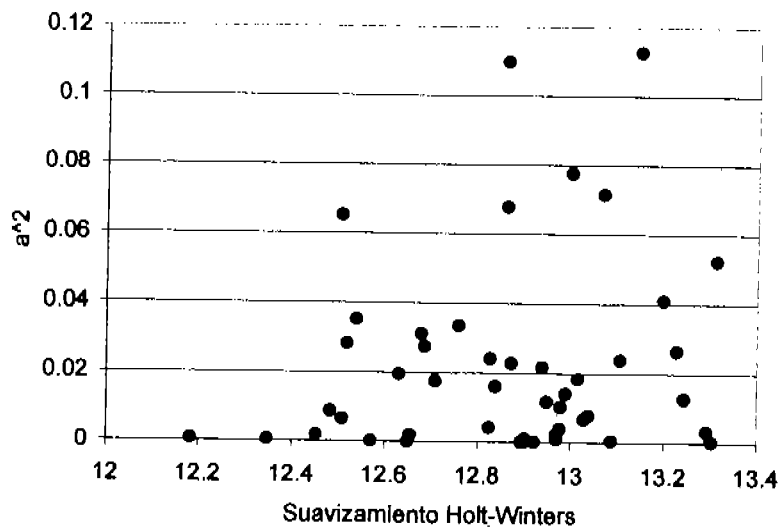


Figura 19: Método gráfico para homoscedasticidad

cumple con este supuesto, ya que los coeficientes no son significativamente distintos de cero y además se encuentran patrones en la gráfica (ver gráfica 20).

Por lo tanto, aunque al juzgar por el método gráfico y el estadístico T se puede considerar que los residuales del modelo tienen media cero y son homoscedásticos, es evidente que existe autocorrelación, por lo tanto no se trata de una serie de ruido blanco. El problema de la autocorrelación se asocia a una mayor varianza en los estimadores. En este caso se toma el pronóstico puntual. Evítese hacer inferencia relacionada con el ancho de las bandas de

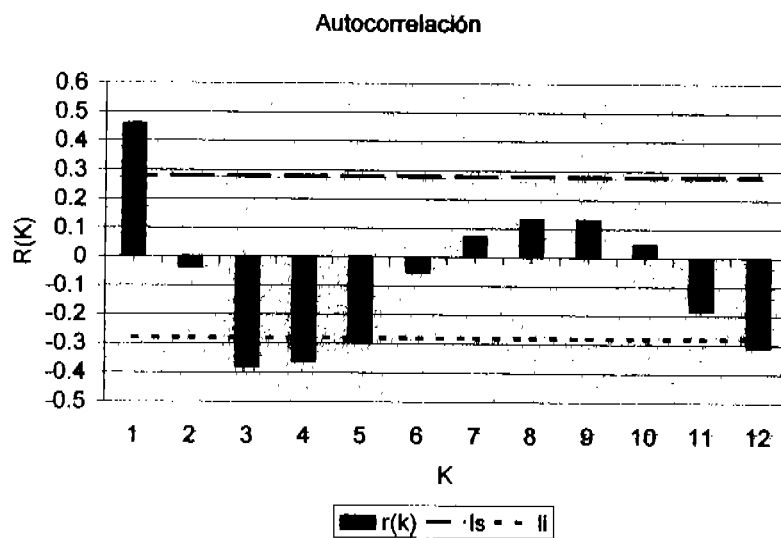


Figura 20: Correlograma de residuales Holt-Winters multiplicativo
 confianza. Después de todo, el pronóstico sigue siendo insesgado, y al menos
 en este ejercicio, da buenos resultados.

4.3. Mínimos cuadrados con variables artificiales para estacionalidad en $\log(W_t)$

Considérese el siguiente modelo:

$$W_{t+h} = T_{t+h} \exp(E_{t+h} + a_{t+h}),$$

T_{t+h} : = tendencia determinista,

$$E_{t+h} = \sum_{q=1}^S \delta_q D_{t+h,q},$$

$$D_{t+h,q} = \begin{cases} 1 & \text{si } t+h = iS + q \text{ para alguna } i = 0, 1, \dots, m, \\ 0 & \text{en otro caso.} \end{cases}$$

Supuestos: $\{a_t\}$ es ruido blanco normal.

Método: Mínimos cuadrados con variables artificiales para estacionalidad en $\log(W_t)$.

Aplíquese la función logaritmo al modelo para obtener:

$$\log(W_{t+h}) = \log(T_{t+h}) + E_{t+h} + a_{t+h},$$

cuya solución se ha de obtener resolviendo el siguiente modelo por medio de mínimos cuadrados, bajo el supuesto de que el logaritmo de la tendencia es lineal:

$$\log(W_t) = \beta_0 + \beta_1 t + \sum_{q=1}^{S-1} \delta_q D_{t,q} + a_t,$$

o bien

$$\log(W_t) = \beta_1 t + \sum_{q=1}^S \delta_q D_{t,q} + a_t,$$

donde b_0 y b_1 son las estimaciones mínimo-cuadráticas de β_0 y β_1 respectivamente. Como medida de bondad de ajuste del modelo se utiliza la suma de los cuadrados de los residuales del modelo.

Para obtener el pronóstico calcúlese:

$$(\widehat{W}_N(h)) = \exp\left(b_0 + b_1(N+h) + \sum_{q=1}^{S-1} \widehat{\delta}_q D_{N+h,q}\right) \exp\left(\widehat{Var}[e_N(h)]/2\right)$$

o bien,

$$(\widehat{W}_N(h)) = \exp\left(b_1(N+h) + \sum_{q=1}^S \widehat{\delta}_q D_{N+h,q}\right) \exp\left(\widehat{Var}[e_N(h)]/2\right),$$

donde $\exp\left(\widehat{Var}[e_N(h)]/2\right)$ es el factor de corrección aproximada del sesgo generado por la transformación del modelo original. El intervalo de predicción del $100(1-\eta)\%$ de probabilidad para $\log(W_{N+h})$ se calcula mediante la fórmula:

$$\exp\left(\log(\widehat{W}_N(h)) \pm t_{\eta/2} \sqrt{\widehat{Var}[e_N(h)]}\right) \exp\left(\widehat{Var}[e_N(h)]/2\right)$$

con

$$\begin{aligned} \widehat{Var} [e_N(h)] &= \widehat{\sigma}_a^2 \left[1 + \widehat{Var}(b_0) + (N+h)^2 \widehat{Var}(b_1) \right. \\ &\quad + \sum_{q=1}^{S-1} \widehat{Var}(\widehat{\delta}_q) D_{N+h,q} + 2(N+h) \widehat{Cov}(b_0, b_1) \\ &\quad + 2 \sum_{q=1}^{S-1} \widehat{Cov}(b_0, \widehat{\delta}_q) D_{N+h,q} \\ &\quad \left. + 2(N+h) \sum_{q=1}^{S-1} \widehat{Cov}(b_1, \widehat{\delta}_q) D_{N+h,q} \right] \end{aligned}$$

o bien,

$$\begin{aligned} \widehat{Var} [e_N(h)] &= \widehat{\sigma}_a^2 \left[1 + (N+h)^2 \widehat{Var}(b_1) \right. \\ &\quad + \sum_{q=1}^S \widehat{Var}(\widehat{\delta}_q) D_{N+h,q} \\ &\quad \left. + 2(N+h) \sum_{q=1}^{S-1} \widehat{Cov}(b_1, \widehat{\delta}_q) D_{N+h,q} \right] \end{aligned}$$

4.3.1. Aplicación a la serie encendido

Al aplicar este método a los datos que conciernen a este trabajo, se obtiene $SCR = 2.75$ (ver figura 21). Nuevamente se encuentra que los valores observados se encuentran dentro del intervalo de confianza del 95 %, suponiendo válidos los supuestos del modelo.

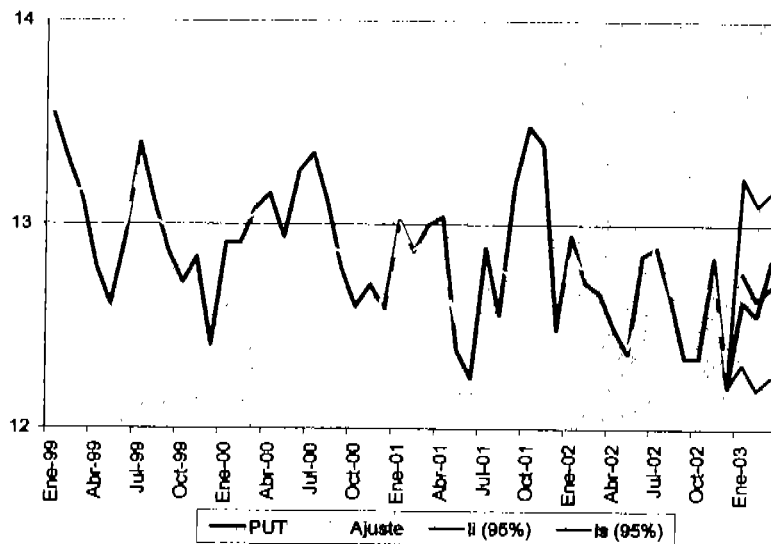


Figura 21: Suavizamiento y pronóstico MC con VA para Est en log

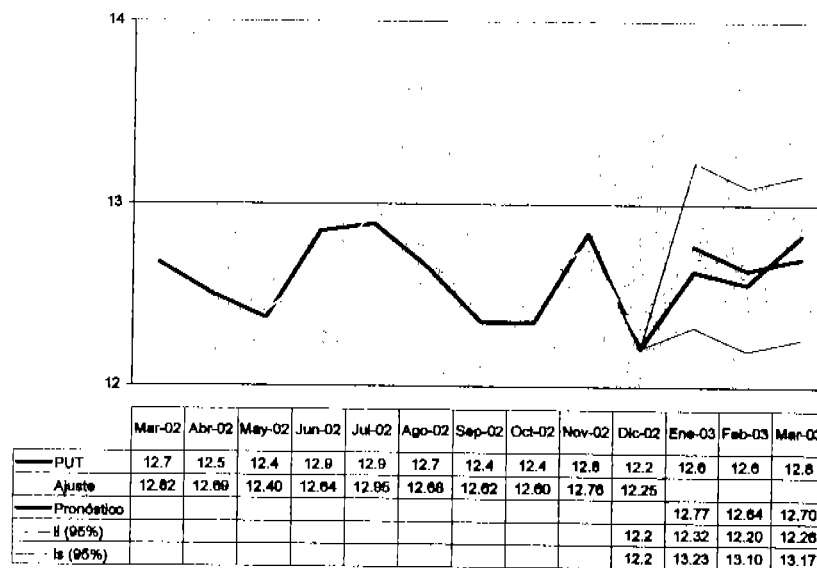


Figura 22: Acercamiento pronóstico MC con VA para Est en log

Variable	Coefficiente	Desv Est	T-Stat	Valor p
1. TENDENCIA	-0.0008	0.0002	-3.9	0.00
2. ENE	2.5848	0.0106	244.2	0.00
3. FEB	2.5753	0.0107	241.1	0.00
4. MAR	2.5813	0.0108	239.4	0.00
5. ABR	2.5719	0.0113	227.2	0.00
6. MAY	2.5495	0.0114	223.7	0.00
7. JUN	2.5693	0.0115	223.8	0.00
8. JUL	2.5942	0.0116	224.3	0.00
9. AGO	2.5738	0.0117	220.8	0.00
10. SEP	2.5700	0.0117	218.8	0.00
11. OCT	2.5696	0.0118	217.0	0.00
12. NOV	2.5829	0.0119	216.4	0.00
13. DIC	2.5428	0.0120	211.3	0.00

