

UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA



PROYECCIÓN DE DEMANDA UTILIZANDO SERIES DE TIEMPO

INFORME DE SUFICIENCIA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ELECTRICISTA

PRESENTADO POR:

ARTURO ISAAC ROMERO SUYO

PROMOCIÓN

2002 - II

LIMA – PERÚ

2008

PROYECCIÓN DE DEMANDA UTILIZANDO SERIES DE TIEMPO

DEDICATORIA

*A mis padres...
por su dedicación y
sus enseñanzas*

*A mis tíos y hermanos...
por su apoyo incondicional*

SUMARIO

Conocer el comportamiento de la demanda eléctrica a futuro es muy importante para las empresas de distribución de energía eléctrica, ya que esto les permitirá orientar y optimizar sus inversiones, para así garantizar un servicio continuo, confiable y de calidad a todos sus clientes.

Es por esta razón que la proyección de demanda es parte fundamental del planeamiento de las empresas de distribución de energía eléctrica.

En cuanto a los métodos de proyección, podemos encontrar dos grandes grupos denominados “métodos cualitativos” y “métodos cuantitativos”.

Este informe desarrolla los métodos de proyección de demanda basados en series de tiempo, dichos métodos pertenecen al grupo de métodos cuantitativos.

Los métodos de proyección basados en series de tiempo son herramientas empleadas en el planeamiento de los sistemas de distribución a nivel de media y baja tensión.

Se muestra una aplicación utilizando software E-Views con el método Arima, para proyectar la demanda de un alimentador de media tensión.

ÍNDICE

INTRODUCCIÓN	1
CAPITULO I	
REGRESIÓN LINEAL	2
1.1 Ajuste por Mínimos Cuadrados	3
1.2 Coeficiente de Correlación (r)	4
CAPITULO II	
DESCOMPOSICIÓN DE SERIES TEMPORALES	6
2.1 Modelo de Holt.	7
2.2 Modelo de Holt-Winters	9
2.3 Modelo Clásico de Descomposición	13
CAPITULO III	
MODELO ARIMA	20
3.1 Modelo Auto Regresivo (AR).	20
3.2 Modelo de Medias Móviles (MA).	21
3.3 Relación Entre Modelos AR y MA.	22
3.4 Estacionariedad de las Series Temporales	23
3.5 Orden de Integrabilidad de las Series Temporales	25
3.5.1 Procedimientos de Detección de Raíces Unitarias	26
3.6 Identificación del Modelo Arima	29
3.6.1 Funciones de Auto correlación (FAC)	34
3.6.2 Funciones de Auto correlación Parcial (FACP)	35
3.7 Estimación Del Modelo Arima	35
CAPITULO IV	
APLICACIÓN	37
4.1 Segmentación del Mercado	37
4.2 Proyección de Demanda	38
4.2.1 Datos del Alimentador M-01	38
4.2.2 Estacionariedad de la Serie	38
4.2.3 Transformación de la Serie	41
4.2.4 Identificación del Modelo	42

4.2.5 Proyección	44
CONCLUSIONES	46
ANEXOS	47
ANEXO A	
MANUAL DE INTRODUCCIÓN AL SOFTWARE EIEWS	48
BIBLIOGRAFÍA	83

INTRODUCCIÓN

Parte fundamental del planeamiento de los sistemas de distribución de energía eléctrica consiste en determinar las necesidades futuras de dicho sistema, para lo cual la proyección de la demanda es de vital importancia.

Las principales empresas de distribución de energía eléctrica del país cuentan con más de 250 alimentadores en media tensión, los cuales presentan sus propias características de carga haciéndolos en muchos casos únicos y para los cuales es necesario predecir su comportamiento a futuro.

En la actualidad se conocen muchos métodos de proyección que pueden denominarse cualitativos o cuantitativos según corresponda.

Este trabajo se enfoca en los métodos cuantitativos, específicamente en los métodos basados en series de tiempo, métodos que nos permiten realizar proyecciones en base a los valores históricos de la variable.

Se explica algunos métodos desde el mas sencillo y fácil de aplicar hasta los mas sofisticados que requieren de un software informático para su aplicación.

También se muestra una aplicación donde se proyecta la demanda para un alimentador de media tensión basado en el método ARIMA.

CAPITULO I REGRESIÓN LINEAL

La técnica de regresión lineal permiten hacer predicciones sobre los valores de cierta variable “Y” (*dependiente*), a partir de los valores de otra variable “X” (*independiente*), entre las que intuimos que existe una relación.

A partir de una muestra de tamaño “n”, para los que se dispone de los valores de ambas variables, $\{(X_i, Y_i), i = 1, \dots, n\}$, se puede visualizar gráficamente la relación existente entre ambas mediante un gráfico de dispersión, en el que los valores de la variable X se disponen en el eje horizontal y los de Y en el vertical.

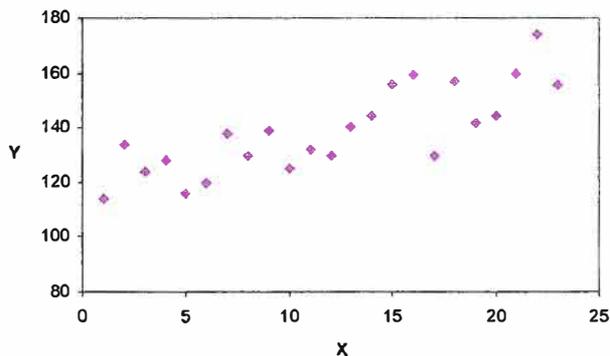


Fig. 1.1 Curva de Dispersión

El problema que subyace a la metodología de la regresión lineal es el de encontrar una recta que ajuste a la dispersión de puntos del diagrama así dibujado, y que pueda ser utilizada para predecir los valores de Y a partir de los de X. La ecuación general de la recta de regresión será entonces de la forma:

$$Y = a + bX + u \quad \dots\dots\dots (1.1)$$

Donde:

Y: Variable dependiente.

X: Variable Independiente.

a: Terminio independiente.

b: Coeficiente de la variable independiente.

u: Termino de error.

Para encontrar aquella recta que mejor se ajuste a los datos se recurre al método de mínimos cuadrados, que elige como recta de regresión a aquella que minimiza las distancias verticales de los puntos a la recta.

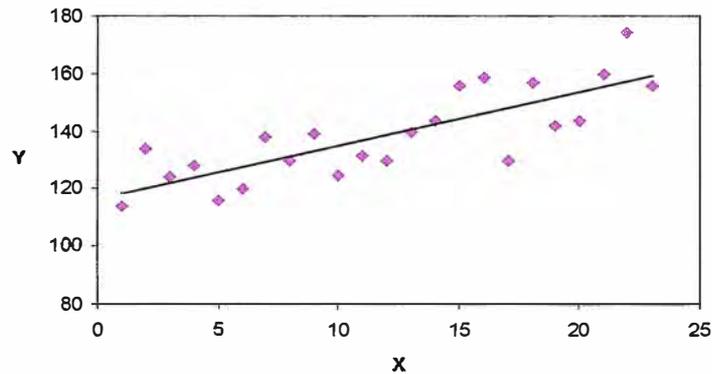


Fig. 1.2 Recta de Regresión

1.1 Ajuste por Mínimos Cuadrados

Supóngase que el conjunto de datos consiste en los puntos (x_i, y_i) siendo $i = 1, 2, \dots, n$. Queremos encontrar una función $Y = a + bX$, tal que Y sea la recta que mejor represente a los datos. Sin embargo, los errores experimentales siempre presentes hacen que no se hallen perfectamente alineados con los datos. El método de mínimos cuadrados permite determinar los valores de los parámetros a y b de la recta que mejor se ajusta a los datos experimentales.

Para encontrar los valores de a y b , que minimicen la suma de los cuadrados de los residuos $E(a, b)$, resolvemos este problema mediante un cálculo de diferenciación, con lo que se obtienen los estimadores mínimos cuadráticos de los coeficientes de la recta de regresión:

$$E(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2 \quad \dots\dots\dots (1.2)$$

$$\frac{\partial E}{\partial a} = 0 \quad \frac{\partial E}{\partial b} = 0 \quad \dots\dots\dots (1.3)$$

De donde:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad \dots\dots\dots (1.4)$$

$$b = \frac{\sum_{i=1}^n y_i - a \sum_{i=1}^n x_i}{n} \quad \dots\dots\dots (1.5)$$

Con los valores anteriores de a y b se dispone ya de la "mejor" recta de predicción, pero puede suceder que incluso esa mejor recta recoja una relación que no sea significativa, con lo que las predicciones serían inconsistentes. Necesitamos pues algún método para valorar la bondad del ajuste entre la serie y el modelo a ella ajustado.

1.2 Coeficiente de Correlación (r)

Se define por la siguiente expresión:

$$r = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{[n(\sum x_i^2) - (\sum x_i)^2][n(\sum y_i^2) - (\sum y_i)^2]}} \quad \dots\dots\dots (1.6)$$

El coeficiente de correlación, además de ser independiente de las unidades de medida de las variables, se caracteriza por tomar valores dentro del intervalo cerrado [-1, 1]

La interpretación del coeficiente de correlación depende del valor y del signo que tome, el valor absoluto del coeficiente de correlación indica la fuerza de la relación lineal. Un coeficiente de correlación muy cercano a uno en valor absoluto indica que la relación entre las variables es muy fuerte, mientras que si es muy cercano a cero, indica que la relación es muy débil. El siguiente cuadro muestra las posibles interpretaciones del coeficiente de correlación.

TABLA N°1.1 Coeficiente de Correlación

Coeficiente de Correlación	Interpretación
$0 < r < 1$ y $r \rightarrow 1$	Relación lineal positiva y fuerte
$0 < r < 1$ y $r \rightarrow 0$	Relación lineal positiva y débil
$r = 0$	No existe relación lineal
$-1 < r < 0$ y $r \rightarrow -1$	Relación lineal negativa y fuerte
$-1 < r < 0$ y $r \rightarrow 0$	Relación lineal negativa y débil

El coeficiente de correlación, no solamente mide el sentido de la relación entre las variables sino también la fuerza de la relación lineal o grado de asociación lineal. La siguiente figura relaciona el grado de asociación lineal con diversos valores del coeficiente de correlación r .

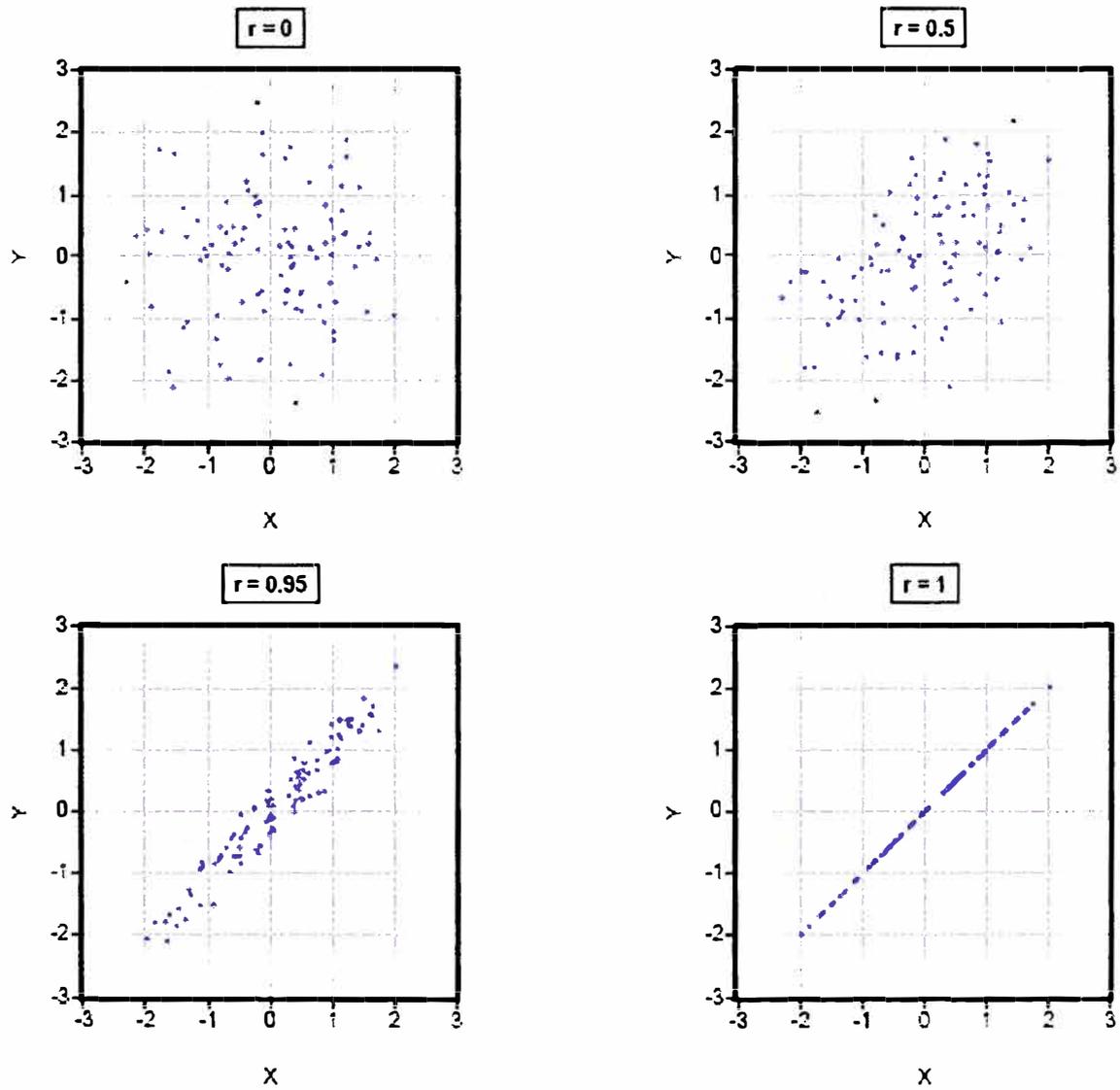


Fig. 1.3 Coeficientes de Correlación

CAPITULO II

DESCOMPOSICIÓN DE SERIES TEMPORALES

Una serie temporal es un conjunto de observaciones tomadas en instantes específicos, generalmente a intervalos iguales. Matemáticamente, una serie temporal se define como una función de la variable de tiempo. Los métodos de descomposición establecen que la serie temporal se compone de cuatro movimientos característicos: tendencia, estacionalidad, ciclicidad y aleatorio o Irregular.

- **Tendencia T(t)** Representa la evolución a largo plazo de la serie.
- **Ciclo C(t)** Movimientos oscilatorios en torno a la tendencia. La distinción entre tendencia y ciclo, resulta muchas veces problemática. Por otra parte muchos de los factores que afectan a la tendencia son responsables también del comportamiento cíclico, de forma que no es conveniente ni posible imponer una distinción clara, por esta razón se suele manejar habitualmente un componente de ciclo-tendencia compuesto por ambos.
- **Estacionalidad E(t)** Se trata de un movimiento periódico o cuasi periódico de duración inferior o igual al año. Viene determinado, principalmente, por factores institucionales, climáticos y técnicos.
- **Irregularidad, I(t)** Movimientos esporádicos y sin un patrón determinado. Dado que no contienen información relevante es necesaria su eliminación a fin de interpretar adecuadamente la evolución de la variable.

No todos los modelos que hacen uso de la descomposición de series tienen en cuenta los cuatro factores anteriores. Tenemos modelos desde el que emplea sólo los factores de tendencia y aleatorio (modelo de Holt), pasando por el que emplea los factores tendencial, estacional y aleatorio (modelo de Winters), hasta el que los tiene todos en cuenta (modelo clásico de descomposición).

2.1 Modelo de Holt.

El modelo de Holt pretende identificar la tendencia de la serie de un modo tal que permita que esa tendencia pueda variar a lo largo del tiempo, produciéndose un ajuste automático del modelo.

La expresión que nos permite obtener predicciones es muy sencilla:

$$Y'_{(t+m)} = S_{(t)} + m \times b_{(t)} \quad \dots\dots\dots (2.1)$$

Donde:

- $Y'_{(t+m)}$ es la previsión obtenida para el instante t+m, desde el instante t.
- $S_{(t)}$ es el nivel suavizado en que se encuentra la serie en el instante t.
- $b_{(t)}$ es la tendencia que presentaba la serie en instante t.

Por analogía con el modelo de regresión lineal, $b_{(t)}$ sería la pendiente de la recta.

Así pues, la previsión a "m" periodos consiste en incrementar el valor suavizado (Las técnicas de suavizado permiten eliminar las oscilaciones bruscas de la curva) que presenta la serie mas "m" veces la tendencia.

En el caso particular, pero importante por la frecuencia de su uso, de la previsión a un periodo ($m=1$), la expresión anterior quedaría:

$$Y'_{(t+1)} = S_{(t)} + b_{(t)} \quad \dots\dots\dots (2.2)$$

Obsérvese que el término correspondiente a la pendiente o tendencia de la serie va afectado por el subíndice temporal, indicando que puede variar con el tiempo. El cálculo de ambos componentes, nivel (S_t) y tendencia (b_t) se hará mediante las siguientes expresiones recurrentes:

$$S_{(t)} = \alpha \times Y_{(t)} + (1 - \alpha) \times (S_{(t-1)} + b_{(t-1)}) \quad 0 < \alpha < 1 \quad \dots\dots (2.3)$$

$$b_{(t)} = \beta \times (S_{(t)} - S_{(t-1)}) + (1 - \beta) \times b_{(t-1)} \quad 0 < \beta < 1 \quad \dots\dots (2.4)$$

Donde α y β son dos constantes de suavizado cuyo valor idóneo debe ser determinado en cada caso. Al elegir valores de inicialización para las fórmulas anteriores, es decir valores para S_1 y b_1 , tenemos varias opciones.

- En primer lugar podríamos aplicar un ajuste por regresión lineal y tomar como valor de S1 el término independiente y de b1 la pendiente.
- Otra posibilidad más sencilla, consiste en tomar para S1 el valor de Y1 y para b1 una aproximación de la pendiente inicial, por ejemplo: $b1=Y2-Y1$, $b1=(Y3-Y1)/2$, $b1=(Y4-Y1)/3$.

La mecánica de la obtención de previsiones es bien sencilla y puede resolverse sin complicaciones. El problema crucial es la selección de unos valores de α y β adecuado a cada caso. Aunque, como se dijo arriba, estas constantes toman valores entre cero y uno, diversos autores consideran que los resultados óptimos se suelen alcanzar en el intervalo $[0,1; 0,3]$. El criterio de selección debe ser la búsqueda de aquella pareja de valores que haga mínimo el error cuadrático medio de las previsiones.

Veamos una aplicación basándonos en las cifras anuales de las pérdidas de energía para una empresa de distribución en un periodo de 10 años.

TABLA Nº 2.1 Pérdidas De Energía

Año	Y(t)	S(t)	b(t)	Y'(t)
1	636.892	642.856	-23.624	
2	601.992	614.060	-24.917	619.232
3	571.018	583.706	-26.276	589.143
4	533.088	550.127	-28.102	557.429
5	515.706	520.129	-28.576	522.025
6	485.352	489.693	-29.041	491.553
7	473.281	464.441	-28.094	460.652
8	456.298	442.332	-26.597	436.347
9	434.490	421.361	-25.191	415.735
10	421.098	403.649	-23.321	396.171
11				380.328

Donde: $Y(t)$ = Pérdidas de energía (kWh-año)

En este ejemplo se han empleado como valores de inicialización el término independiente (para S1) y la pendiente (para b1) de la recta de regresión. Los valores de los coeficientes α y β han sido respectivamente 0,3 y 0,25.

A continuación se muestra gráficamente el resultado modelo de Holt.