Figura 23: Coeficientes regresión VA para estacionalidad en log

Como se muestra a continuación en la tabla de coeficientes generados por el método de Mínimos Cuadrados (ver tabla 23), la opción elegida para este ejercicio fue incluir una variable artificial para todos y cada uno de los meses y suprimir el término independiente. Obsérvese que todos los coeficientes son significativos con $\alpha = 5\%$.

Pruebas de hipótesis

Normalidad de los residuos. El estadístico T_1 de Lilliefors resultó ser $0.10 < 0.13 = q_{0.95}$ para $n \geq 31$; por lo tanto, no se rechaza la hipótesis de normalidad con $\alpha = 0,05$.

Ruido Blanco. En este ejercicio, se obtuvieron los siguientes resultados:

- $t \approx 0 \in [t_{0.025,47}, t_{0.975,47}] = [-2.01, 2.01]$, por lo tanto no rechazamos $H_0 : E[a] = 0$ con $\alpha = 5\%$;
- $White \approx 0 < 22.36 = \chi_{5\%,13}$, por lo tanto no se rechaza el supuesto de homoscedasticidad;
- d de *Durbin Watson* $\approx 0.82 < 0.97 \approx d_L(n = 50, k' = 13)$ con $\alpha = 5\%$, por lo que nos encontramos en el área de rechazo de no autocorrelación positiva. Al igual que en el modelo anterior, la implicación de esto es un incremento en la varianza del estimador. Considérese el pronóstico puntual.

4.4. Mínimos cuadrados y funciones senoidales en $\log(W_t)$

El modelo a seguir es:

$$W_{t+h} = T_{t+h} \exp(E_{t+h} + a_{t+h}) \text{ para } t = 1, \dots, N \text{ y } h = 1, \dots, H,$$

T_{t+h} : = tendencia determinista,

$$E_{t+h} = \sum_{j=1}^S \left[\alpha_j \sin \left((t+h) \frac{2\pi j}{S} \right) + \gamma_j \cos \left((t+h) \frac{2\pi j}{S} \right) \right],$$

$$\text{con } 1 \leq g \leq \frac{S}{2}.$$

Supuestos: $\{a_t\}$ es ruido blanco normal.

Método: Mínimos cuadrados con funciones senoidales, en $\log(W_t)$.

Este método se basa en el ajuste de g armónicos de la serie. Cuando las fluctuaciones estacionales son muy regulares, puede ser suficiente con el ajuste del primer armónico ($g = 1$) y cuando dicha regularidad no es grande, se requieren dos o más armónicos ($g \geq 2$), sin que puedan ser más de $S/2$. Supóngase que el logaritmo de la tendencia es lineal. Se requiere entonces estimar por mínimos cuadrados los parámetros del modelo:

$$\log(W_t) = \beta_0 + \beta_1 t + \sum_{j=1}^S \alpha_j F_{j,t} + \sum_{j=1}^S \gamma_j G_{j,t} + a_t,$$

en donde

$$F_{j,t} = \sin \left(t \frac{2\pi j}{S} \right) \quad \text{y} \quad G_{j,t} = \cos \left(t \frac{2\pi j}{S} \right).$$

La varianza residual se estima mediante la expresión

$$\hat{\sigma}_a^2 = \frac{\sum_{t=1}^N \left(\log(W_t) - b_0 - b_1 t - \sum_{j=1}^S \hat{\alpha}_j F_{j,t} - \sum_{j=1}^S \hat{\alpha}_j G_{j,t} \right)}{N - 2 - 2g}$$

y la suma de los cuadrados de los residuales se obtiene como

$$SCR = (N - 2 - 2g) \hat{\sigma}_a^2.$$

Los intervalos de $100(1 - \eta) \%$ de confianza para las α y γ , consideradas individualmente:

$$\hat{\alpha}_j \pm t_{\eta/2} (N - 2 - 2g) \sqrt{\widehat{Var}(\hat{\alpha})}$$

y

$$\hat{\gamma}_j \pm t_{\eta/2} (N - 2 - 2g) \sqrt{\widehat{Var}(\hat{\gamma})}$$

para $j = 1, \dots, g$.

Una vez estimados los parámetros, se pronostica de la siguiente manera:

$$\widehat{W}_N(h) = \frac{\exp \left(b_0 + b_1 (N + h) + \sum_{j=1}^S \hat{\alpha}_j F_{j,N+h} + \sum_{j=1}^S \hat{\gamma}_j G_{j,N+h} \right)}{\exp \left(\widehat{Var}[e_N(h)] / 2 \right)}$$

el intervalo de predicción de $100(1 - \eta) \%$ de probabilidad para W_{N+h} :

$$\exp \left(\log(\widehat{W}_N(h)) \pm t_{\eta/2} (N - S - 1) \sqrt{\widehat{Var}[e_N(h)]} \right) \exp \left(\widehat{Var}[e_N(h)] / 2 \right)$$

donde

$$\begin{aligned}
\widehat{Var}[e_N(h)] &= \widehat{\sigma}_a^2 \left[1 + \widehat{Var}(b_0) + (N+h)^2 \widehat{Var}(b_1) \right. \\
&\quad + \sum_{j=1}^g \left(F_{j,N+h}^2 \widehat{Var}(\widehat{\alpha}_j) + G_{j,N+h}^2 \widehat{Var}(\widehat{\gamma}_j) \right) \\
&\quad + 2(N+h) \widehat{Cov}(b_0, b_1) \\
&\quad + 2 \sum_{j=1}^g \left(F_{j,N+h} \widehat{Cov}(b_0, \widehat{\alpha}_j) + G_{j,N+h} \widehat{Cov}(b_0, \widehat{\gamma}_j) \right) \\
&\quad + 2(N+h) \sum_{j=1}^g \left(F_{j,N+h} \widehat{Cov}(b_1, \widehat{\alpha}_j) + G_{j,N+h} \widehat{Cov}(b_1, \widehat{\gamma}_j) \right) \\
&\quad \left. + 2 \sum_{j=1}^g \left(F_{j,N+h} G_{j,N+h} \widehat{Cov}(\widehat{\alpha}_j, \widehat{\gamma}_j) \right) \right].
\end{aligned}$$

4.4.1. Aplicación a la serie encendido

Para aplicar este método a la serie PUT se utilizó $g = 3$ debido a que resultó ser la mejor opción en términos de regresión, sin embargo, desde el punto de vista de una regresión, el desempeño de este modelo no resultó ser bueno, ya que los coeficientes que determinan la estacionalidad no son significativos, y esto se refleja en el comportamiento del pronóstico, que si bien comprende al resultado real dentro de las bandas de confianza, la estimación

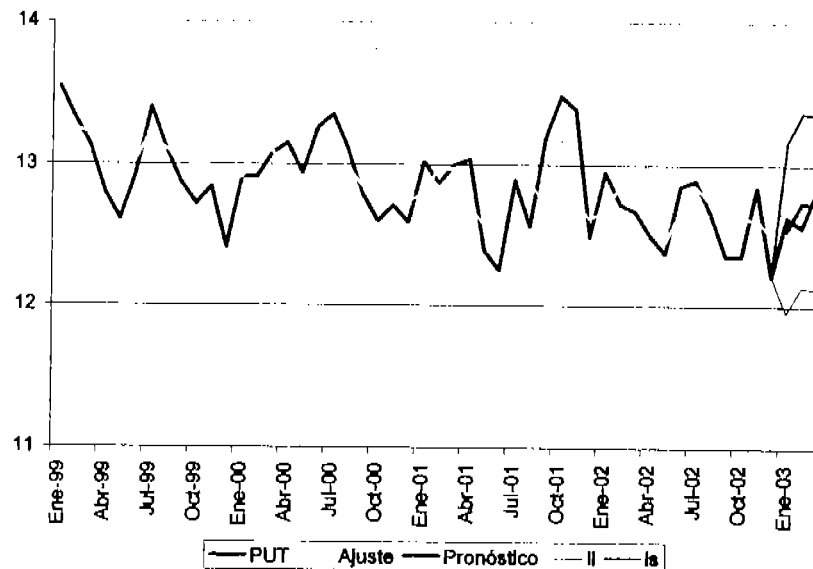


Figura 24: MC con funciones senoidales en log

puntual sigue un patrón que tiene poco que ver con el comportamiento de la serie en los meses de enero a marzo de 2003. En este ejercicio, se obtuvo $SCR = 3.56$ (ver figura 24).