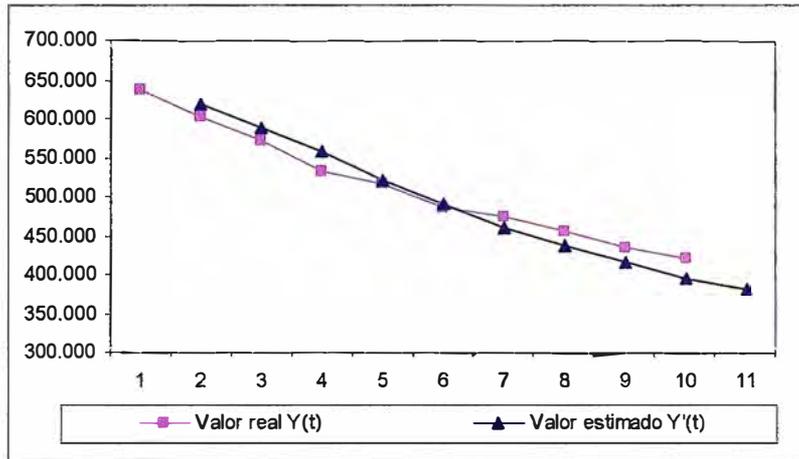


Fig. 2.1 Modelo de Holt

2.2 Modelo de Holt-Winters.

El modelo de Holt-Winters está pensado para series que presenten una tendencia aproximadamente lineal es su evolución. Ahora la expresión de obtención de las previsiones es:

$$Y'_{(t+m)} = T_{(t+m)} \times E_{(t+m-L)} \dots\dots\dots (2.5)$$

en el que se supone que

$$T_{(t+m)} = S_{(t)} + m \times b_{(t)} \dots\dots\dots (2.6)$$

Donde

- $Y'_{(t+m)}$ es la previsión para el periodo "t+m" realizada desde "t".
- $T_{(t+m)}$ es la previsión sin estacionalidad (previsión de la tendencia).
- $E_{(t+m-L)}$ es el factor estacional correspondiente al mismo periodo que se desea prever (t+m), pero en el ciclo anterior
- $S_{(t)}$ es el factor de nivel
- $b_{(t)}$ es factor de tendencia (pendiente)

Sustituyendo en la expresión anterior, resultará:

$$Y'_{(t+m)} = (S_{(t)} + m \times b_{(t)}) \times E_{(t+m-L)} \dots\dots\dots (2.7)$$

Las expresiones correspondientes para el cálculo de los distintos elementos de la previsión, similares a las del método de Holt, son:

$$S_{(t)} = \alpha \times Y_{(t)} + (1 - \alpha) \times (S_{(t-1)} + b_{(t-1)}) \quad 0 < \alpha < 1 \quad \dots\dots\dots (2.8)$$

$$b_{(t)} = \beta \times (S_{(t)} - S_{(t-1)}) + (1 - \beta) \times b_{(t-1)} \quad 0 < \beta < 1 \quad \dots\dots (2.9)$$

$$E_{(t)} = \gamma \times \left(\frac{Y_{(t)}}{S_{(t)}} \right) + (1 - \gamma) \times E_{(t-L)} \quad 0 < \gamma < 1 \quad \dots\dots (2.10)$$

Donde α , β y γ son las constantes de suavizado, a estimar en cada caso, y L es la longitud del ciclo estacional.

Así pues, es ahora necesario estimar tres parámetros, α , β y γ . El primero está relacionado con el factor aleatorio de la serie, el segundo con la tendencia (representa la pendiente) y el tercero con la componente estacional. La mayor capacidad de representación de este modelo frente al anterior se consigue a costa de una mayor complejidad en la estimación, que requiere la exploración de los valores de los tres parámetros usando como criterio de selección el valor de la varianza residual o la DAM (desviación absoluta media), junto con un buen ajuste gráfico entre la serie original y la suavizada. Los valores óptimos para esas constantes parecen encontrarse habitualmente entre 0,05 y 0,40. Si los valores de α y β superan el 0,5, es posible que el uso del método de Holt-Winters no sea adecuado. En la constante γ no son infrecuentes valores superiores a ese 0,5.

Para el arranque de las anteriores expresiones recursivas se tomarán los L primeros valores consecutivos para poder fijar los índices estacionales $E_{(t)}$. Para estimar el término de tendencia se suelen tomar $2L$ valores. Para el término de nivel se puede emplear una media móvil centrada, de longitud el ciclo estacional. Para esta inicialización se tomaría:

$$S_{(L)} = M \times M_{(L)} \quad \dots\dots\dots (2.11)$$

$$b_{(t)} = \frac{1}{L} \times \left(\frac{Y'_{(L+1)} - Y'_1}{L} + \frac{Y'_{(L+2)} - Y'_2}{L} + \dots + \frac{Y'_{(L+L)} - Y'_L}{L} \right) \quad \dots\dots (2.12)$$

$$E_i = \frac{Y_i}{\sum_{i=1}^L Y_i} \times L \quad i=1,2,\dots,L \quad \dots\dots\dots (2.13)$$

Obsérvese que el valor de inicialización de los E_i corresponde a tomar como factores estacionales iniciales el reparto, en tanto por uno, del total del primer ciclo entre los distintos periodos del mismo. La inicialización de $S_{(t)}$ y de $b_{(t)}$ nos dará los valores para el periodo L , con lo cual la primera previsión a un periodo será para el periodo $L+1$. La inicialización de $b_{(t)}$ busca la pendiente promedio de los primeros $2L$ datos, pero emparejando cada dato del primer ciclo con su correspondiente del segundo, a fin de que la oscilación estacional no enmascare la identificación de la pendiente.

Veamos una aplicación basado en la evolución de las pérdidas de energía durante 36 meses. El valor elegido para los parámetros α , β y γ ha sido respectivamente 0,12, 0,14 y 0,6. Los valores de inicialización de $E(t)$ y $b(t)$ han sido calculados en base a las anteriores expresiones, y el de $S(t)$ tomando una media móvil de orden doce, seguida de otra media móvil de orden dos para centrar el valor suavizado.

TABLA N° 2.2 Pérdidas de Energía

Mes	Y(t)	S(t)	b(t)	E(t)	Y'(t)
1	42.472			0,988	
2	39.145			0,911	
3	43.587			1,014	
4	43.797			1,019	
5	46.998			1,094	
6	43.133			1,004	
7	44.357			1,032	
8	42.804			0,996	
9	43.470			1,012	
10	43.131			1,004	
11	41.044			0,955	
12	41.768	41.706	-211	0,996	
13	40.786	41.889	-156	0,949	41.009
14	37.137	41.119	-242	0,948	38.013
15	41.228	40.827	-249	1,014	41.459
16	39.187	40.009	-328	1,025	41.354
17	43.650	40.137	-264	1,054	43.394
18	40.679	39.818	-272	1,026	40.019
19	43.012	39.982	-211	1,044	40.817
20	41.714	39.947	-186	1,031	39.612
21	41.129	39.907	-166	1,020	40.218
22	40.170	40.020	-127	0,984	39.885
23	37.402	39.611	-166	0,965	38.100
24	39.258	39.678	-134	0,973	39.295
25	39.134	39.755	-104	0,970	37.510
26	37.579	39.342	-147	0,979	37.572
27	40.130	39.188	-148	1,025	39.725
28	39.140	38.812	-180	1,027	40.021
29	42.297	38.943	-137	1,062	40.716
30	39.675	38.711	-150	1,033	39.809
31	39.610	38.543	-152	1,029	40.253
32	39.725	38.458	-143	1,028	39.586
33	39.821	38.572	-107	1,013	39.074
34	40.731	38.914	-44	1,014	37.860
35	37.813	38.869	-44	0,973	37.510
36	37.626	38.822	-45	0,969	37.779
37					37.601

La siguiente gráfica muestra la serie original y la serie de previsiones del modelo de Holt Winters, pudiendo apreciarse un buen nivel de ajuste entre ambas, a pesar de ser esta una serie en la que el comportamiento estacional no tiene excesiva fuerza.

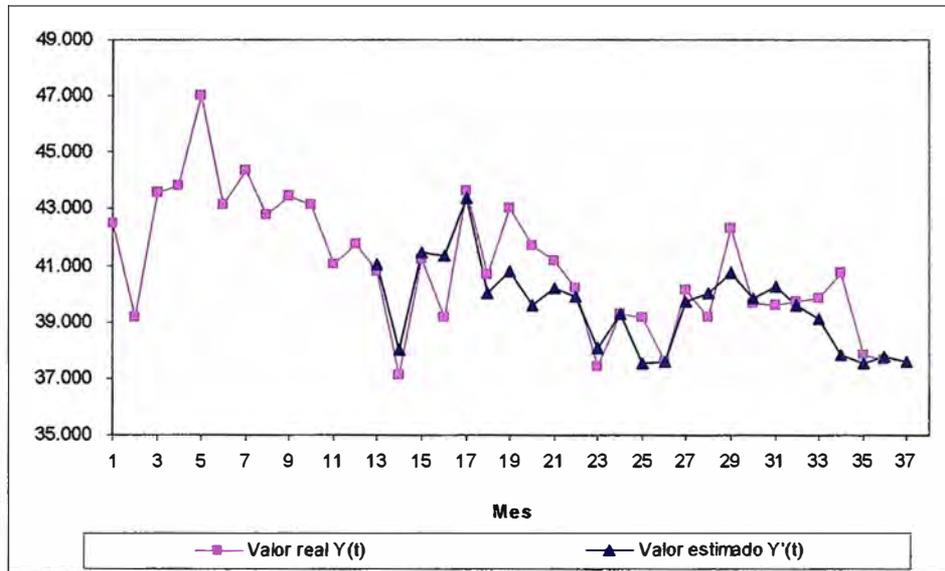


Fig. 2.2 Modelo Holt-Winters

2.3 Modelo Clásico de Descomposición.

Se basa en la consideración de que el comportamiento de la serie es el resultado de la acción de cuatro componentes o factores: tendencia, estacionalidad, ciclo y componente aleatorio, actuando del modo siguiente:

$$Y_{(t)} = T \times E \times C + u \quad \dots\dots\dots (2.14)$$

Donde:

- Y(t) es la serie temporal.
- T es el factor de tendencia.
- E es el factor estacional.
- C es el factor cíclico
- u es el factor aleatorio o ruido que afecta a la serie.

La diferencia que existe en el término E y término C, estacionalidad y ciclo respectivamente, afecta a la longitud del ciclo, y no al carácter repetitivo del comportamiento que representan. El ciclo estacional tiene una longitud inferior al cíclico, que debe ser varias veces el anterior. En una serie mensual, por ejemplo, el factor estacional representaría los ciclos de longitud doce (un año), mientras el factor cíclico

recogería oscilaciones repetitivas de mayor periodo, como los cinco a siete años con que muchas series económicas reflejan los ciclos de expansión- recesión.

El modelo anterior es el denominado multiplicativo, por actuar los distintos factores (salvo el ruido) de tal modo. Es el que seguiremos en los párrafos siguientes. Cabe de todos modos plantearse una notación alternativa, que constituye el modelo llamado aditivo:

$$Y_{(t)} = T + E + C + u \quad \dots\dots\dots (2.15)$$

Nos centrándonos en el primero de ellos y veremos el procedimiento del cálculo de cada uno de ellos, pues una vez que los tengamos, la obtención de la previsión para un periodo "t" es muy sencilla:

$$Y'_{(t)} = T_{(t)} \times E_{(t)} \times C_{(t)} \quad \dots\dots\dots (2.16)$$

Como se aprecia se elimina el elemento aleatorio, impredecible por naturaleza. Según se ve en esta expresión de predicción, para obtener la previsión Y'(t), necesitamos disponer de los valores previstos de cada uno de los tres elementos que la componen.

- El término de tendencia se obtiene mediante la realización de un ajuste de la serie a una recta.
- El cálculo de término estacional es una operación más compleja que requiere el cálculo de medias móviles de longitud un ciclo estacional L, obteniendo para cada instante el valor S_(t) correspondiente. Una vez se tiene esta media móvil S_(t) se calcula el cociente entre cada dato y la media móvil correspondiente:

$$E_{(t)} = \frac{Y_{(t)}}{S_{(t)}} \quad \dots\dots\dots (2.17)$$

A continuación se calcula el promedio de los índices anteriores para cada uno de los periodos del ciclo (por ejemplo, en una serie mensual, el promedio de los eneros, el de los febreros, ...) y este valor se corrige de modo que la suma total sea L.

- Por último, el factor cíclico se estima mediante los cocientes entre la serie suavizada y la tendencia, mediante:

$$C_{(t)} = \frac{S_{(t)}}{T_{(t)}} \quad \dots\dots\dots (2.18)$$

En la elaboración de la previsión a cada periodo se le aplica el coeficiente asociado al periodo del ciclo estacional que corresponda (en el ejemplo mensual, a todos los eneros se les aplicará el coeficiente estacional medio de enero, etc.).

A continuación se muestra una aplicación donde se detalla en una tabla: la serie original, tendencia, serie suavizada, cociente entre original y suavizada, y factor cíclico, así como la previsión obtenida para la serie.

TABLA N° 2.3 Modelo Clásico de Descomposición

Z(t)	S(t)	E(t)	T(t)	C(t)	x(t)
10.997			15.307		
9.053			15.357		
14.479			15.408		
15.016			15.458		
15.995			15.508		
16.428			15.558		
25.921	16.346	1,586	15.609	1,047	22.374
21.575	16.288	1,325	15.659	1,040	23.706
21.514	16.269	1,322	15.709	1,036	22.669
19.088	16.313	1,170	15.759	1,035	18.547
10.022	16.370	0,612	15.809	1,035	10.886
16.067	16.349	0,983	15.860	1,031	16.045
10.298	16.524	0,623	15.910	1,039	9.784
8.824	16.173	0,546	15.960	1,013	8.492
15.003	16.210	0,926	16.010	1,012	14.417
15.707	16.427	0,956	16.061	1,023	14.284
15.742	16.301	0,966	16.111	1,012	16.772
18.531	16.342	1,134	16.161	1,011	17.878
21.704	16.462	1,318	16.211	1,015	22.532
22.020	16.401	1,343	16.262	1,009	23.871
24.113	16.391	1,471	16.312	1,005	22.839
17.586	16.515	1,065	16.362	1,009	18.777
10.514	16.308	0,645	16.412	0,994	10.845
17.500	16.303	1,073	16.462	0,990	16.000
9.569	16.399	0,584	16.513	0,993	9.710
8.698	16.363	0,532	16.563	0,988	8.592
16.495	16.706	0,987	16.613	1,006	14.859
13.226	16.619	0,796	16.663	0,997	14.451
15.681	16.732	0,937	16.714	1,001	17.215
19.678	16.830	1,169	16.764	1,004	18.412
21.270	16.638	1,278	16.814	0,990	22.773
26.146	16.648	1,571	16.864	0,987	24.231
23.070	16.596	1,390	16.914	0,981	23.125
18.933	16.415	1,153	16.965	0,968	18.663
11.698	16.525	0,708	17.015	0,971	10.989
15.194	16.748	0,907	17.065	0,981	16.437
9.689	16.577	0,584	17.115	0,969	9.816
8.070	16.633	0,485	17.166	0,969	8.734
14.326	16.698	0,858	17.216	0,970	14.851
14.547	16.756	0,868	17.266	0,970	14.570
18.356	16.824	1,091	17.316	0,972	17.310
17.630	16.838	1,047	17.366	0,970	18.420
21.940	16.950	1,294	17.417	0,973	23.199
26.920	16.971	1,586	17.467	0,972	24.700
23.765	17.072	1,392	17.517	0,975	23.788
19.753	17.011	1,161	17.567	0,968	19.341
11.862	17.039	0,696	17.618	0,967	11.331
16.536	17.159	0,964	17.668	0,971	16.839
9.941	17.193	0,578	17.718	0,970	10.180
9.284	17.230	0,539	17.768	0,970	9.047
13.596	17.250	0,788	17.819	0,968	15.343
14.889	17.327	0,859	17.869	0,970	15.066
19.786	17.618	1,123	17.919	0,983	18.126
18.045	17.564	1,027	17.969	0,977	19.214
22.378	17.549	1,275	18.019	0,974	24.020
27.169		1,456	18.070	0,974	25.608
24.679		1,393	18.120	0,975	24.610
23.248		1,137	18.170	0,976	20.164
11.213		0,665	18.220	0,977	11.838
16.356		0,981	18.271	0,978	17.547

Se recogen los términos $E_{(t)}$ agrupados por meses, a fin de ilustrar el cálculo de los coeficientes estacionales provisionales E_i' (promedio) y los definitivos E_i^* o ajustados (con $i=1, \dots, 12$). Estos últimos son los empleados en la obtención de la previsión.

TABLA N° 2.4 Coeficientes Estacionales

MES	E(t)					E'	E*
	'02	'03	'04	'05	'06		
E		0,623	0,584	0,584	0,578	0,592	0,592
F		0,546	0,532	0,485	0,539	0,526	0,525
M		0,926	0,987	0,858	0,788	0,890	0,889
A		0,956	0,796	0,868	0,859	0,870	0,869
M		0,966	0,937	1,091	1,123	1,029	1,029
J		1,134	1,169	1,047	1,027	1,094	1,094
J	1,586	1,318	1,278	1,294		1,369	1,369
A	1,325	1,343	1,571	1,586		1,456	1,456
S	1,322	1,471	1,390	1,392		1,394	1,393
O	1,170	1,065	1,153	1,161		1,137	1,137
N	0,612	0,645	0,708	0,696		0,665	0,665
D	0,983	1,073	0,907	0,964		0,982	0,981
SUMA						12,004	12,000

A continuación se muestra un gráfico con el factor estacional, apreciándose el diferente comportamiento de los distintos meses. El valle que corresponde al mes de febrero es debido en parte al hecho de que se trata de un mes más corto que los otros, pues tal factor produce un descenso del número de días disponibles cercano al 10%. Otro factor que influye sin duda en el comportamiento de la serie es el número de fines de semana que tenga cada mes (valor que oscila entre cuatro y cinco). También la posición de la Semana Santa, y de las fiestas móviles que de ella dependen, influye con seguridad en la serie, tal como se aprecia en los factores estacionales de los meses de marzo y abril de la tabla anterior.

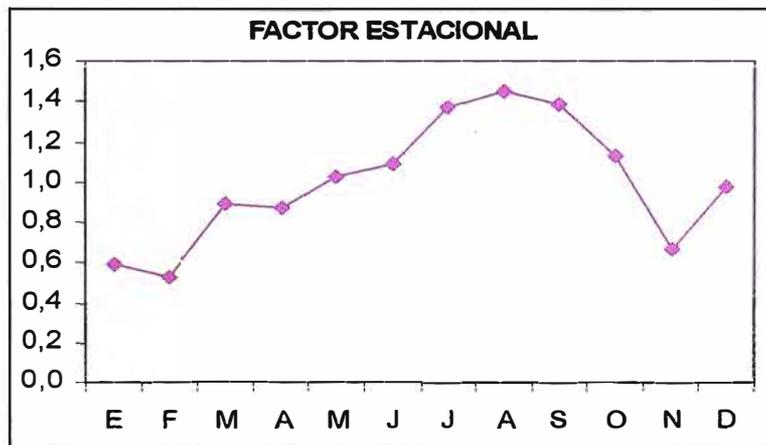


Fig. 2.3 Factor Estacional

A continuación se muestra un gráfico con el factor cíclico, donde se observa lo que podría ser una parte de un ciclo, aunque con el número de años empleados no se puede apreciar si existe comportamiento cíclico o no. Caso de haberlo, parece que tendría un periodo de unos siete años.

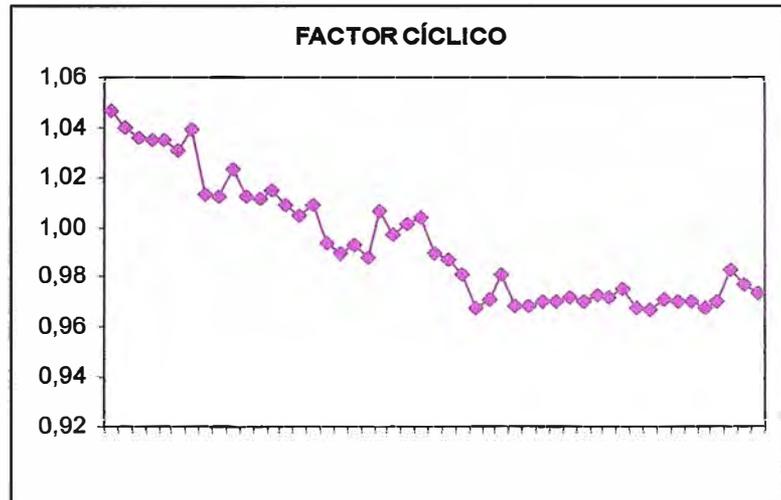


Fig. 2.4 Factor Cíclico

La siguiente gráfica muestra el excelente nivel de ajuste que esta técnica consigue, al menos en lo que se refiere a los datos pasados. De cara a la obtención de previsiones, el mayor problema que se presenta es la dificultad en extrapolar valores del factor cíclico, sobre todo cuando no está clara la longitud del ciclo, cosa que ocurre en esta aplicación y en general con bastante frecuencia.

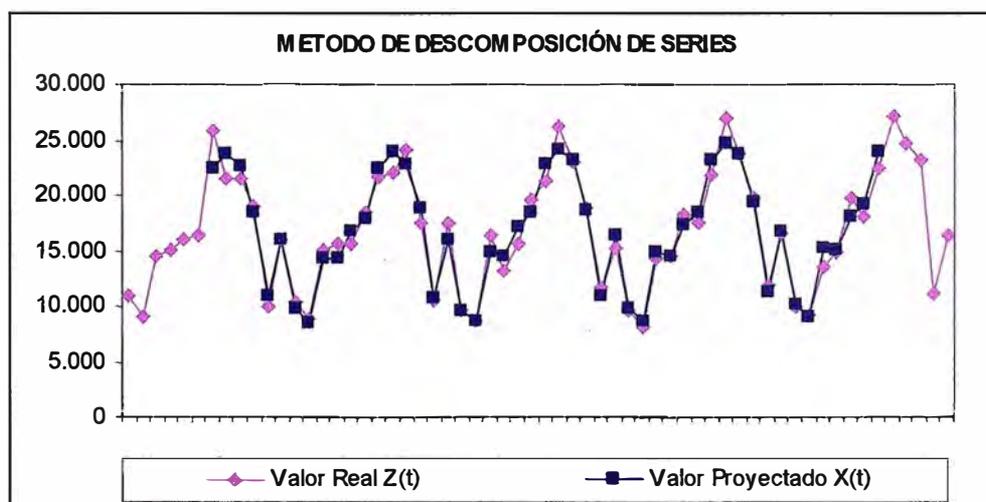


Fig. 2.5 Método Clásico de Descomposición

Resumiendo, entre las ventajas del método clásico de descomposición de series temporales está su sencillez conceptual y su buen comportamiento en series que tengan

una estacionalidad homogénea (estable). Como inconvenientes tenemos la necesidad de disponer de series de al menos tres años si decidimos no dar importancia al factor cíclico, y de por lo menos el doble si esperamos identificar este factor.

CAPITULO III MODELO ARIMA

La palabra ARIMA significa Modelos Autorregresivos Integrados de Medias Móviles, los cuales analizaremos por separado.

3.1 Modelo Auto Regresivo (AR).

Definimos un modelo como autorregresivo si la variable endógena $Y_{(t)}$ de un período t es explicada por las observaciones de ella misma correspondientes a períodos anteriores añadiéndose un término de error.

Los modelos auto regresivos se abrevian con la palabra AR(p) donde (p) indica el orden del modelo: AR(1), AR(2),....etc.

El orden del modelo expresa el número de observaciones retasadas de la serie temporal analizada que intervienen en la ecuación. Así, por ejemplo, un modelo AR(1) tendría la siguiente expresión:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + a_t \quad \dots\dots\dots (3.1)$$

El término de error de los modelos de este tipo se denomina generalmente *ruido blanco* cuando cumple las tres hipótesis básicas tradicionales:

- Media nula
- Varianza constante
- Covarianza nula entre errores correspondientes a observaciones diferentes

La expresión genérica de un modelo autorregresivo, no ya de un AR(1) sino de un AR(p) sería la siguiente:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots\dots + \phi_p Y_{t-p} + a_t \quad \dots\dots\dots (3.2)$$

Pudiéndose escribir de forma abreviada como:

$$\phi_p(L)Y_t = \phi_0 + a_t \quad \dots\dots\dots (3.3)$$

Donde $\phi_p(L)$ es lo que se conoce como operador polinomial de retardos:

$$\phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad \dots\dots\dots (3.4)$$

y donde, a su vez, el término L es lo que se conoce como operador retardo tal que, aplicado al valor de una variable en t, dé como resultado el valor de esa misma variable en t-1, es decir:

$$LY_t = Y_{t-1} \quad \dots\dots\dots (3.5)$$

y aplicado sucesivamente p veces retarda el valor en p períodos

$$L^p Y_t = Y_{t-p} \quad \dots\dots\dots (3.6)$$

Normalmente, se suele trabajar con modelos autorregresivos de órdenes bajos: AR(1) o AR(2), o bien con órdenes coincidentes con la periodicidad de los datos de la serie analizada (si es trimestral AR(4), si es mensual AR(12)....).

3.2 Modelo de Medias Móviles (MA)

Un modelo denominado de medias móviles es aquel que explica el valor de una determinada variable en un período t en función de un término independiente y una sucesión de errores correspondientes a períodos precedentes, ponderados convenientemente. Estos modelos se denotan normalmente con las siglas MA(q), donde (q) como en el caso de los modelos autor regresivos, es el orden. Así, un modelo con q términos de error MA(q) respondería a la siguiente expresión:

$$Y_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad \dots\dots\dots (3.7)$$

Que de nuevo puede abreviarse utilizando el polinomio de retardos (como en caso de los modelos AR):

$$Y_t = \theta_q(L) a_t + \mu \quad \dots\dots\dots (3.8)$$

Al igual que en el caso de los modelos autorregresivos, el orden de los modelos de medias móviles suele ser bajo MA(1), MA(2) o corresponderse con la periodicidad de los datos analizados MA(4), para series trimestrales, o MA(12) para series mensuales.

3.3 Relación Entre Modelos AR y MA

Bajo ciertas condiciones los modelos AR y los MA pueden relacionarse. Estas condiciones se denominan de invertibilidad y de estacionariedad. La explicación de tales características requiere de un instrumental matemático específico, por lo que tan solo haremos mención a la intuición de las mismas, sabiendo que se trata de características, por lo demás, fáciles de encontrar desde el punto de vista práctico en las series a modelizar.

Por ejemplo, supongamos un modelo AR(1) sin término independiente como:

$$Y_t = \phi Y_{t-1} + a_t \quad \dots\dots\dots (3.9)$$

Puesto que

$$Y_{t-1} = \phi Y_{t-2} + a_{t-1} \quad \dots\dots\dots (3.10)$$

Reemplazando (3.9) en (3.10)

$$Y_t = a_t + \phi a_{t-1} - \phi^2 Y_{t-2} \quad \dots\dots\dots (3.11)$$

podríamos llegar, por sustituciones sucesivas a una expresión como:

$$Y_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + \dots + \phi^j a_{t-j} + \dots \quad \dots\dots (3.12)$$

Es decir, un proceso autorregresivo de primer orden es equivalente a una media móvil de infinitos términos con una ponderación decreciente en forma exponencial, cuando

$$0 < |\phi| < 1 \quad \dots\dots\dots (3.13)$$

Este resultado es generalizable, y puede demostrarse que bajo las condiciones de *estacionariedad* un modelo AR de orden reducido puede transformarse en un modelo MA de orden elevado o, teóricamente, infinito. De igual forma, bajo las condiciones de *invertibilidad*, modelos MA de orden reducido pueden aproximarse por modelos AR de un número suficientemente elevado de términos.

La utilidad de esta relación estriba en que, por aplicación de un principio de parsimonia o economicidad (y sin olvidar que, en la práctica, puede favorecer la estimación del modelo) es preferible un modelo sencillo, con el menor número posible de términos y, por lo tanto, de parámetros a estimar, frente a un modelo con un gran número de coeficientes, siempre que, por supuesto, nos conduzca a resultados similares.

3.4 Estacionariedad de las Series Temporales

Para la obtención de estimaciones con propiedades estadísticas adecuadas de los parámetros de un modelo ARMA, es necesario que la serie muestral que utilizamos para la estimación sea estacionaria en media y varianza. En un sentido laxo del término diríamos que precisamos que la serie no tenga tendencia, y que presente un grado de dispersión similar en cualquier momento de tiempo.

Cuando, queremos inferir a partir de una serie temporal (muestra) la estructura del proceso estocástico mediante modelos AR ó MA, debemos cubrir dos etapas:

Etapa 1, Asegurarnos de que la serie temporal, como muestra del proceso estocástico, es estacionaria y, si no lo es,

Etapa 2, Transformar la serie temporal original de forma que la nueva serie transformada si lo sea.

¿Qué entendemos por estacionariedad de una serie?

La estimación de los parámetros de un modelo ARIMA exige que el proceso estocástico subyacente que generó la serie sea invariable con respecto al tiempo. Si las características del proceso estocástico cambian con el tiempo, es decir, si el proceso no es estacionario, no resultaría posible la obtención de una estimación con propiedades estadísticas razonables.

Para definir la estacionariedad de una serie, disponemos de dos posibilidades:

Por una parte se dice que un proceso es *estrictamente estacionario* si sus propiedades, como se dijo antes, no se ven afectadas por cambios de origen temporal, esto es, cuando al realizar un mismo desplazamiento en el tiempo de todas las variables de cualquier distribución conjunta finita, resulta que esta distribución no varía. Considerando la función de distribución conjunta.

$$F(Y_{t1}, Y_{t2}, \dots, Y_{tk}) \dots\dots\dots (3.14)$$

Si se adopta el supuesto de que todos los elementos de la anterior distribución se desplazan m periodos, la nueva distribución de probabilidad será

$$F(Y_{t1+m}, Y_{t2+m}, \dots, Y_{tk+m}) \dots\dots\dots (3.15)$$

Pues bien, el proceso será estocástico en sentido estricto si se verifica que

$$F(Y_{t1}, Y_{t2}, \dots, Y_{tk}) = F(Y_{t1+m}, Y_{t2+m}, \dots, Y_{tk+m}) \quad \dots\dots (3.16)$$

y ello para cualquier distribución conjunta finita.

Por ejemplo, para $k=1$ se tendría que la distribución de probabilidad marginal en el momento t sería la misma que la distribución de probabilidad marginal de cualquier otro punto en el tiempo, $p(Y_t) = p(Y_{t+m})$, lo cual implica que la media y la varianza de $Y_{(t)}$ deben ser constantes:

$$E(Y_1) = E(Y_2) = \dots = E(Y_t) = \int_{-\infty}^{\infty} y \cdot p(y) dy \quad cte. \quad \dots\dots (3.17)$$

$$\text{var}(Y_1) = \text{var}(Y_2) = \dots = \text{var}(Y_t) = \sigma_y^2 = \int_{-\infty}^{\infty} (y - \mu)^2 p(y) dy \quad cte. \quad \dots (3.18)$$

Si $k=2$, la estacionariedad implica que ninguna de las distribuciones bivariantes $p(Y_t, Y_{t-m})$ dependen del tiempo t , lo cual a su vez implica que las covarianzas son función solamente del intervalo m , y no del tiempo t , y ello para todo m .

$$\begin{aligned} \text{cov}(Y_1, Y_{1+m}) &= \text{cov}(Y_2, Y_{2+m}) = \dots = \text{cov}(Y_n, Y_{n+m}) \\ &= \text{cov}(Y_t, Y_{t-m}) \end{aligned} \quad \dots\dots (3.19)$$

Por tanto, la hipótesis de estacionariedad implica que la media y la varianza del proceso sean constantes y que las auto covarianzas:

$$\gamma_k = \text{cov}(Y_t, Y_{t-k}) = E[(Y_t - \mu) \cdot (Y_{t-k} - \mu)] \quad \dots\dots (3.20)$$

y las auto correlaciones:

$$\rho_k = \frac{\text{cov}(Y_t, Y_{t-k})}{\sqrt{\text{var}(Y_t) \cdot \text{var}(Y_{t-k})}} = \frac{\gamma_k}{\gamma_0} \quad \dots\dots\dots (3.21)$$

dependen solamente del retardo k , y no del momento del tiempo.

Ahora bien, la estacionariedad en sentido estricto conlleva, entre otras cuestiones, la necesidad de conocer la función de distribución del proceso estocástico, lo cual no es posible. Por ello el concepto de estacionariedad que se asume es en un *sentido débil*, o de segundo orden, el cual se obtiene cuando la media y varianza son constantes e independientes del tiempo, y que la auto covarianza y, por tanto, los coeficientes de autocorrelación sólo dependen del desfase, pero no del momento de tiempo.

Lo cierto es que una serie estacionaria en sentido estricto, lo será también en sentido débil. Además, cuando el proceso estacionario en sentido débil se distribuye como una normal, entonces se demuestra que también es estacionario en sentido estricto. Todo ello es suficiente para que a efectos prácticos nos conformemos con este concepto amplio de estacionariedad.

La clave de la importancia de la estacionariedad en series económicas estriba en que, en economía, no se dispone de la capacidad de experimentación, sino de la de observación. Ello implica que las series económicas que queremos modelizar se constituyen como muestras de tamaño 1. Efectivamente, aunque disponemos de n datos (uno por año), esto es sólo *una* realización (una serie) del proceso estocástico subyacente. En principio ello nos haría imposible la estimación, puesto que contaríamos con más parámetros (medias, varianzas, covarianzas... para cada año) que datos de la serie. Ahora bien, ello sí es posible cuando la serie es estacionaria y ergódica, propiedad que garantiza la consistencia de los estimadores.

Aunque la mayor parte de las series económicas presentan la característica de la ergodicidad (por lo que, a efectos prácticos no suele ser objeto de atención), lo cierto es que también presentan un grado de dispersión variable en unos y otros momentos del tiempo y, en general, acusada tendencia. Estas circunstancias son una señal inequívoca del incumplimiento de la condición de estacionariedad. Ahora bien, afortunadamente existen sencillos tratamientos (diferenciación) que, aplicados a los datos, pueden solventar razonablemente bien estos problemas.

3.5 Orden de Integrabilidad de las Series Temporales

Para la determinación del orden de integrabilidad, esto es, para determinar el número de veces que será necesario diferenciar la serie para hacerla estacionaria en media, existen dos procedimientos fundamentales de *detección del número de raíces unitarias* (aparte de la simple representación gráfica, que no es más que un método intuitivo), como son el Test Dickey-Fuller, o Test Dickey-Fuller Aumentado, (abreviadamente DF o ADF respectivamente) y el Test de Phillips-Perron (Test PP).

3.5.1 Procedimientos de Detección de Raíces Unitarias

Tal y como se ha señalado repetidamente, la aplicación de la metodología ARIMA precisa la utilización de series económicas estacionarias en varianza y en media.

Cuando una serie no es estacionaria en media, o lo que es lo mismo, cuando no es integrada de orden cero $I(0)$, se dice que presenta al menos una raíz unitaria. Cuando esto ocurre, sabemos que es posible la obtención de una serie estacionaria mediante una sencilla transformación de la serie original, como es la diferenciación adecuada. Pues bien, el número de diferencias que habrá que tomar en la serie para convertirla en estacionaria en media viene dado, justamente, por el número de raíces unitarias que la serie original presente. Pues bien, el número de raíces unitarias de la serie se determina formalmente mediante los procedimientos de Dickey-Fuller (con su versión ampliada) y de Phillips Perron.

El planteamiento más sencillo de Dickey y Fuller es el siguiente. Plantear un modelo autorregresivo AR(1) como:

$$Y_t = \mu + \rho Y_{t-1} + \varepsilon_t \quad \dots\dots\dots (3.22)$$

Donde μ y ρ son parámetros a estimar y ε_t es un término de error para el que se asume que cumple las propiedades de ruido blanco. Dado que ρ es un coeficiente de autocorrelación que, por tanto, toma valores entre $-1 < \rho < 1$, si $\rho = 1$, la serie Y_t es no estacionaria. Si el valor de ρ es mayor que la unidad, entonces se dice que la serie es explosiva. De esta forma, la hipótesis de estacionariedad de una serie puede ser evaluada analizando si el valor absoluto de ρ es estrictamente menor que 1. Pues bien, el test DF plantea, sencillamente, contrastar estadísticamente si $\rho = 1$. Puesto que en economía las series explosivas no tienen mucho sentido, esta hipótesis nula se analiza frente a la alternativa que establece que $H_1: \rho < 1$.

Se puede demostrar que una expresión equivalente a la anterior, mediante la toma de diferencias, viene dada por:

$$\Delta Y_t = \mu + \gamma Y_{t-1} + \varepsilon_t \quad \dots\dots\dots (3.23)$$

donde $\gamma = \rho - 1$ y las hipótesis nula y alternativa son, respectivamente, $H_0: \gamma = 0$ y $H_1: \gamma < 0$. Aunque el test calcula un estadístico t sobre el valor estimado de γ , bajo la hipótesis nula de existencia de una raíz unitaria, este estadístico no tiene la distribución convencional de la *t-Student*. Por ello, Dickey y Fuller simulaban los valores críticos para una selección de distintos tamaños muestrales, simulación ampliada posteriormente por MacKinnon. Estas son las tablas que se utilizan para determinar los valores críticos del test DF.

El problema del test DF simple es que asume que el proceso estocástico subyacente a los datos sigue un AR(1). Por tanto, cuando el proceso siga otro esquema, la estimación de la regresión auxiliar del test nos arrojará un incumplimiento de la condición de ruido blanco para los residuos, debido a una mala especificación. Por lo tanto, aunque el primer paso para la detección de raíces unitarias debe ser especificar un test DF simple debe, a continuación contrastar estadísticamente las características de ruido blanco de los residuos (estadístico Q de Ljung-Box).

Para estos casos disponemos del test ampliado de Dickey-Fuller (ADF), el cual permite considerar otros esquemas de autocorrelación. Para ello, el test ADF añade términos diferenciados de la variable dependiente y en el lado derecho de la regresión:

$$\Delta Y_t = \mu + \gamma Y_{t-1} + \delta_1 \Delta Y_{t-1} + \delta_2 \Delta Y_{t-2} + \dots + \delta_{p-1} \Delta Y_{t-p+1} + \varepsilon_t \quad \dots\dots\dots (3.24)$$

Esta especificación aumentada del test se basa en las siguientes hipótesis nula y alternativa:

$$H_0: \gamma=0 \text{ y } H_1: \gamma<1 \quad \dots\dots\dots (3.25)$$

Un resultado importante obtenido por Fuller es que la distribución asintótica del estadístico t sobre γ es independiente del número de retardos de la serie en primeras diferencias incluidos en la regresión del test ADF. Además, aunque la asunción de que y sigue un proceso autorregresivo puede parecer restrictiva, lo cierto es que Said y Dickey demostraron que el test ADF continúa siendo válido incluso cuando la serie presenta un componente de medias móviles (MA), dado que se aumentan en la regresión el número suficiente de términos de diferencias retardadas.

Un aspecto igualmente importante de estos test DF y ADF pasa por la inclusión o no de términos adicionales en la regresión como un término independiente (como hemos hecho en la explicación hasta el momento), una constante de tendencia o ambos. Esto es un asunto importante, puesto que la distribución asintótica del estadístico del test bajo la hipótesis nula depende de la inclusión o no que hagamos de estos términos deterministas. Así, si incluimos una constante en el test de regresión, el estadístico t tiene una distribución no estándar si el proceso subyacente contiene una raíz unitaria con una constante igual a cero. Si incluimos un término constante más una tendencia lineal en la regresión, entonces el estadístico t tiene una distribución no estándar si el proceso subyacente contiene una raíz unitaria con una tendencia lineal cero. La distribución asintótica varía cuando estos planteamientos no son satisfechos. Por ejemplo, si incluimos una constante en el test de regresión y si el proceso subyacente contiene una

raíz unitaria con una constante distinta de cero, entonces el estadístico t tiene una distribución asintótica normal estándar bajo la hipótesis nula de raíz unitaria.