Pruebas de hipótesis

Normalidad de los residuos. El estadístico T_1 de Lilliefors = $0.14 > 0.13 =$

$q_{0.95}$ para $n \geq 31$; por lo tanto se rechaza la hipótesis de normalidad con $\alpha = 0.05$. La normalidad de la distribución del error es útil esencialmente

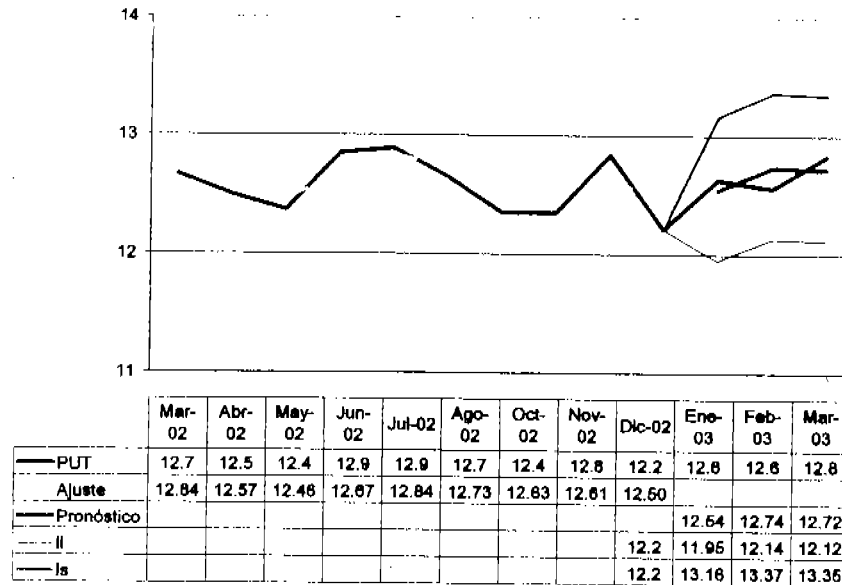


Figura 25: Acercamiento pronóstico MC con func senoidales en log

Variable	Coficiente	Desv Est	T-Stat	Valor p
Constant	2.5724	0.0062	414.1118	0.0000
TENDENCIA	-0.0008	0.0002	-3.7756	0.0005
F1,1	-0.0008	0.0045	-0.1834	0.8554
F1,2	0.0077	0.0044	1.7256	0.0916
F1,3	0.0039	0.0045	0.8737	0.3871
F2,1	-0.0011	0.0044	-0.2447	0.8079
F2,2	0.0050	0.0044	1.1162	0.2705
F2,3	-0.0083	0.0044	-1.8775	0.0672

Figura 26: Coeficientes regresión con funciones senoidales en log

para establecer los niveles de confianza en la estimación de parámetros por intervalo, los niveles de significación en las pruebas de hipótesis y los niveles de probabilidad en los intervalos de predicción, sin embargo esto no afecta a la estimación a nivel puntual.

Ruido Blanco.

- $t \approx 0 \in [t_{0.025,47}, t_{0.975,47}] = [-2.01, 2.01]$, por lo tanto no se rechaza $H_0 : E[a] = 0$ con $\alpha = 5\%$;
- $White \approx 0 < 49.7 = \chi_{0.05,35}$, \therefore no se rechaza el supuesto de homoscedasticidad;
- d de *Durbin Watson* $= 1.35 \in (1.246, 1.875) = (d_L, d_U)$ con $n = 50$, $k' = 7$ y $\alpha = 5\%$, por lo que se encuentra en el área de indecisión autocorrelación positiva.

4.5. Promedios móviles simples en $\nabla\nabla_S \log(W_t)$

El modelo es:

$$W_{t+h} = T_{t+h} E_{t+h} \varepsilon_{t+h} \text{ para } t = 1, \dots, N \text{ y } h = 1, \dots, H,$$

$$T_t = T_{t-1} \beta b_t,$$

$$E_{t+h} = E_{t+h-S}.$$

Supuestos: $W_t > 0$; $\{\varepsilon_t\}, \{b_t\}$ errores aleatorios positivos, $E(\varepsilon_t) = E(b_t) =$

1.

Método: Promedios móviles simples en $\nabla\nabla_S \log(W_t)$.

En este caso, al aplicar las diferencias regular y estacional al logaritmo de la serie, se obtiene

$$\nabla\nabla_S \log(W_{t+h}) = \nabla_S \log(b_t) + \nabla\nabla_S \log(\varepsilon_{t+h})$$

con

$$E[\nabla_S \log(b_t) + \nabla\nabla_S \log(\varepsilon_{t+h})] = 0.$$

La idea de este método consiste en realizar cálculos sencillos, del tipo de promedios aritméticos móviles asimétricos,

$$PAMA_k [\nabla\nabla_S \log(W_t)].$$

Estos promedios sólo tienen en cuenta observaciones previas al momento t y el número de términos involucrados satisface $1 \leq k \leq N$. La expresión correspondiente es

$$PAMA_k[\nabla\nabla_S \log(W_t)] = \frac{1}{k} \sum_{i=1}^k \nabla\nabla_S \log(W_{t-i}).$$

Si $k \rightarrow 1$, entonces el pronóstico reaccionará con rapidez a cambios repentinos de la serie, mientras que al hacer $k \rightarrow N$, dicha reacción se vuelve lenta.

El pronóstico se realiza aplicando la siguiente fórmula:

$$\widehat{W_N}(h) = \frac{\widehat{W_N}(h-1)\widehat{W_N}(h-S)}{\widehat{W_N}(h-S-1)} \prod_{i=1}^k \left(\frac{W_{N+1-i}}{W_{N-i}} \right)^{1/k} \left(\frac{W_{N-S-i}}{W_{N+1-S-i}} \right)^{1/k}$$

con $\widehat{W_N}(h) = W_{N+j}$ si $j \leq 0$.

4.5.1. Aplicación a la serie encendido

A diferencia de los métodos mostrados anteriormente en este capítulo, aquí se pronostica puntualmente. Se utilizó una $g = 3$, obteniendo $SCR = 5.07$. Nótese que después del impacto causado por los atentados del 11 de septiembre de 2001 y la subsecuente guerra en Afganistán en la audiencia de Televisión, la varianza de los errores del modelo creció considerablemente hasta

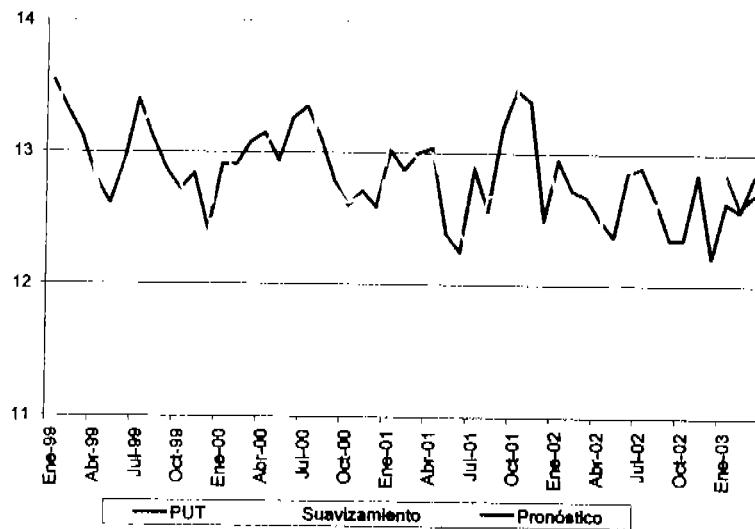


Figura 27: Promedios móviles simples en $\nabla\nabla_S \log(W_t)$

finales de 2002, mientras que durante el período de pronóstico la varianza parece estabilizarse (ver figura 27).

El pronóstico obtenido mediante este método resultó bastante bueno, no sólo en términos de la distancia entre los valores esperados y los obtenidos, sino también en comportamiento estacional, como puede observarse en la gráfica (ver figura 28).

Prueba de hipótesis

La hipótesis que es preciso probar al utilizar este método es:

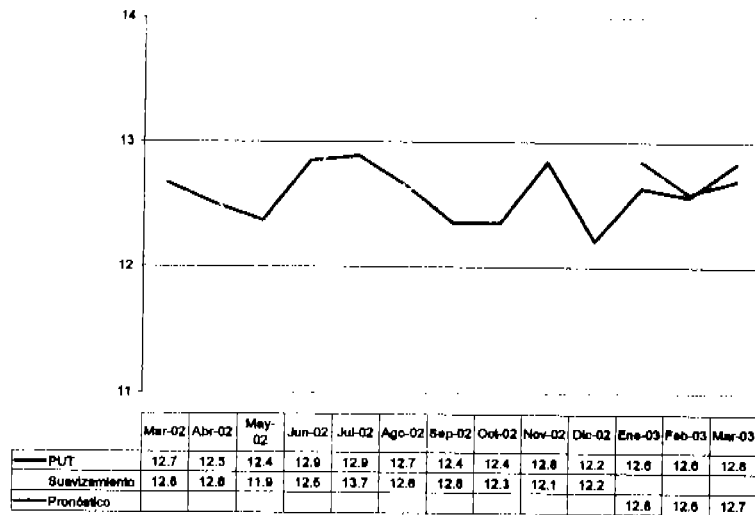


Figura 28: Pronóstico promedios móviles simples en $\nabla\nabla_S \log(W_t)$

$$H_0 : E[\nabla_S \log(b_t) + \nabla\nabla_S \log(\varepsilon_{t+h})] = 0$$

Al calcular el estadístico

$$T = \frac{\nabla_S \log(b_t) + \nabla\nabla_S \log(\varepsilon_{t+h})}{S/\sqrt{n}}$$

se obtuvo $t \approx 0.134 \in [t_{0.025,34}, t_{0.975,34}] = [-2.03, 2.03]$, por lo tanto no rechazamos H_0 con $\alpha = 5\%$.

4.6. La elección del modelo

En este capítulo, se han expuesto cuatro técnicas extrapolativas estocásticas para series multiplicativas: Suavizamiento Exponencial de Holt-Winters

	SCR		Supuestos no cumplidos	Adaptividad	Observaciones
	Modelo	Pronóstico			
HW	1.06	0.040	Autocorrelación	Sí	
VA	2.75	0.043	Autocorrelación	No	
FS	3.56	0.053		No	Zona de indecisión en autocorrelación (Durbin Watson)
PM	5.07	0.065		Sí	Estimación Puntual

Figura 29: Tabla comparativo modelos

Multiplicativo (HW), Mínimos Cuadrados con Variables Artificiales para Estacionalidad en Logaritmos (VA), Mínimos Cuadrados con Funciones Senoidales en Logaritmos (FS) y Promedios Móviles Simples en Primera Diferencia de la Primera Diferencia Estacional en Logaritmos (PM). Es una práctica común elegir el modelo cuya suma de los cuadrados de los residuales sea mínima. sin embargo, como ya se ha mencionado, el hecho de que un modelo ajuste mejor que otro, no implica que produzca mejores pronósticos.

De los modelos revisados, sólo HW y PM son adaptivos, es decir se pueden adaptar a un cambio estructural o a uno de parámetros a nivel local. Esto es muy relevante, ya que existe un antecedente de cambio estructural muy evidente a partir de enero de 1999 y no hay razón para pensar que una situación semejante no se volverá a presentar en el futuro. En cuanto a la

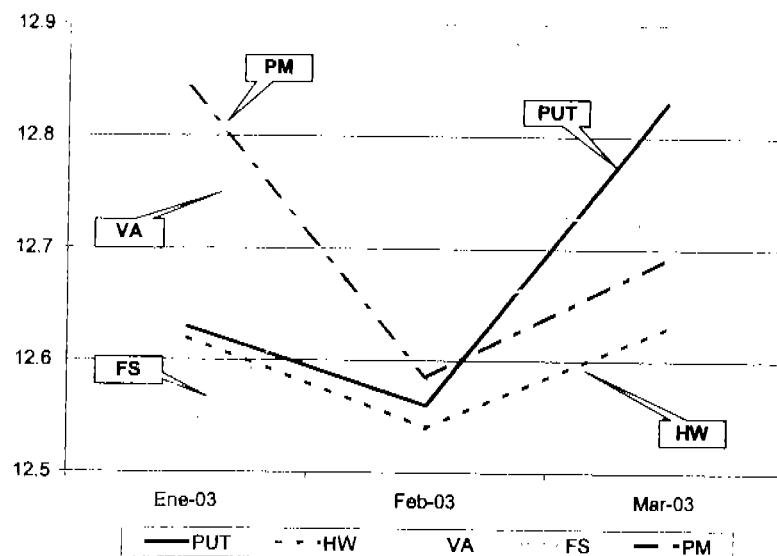


Figura 30: Comparativo Pronóstico Modelos

elección entre HW y PM, tanto el ajuste del modelo como el resultado del pronóstico señalan a HW como el modelo que mejor desempeño tiene para estos datos, a pesar de no cumplir con el supuesto de no autocorrelación, lo cual como ya se mencionó repetidas veces, tiene impacto en la varianza del pronóstico ensanchando las bandas de confianza.

5. Procesos ARMA

En este capítulo se estudia la familia de modelos econométricos estacionarios autoregresivos y de promedios móviles. Su utilización se ha vuelto popular, en parte gracias al desarrollo de paquetería que facilita su ejecución.

5.1. Conceptos fundamentales

5.1.1. Proceso estacionario

Una serie de tiempo es estrictamente estacionaria si la distribución conjunta de $x_{t_1}, x_{t_2}, \dots, x_{t_n}$ es la misma que la de $x_{t_1+\tau}, x_{t_2+\tau}, \dots, x_{t_n+\tau}$ para cualesquiera t_1, t_2, \dots, t_n y τ . Es decir, los corrimientos en τ unidades de tiempo no tienen efecto en la distribución conjunta. Al definir la estacionariedad en términos de la función de distribución conjunta, es difícil verificarla. Debido a esto, se define en términos de momentos, de la siguiente manera: un proceso estocástico es estacionario si $E[x_t] = \mu$, $\sigma_{x_t}^2 = \sigma^2$ y $Cov[x_t, x_{t+\tau}] = \gamma(\tau)$ para todos los valores de t . Nótese que la covarianza sólo depende del desplazamiento τ y no del tiempo.

5.1.2. Rezago

Sea $\{x_t\}_{t=-\infty}^{\infty}$ una serie de tiempo. El *rezago*, al cual se denotará como L , opera sobre los elementos de la serie de la siguiente manera:

$$Lx_t \equiv x_{t-1}.$$

En general, para cualquier entero k ,

$$L^k x_t = x_{t-k}.$$

El operador rezago cumple con las siguientes características:

- Es conmutativo con el operador producto escalar: $L(\beta x_t) = \beta Lx_t$,
- Es distributivo sobre el operador suma de series de tiempo: $L(x_t + w_t) = Lx_t + Lw_t$,
- Sobre una serie constante $x_t = c$ para toda t , produce la misma serie de constantes: $Lc = c$.

Una expresión del tipo $\sum_{i=0}^n a_i L^i$ es llamada *polinomial en operadores de rezago*. Es algebraicamente similar a un polinomio $\sum_{i=0}^n a_i z^i$ donde z es un escalar. La diferencia radica en que este último se refiere a un número en particular, y el polinomio en operadores de rezago se refiere a un operador

que se aplica a una serie de tiempo $\{x_t\}_{t=-\infty}^{\infty}$ para producir una nueva serie $\{y_t\}_{t=-\infty}^{\infty}$.

5.2. Procesos estacionarios de promedios móviles

5.2.1. Proceso de promedios móviles de primer orden

Considérese el siguiente proceso:

$$Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1},$$

donde $\{\varepsilon_t\}$ es ruido blanco, y tanto μ como θ son constantes. A esta serie se le llama proceso de promedios móviles de primer orden, y se denota $MA(1)$. La esperanza de Y_t está dada por:

$$E[Y_t] = E[\mu + \varepsilon_t + \theta\varepsilon_{t-1}] = \mu + E[\varepsilon_t] + \theta E[\varepsilon_{t-1}] = \mu.$$

La varianza de Y_t es:

$$\begin{aligned} E[Y_t - \mu]^2 &= E[\varepsilon_t + \theta\varepsilon_{t-1}]^2 \\ &= (1 - \theta^2)\sigma_\varepsilon^2. \end{aligned}$$

La primera autocovarianza es:

$$\begin{aligned} E(Y_t - \mu)(Y_{t-1} - \mu) &= E(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2}) \\ &= E(\varepsilon_t\varepsilon_{t-1} + \theta\varepsilon_{t-1}^2 + \theta\varepsilon_t\varepsilon_{t-2} + \theta^2\varepsilon_{t-1}\varepsilon_{t-2}) \\ &= 0 + \theta\sigma_\varepsilon^2 + 0 + 0. \end{aligned}$$

Las demás autocovarianzas son cero:

$$E(Y_t - \mu)(Y_{t-j} - \mu) = E(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-j} + \theta\varepsilon_{t-j-1}) = 0$$

para $j > 1$.

La j -ésima autocorrelación se obtiene de la siguiente manera:

$$\begin{aligned} \text{Corr}(Y_t, Y_{t-j}) &= \frac{\text{Cov}(Y_t, Y_{t-j})}{\sqrt{\text{Var}(Y_t)}\sqrt{\text{Var}(Y_{t-j})}} \\ &= \frac{\gamma_j}{\sqrt{\gamma_0}\sqrt{\gamma_0}} = \frac{\gamma_j}{\gamma_0} = \rho_j. \end{aligned}$$

Obsérvese que en particular, la primera autocorrelación está dada por la ecuación:

$$\rho_j = \frac{\theta\sigma^2}{(1 - \theta^2)\sigma_\varepsilon^2} = \frac{\theta}{1 - \theta^2}.$$

Tal como se mencionó en el capítulo anterior, ρ_j puede ser graficada como función de j , de lo cual se obtiene el correlograma, mismo que será de gran utilidad para la buena aplicación de los métodos descritos en este capítulo.

5.2.2. Proceso de promedios móviles de orden q

Un proceso de promedios móviles de orden q se denota $MA(q)$ y está caracterizado por:

$$Y_t = \mu + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i},$$

donde $\{\varepsilon_t\}$ es ruido blanco y $\{\theta_i\}_{i=1,2,\dots,q}$ son números reales. La esperanza de Y_t es:

$$E[Y_t] = \mu + E[\varepsilon_t] + \sum_{i=1}^q \theta_i E[\varepsilon_{t-i}] = \mu.$$

La varianza está dada por la siguiente ecuación:

$$\begin{aligned} \gamma_0 &= E[Y_t - \mu]^2 \\ &= E[\varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}]^2 \\ &= \sigma_\varepsilon^2 (1 + \sum_{i=1}^q \theta_i^2). \end{aligned}$$

La autocovarianza para $j = 1, 2, \dots, q$ se calcula de la siguiente manera:

$$\begin{aligned} \gamma_j &= E(\varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i})(\varepsilon_{t-j} + \sum_{k=1}^q \theta_k \varepsilon_{t-j-k}) \\ &= E(\sum_{i=0}^{q-j} \theta_{j-i} \theta_i \varepsilon_{t-j-i}^2). \end{aligned}$$

Los términos que involucran ε de diferentes fechas se eliminan, debido a que su producto tiene esperanza cero, y θ_0 se define igual a 1. Para $j > q$ no

existen ε con fechas comunes, de tal manera que la esperanza es cero. Por lo tanto:

$$\gamma_j = \begin{cases} \sum_{i=0}^{q-j} \theta_{j-i} \theta_i \sigma_\varepsilon^2 & \text{si } j = 1, 2, \dots, q, \\ 0 & \text{si } j > q. \end{cases}$$

5.2.3. Proceso de promedios móviles de orden infinito

Considérese el siguiente proceso a cuya clase se denota como $MA(\infty)$:

$$Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},$$

donde $\{\varepsilon_t\}$ es ruido blanco y $\{\psi_i\}$ son números reales. Por convención ψ denota a los coeficientes de un proceso de promedios móviles de orden infinito y θ a los de uno de orden finito. Una característica de esta clase de procesos es que:

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty \Rightarrow Y_t \text{ es débilmente estacionario.}$$

Es conveniente trabajar con la siguiente condición que es más fuerte que la anterior:

$$\sum_{j=0}^{\infty} |\psi_j| < \infty,$$

ya que permite calcular la esperanza y las autocovarianzas como una extrapolación de los resultados de un proceso $MA(q)$:

$$\begin{aligned} E[Y_t] &= \lim_{T \rightarrow \infty} E\left[\mu + \sum_{i=0}^T \psi_i \varepsilon_{t-i}\right] \\ &= \mu, \end{aligned}$$

$$\begin{aligned} \gamma_0 &= E[Y_t - \mu]^2 \\ &= \lim_{T \rightarrow \infty} E\left[\sum_{i=0}^T \psi_i \varepsilon_{t-i}\right]^2 \\ &= \lim_{T \rightarrow \infty} \sum_{i=0}^T \psi_i^2 \sigma_\varepsilon^2, \end{aligned}$$

$$\begin{aligned} \gamma_j &= E[Y_t - \mu][Y_{t-j} - \mu] \\ &= \lim_{T \rightarrow \infty} \sigma_\varepsilon^2 \left(\sum_{i=0}^T \psi_{j+i} \psi_i\right). \end{aligned}$$