Dado que se presenta la duda sobre si incluir o no un término independiente acompañado o no de un término de tendencia en la regresión del test, podemos plantearnos a priori incluir ambos términos (constante y tendencia) y considerar que los otros casos son situaciones particulares de esta especificación general. Sin embargo, hay que considerar que si se incluyen regresores irrelevantes en el test se reduce el poder de contraste del mismo, concluyendo posiblemente en la existencia de una raíz unitaria cuando, en realidad, no existe.

En este sentido, podemos aplicar las siguientes normas de carácter general. Si la serie original presenta tendencia se deberían incluir ambos términos en la regresión, es decir, considerar como regresores al término independiente y al término de tendencia lineal. Si la serie no parece presentar tendencia y tiene un valor medio distinto de cero, deberíamos incluir un término constante en la regresión. Finalmente, si la serie parece fluctuar en torno al valor medio cero, no se considera necesario incluir ningún regresor adicional en la regresión, es decir, no incluimos ni constante ni término de tendencia.

El test de Phillips Perron (PP) es un método no paramétrico para controlar la correlación serial de orden elevado en una serie. El test de regresión contenido en el test PP es el proceso autorregresivo AR(1):

$$\Delta y_t = \alpha + \beta y_{t-1} + \varepsilon_t \quad \dots\dots\dots (3.26)$$

Mientras que el test ADF corrige la correlación serial de orden elevado añadiendo más retardos del término diferenciado de la serie original en el lado derecho de la ecuación, el test PP realiza una corrección del estadístico t sobre el coeficiente y en la regresión AR(1) para considerar la correlación serial en el término ε .

La distribución asintótica del estadístico t del test PP es la misma que la del estadístico t del test ADF y se contrastan los resultados del test con los valores críticos de MacKinnon. Igual que en el test ADF tenemos que especificar si incluimos o no una constante, un término de tendencia o ambos en la regresión. Para el test PP, además, hay que especificar el número de periodos de correlación serial a incluir.

La corrección que realiza este test es no paramétrica debido a que utiliza una estimación del espectro del término ε en la frecuencia cero que es robusta para una forma no conocida de heteroscedasticidad y autocorrelación. Utiliza la corrección conocida como de Newey-West para la heteroscedasticidad y auto correlación.

3.6 Identificación del Modelo Arima

Para identificar el modelo ARIMA generador de la serie se utilizan las funciones *fac* y *facp* muestrales. La representación gráfica de ambas, en función de los distintos retardos se denomina *correlograma*. Pues bien, la identificación se realiza mediante la comparación de las características observadas en el correlograma y las características que cabe esperar para los distintos, y posibles, modelos teóricos.

Para clarificar este aspecto, se presentan los siguientes ejemplos:

En primer lugar, se puede comprobar que un modelo AR(1) podría presentar uno de los dos siguientes correlogramas, en los que se recogen, por ejemplificar con una cifra, los 10 hipotéticos primeros retardos:

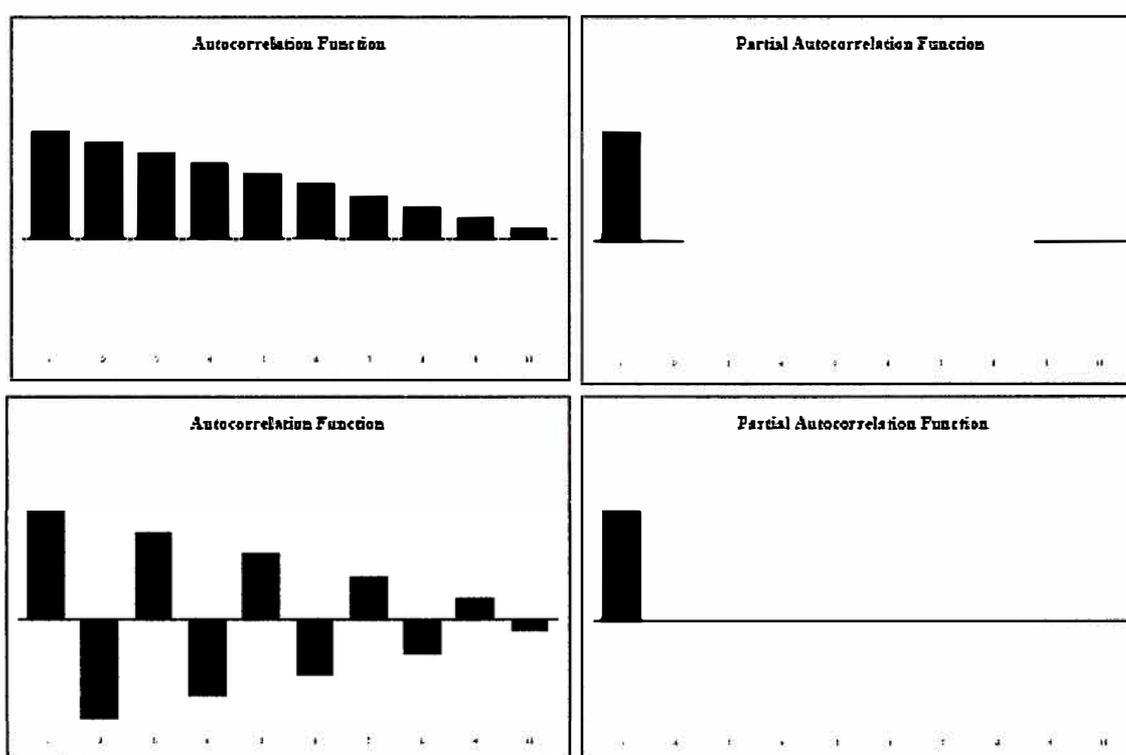


Fig. 3.1 Correlograma AR(1)

Es decir, en un modelo AR(1) o, lo que es lo mismo ARMA(1,0), encontramos que la función de autocorrelación decrece rápidamente hacia cero, bien sea de forma regular (primer correlograma), o alternando valores positivos y negativos (segundo correlograma). Aunque en nuestros gráficos, para ejemplificar, se haría cero para valores superiores al 11 retardo, el modelo teórico puede presentar tales valores para retardos previos. A su vez, la función de autocorrelación parcial presenta un único valor distinto de cero, es decir, tantos como el orden del proceso AR. Estas características son

generalizables para los modelos ARMA, con orden de medias móviles igual a cero. Es decir, un modelo ARMA(2,0) podría presentar los siguientes correlogramas teóricos:

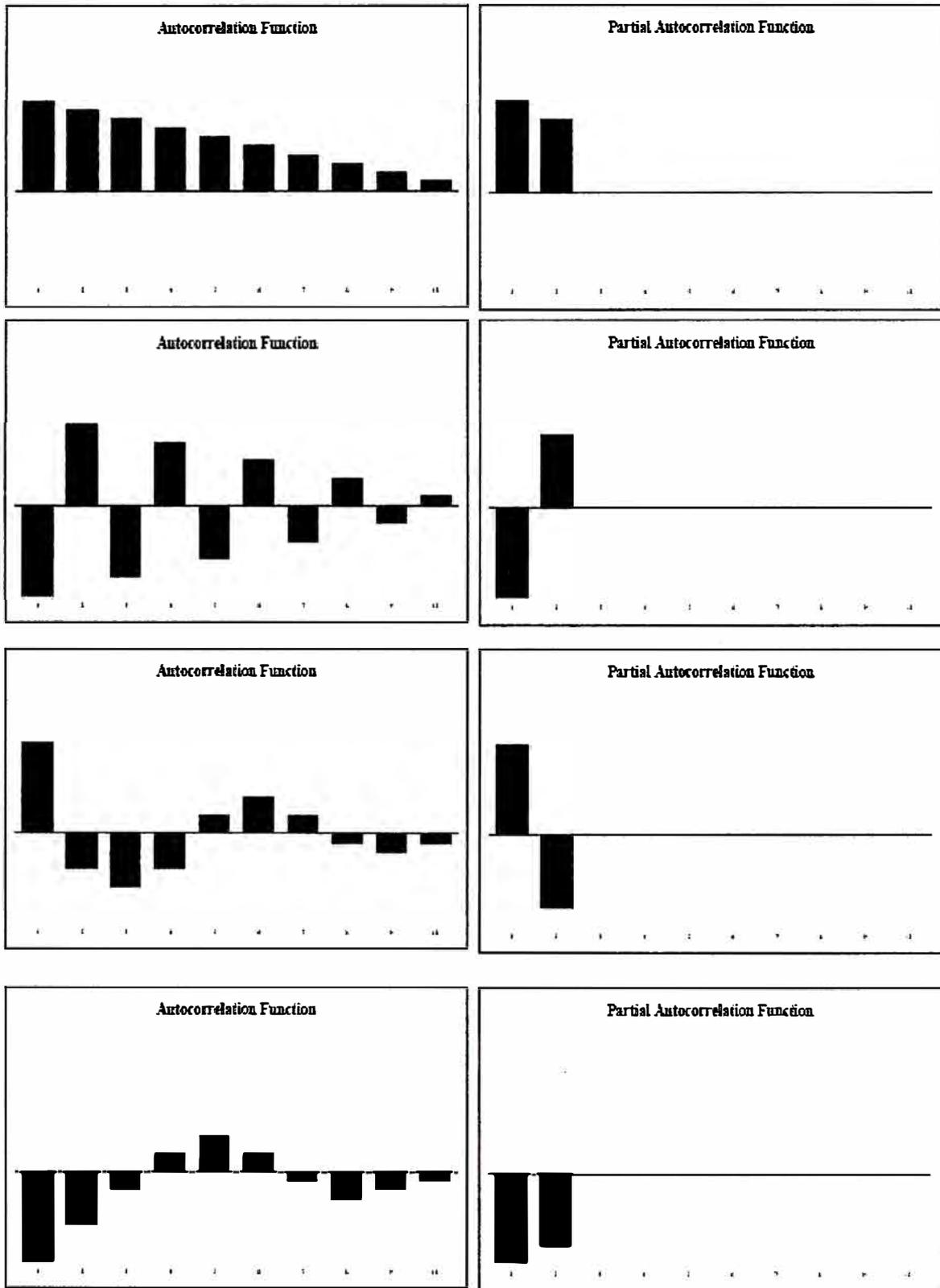


Fig. 3.2 Correlograma AR(2)

Por tanto, a efectos prácticos debemos quedarnos con la siguiente regla:

- Un modelo $ARMA(p,0)$ presenta una función de autocorrelación que decrece rápidamente hacia cero, bien de forma regular, sinusoidal o alternando valores positivos y negativos, junto con una función de autocorrelación parcial con tantos valores distintos de cero como orden del autorregresivo.
- Un modelo $ARMA(0,q)$ presenta una función de autocorrelación parcial que decrece rápidamente hacia cero, bien de forma regular, sinusoidal o alternando valores positivos y negativos, junto con una función de autocorrelación con tantos valores distintos de cero como orden de la media móvil.

Los modelos más habituales de este estilo serían los $MA(1)$ y $MA(2)$. Por ejemplo, un modelo $ARMA(0,2)$ podría presentar cualquiera de los siguientes correlogramas teóricos:

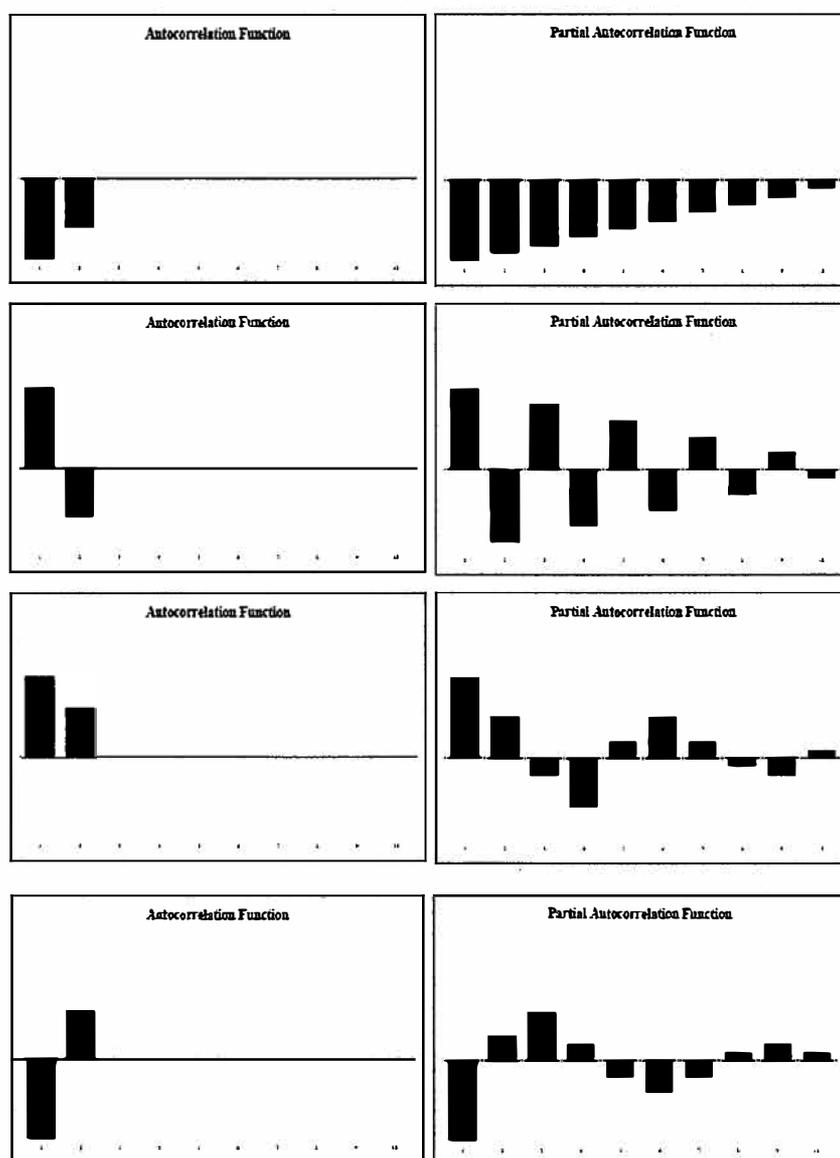


Fig. 3.3 Correlograma $MA(2)$

Por lo que respecta a los modelos $ARIMA(p,d,q)$, decir que el más habitual es el $ARIMA(1,d,1)$. Pues bien, en este caso la regla de identificación sería, sencillamente, observar que ambas funciones decrecen, admitiendo varias posibilidades de la combinación (sinusoidal, regular y alternando positivos y negativos) de ambas funciones. Cualquiera de los siguientes correlogramas sería representativo de un modelo $ARIMA(1,d,1)$:

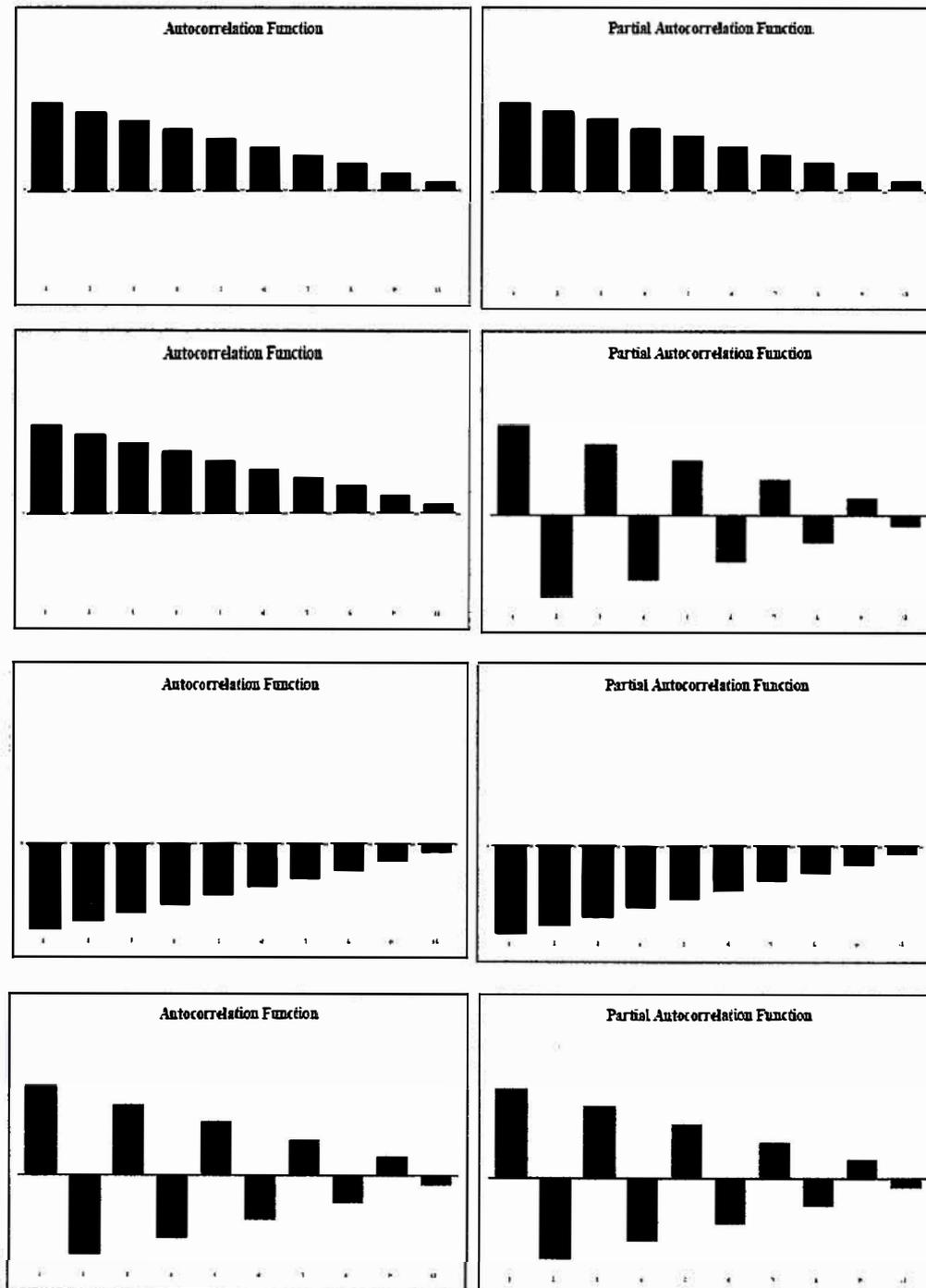


Fig. 3.4 Correlograma $ARIMA(1,d,1)$

En definitiva, para modelos sin estacionalidad, podemos adoptar la siguiente regla:

TABLA N° 3.1 Regla de Correlogramas

<i>Facp</i>	Decrece	Un coeficiente significativo	Dos coeficientes significativos
Decrece	ARMA(1,1)	AR(1)	AR(2)
Un coeficiente significativo	MA(1)	--	--
Dos coeficientes significativos	MA(2)	--	--

El problema de esta fase de identificación, estriba en que la estimación muestral de estas funciones para nuestra serie se aleja, en mayor o menor grado, de cualquiera de estas representaciones teóricas. En decir, con frecuencia, las estimaciones muestrales dan una visión distorsionada de las correspondientes funciones teóricas. Y en estas circunstancias, optar por uno u otro modelo generador de los datos depende más que nada de la pericia del analista.