Más aún, si los valores absolutos de un proceso $MA(\infty)$ son sumables, entonces los valores absolutos de las autocovarianzas también lo son:

$$\sum_{j=0}^{\infty} |\gamma_j| < \infty.$$

5.3. Procesos estacionarios autoregresivos

5.3.1. Proceso autoregresivo de primer orden

Una autoregresión de primer orden, denotada $AR(1)$, satisface la siguiente ecuación en diferencias:

$$Y_t = c + \phi Y_{t-1} + \varepsilon_t.$$

Nuevamente, $\{\varepsilon_t\}$ es un proceso de ruido blanco. Si $|\phi| \geq 1$, entonces las consecuencias de las ε para Y se acumulan en lugar de desvanecerse a lo largo del tiempo, bajo estas circunstancias el proceso no es estacionario. Por otro lado, si $|\phi| < 1$, entonces se tiene un proceso estacionario débil que satisface la siguiente ecuación:

$$\begin{aligned} Y_t &= \lim_{T \rightarrow \infty} \sum_{i=0}^T \phi^i (c + \varepsilon_{t-i}) \\ &= \frac{c}{1 - \phi} + \lim_{T \rightarrow \infty} \sum_{i=0}^T \phi^i \varepsilon_{t-i}. \end{aligned}$$

Esto puede ser visto como un proceso $MA(\infty)$ cuyos coeficientes ψ_j están dados por ϕ^j . De aquí en adelante, se asume que $|\phi| < 1$. Bajo este supuesto, la esperanza, varianza y j -ésima autocovarianza de un proceso $AR(1)$ son los

siguientes:

$$\begin{aligned} E[Y_t] &= E \left[\frac{c}{1-\phi} + \lim_{T \rightarrow \infty} \sum_{i=0}^T \phi^i \varepsilon_{t-i} \right] \\ &= \frac{c}{1-\phi}, \end{aligned}$$

$$\begin{aligned} \gamma_0 &= E[Y_t - \mu]^2 \\ &= E \left[\lim_{T \rightarrow \infty} \sum_{i=0}^T \phi^i \varepsilon_{t-i} \right]^2 \\ &= \lim_{T \rightarrow \infty} \sum_{i=0}^T \phi^{2i} \sigma_\varepsilon^2 \\ &= \frac{\sigma_\varepsilon^2}{1-\phi^2}, \end{aligned}$$

$$\begin{aligned} \gamma_j &= E[Y_t - \mu][Y_{t-j} - \mu] \\ &= E \left(\lim_{T \rightarrow \infty} \sum_{i=0}^T \phi^i \varepsilon_{t-i} \right) \left(\lim_{T \rightarrow \infty} \sum_{i=0}^T \phi^i \varepsilon_{t-j-i} \right) \\ &= \lim_{T \rightarrow \infty} \sum_{i=0}^T \phi^{j+2i} \sigma_\varepsilon^2 \\ &= \phi^j \lim_{T \rightarrow \infty} \sum_{i=0}^T \phi^{2i} \sigma_\varepsilon^2 \\ &= \frac{\phi^j}{1-\phi^2} \sigma_\varepsilon^2. \end{aligned}$$

Nótese que el coeficiente de autocorrelación está dado por:

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \phi^j,$$

por lo tanto, dado que $\phi < 1$, el autocorrelograma sigue un patrón de decaimiento geométrico.

5.3.2. Proceso autoregresivo de segundo orden

Una autoregresión de segundo orden se denota $AR(2)$, y satisface la siguiente ecuación:

$$Y_t = c + \sum_{i=1}^2 \phi_i Y_{t-i} + \varepsilon_t,$$

o en operadores de rezago:

$$(1 - \phi_1 L - \phi_2 L^2) Y_t = c + \varepsilon_t.$$

Si las raíces del polinomio $(1 - \phi_1 z + \phi_2 z^2) = 0$ se encuentran fuera del círculo unitario en el plano imaginario, entonces el proceso $AR(2)$ es débilmente estacionario y el inverso del operador autoregresivo está dado por:

$$\psi(L) = (1 - \phi_1 L - \phi_2 L^2)^{-1} = \lim_{T \rightarrow \infty} \sum_{i=0}^T \psi_i L^i,$$

por lo tanto:

$$\begin{aligned} Y_t &= \psi(L)c + \psi(L)\varepsilon_t \\ &= \frac{c}{1 - \phi_1 - \phi_2} + \psi(L)\varepsilon_t, \end{aligned}$$

lo cual es un proceso $MA(\infty)$ con coeficientes sumables en valor absoluto con media:

$$\mu = \frac{c}{1 - \phi_1 - \phi_2}.$$

La varianza del proceso $AR(2)$ se puede obtener partiendo del cálculo de la autocorrelación:

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2}$$

$$\frac{\gamma_j}{\gamma_0} = \frac{\phi_1 \gamma_{j-1}}{\gamma_0} + \frac{\phi_2 \gamma_{j-2}}{\gamma_0}$$

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2}.$$

En particular:

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

$$= \frac{\phi_1}{1 - \phi_2}$$

y

$$\rho_2 = \phi_1 \rho_1 + \phi_2.$$

Ahora bien:

$$\begin{aligned}
 Y_t - \mu &= \sum_{i=1}^2 \phi_i (Y_{t-i} - \mu) + \varepsilon_t, \\
 (Y_t - \mu)^2 &= \left(\sum_{i=1}^2 \phi_i (Y_{t-i} - \mu) + \varepsilon_t \right) (Y_t - \mu), \\
 \gamma_0 &= \sum_{i=1}^2 \phi_i \gamma_i + E(\varepsilon_t) (Y_t - \mu) \\
 &= \sum_{i=1}^2 \phi_i \gamma_i + \sigma_\varepsilon^2 \\
 &= \sum_{i=1}^2 \phi_i \rho_i \gamma_0 + \sigma_\varepsilon^2 \\
 &= \left[\frac{\phi_1^2}{1 - \phi_2} + \frac{\phi_2 \phi_1^2}{1 - \phi_2} + \phi_2^2 \right] \gamma_0 + \sigma_\varepsilon^2 \\
 &= \frac{(1 - \phi_2) \sigma_\varepsilon^2}{(1 - \phi_2) ((1 - \phi_2)^2 - \phi_1^2)}.
 \end{aligned}$$

5.3.3. Proceso autoregresivo de orden p

Un proceso autoregresivo de orden p se denota $AR(p)$ y satisface:

$$Y_t = c + \sum_{i=1}^p \phi_i Y_{t-i} + \varepsilon_t.$$

Si las raíces del polinomio en z :

$$1 - \sum_{i=1}^p \phi_i z^i = 0,$$

se encuentran fuera del círculo unitario en el plano imaginario. En este caso $\{Y_t\}$ es un proceso débilmente estacionario y existe una representación de la serie en forma de un proceso de promedios móviles:

$$Y_t = \mu + \psi(L)\varepsilon_t,$$

donde

$$\psi(L) = \left(1 - \sum_{i=1}^p \phi_i L^i\right)^{-1}$$

y $\sum_{j=0}^{\infty} |\psi_j| < \infty$. Suponiendo que el proceso es estacionario, una forma de encontrar la esperanza de Y_t es:

$$\begin{aligned} \mu &= c + \sum_{i=1}^p \phi_i \mu \\ &= \frac{c}{1 - \sum_{i=1}^p \phi_i}. \end{aligned}$$

De este par de ecuaciones se obtiene que:

$$\begin{aligned} Y_t &= c + \sum_{i=1}^p \phi_i Y_{t-i} + \varepsilon_t \\ \Rightarrow Y_t - \mu &= \sum_{i=1}^p \phi_i (Y_{t-i} - \mu) + \varepsilon_t. \end{aligned}$$

Las autocovarianzas se encuentran multiplicando ambos miembros de esta ecuación por $(Y_{t-j} - \mu)$ y tomando las esperanzas:

$$\gamma_j = \begin{cases} \sum_{i=1}^p \phi_i \gamma_{j-i} & \text{si } j > 0 \\ \sum_{i=1}^p \phi_i \gamma_i + \sigma^2 & \text{si } j = 0 \end{cases}$$

Al dividir γ_j entre γ_0 se obtiene lo que se conoce como “ecuaciones de Yule-Walker”:

$$\rho_j = \sum_{i=1}^p \phi_i \rho_{j-i}.$$

Las soluciones para las autocovarianzas son de la forma:

$$\gamma_j = \sum_{i=1}^p g_i \lambda_i^j,$$

donde los valores propios $(\lambda_i)_{i=1}^p$ son las soluciones del polinomio

$$\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \dots - \phi_p = 0.$$

5.4. Procesos estacionarios mixtos autoregresivos y de promedios móviles

Un proceso $ARMA(p, q)$ incluye tanto el término autoregresivo como el de promedios móviles:

$$Y_t = c + \sum_{i=1}^p \phi_i Y_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$$

o en forma de operadores de rezago,

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) Y_t = c + \left(1 + \sum_{j=1}^q \theta_j L^j\right) \varepsilon_t.$$

Bajo el supuesto de que las raíces del polinomio

$$1 - \sum_{i=1}^p \phi_i z^i = 0$$

se encuentran fuera del círculo unitario, se puede dividir la ecuación por

$(1 - \sum_{i=1}^p \phi_i L^i)$ obteniendo así:

$$Y_t = \mu + \psi(L)\varepsilon_t,$$

donde

$$\psi(L) = \frac{1 + \sum_{j=1}^q \theta_j L^j}{1 - \sum_{i=1}^p \phi_i L^i},$$

$$\sum_{i=0}^{\infty} |\psi_j| < \infty$$

y

$$\mu = \frac{c}{1 - \sum_{i=1}^p \phi_i}.$$

Así, la estacionariedad de un proceso *ARMA* depende exclusivamente de los parámetros autoregresivos.

A menudo resulta conveniente escribir el proceso *ARMA* en términos de desviaciones alrededor de la media:

$$Y_t - \mu = \sum_{i=1}^p \phi_i (Y_{t-i} - \mu) + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}.$$

La autocovarianza se calcula multiplicando ambos lados de esta ecuación por $(Y_{t-j} - \mu)$ y tomando la esperanza. Para $j > q$, la ecuación resultante

toma la forma

$$\gamma_j = \sum_{i=1}^p \phi_i \gamma_{j-i}.$$

De esta forma, después de q rezagos, la función de autocovarianza γ_j , así como la de autocorrelación ρ_j , dependen de los parámetros autoregresivos. Nótese que esta expresión no es válida para $j \leq q$, debido a la correlación entre $\theta_j \varepsilon_{t-j}$ y Y_{t-j} . Así, un proceso $ARMA(p, q)$ tendrá una función de autocovarianza más complicada para los rezagos 1 a q .