Algunas circunstancias que se nos pueden presentar son las siguientes;

1) En primer lugar, estamos trabajando con un estadístico (con su función de probabilidad). Por ello, los valores obtenidos están "asociados" a una determinada probabilidad. Para saber si un determinado coeficiente de autocorrelación es significativamente distinto de cero necesitamos de un contraste estadístico, que nos establecerá los intervalos de confianza en los cuales los coeficientes resultan significativos con una determinada probabilidad. La teoría nos dice que la varianza de cada coeficiente se obtiene a partir de la expresión:

$$\text{var}(r_k) \cong \frac{1}{N} \left(1 + 2 \sum_{i=1}^{k-1} r_i^2 \right) \dots\dots\dots (3.27)$$

Ahora bien, esta expresión se puede aproximar para todos los coeficientes por:

$$\frac{1}{N} \dots\dots\dots (3.28)$$

siendo *N* el número total de datos. Admitida la normalidad de la distribución, un coeficiente no significativamente distinto de cero deberá estar, en el 95% de los casos, comprendido entre:

$$-2\sqrt{1/N} \leq r_k \leq +2\sqrt{1/N} \dots\dots\dots (3.29)$$

Por tanto, para que consideremos que un coeficiente estimado es significativamente distinto de cero debe superar el intervalo anterior.

2) Ahora bien, no podemos olvidar que estamos trabajando con un intervalo de confianza del 95 por 100. Quiere ello decir, que en el 5 por 100 de los casos estaremos rechazando la hipótesis nula de coeficiente igual a cero aun siendo esta correcta. En otras palabras, uno de cada veinte coeficientes de autocorrelación puede superar la banda de confianza sin que sea realmente significativo. Igualmente un coeficiente teórico no nulo puede no alcanzar en la práctica un valor suficientemente alto como para tomarle en consideración.

Todo ello provoca que, con frecuencia, se tengan dudas razonables sobre el modelo más adecuado. En esos casos se recomienda la selección de varios modelos alternativos para analizar posteriormente cuál de ellos resulta finalmente más idóneo. Completaremos esta etapa de la identificación posteriormente con el estudio de series con estacionalidad.

3.6.1 Funciones de Auto Correlación (FAC)

La función de autocorrelación (fac) y la función de autocorrelación parcial (facp) miden la relación estadística entre las observaciones de una serie temporal. Por ejemplo, el coeficiente de autocorrelación entre la variable Y_t y la misma variable un período antes, Y_{t-1} , al que denominaremos coeficiente de autocorrelación de primer orden, se formula como:

$$\rho_1 = \frac{\text{cov}(y_t, y_{t-1})}{\sqrt{\text{var}(y_t) \cdot \text{var}(y_{t-1})}} \quad \dots\dots\dots (3.30)$$

Dado el supuesto de estacionariedad, se tiene que $\text{var}(Y_t) = \text{var}(Y_{t-1})$, por lo que

$$\rho_1 = \frac{\text{cov}(y_t, y_{t-1})}{\text{var}(y_t)} \quad \dots\dots\dots (3.31)$$

En general, para un desfase de k períodos se tiene que:

$$\rho_k = \frac{\text{cov}(y_t, y_{t-k})}{\text{var}(y_t)} \quad \dots\dots\dots (3.32)$$

y cuando $k=0$,

$$\rho_0 = \frac{\text{cov}(y_t, y_t)}{\text{var}(y_t)} = \frac{\text{var}(y_t)}{\text{var}(y_t)} = 1 \quad \dots\dots\dots (3.33)$$

A efectos de la identificación del modelo, debemos comparar el valor que esta función presentaría para los distintos modelos teóricos, con una estimación de la misma para

nuestra serie. Pues bien, el estimador muestral de la *fac*, para el que utilizaremos la expresión r_k , viene dado:

$$r_k = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2} \dots\dots\dots (3.34)$$

3.6.2 Funciones de Auto Correlación Parcial (FACP)

La función de autocorrelación parcial mide la “aportación” que a las variaciones de una variable como Y_t tiene otra variable, digamos Y_{t-2} , aislados los efectos de las posibles restantes variables, por ejemplo Y_{t-1} . Por el contrario, la función de autocorrelación ignora el hecho de que parte de la correlación que pueda existir entre, por ejemplo Y_t y Y_{t-2} , se debe a que ambas están correlacionadas con Y_{t-1} . Pues bien, los distintos coeficientes de

autocorrelación parcial de los modelos teóricos se denotan como ϕ_{kk} , y los estimados para nuestra muestra como $\hat{\phi}_{kk}$. La utilidad de los mismos se deriva de que en determinadas ocasiones el simple conocimiento de la *fac* muestral no sería suficiente para la determinación del verdadero proceso generador de la serie.

Aunque no entraremos en el detalle de la formulación teórica sí es importante conocer algunas propiedades de la *facp*. En primer lugar, por definición la *facp* de orden 1 es igual a la *fac* de orden 1. En segundo lugar, en el modelo teórico, el coeficiente de autocorrelación parcial coincide con el último coeficiente autorregresivo de un modelo AR.

En otros términos, el coeficiente de correlación parcial de orden 1 será el valor de ϕ_1 en un AR(1); el de orden 2 coincidirá con ϕ_2 en un AR(2), y así sucesivamente. En la práctica, los coeficientes de autocorrelación parcial calculados no son buenos estimadores de los parámetros correspondientes, aunque pueden servir como valores iniciales para el proceso iterativo de cómputo que ha de seguirse.

3.7 Estimación Del Modelo Arima

La fase de la estimación consiste en la obtención de unos valores numéricos para los parámetros del modelo previamente identificado, de forma que contengan buenas propiedades estadísticas. Para ello, la metodología ARIMA realiza un proceso de

búsqueda iterativo de tales valores, de prueba y error, hasta dar con los valores óptimos, puesto que las ecuaciones de resolución no son lineales.

La estimación presupone el conocimiento de valores no disponibles de la variable a analizar, así como de la variable de error, para lo cual los programas de estimación plantean distintas alternativas.

Además, en la búsqueda iterativa de las estimaciones, se debe partir de un valor inicial. Todos los programas lo realizan de forma automática pero existen ciertas relaciones teóricas que es importante conocer.

Una vez estimado el modelo podemos interpretar los parámetros del mismo.

A la hora de realizar la etapa de validación (o contraste) del modelo podemos recurrir a diversos criterios y contrastes estadísticos. Todos los programas de computadora recogen estos test.

Además, para validar el modelo podemos estudiar el gráfico de los residuos, que nos proporciona una visión de conjunto de la cuantía de los errores, sesgos sistemáticos y puntos de errores excepcionales (llamados outliers) que deben ser analizados con especial atención.

Seleccionado el modelo, puede pasarse a la etapa de la predicción. La verdadera predicción será la realizada a partir del último dato del período muestral, aunque también puede resultar útil analizar cómo se habría comportado el modelo si hubiera tenido que hacer una predicción dentro del período histórico ya conocido, y que ha servido de base a su estimación y contraste. Para ello disponemos de dos alternativas: la estática y la dinámica.

CAPITULO IV APLICACIÓN

Para un mejor entendimiento del proceso de proyección se detalla la aplicación del método Arima para estimar la demanda futura de un alimentador de media tensión de una empresa de distribución eléctrica del País.

4.1 Segmentación del Mercado

Debido al dinamismo de las redes de media tensión es necesario definir nuestro mercado a analizar para lo cual aplicamos el concepto de frontera eléctrica.

Se selecciona uno de los meses más estables del año donde el área geográfica que atiende el alimentador determinará la frontera eléctrica.

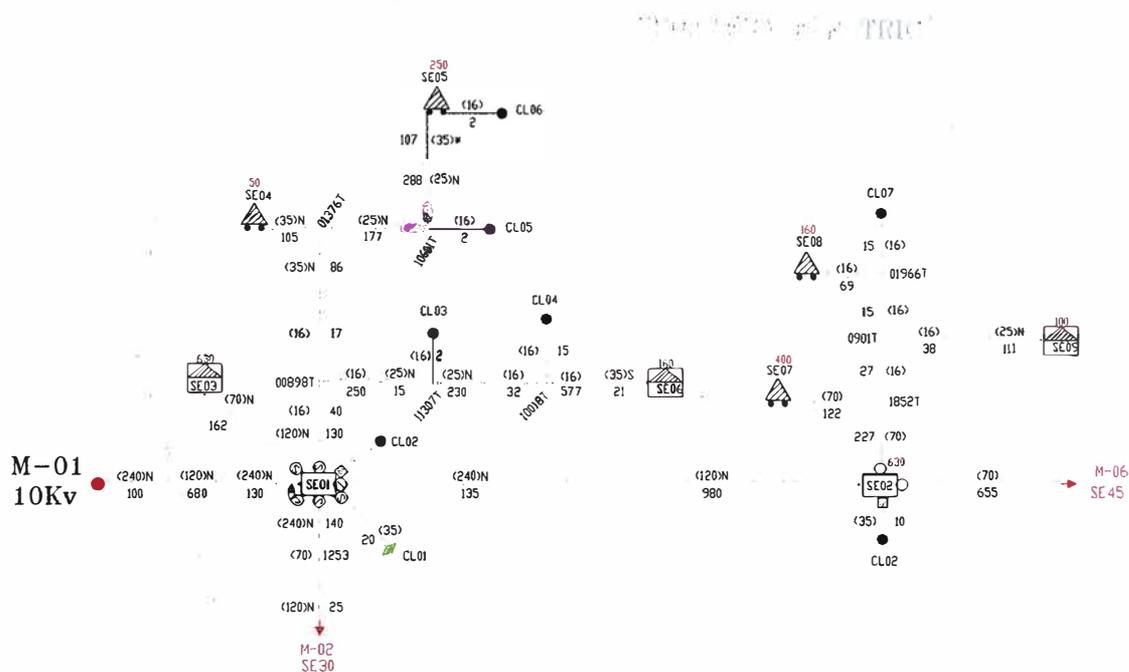


Fig. 4.1 Esquema Unifilar, Alimentador M-01

Determinado la frontera eléctrica del alimentador se identifica los suministros que le pertenecen, los cuales nos permitirán construir la serie de tiempo del alimentador (valores históricos).

Definida la serie de tiempo podemos realizar las proyecciones necesarias. Para esta aplicación se utilizará el método Arima.

4.2 Proyección de Demanda

Contamos con una serie que recoge la evolución mensual del consumo de energía eléctrica del alimentador de media tensión denominado "M-01", para el cual se estimará y contrastará un modelo ARIMA utilizando el software EViews.

4.2.1 Datos Del Alimentador M-01

TABLA N° 4.1 Consumo de Energía, Alimentador M-01

MES	CONSUMO DE ENERGÍA ELÉCTRICA (MWH) - ALIMENTADO M01								
	1997	1998	1999	2000	2001	2002	2003	2004	2005
ENE	1.189	1.231	1.351	1.372	1.461	1.613	1.642	1.810	1.808
FEB	1.128	1.264	1.247	1.322	1.485	1.515	1.646	1.705	1.869
MAR	1.011	1.246	1.217	1.331	1.477	1.488	1.697	1.817	1.869
ABR	1.100	1.338	1.375	1.392	1.501	1.614	1.594	1.793	1.881
MAY	1.163	1.203	1.225	1.338	1.359	1.493	1.523	1.650	1.871
JUN	1.071	1.139	1.211	1.355	1.350	1.449	1.530	1.702	1.818
JUL	1.078	1.177	1.263	1.330	1.426	1.576	1.531	1.679	1.856
AGO	1.251	1.191	1.268	1.314	1.418	1.512	1.616	1.700	1.851
SET	1.119	1.147	1.244	1.419	1.381	1.432	1.609	1.732	1.897
OCT	1.160	1.214	1.300	1.280	1.405	1.586	1.650	1.726	1.855
NOV	1.164	1.142	1.233	1.419	1.354	1.511	1.565	1.868	1.977
DIC	1.129	1.197	1.306	1.348	1.388	1.556	1.668	1.810	2.038

4.2.2 Estacionariedad de la Serie

Para identificar la estacionariedad de la serie en media y varianza podemos utilizar la representación gráfica de la serie. Dado que la serie muestra una clara tendencia al alza, podemos suponer ya a priori, que no será estacionaria y que, por lo tanto, presentará al menos una raíz unitaria.

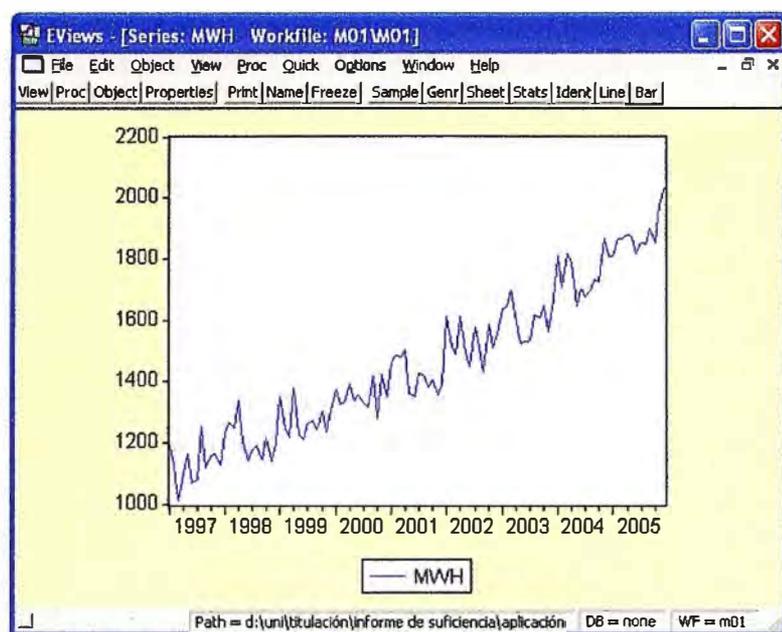


Fig. 4.2 Consumo de Energía

Para comprobar la posible existencia de una raíz unitaria en la serie podemos inspeccionar el correlograma de la misma.

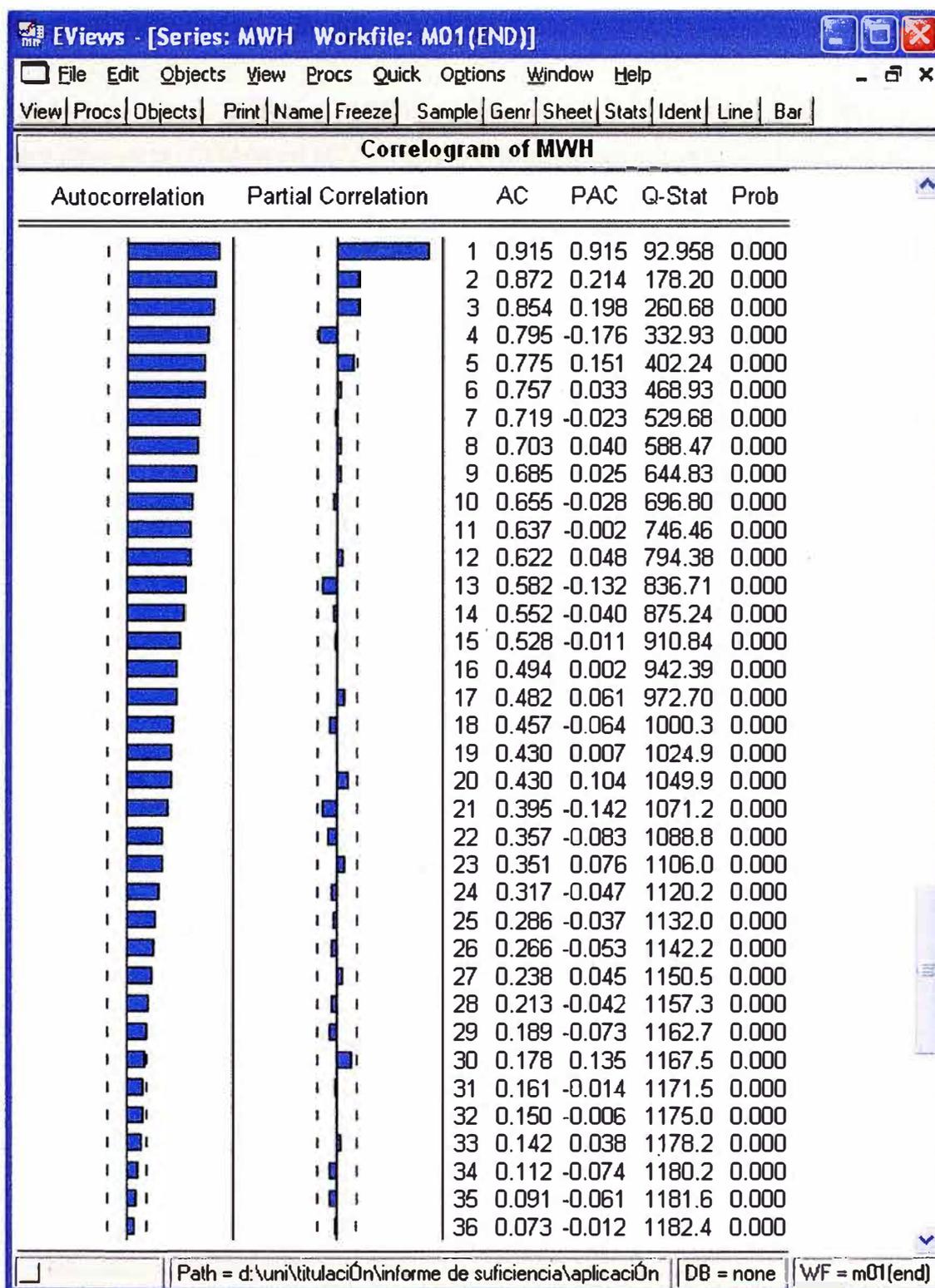


Fig. 4.3 Correlograma del Consumo de Energía

Como podemos apreciar, la función de auto correlación decrece de forma lenta, mientras que la función de auto correlación parcial presenta un valor significativo en el retardo uno, con un coeficiente de auto correlación 0,869. Este gráfico puede considerarse como indicativo de la no estacionariedad de la serie, es decir, presenta una raíz unitaria.

4.2.3 Transformación de la Serie

Tomando logaritmos ($LMWH = \text{LOG}(MWH)$) se reduce la dispersión de la serie y con la primera diferencia ($DLMWH = LMWH - LMWH(-1)$) se observa que la tendencia se anula.

El cálculo de ambas series da lugar a los siguientes gráficos, donde se aprecia la eliminación de la tendencia de la serie con la obtención de DLMWH, es decir, esta última serie es estacionaria en media.

Una identificación previa del correlograma de la serie en logaritmos (LMWH), transformación que la convierte en estacionaria en varianza, nos habría indicado la existencia de una raíz unitaria y, por tanto, la no estacionariedad en media.