5.5. Pronóstico

5.5.1. Conceptos fundamentales

Sea $Y_{t+1|t}^*$ el pronóstico de Y_{t+1} con base en un conjunto de variables X_t observadas a la fecha t . Para este fin, se utiliza el error cuadrático medio (MSE) con el fin de minimizar la inexactitud del pronóstico

$$MSE(Y_{t+1|t}^*) \equiv E(Y_{t+1} - Y_{t+1|t}^*)^2.$$

El pronóstico con el menor MSE es la esperanza de Y_{t+1} condicionada a X_t :

$$Y_{t+1|t}^* = E(Y_{t+1} | X_t).$$

Considere el caso en el cual $Y_{t+1|t}^*$ es una función lineal de X_t :

$$Y_{t+1|t}^* = \alpha' X_t.$$

Supóngase que se quiere encontrar los valores de α tales que no haya correlación entre X_t y el error $(Y_{t+1} - \alpha' X_t)$:

$$E[(Y_{t+1} - \alpha' X_t) X_t'] = 0.$$

Si esto se cumple, entonces el pronóstico $\alpha' X_t$ es llamado proyección lineal de Y_{t+1} en X_t . La proyección lineal produce el *MSE* más pequeño dentro de la clase de pronósticos lineales:

$$MSE(\alpha' X_t) = E(Y_{t+1})^2 - E(Y_{t+1} X_t') [E(X_t X_t')]^{-1} E(X_t Y_{t+1}).$$

Los coeficientes de la proyección α en términos de Y_{t+1} y X_t se calculan mediante la ecuación:

$$\alpha' = E(Y_{t+1} X_t') [E(X_t X_t')]^{-1},$$

asumiendo que $E(X_t X_t')$ es una matriz no singular. Si fuese singular, entonces el vector de coeficientes α no quedaría determinado de manera única.

5.5.2. Pronóstico de un proceso ARMA(p,q)

Considérese el pronóstico de un proceso $ARMA(p, q)$ estacionario e invertible:

$$\left(1 - \sum_{i=1}^p \phi_i L^i\right) (Y_t - \mu) = \left(1 + \sum_{j=1}^q \theta_j L^j\right) \varepsilon_t.$$

De aquí, el pronóstico para un período posterior a la última observación estaría dado por:

$$\left(\widehat{Y}_{t+1|t} - \mu\right) = \sum_{i=1}^p \phi_i (Y_{t+1-i} - \mu) + \sum_{j=1}^q \theta_j \varepsilon_{t+1-j},$$

con $\{\widehat{\varepsilon}_j\}$ generada recursivamente de la siguiente manera:

$$\widehat{\varepsilon}_t = Y_t - \widehat{Y}_{t|t-1}.$$

Para pronosticar s períodos hacia adelante se procede de acuerdo con la siguiente ecuación:

$$\begin{aligned} & \left(\widehat{Y}_{t+1|t} - \mu\right) \\ = & \begin{cases} \sum_{i=1}^p \phi_i \left(\widehat{Y}_{t+s-i|t} - \mu\right) + \sum_{j=0}^{q-s} \theta_{s+j} \varepsilon_{t-j} & \text{para } s = 1, 2, \dots, q, \\ \sum_{i=1}^p \phi_i \left(\widehat{Y}_{t+s-i|t} - \mu\right) & \text{para } s = q + 1, q + 2, \dots \end{cases} \end{aligned}$$

donde

$$\widehat{Y}_{\tau|t} = Y_{\tau} \text{ para } \tau \leq t.$$

De esta forma, para un horizonte de pronóstico s mayor que el orden q del componente de promedios móviles, el pronóstico está dado por una ecuación diferencial de orden p determinada exclusivamente por los parámetros autoregresivos.

5.5.3. Pronóstico con base en un número finito de observaciones

Aproximación a un pronóstico óptimo. Una manera de pronosticar con base en un número finito de observaciones es suponer que la premuestra de ε son iguales a cero. La idea es utilizar la siguiente aproximación:

$$\begin{aligned} & \widehat{E}(Y_{t+s}|Y_t, Y_{t-1}, \dots) \\ & \cong \widehat{E}(Y_{t+s}|Y_t, Y_{t-1}, \dots, Y_{t-m+1}, \varepsilon_{t-m} = 0, \varepsilon_{t-m-1} = 0, \dots). \end{aligned}$$

Por ejemplo, considérese un proceso $MA(q)$. La recursión puede empezar estableciendo

$$\widehat{\varepsilon}_{t-m} = \widehat{\varepsilon}_{t-m-1} = \dots = \widehat{\varepsilon}_{t-m-q+1} = 0$$

y después iterando para generar $\widehat{\varepsilon}_{t-m+1}, \widehat{\varepsilon}_{t-m+2}, \dots, \widehat{\varepsilon}_t$ de la siguiente manera:

$$\begin{aligned} \widehat{\varepsilon}_{t-m+1} &= (Y_{t-m+1} - \mu), \\ \widehat{\varepsilon}_{t-m+2} &= (Y_{t-m+2} - \mu) - \theta_1 \widehat{\varepsilon}_{t-m+1}, \\ \widehat{\varepsilon}_{t-m+3} &= (Y_{t-m+3} - \mu) - \theta_1 \widehat{\varepsilon}_{t-m+2} - \theta_2 \widehat{\varepsilon}_{t-m+1}, \end{aligned}$$

y así sucesivamente. Los valores resultantes $\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+s}$ son sustituidos directamente dentro de la expresión del proceso $MA(q)$ y se procede a pronosticar. Por ejemplo, para $s = q = 1$ el pronóstico sería

$$\widehat{Y}_{t+1|t} = \mu + \sum_{i=1}^m (-1)^{i-1} \theta^i (Y_{t-i+1} - \mu).$$

Para m grande y $|\theta|$ pequeña, este procedimiento da una aproximación. Si $|\theta|$ es cercana a 1, la aproximación no será tan buena. Obsérvese que si el operador de promedios móviles es no invertible, entonces esta forma de pronóstico es inapropiada y no debe ser utilizada.

Pronóstico exacto. Sea

$$\mathbf{X}_t \equiv \begin{bmatrix} 1 \\ Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-m+1} \end{bmatrix}.$$

Se busca un pronóstico lineal de la forma

$$\alpha^{(m)'} \mathbf{X}_t = \alpha_0^{(m)} + \sum_{i=1}^m \alpha_i^{(m)} Y_{t-i+1}.$$

El coeficiente que relaciona a Y_{t+1} con Y_t en una proyección de Y_{t+1} con base en las m observaciones más recientes de Y se denota $\alpha_1^{(m)}$. En general,

este número es diferente de el coeficiente que relaciona a Y_{t+1} con Y_t en una proyección de Y_{t+1} con base en las $m + 1$ observaciones más recientes de Y , $\alpha_1^{(m+1)}$.

Si la serie $\{Y_t\}$ es débilmente estacionaria, entonces $E(Y_t Y_{t-j}) = \gamma_j + \mu^2$.

Por proyección lineal, se tiene que

$$\begin{aligned} \alpha^{(m)'} &\equiv \begin{bmatrix} \alpha_0^{(m)} & \alpha_1^{(m)} & \alpha_2^{(m)} & \cdots & \alpha_m^{(m)} \end{bmatrix} \\ &= \begin{bmatrix} \mu & (\gamma_1 + \mu^2) & (\gamma_2 + \mu^2) & \cdots & (\gamma_m + \mu^2) \end{bmatrix} \\ &\quad \times \begin{bmatrix} 1 & \mu & \mu & \cdots & \mu \\ \mu & \gamma_0 + \mu^2 & \gamma_1 + \mu^2 & \cdots & \gamma_{m-1} + \mu^2 \\ \mu & \gamma_1 + \mu^2 & \gamma_0 + \mu^2 & \cdots & \gamma_{m-2} + \mu^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \mu & \gamma_{m-1} + \mu^2 & \gamma_{m-2} + \mu^2 & \cdots & \gamma_0 + \mu^2 \end{bmatrix}^{-1}. \end{aligned}$$

Cuando un término constante es incluido en X_t , resulta más conveniente expresar las variables como desviaciones de la media. Entonces se podría calcular la proyección de $(Y_{t+1} - \mu)$ en

$$X_t = [(Y_t - \mu), (Y_{t-1} - \mu), \dots, (Y_{t-m+1} - \mu)]' :$$

$$\hat{Y}_{t+1|t} - \mu = \sum_{i=1}^m \alpha_i^{(m)} (Y_{t-i+1} - \mu).$$

Para esta definición de X_t los coeficientes pueden ser calculados de la siguiente manera:

$$\begin{bmatrix} \alpha_1^{(m)} \\ \alpha_2^{(m)} \\ \vdots \\ \alpha_m^{(m)} \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{m-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{m-2} \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{m-1} & \gamma_{m-2} & \cdots & \gamma_0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_m \end{bmatrix}.$$

Para generar un pronóstico s periodos hacia adelante $\hat{Y}_{t+s|t}$, se utilizaría

$$\hat{Y}_{t+s|t} = \mu + \sum_{i=1}^m \alpha_i^{(m,s)} (Y_{t-i+1} - \mu),$$

donde

$$\begin{bmatrix} \alpha_1^{(m,s)} \\ \alpha_2^{(m,s)} \\ \vdots \\ \alpha_m^{(m,s)} \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{m-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{m-2} \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{m-1} & \gamma_{m-2} & \cdots & \gamma_0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma_s \\ \gamma_{s+1} \\ \vdots \\ \gamma_{s+m-1} \end{bmatrix}.$$

5.5.4. Procedimiento de Box-Jenkins

Box y Jenkins (1976) fueron partidarios del uso de la menor cantidad posible de parámetros. Ellos observaron que en la práctica, los analistas sustituyen a final de cuentas los operadores verdaderos $\theta(L)$ y $\phi(L)$ por las esti-

maciones $\hat{\theta}(L)$ y $\hat{\phi}(L)$ con base en los datos y que existe mayor espacio para equivocarse cuando se tienen más parámetros a estimar.

La manera de abordar el pronóstico propuesta por Box y Jenkins[2] se divide en 4 pasos según Hamilton⁶:

1. Transformar los datos, si es necesario, para que el supuesto de estacionariedad débil sea razonable.
2. Intentar con valores pequeños de p y q para un modelo $ARMA(p, q)$ que pueda describir el comportamiento de la serie transformada.
3. Estimar los parámetros de $\theta(L)$ y $\phi(L)$.
4. Realizar un diagnóstico para confirmar que el modelo es consistente con el comportamiento de los datos observados.

Una vez llevados a cabo los pasos anteriores se procede con el pronóstico, que es el uso que se le da al modelo en el presente trabajo.

Para la elección de una q inicial, recuérdese que si los datos siguen un proceso $MA(q)$, entonces la autocorrelación ρ_j será cero para $j > q$.

⁶Otros autores desarrollan el procedimiento de manera análoga dividiendo el procedimiento en los siguientes pasos: identificación, estimación, pruebas de diagnóstico y usos o pronóstico.

Otra medida útil es la *autocorrelación parcial*. La m -ésima autocorrelación parcial, denotada $\alpha_m^{(m)}$, se define como el último coeficiente en una proyección lineal de Y sobre sus m más recientes valores:

$$Y_{t+1|t} - \mu = \sum_{i=1}^m \alpha_i^{(m)} (Y_{t+1-i} - \mu),$$

el vector $\alpha^{(m)}$ puede ser calculado como

$$\begin{bmatrix} \alpha_1^{(m)} \\ \alpha_2^{(m)} \\ \vdots \\ \alpha_m^{(m)} \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{m-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{m-2} \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{m-1} & \gamma_{m-2} & \cdots & \gamma_0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_m \end{bmatrix}.$$

Si los datos estuviesen generados realmente por un proceso $AR(p)$, entonces únicamente los p valores más recientes de Y serían útiles para el pronóstico. En este caso, los coeficientes de la proyección en Y para períodos más allá del p -ésimo anterior serían igual a cero.

En conclusión, para una aproximación inicial, tómese como q el valor donde la autocorrelación se acerca súbitamente a cero, y como p cuando esto mismo ocurre con la autocorrelación parcial.

5.6. Aplicación la metodología ARMA a la serie encendido de TV

Los valores iniciales a intentar para p y q generalmente se determinan como se hace para un modelo $AR(p)$ o para un $MA(q)$ por separado. Se deben hacer ajustes al modelo de modo iterativo hasta que idealmente se cumpla que:

1.
 - La serie utilizada sea estacionaria.
 - Todos los parámetros seleccionados sean significativos.
 - Las correlaciones entre cualquier par de estimadores de parámetros sean pequeñas.
 - Los resultados del modelo ajustado sean ruido blanco.
 - El modelo ajustado sea mejor que otros posibles.

A continuación se presentan los pasos a seguir para un análisis *ARMA*:

1. Verificar que la serie no contenga huecos.
2. Graficar la serie de tiempo original.
3. Realizar alguna prueba de existencia de raíz unitaria.

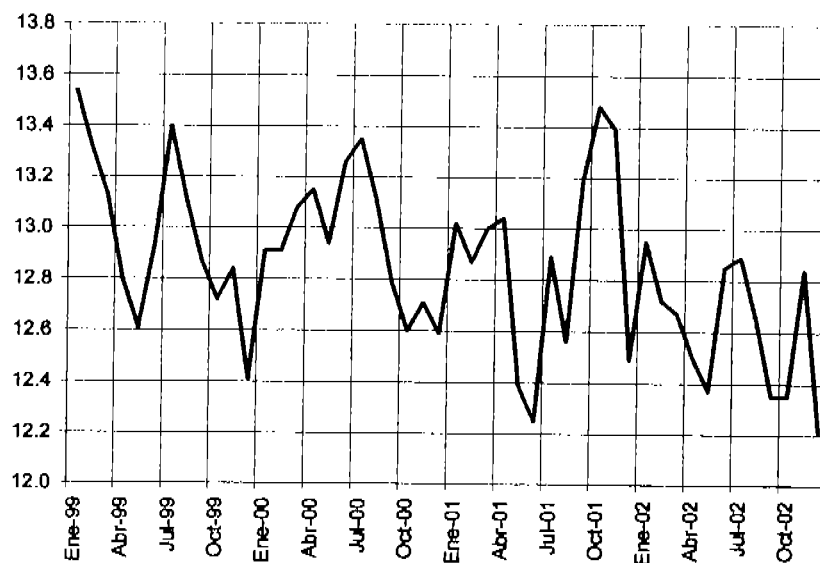


Figura 31: Serie PUT original

Dickey-Fuller Test = -29.74798	
Probabilidad de obtener un valor a	
x es 1% para n=50	
Sin constante ni tiempo	-2.62
Con constante	-3.58
Con constante y tiempo	-4.15

Figura 32: Prueba de Dickey-Fuller para existencia de raíces unitarias

En este caso, al ser el estadístico menor que los cuantiles para 50 observaciones se rechaza la hipótesis de existencia de raíces unitarias.

4. Remover cualquier tendencia y periodicidad existente.

Primero se aplica alguna prueba de tendencia. En este ejercicio se utiliza la prueba de discordancias para series estacionales. Se define para la serie y_t :

$$q_{ij} = \begin{cases} 1 & y_i > y_j, i < j, \\ 0 & \text{otro caso.} \end{cases}$$

$$Q_j = \sum_{i < j} q_{ij},$$

$$\tau_j = 1 - \frac{4 \sum Q_j}{rm(m-1)}, \text{ con } \begin{cases} r = \# \text{ de datos por año,} \\ m = \# \text{ de años.} \end{cases}$$

$$\tau = \frac{1}{r} \sum_{j=1}^r \tau_j,$$

$$z = \frac{\tau}{\sqrt{\frac{2(2m+5)}{9rm(m-1)}}},$$

y se rechazará la hipótesis nula en favor de la existencia de una tenden-

cia si $|z| > z_{1-\alpha/2}$, tal como ocurre en este ejercicio con $|z| = 5.03 > 1.96 =$

$z_{1-5\%/2}$. Por lo tanto, estamos en el escenario de tendencia sin raíz

unitaria, en este caso se sugiere eliminar la tendencia por medio de

	<i>Coefficients</i>	<i>P-value</i>
Intercept	13.28	0.00
t	-0.01	0.00
feb	-0.14	0.48
mar	-0.12	0.56
abr	-0.20	0.31
may	-0.49	0.02
jun	-0.23	0.24
jul	0.09	0.67
ago	-0.18	0.37
sep	-0.23	0.26
oct	-0.23	0.25
nov	-0.06	0.75
dic	-0.57	0.01

Figura 33: Coeficientes regresión

regresiones y no por diferenciación, pues esta última opción daría como resultado un proceso no invertible. A continuación se muestran los coeficientes de la regresión utilizada para eliminar tendencia y estacionalidad:

El resultado es una serie residual de la regresión que no muestra tendencia ni estacionalidad (ver figura 34).

5. Obtener correlograma y correlación parcial de la serie transformada (ver figura 35).

Obsérvese que la primera y segunda correlación parcial es significa-

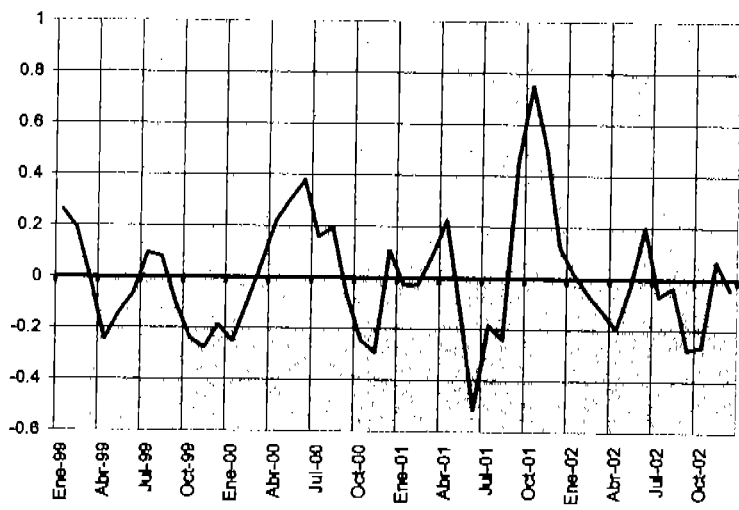


Figura 34: Serie transformada

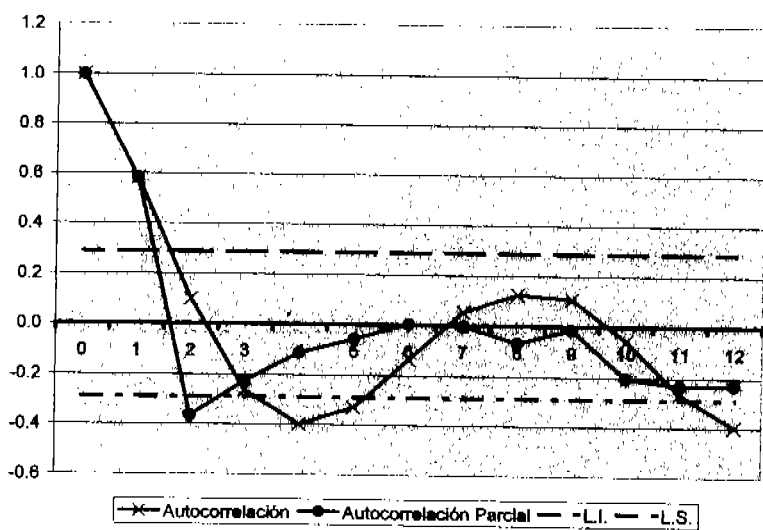


Figura 35: Autocorrelograma de serie transformada

Q(11-2) 3.906798
 Significance level of Q 0.917439

Variable	Coeff	Std Error	T-Stat	Signif
AR{1}	0.8012	0.1400	5.7212	0.0000
AR{2}	-0.3697	0.1384	-2.6719	0.0105

Covariance\Correlation Matrix of Coefficients		
	AR{1}	AR{2}
AR{1}	0.019613126	-0.592574
AR{2}	-0.011482879	0.019146

Figura 36: AR(2)

tivamente distinta de cero, también se observa un patrón dentro del correlograma.

6. Ajustar distintos modelos *ARMA* y reportar el mejor (ver figura 36).

El modelo del cual se obtuvo el mejor resultado fue *AR(2)* ya que es el único cuyos parámetros en su totalidad son significativos. La matriz de covarianza-correlación que se muestra en la tabla se lee como covarianza de la diagonal hacia abajo y correlación arriba de la diagonal. La correlación no es tan pequeña y el alto valor de significancia del estadístico *Q* indica que hasta el momento no hay evidencia de correlación en los residuales.