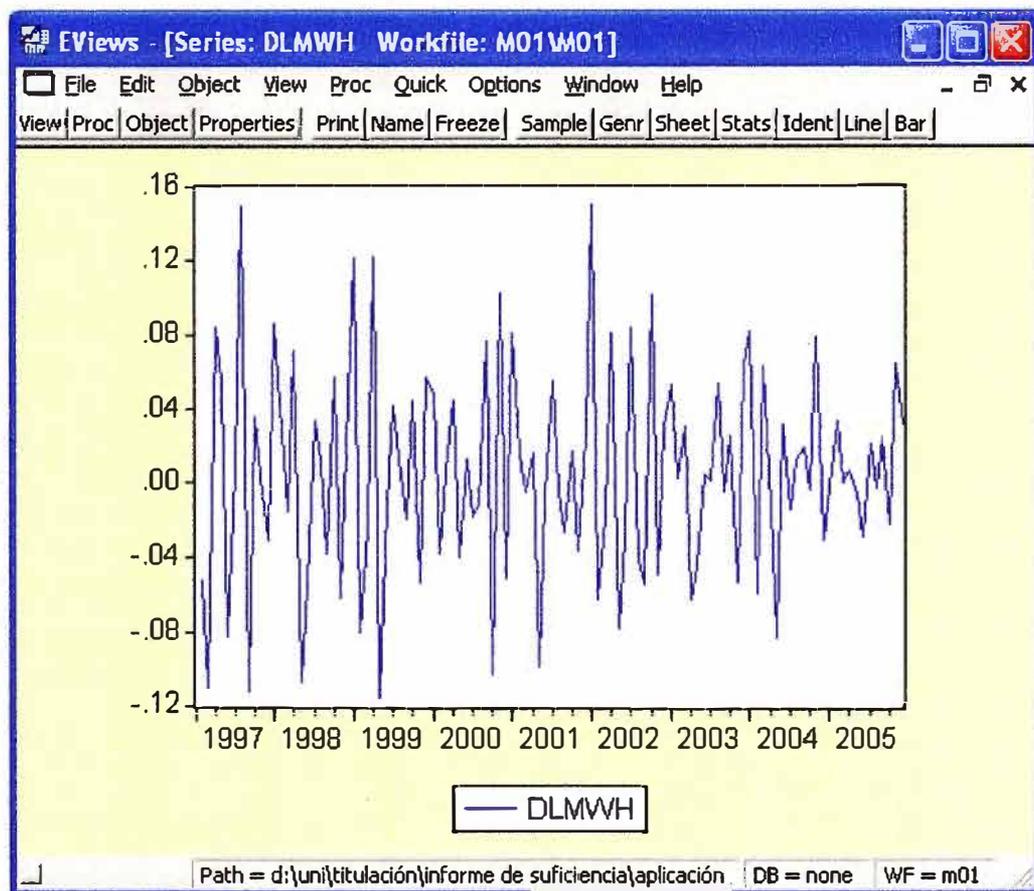


Fig. 4.4 Derivada del Logaritmo del Consumo de Energía

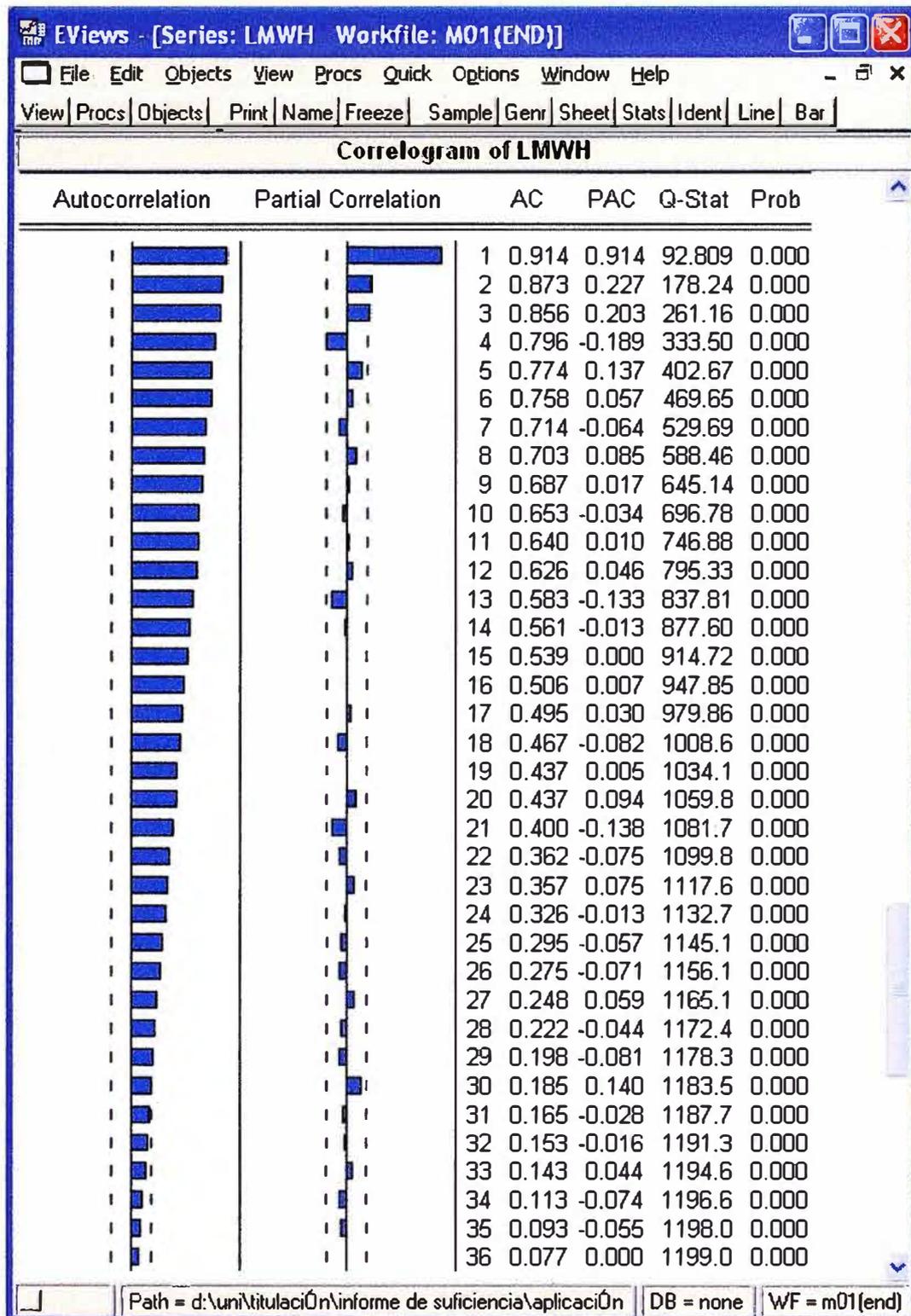


Fig. 4.5 Correlograma del Logaritmo del Consumo de Energía

4.2.4 Identificación del Modelo

Volviendo a nuestra serie ya transformada (DLIPC), procedemos a la obtención en EViews de las funciones de auto correlación y auto correlación parcial.

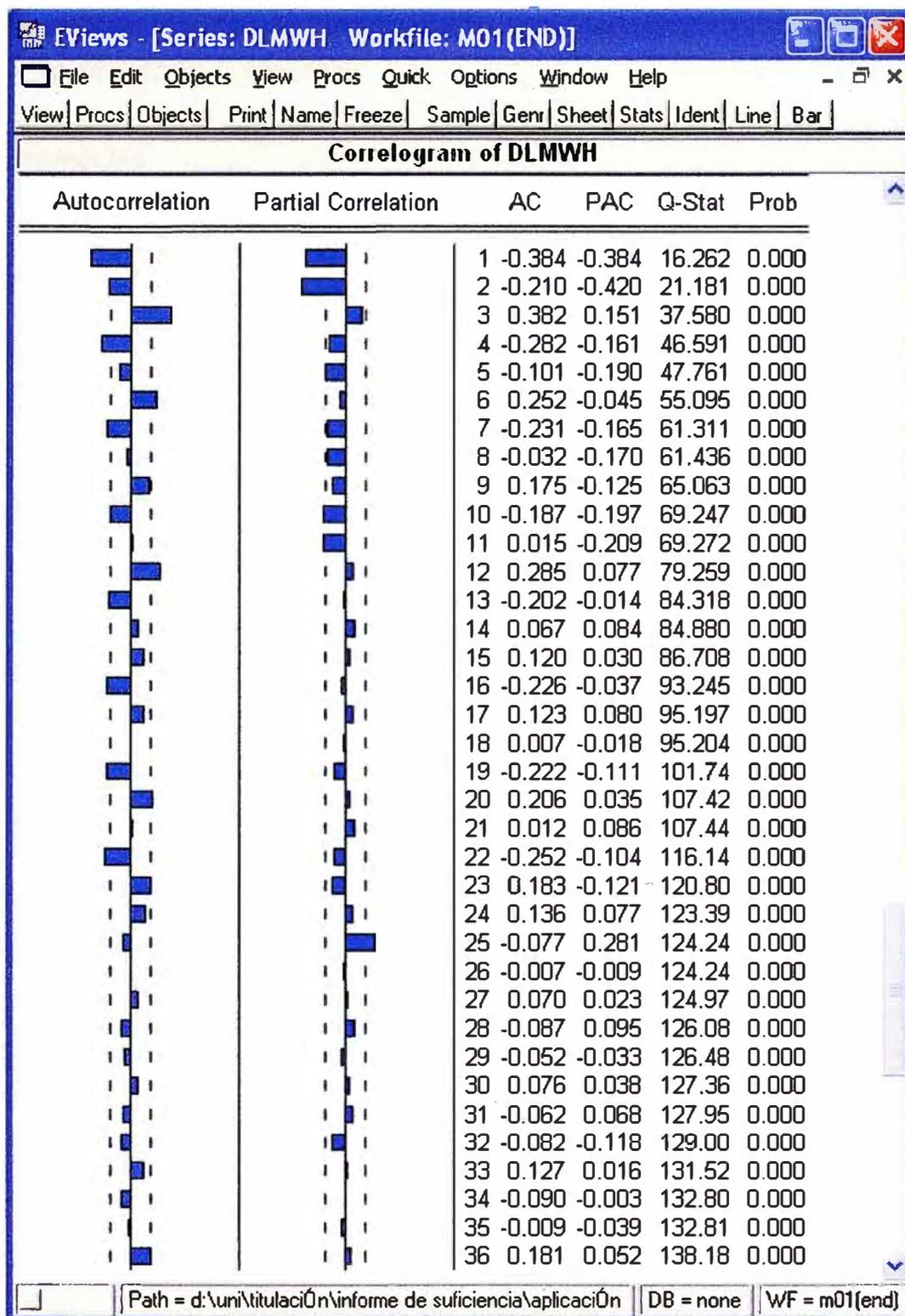


Fig. 4.6 Correlograma de la Derivada del Logaritmo del Consumo de Energía

Al observar este correlograma apreciamos, en la función de auto correlación, y la función de auto correlación parcial, identificamos a los procesos AR(2) y MA(1).

Con los procesos identificados procedemos a estimar (predecir) los valores de demanda de energía eléctrica en el año 2006.

4.2.5 Proyección

Para realizar la proyección es necesario utilizar la función FORECAST, en el cual hay que especificar el periodo de predicción que deseamos analizar.

Forecast

Forecast of
 MWH
 DLOG(MWH)

Series names
Forecast name: mwhf
S.E. (optional):
GARCH(optional):

Forecast sample
2006:01 2006:12

Method
 Dynamic
 Static
 Structural (ignore ARMA)

Output
 Do graph
 Forecast evaluation

Insert actuals for out-of-sample

OK Cancel

Fig. 4.7 Función Forecast

Como resultado se obtiene la predicción para los 12 meses del 2006.

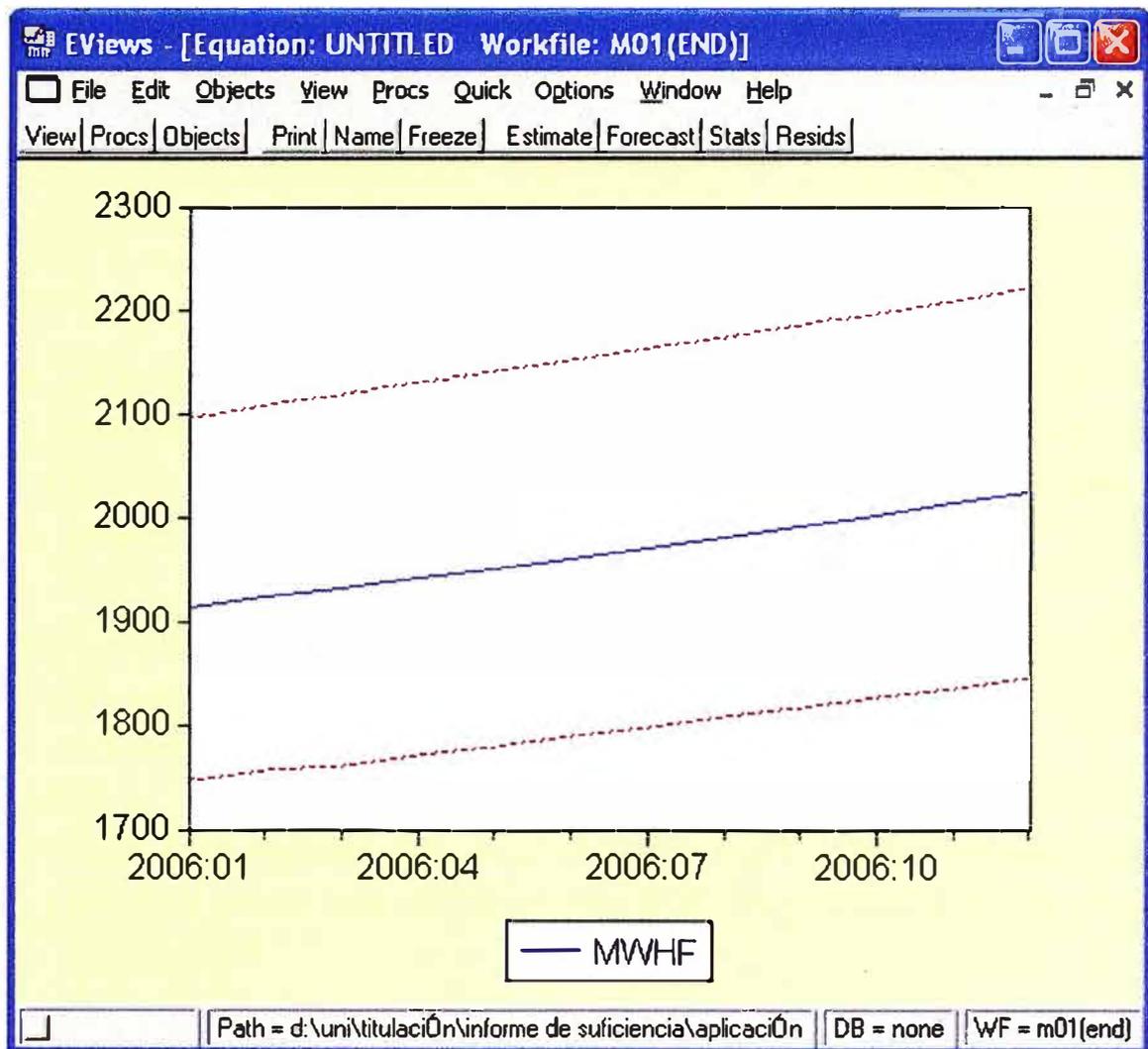


Fig. 4.8 Predicci3n a 12 Meses del Consumo de Energ3a

CONCLUSIONES

1. La proyección de demanda es una de las tareas mas importantes y de mayor relevancia en el planeamiento de sistemas de distribución eléctrica, por lo que además de la evolución histórica debe tenerse en cuenta otros factores como son: geográficos, comerciales y poblacionales entre otros
2. El análisis basado en series de tiempo es valido si es que no se dan otros factores que puedan influenciar de manera significativa en la tendencia y que distorsione su comportamiento, por ende la predicción.
3. La regresión lineal simple es la técnica mas sencilla y accesible para determinar la tendencia de una serie de tiempo, en el cual cobra relevancia el coeficiente de correlación "r".
4. El campo de aplicación del modelo ARIMA fundamentalmente lo constituye la predicción a corto plazo y en series con componente estacional, tanto más útil cuanto de mayor frecuencia. Es ahí donde la metodología ARIMA alcanza sus mayores ventajas a efectos de predicción.
5. Para realizar la proyección de demanda en alimentadores de media tensión no es suficiente el análisis de la evolución histórica, también es importante conocer los compromisos de carga que puedan distorsionar la tendencia.

ANEXOS

ANEXO A
MANUAL DE INTRODUCCIÓN AL SOFTWARE EIEWS

BREVE INTRODUCCION AL MANEJO DEL EIEWS

*Autor: Dr. Gustavo A. Marrero Díaz
Dpto. Fundamentos del Análisis Económico y Economía Cuantitativa II
Facultad Ciencias Económicas y Empresariales, UCM
Noviembre - 2000*

EViews (EV), es un programa diseñado para el análisis estadístico y econométrico de todo tipo de datos, especialmente datos de series temporales. Sus posibilidades son amplias y muy variadas. EV utiliza las características visuales del software de Windows moderno, lo que hace relativamente sencillo su manejo, ya que funciona a base de menús y cajas de diálogo.

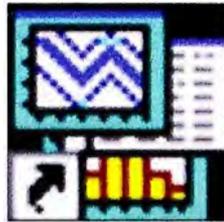
A continuación, se listan algunas de las funciones más importantes que EV puede realizar en el tratamiento de datos.

- Análisis descriptivo y gráfico de datos.
 - Estadísticos descriptivos e histogramas de frecuencia.
 - Covarianzas, correlaciones y correlaciones cruzadas.
 - Funciones de autocorrelación simple y parcial.
 - Gráfico de línea, de barras, diagrama de dispersión, gráficos de doble escala, gráficos de series normalizadas, etc.
- Estimación de modelos estadísticos:
 - Estimación por Mínimos Cuadrados Ordinarios (regresión simple y múltiple).
 - Estimación por mínimos cuadrados corregidos de autocorrelación. Mínimos Cuadrados en dos etapas.
 - Estimación por Mínimos Cuadrados No Lineales.
 - Estimación de modelos de elección binaria: Logit y Probit.
 - Estimación de sistemas de ecuaciones lineales y no lineales.
 - Estimación y previsión de modelos ARCH y GARCH.
- Contraste de hipótesis
 - Test de Wald.
 - Test de Chow.
 - Contraste de significación individual y conjunta.
- Análisis de Series Temporales
- Previsión basada en regresión.

Esta guía pretende ser autodidáctica, de tal manera que el lector pueda ser capaz de seguirla y aprender las principales capacidades de EV por su cuenta. Para hacer la guía más didáctica, se desarrollará a lo largo de ella un ejemplo muy sencillo, extraído del libro de “Estadística y Econometría” de A. Novales (ed. McGraw Hill, 1995). En el ejemplo se plantea estudiar la relación entre el volumen de ventas de una determinada empresa a lo largo de 10 períodos, el gasto destinado en publicidad y el precio medio anual de su producto.