7. Mostrar residuales, obtener correlograma y correlación parcial y hacer

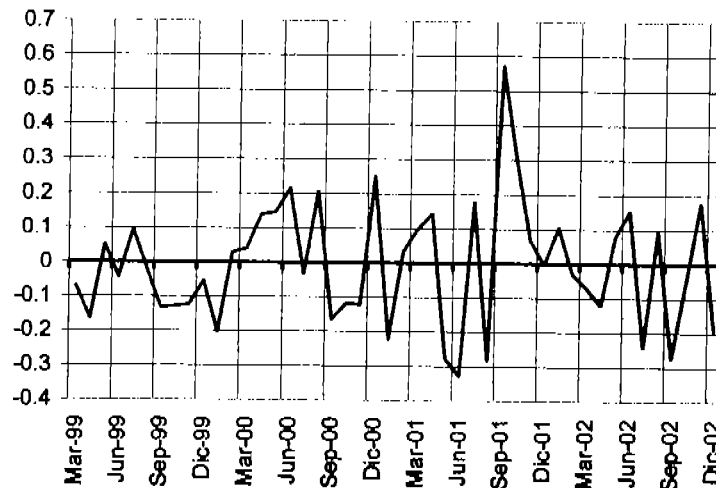


Figura 37: Residuales AR(2)

una prueba de aleatoriedad para comprobar que se trate de ruido blanco (ver figuras 37 y 38).

Obsérvese que los coeficientes de autocorrelación y autocorrelación parcial ya se encuentran dentro de las bandas de confianza, también se diluyó, hasta cierto punto, el patrón que mostraban los datos transformados en el autocorrelograma.

La aleatoriedad de los residuales se probó mediante puntos de cambio. Una desviación de la aleatoriedad se da cuando una serie exhibe pocos o muchos puntos de cambio bien definidos como "crestas" o "valles".

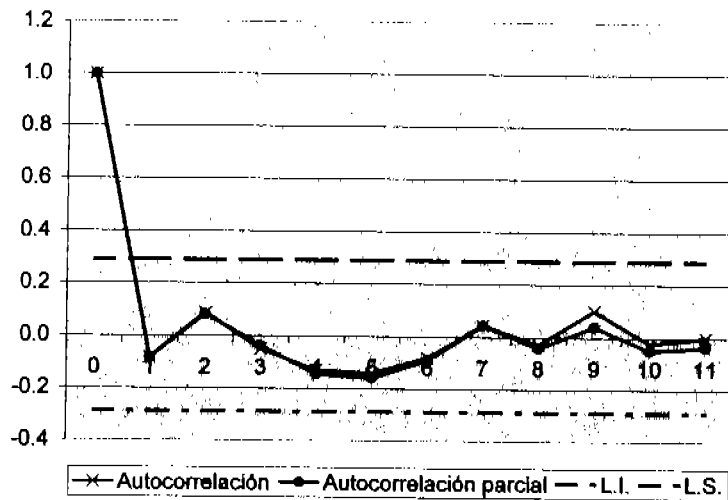


Figura 38: Correlograma residuales AR(2)

Una observación y_t es una "cresta" si el valor observado es mayor que el inmediato anterior y el inmediato posterior y es un "valle" si es menor.

Se define

$$u_i = \begin{cases} 1 & \text{si } y_i \text{ es punto de cambio,} \\ 0 & \text{en otro caso.} \end{cases}$$

$$p = \sum_{i=2}^{n-1} u_i.$$

La variable p converge en distribución rápidamente a

$$N\left(\frac{2(n-2)}{3}, \frac{16n-29}{40}\right),$$

por lo cual, el estadístico de prueba es

$$z = \frac{p - \frac{2(n-2)}{3}}{\sqrt{\frac{16n-29}{40}}},$$

que se distribuye $N(0, 1)$. Se rechaza la hipótesis de aleatoriedad si $|z| > z_{1-\alpha/2}$. En este ejercicio $|z| = 0.32 < 1.96 = z_{1-5\%/2}$, por lo tanto no se puede rechazar la hipótesis de aleatoriedad con $\alpha = 5\%$.

8. Realizar el pronóstico (3 meses), reincorporar todos los ajustes hasta llegar a un modelo en términos de los datos originales.

Aplicando los coeficientes de la regresión a los datos estimados por el modelo se obtiene la serie que se muestra (ver figura 39) comparada con los datos reales.

El error cuadrático medio obtenido dentro del período de datos disponibles (enero 1999-diciembre 2002) es 2.58, y dentro del período de pronóstico (enero 2003-marzo 2003) 0.07.

El resultado es satisfactorio tanto en términos de ajuste como de cumplimiento de supuestos.

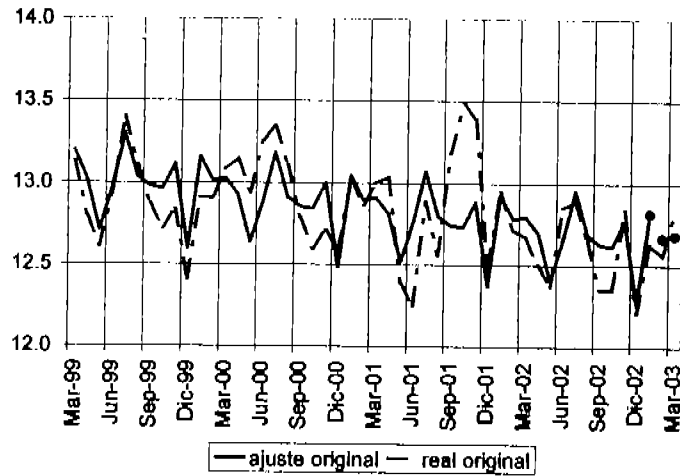


Figura 39: Ajuste AR(2)

Conclusión

El análisis de la serie de tiempo “Encendido de Televisión” puede ser realizado a través del enfoque clásico, en el cual salen a relucir los elementos o patrones de la serie, esto es tendencia-ciclo, estacionalidad e irregularidad. La manera en la cual se relacionan entre si estos elementos es aditiva o multiplicativa. Para la serie en estudio se utilizó el esquema multiplicativo, aunque las pruebas sugeridas para definir este esquema así lo indicaban, en realidad no lo hacían de manera muy clara. Si bien el análisis clásico da un buen resultado en el pronóstico, es necesario encontrar un modelo flexible debido a la naturaleza adaptiva de la serie PUT que muestra de manera muy

clara un cambio estructural a partir de enero de 1999 y podría eventualmente sufrir otro cambio de este tipo. En este sentido, dentro de la clase de técnicas extrapolativas de pronóstico, se encontraron dos modelos que resultan interesantes: Holt-Winters mixto y promedios móviles, de los cuales el primero resultó dar el mejor ajuste tanto en el período de observaciones como en el de pronósticos a pesar de no cumplir con el supuesto de no autocorrelación. Finalmente, dentro de la clase de modelos ARMA, AR(2) también dio un resultado satisfactorio y pasó todas las pruebas de hipótesis sugeridas, sin embargo, Holt-Winters sigue dando mejores resultados gracias a su flexibilidad, recuérdese que los coeficientes de este modelo cambian con el tiempo. La implicación de este resultado, es que el Encendido de Televisión es un fenómeno adaptivo, tiene comportamientos locales en términos de tendencia y estacionalidad.

La línea de investigación a futuro que se sugiere a partir de los resultados obtenidos en este trabajo se encuentra en el terreno de las ciencias sociales. Entender a qué fenómeno social obedecen los movimientos de la audiencia, por ejemplo, cómo afecta a la audiencia la masificación de nuevas alternativas mediáticas, como Internet; o el papel que juega el desempeño económico de

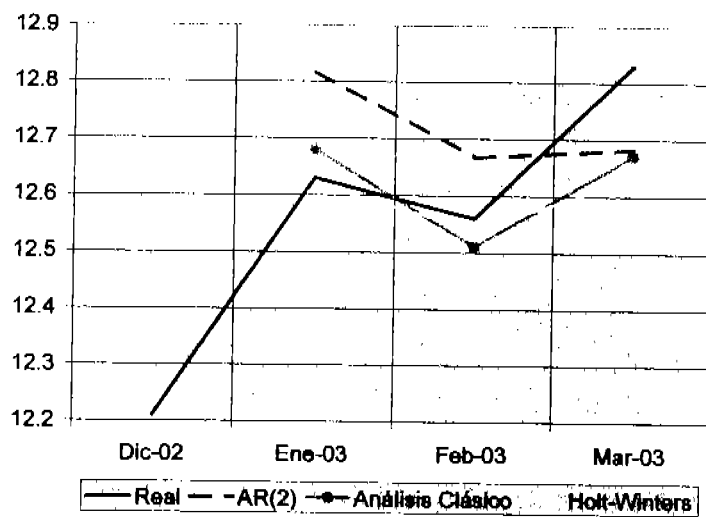


Figura 40: Pronóstico PUT 1/2003-3/2003

una sociedad en el consumo de los medios de comunicación.

Referencias

- [1] Rodríguez Morilla, Carmen. Análisis de Series Temporales. Cuadernos de Estadística. Madrid: Editorial La Muralla, S.A., 2000.
- [2] Hamilton, James D. Time Series Analysis. Princeton University Press, 1994.
- [3] INEGI. XII Censo General de Población y Vivienda, 2000.
- [4] IBOPE AGB México. Establishment Survey. México, 2000.
- [5] IBOPE AGB México. Ratings. Sistema de consulta Telereport.
- [6] Karlin, Samuel; Taylor, Howard M. A First Course in Stochastic Processes. Segunda Edición. San Diego: Academic Press, 1975.
- [7] Guerrero, Víctor M. Modelos Estadísticos de Pronóstico para Series de Tiempo (versión preliminar). Instituto Tecnológico Autónomo de México, 1998.
- [8] Montgomery, D. C. and Johnson, L. A. Forecasting and Time Series Analysis. New York: McGraw-Hill Book Co, 1976.

- [9] Conover, W. J. Practical Nonparametric Statistics. Third Edition. John Wiley & Sons, Inc, 1999.
- [10] Canavos, George C. Probabilidad y Estadística. Aplicaciones y Métodos. México: McGraw-Hill, 1988.
- [11] Montgomery, Douglas C. and Peck, Elizabeth A. Introduction to Linear Regression Analysis. Second Edition. John Wiley & Sons, Inc, 1992.