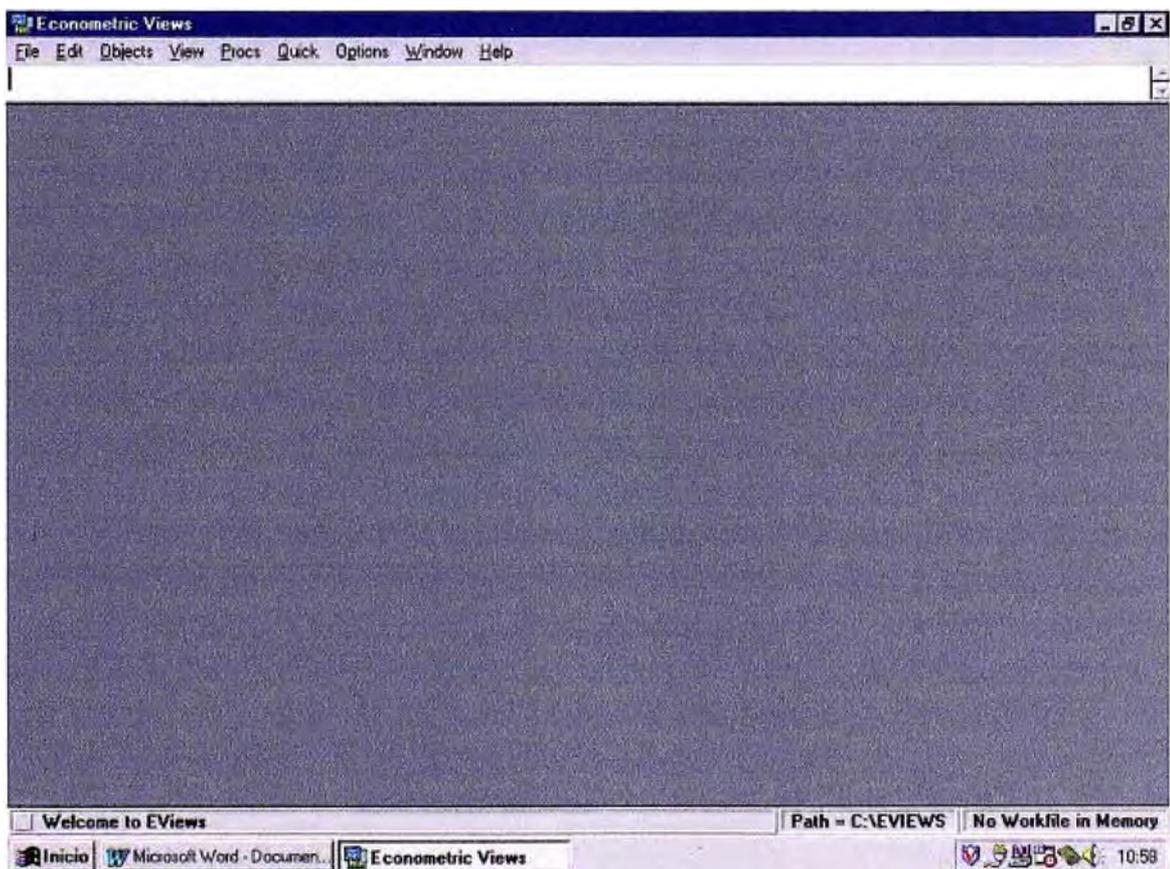
IPONER EN MARCHA EV

El Icono del programa EV será, generalmente, el siguiente:



Eviews.Ink

Por tanto, para iniciar el programa, una posibilidad es hacer doble click en dicho Icono. La pantalla inicial es la siguiente:



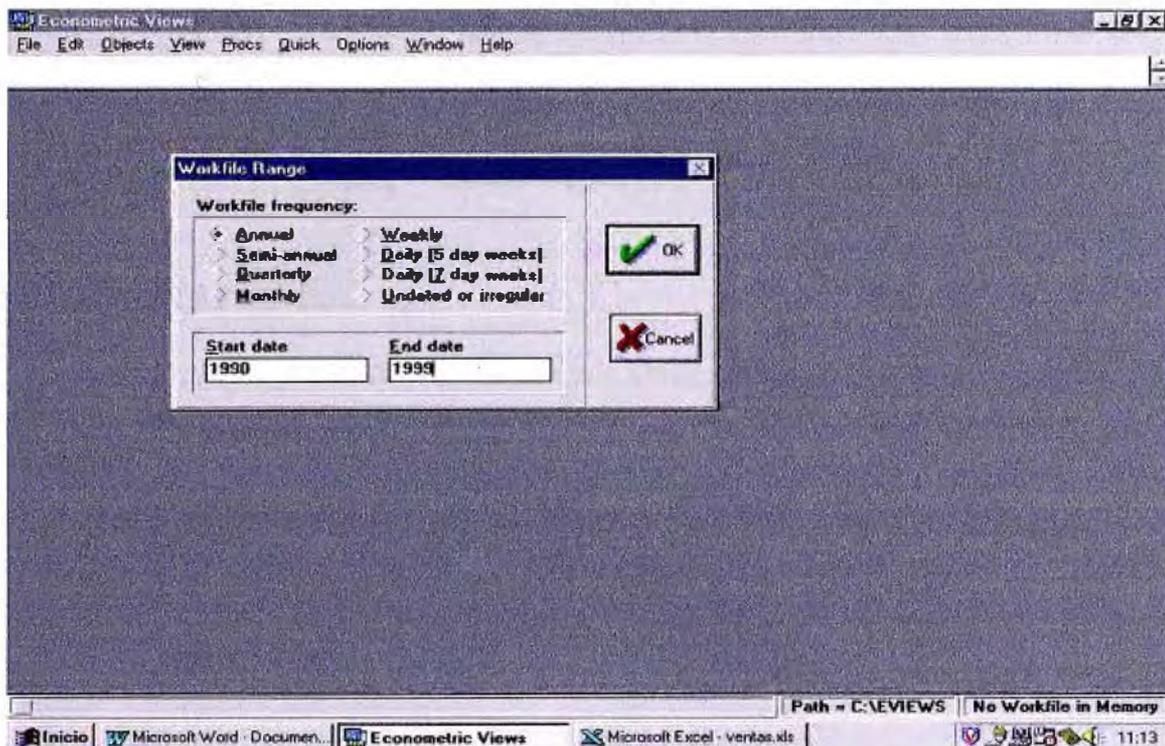
Como se puede apreciar, en la parte inferior derecha está escrito *No Workfile in Memory*. Esto quiere decir que no se encuentra cargado ningún libro de trabajo.

II. CREAR UN LIBRO DE TRABAJO

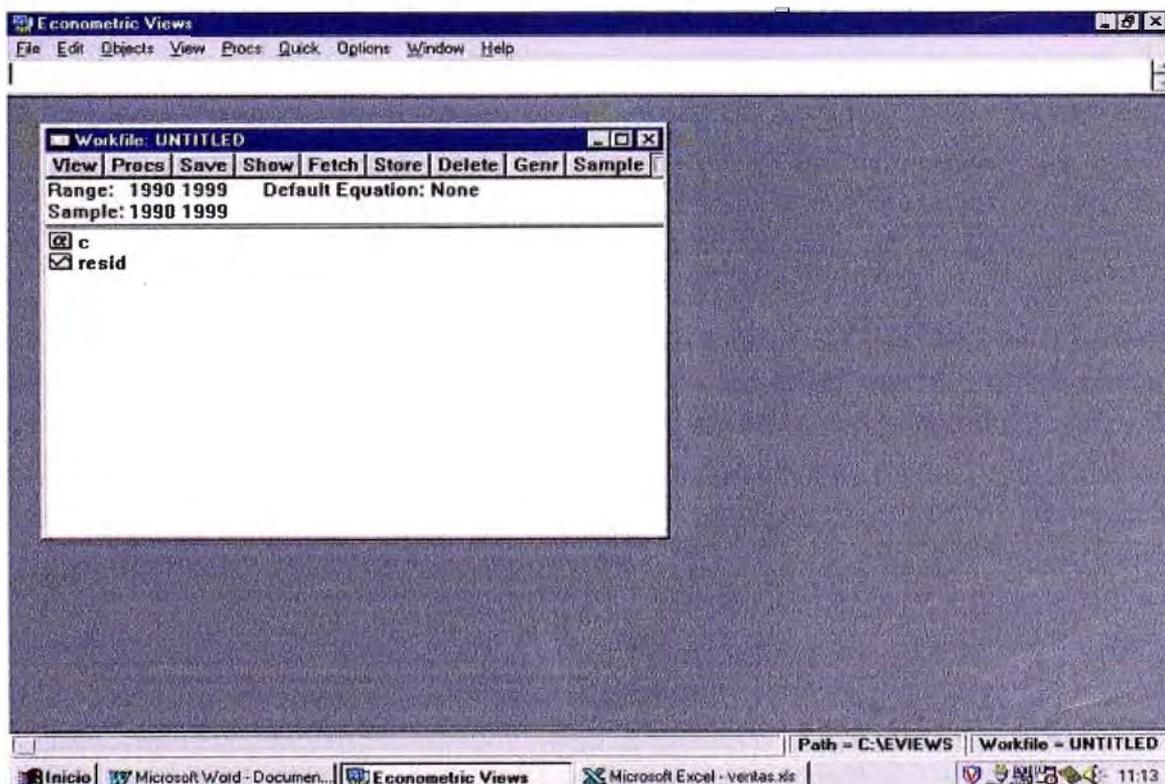
Para empezar a trabajar con EV se tiene que “cargar” un libro ya existente (en el menú *File-Open*) o bien “crear” uno nuevo. El procedimiento a seguir para generar un área de trabajo nuevo es el siguiente: Seleccionar la opción *File*, situada en la parte superior de la pantalla. Posteriormente, elegir *New* y después, situarse sobre la opción *Workfile*.



Con esto, le hemos dado la instrucción de crear un “nuevo” “libro de trabajo”. A continuación, aparecerá una pantalla similar a la siguiente:



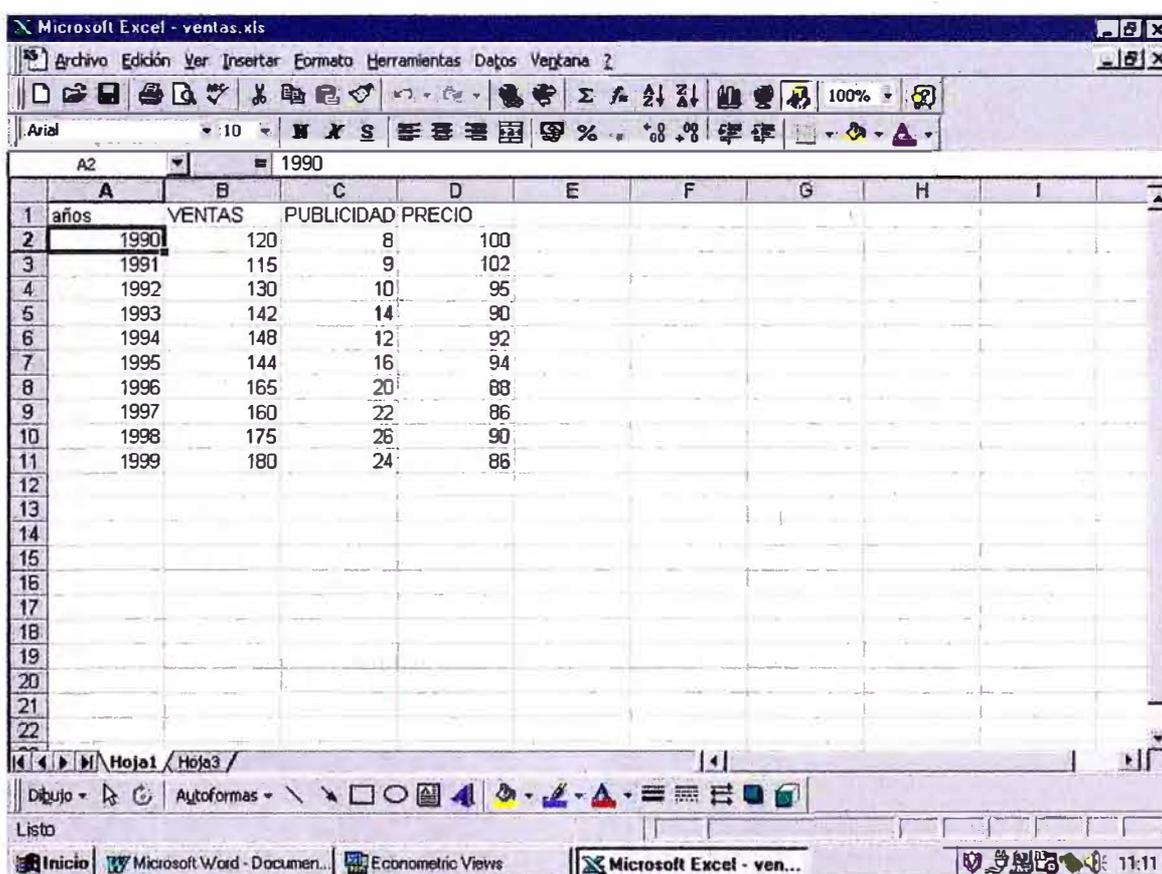
En esta pantalla el usuario debe introducir información sobre la periodicidad de los datos con los que va a trabajar. Para ello debe marcar alguna de las opciones que se muestran en la pantalla (anuales, mensuales,..., o irregular si, por ejemplo, estamos tratando con series de sección cruzada). También se tiene que escribir la fecha de comienzo (*Start Data*) y final (*End Data*) de los datos utilizados. En nuestro ejemplo, son datos anuales que comienzan en 1990 y terminan en 1999. Presionar el botón *OK*. Al hacerlo, el programa desplegará la siguiente pantalla:



Por el momento, el *Workfile* no tiene nombre (*Untitled*). Por defecto, EV crea dos Iconos: una con una “C” que refleja la constante del modelo de regresión y una serie, con el nombre “*resid*”, donde se acumularán los residuos de la última regresión que se haga con el programa. Una vez creada el área de trabajo, hay que importar los datos.

III. IMPORTAR DATOS.

Lo más cómodo es que los datos estén almacenados en una hoja de cálculo, por ejemplo, de Excel. La forma más eficiente para que los datos sean cargados por EV es la siguiente:

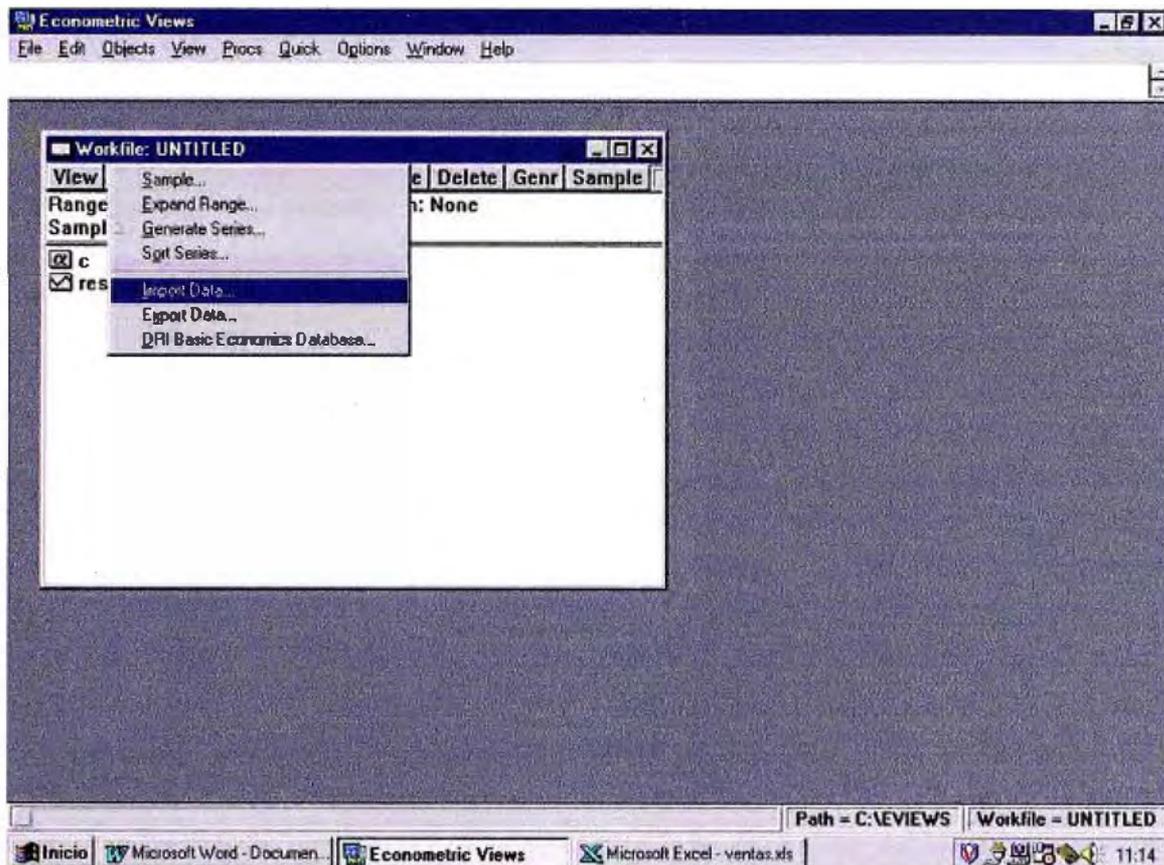


The screenshot shows a Microsoft Excel window titled 'Microsoft Excel - ventas.xls'. The spreadsheet contains the following data:

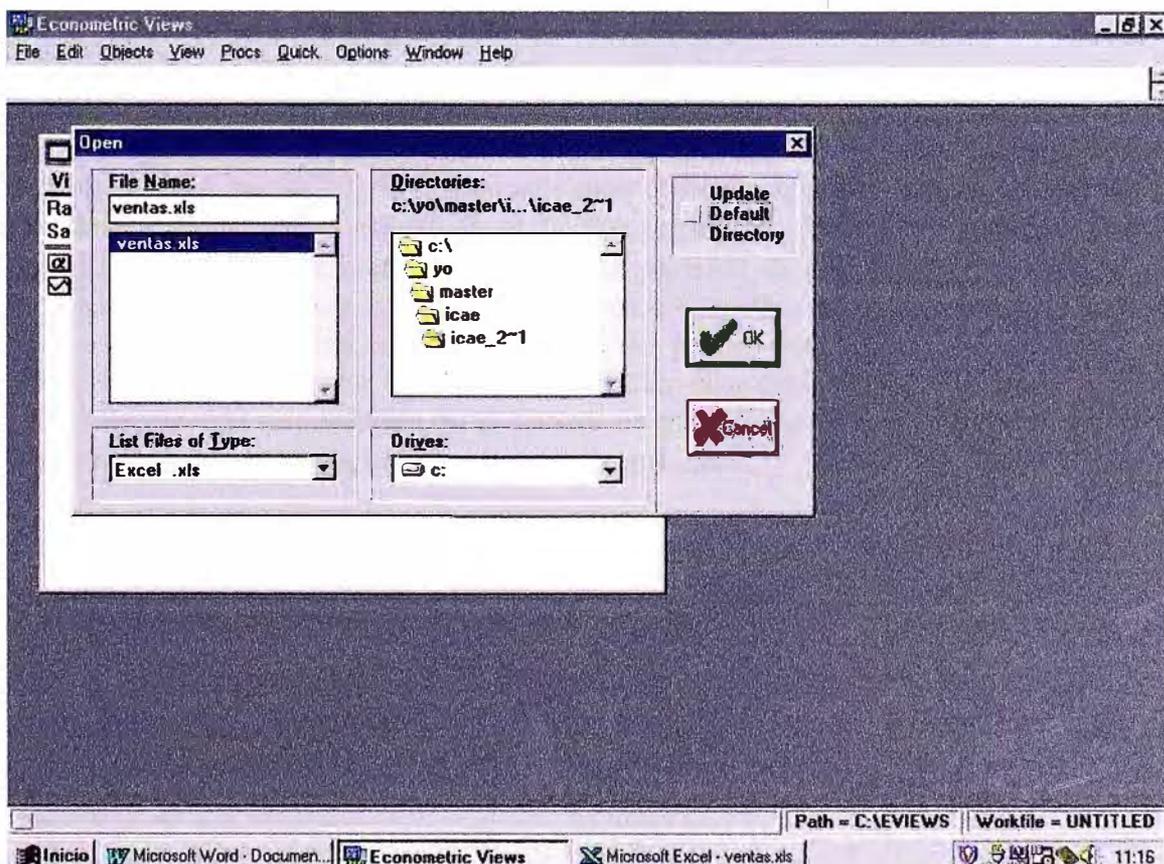
1	A	B	C	D	E	F	G	H	I
1	años	VENTAS	PUBLICIDAD	PRECIO					
2	1990	120	8	100					
3	1991	115	9	102					
4	1992	130	10	95					
5	1993	142	14	90					
6	1994	148	12	92					
7	1995	144	16	94					
8	1996	165	20	88					
9	1997	160	22	86					
10	1998	175	26	90					
11	1999	180	24	86					
12									
13									
14									
15									
16									
17									
18									
19									
20									
21									
22									

Nótese que el primer dato se encuentra en la celda B2 y encima de cada serie hemos puesto un nombre distintivo de cada una de ellas (hay nombres que son prohibidos por EV, debido a que el programa los usa como código propio, como “C”, “*resid*”, ...).

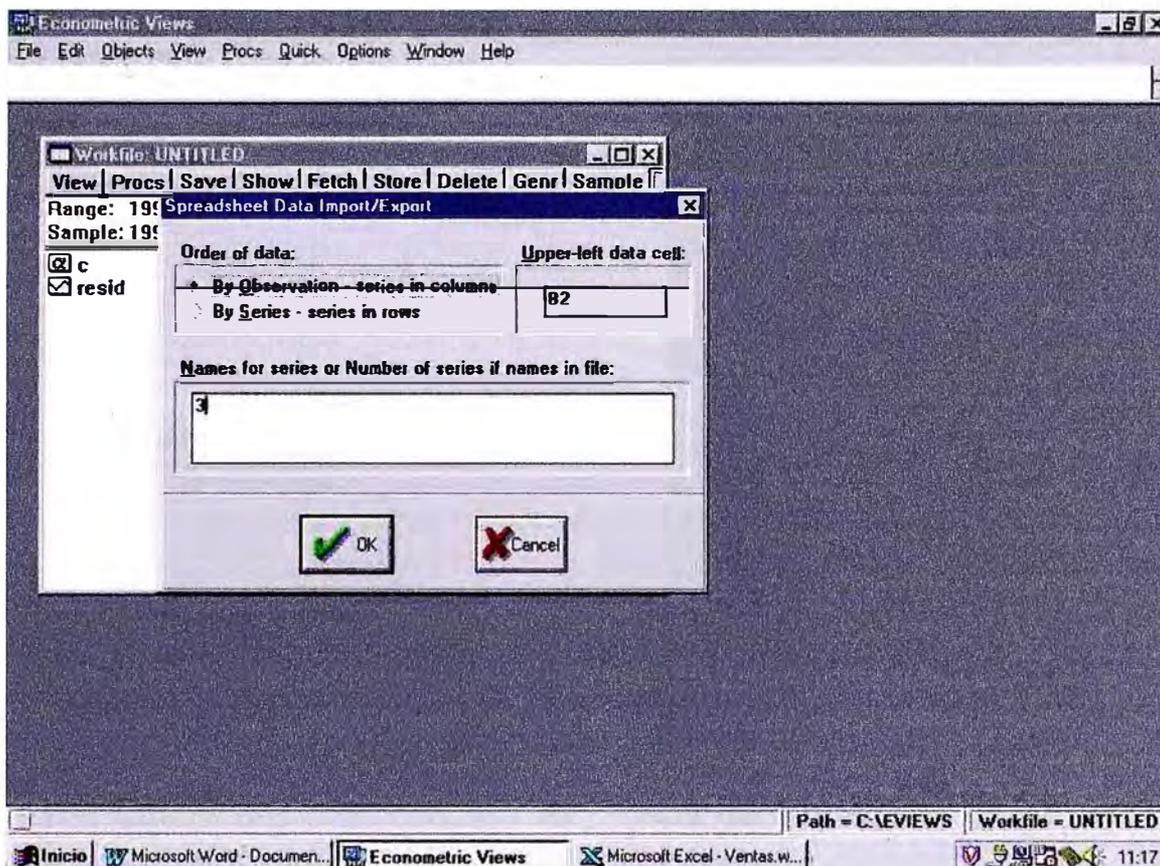
Para importar datos situar el ratón en la opción *Procs*, en la parte superior del fichero de trabajo. Posteriormente, seleccionar la opción *Import Data*.



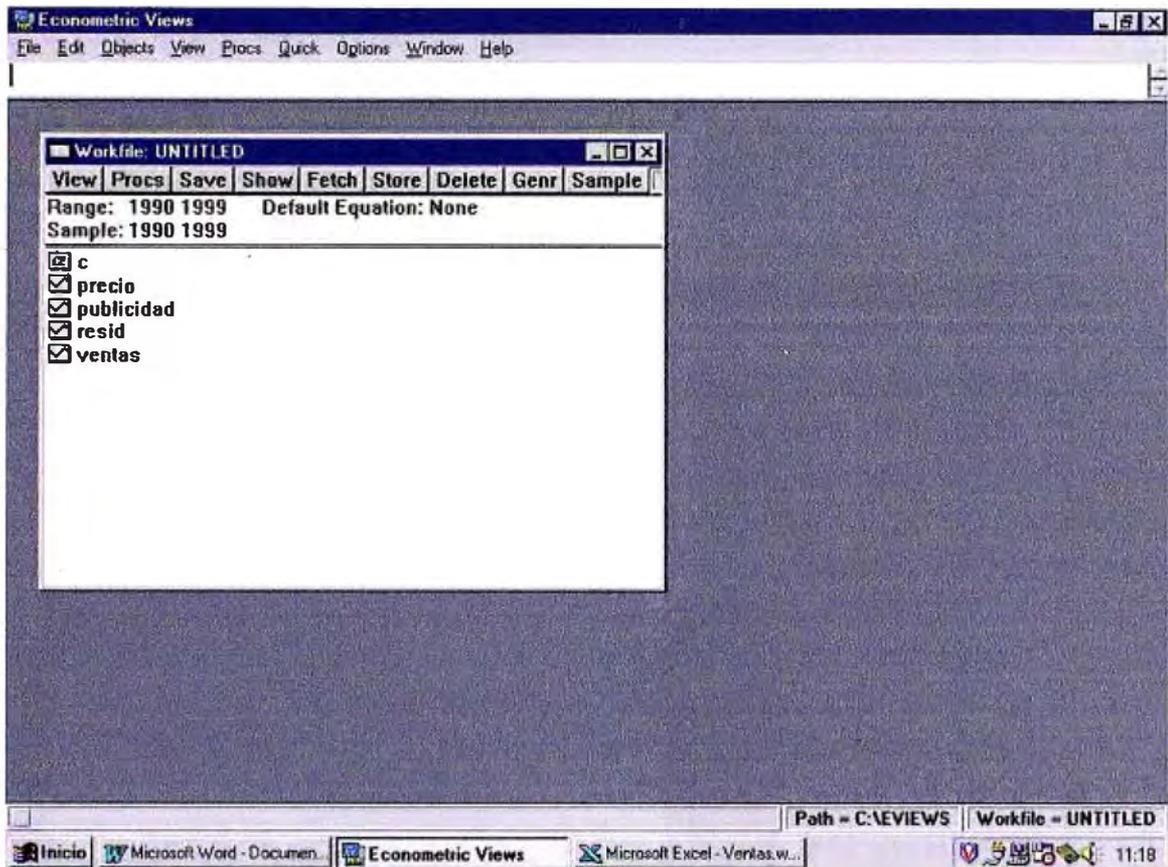
El programa desplegará una ventana donde el usuario debe seleccionar la unidad y el directorio donde están los datos con los que se quiere trabajar.



Una vez seleccionada la hoja de cálculo en la que se tienen almacenados los datos, pulsamos *OK*. A continuación, el programa desplegará la siguiente pantalla:

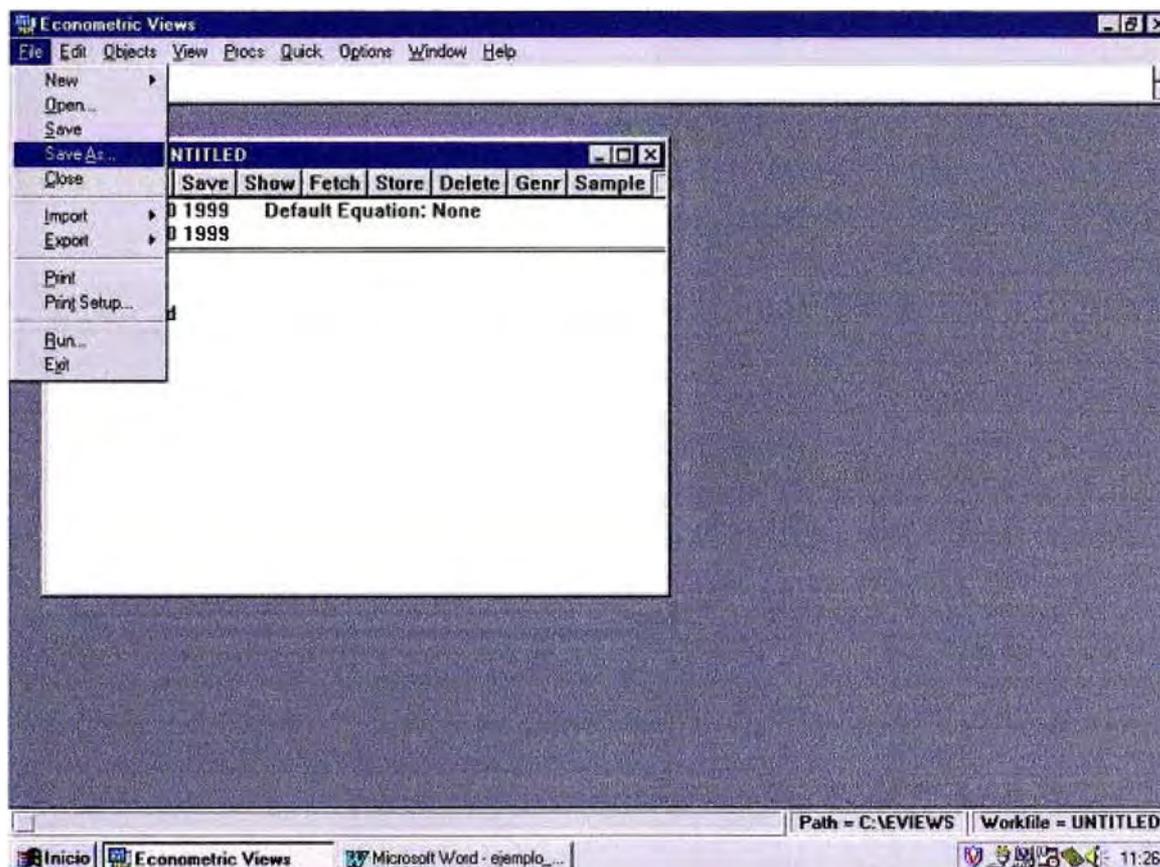


En el recuadro en blanco, se debe escribir el nombre que se asigna a los datos. Por ejemplo, si se han seleccionado datos de consumo, es razonable identificar dicha variable con el nombre "Consumo". Sin embargo, si hemos cargado los datos de una hoja de cálculo con el diseño comentado anteriormente, tan sólo habría que indicar el número de series que se cargan (3 en nuestro ejemplo). En la parte superior de esta pantalla hay que seleccionar con el ratón la posición de los datos en el fichero de lectura, es decir, indicar si éstos están situados en filas o en columnas. También hemos de señalar la celda superior izquierda del primer dato (B2 en nuestro ejemplo). Hacer click sobre el botón *OK*. Inmediatamente, el usuario puede comprobar que la serie seleccionada se encuentra en el fichero de trabajo que creó inicialmente. En caso contrario, algún problema habrá ocurrido con los datos.



IV. GUARDAR FICHERO DE TRABAJO.

Para guardar el fichero de trabajo situar el ratón sobre la opción *File* del menú principal de EV. Seleccionar la opción *Save As* y escribir el nombre con el que se quiere guardar el fichero de trabajo, así como la unidad y el directorio. La extensión de los documentos de trabajo de EV es *wfl*.

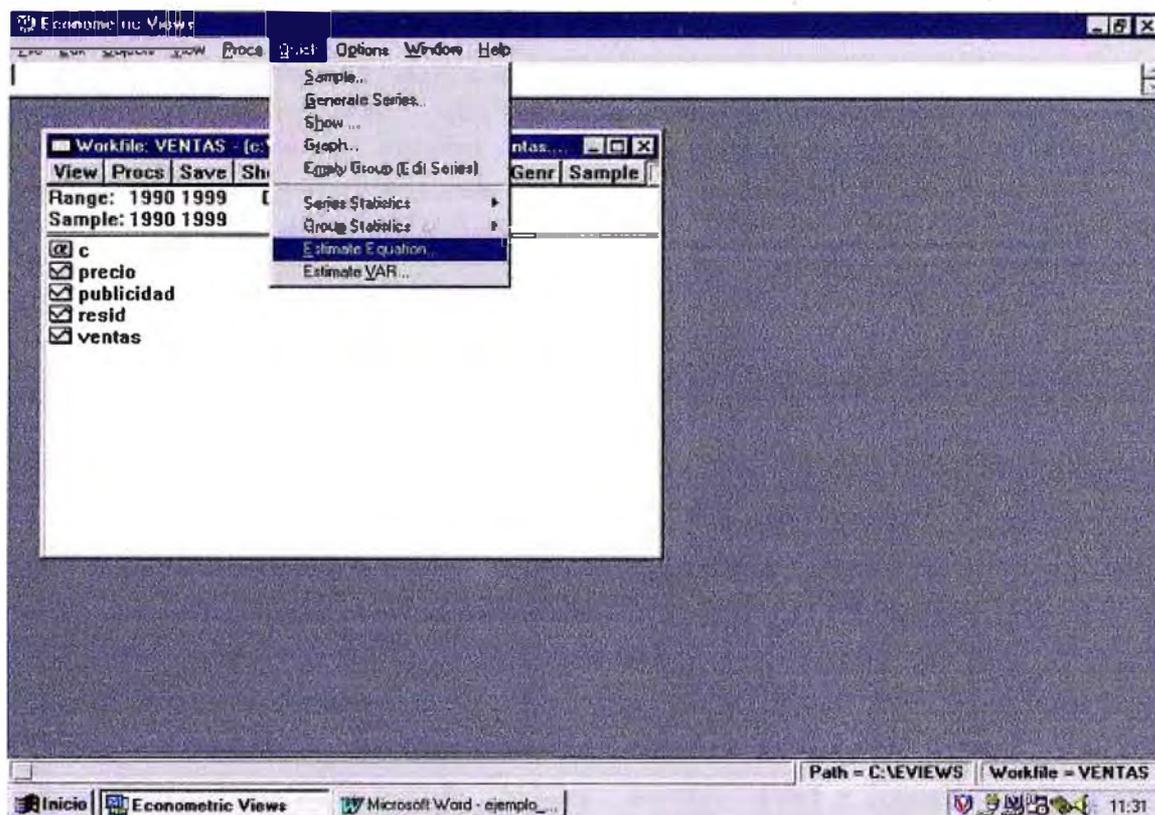


V. SALIR DE EV

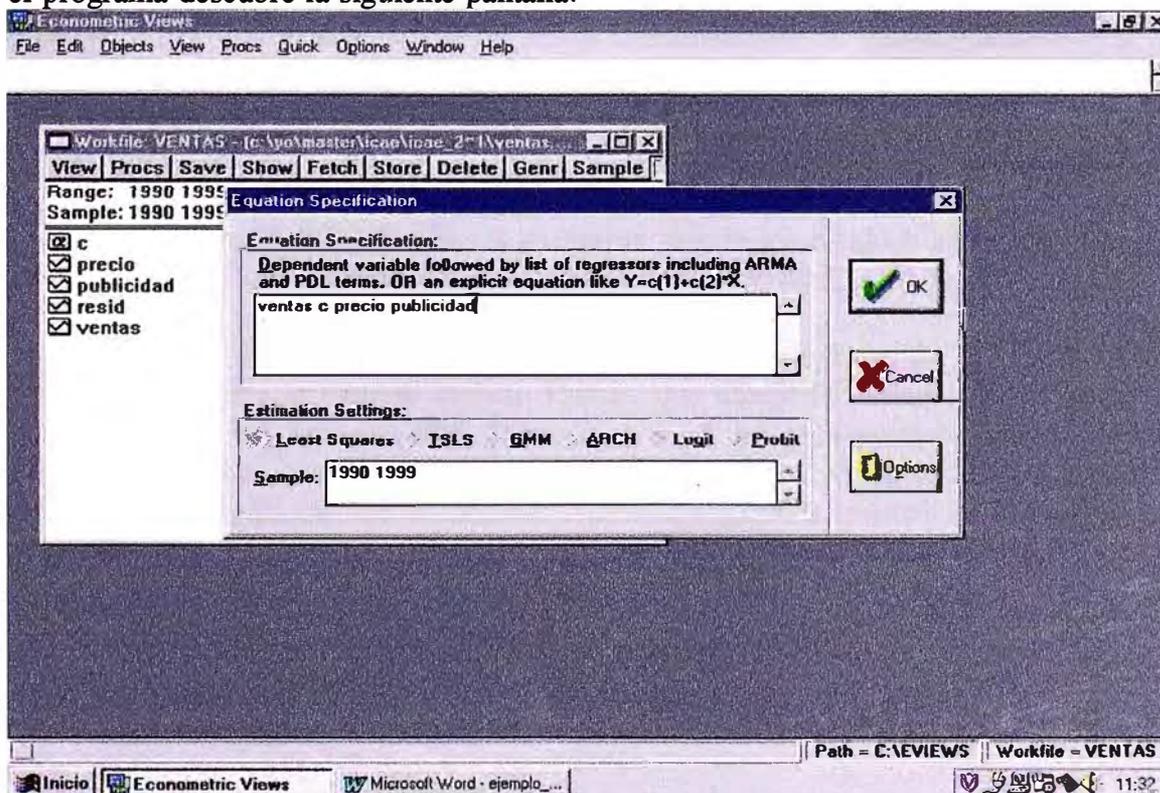
Para salir de EV, seleccionar la opción *File* en el menú principal y elegir la opción *Exit*.

VI. ESTIMACIÓN POR MÍNIMOS CUADRADOS ORDINARIOS.

La manera más directa para estimar un modelo econométrico es usar el menú *Quick* de la barra de herramientas principal y seleccionar *Estimate Equation*:



Alternativamente, se podrían haber seguido los siguientes pasos: seleccionar la opción *Objects* que se encuentra en la parte superior del menú principal (no del fichero de trabajo, donde aparece también esta opción). Seleccionar la opción *New Object* y a continuación la opción *Equation* y pulsar *OK*. En cualquiera de los casos, el programa descubre la siguiente pantalla:



En el recuadro en blanco introducir el nombre de la variable dependiente (*ventas* en nuestro ejemplo), *c* que simboliza el término constante de la regresión (lo cual es opcional) y el nombre de la(s) variable(s) independiente(s) (*precio* y *publicidad*). En la parte inferior de la pantalla el usuario debe seleccionar el método de estimación. Por defecto, EV tiene seleccionada la opción *Least Squares* (Mínimos Cuadrados Ordinarios). Pulsar el botón *OK*. Los resultados de la estimación se ofrecen en una pantalla del siguiente tipo:

Workfile: VENTAS - To: \yo\master\Nico\ev\oe_2-1\ventas...

View Procs Save Show Fetch Store Delete Genr Sample

Equation: UNTITLED Workfile: VENTAS

View Procs Objects Print Name Freeze Estimate Forecast Stats Resids

LS // Dependent Variable is VENTAS
Date: 10/22/00 Time: 11:33
Sample: 1990 1999
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	247.5675	67.35953	3.675315	0.0079
PRECIO	-1.464234	0.648685	-2.257233	0.0586
PUBLICIDAD	2.203809	0.545412	4.040634	0.0049

R-squared 0.943494 Mean dependent var 147.9000
Adjusted R-squared 0.927349 S.D. dependent var 22.19835
S.E. of regression 5.983315 Akaike info criterion 3.821275
Sum squared resid 250.6004 Schwarz criterion 3.912050
Log likelihood -30.29576 F-statistic 58.43983
Durbin-Watson stat 2.549146 Prob(F-statistic) 0.000043

Path = C:\EVIEWS | Workfile = VENTAS

Inicio Econometric Views Microsoft Word - ejemplo... 11:33

Al ejecutar esta operación, el programa guarda los residuos de la regresión en la variable *Resid* que aparece por defecto al crear el fichero de trabajo.

Algunas de las opciones que aparecen en la parte superior del cuadro son: *Print*, que imprime la pantalla activa; *Resids*, que dibuja la variable dependiente, el ajuste del modelo y los residuos; *Forecast*, que se usa para predecir con el modelo estimado; *Estimate*, vuelve a aparecer la ventana en la que se introducen los nombres de las variables para estimar el modelo; *Name*, que permite guardar el *Workfile*, con el nombre que se quiera (por ejemplo, “conjunta”):

The screenshot shows the 'Econometric Views' application window. A dialog box titled 'Object Name' is open, prompting the user to 'Enter name of 16 or fewer characters:'. The name 'conjunta' has been entered. Below the dialog box, a table displays regression statistics for the new object.

	Std. Error	t-Statistic	Prob.
	67.35953	3.675315	0.0079
	0.648685	-2.257233	0.0586
	0.545412	4.040634	0.0049

R-squared	0.943494	Mean dependent var	147.9000
Adjusted R-squared	0.927349	S.D. dependent var	22.19835
S.E. of regression	5.983315	Akaike info criterion	3.821275
Sum squared resid	250.6004	Schwarz criterion	3.912050
Log likelihood	-30.29576	F-statistic	58.43983
Durbin-Watson stat	2.549146	Prob(F-statistic)	0.000043

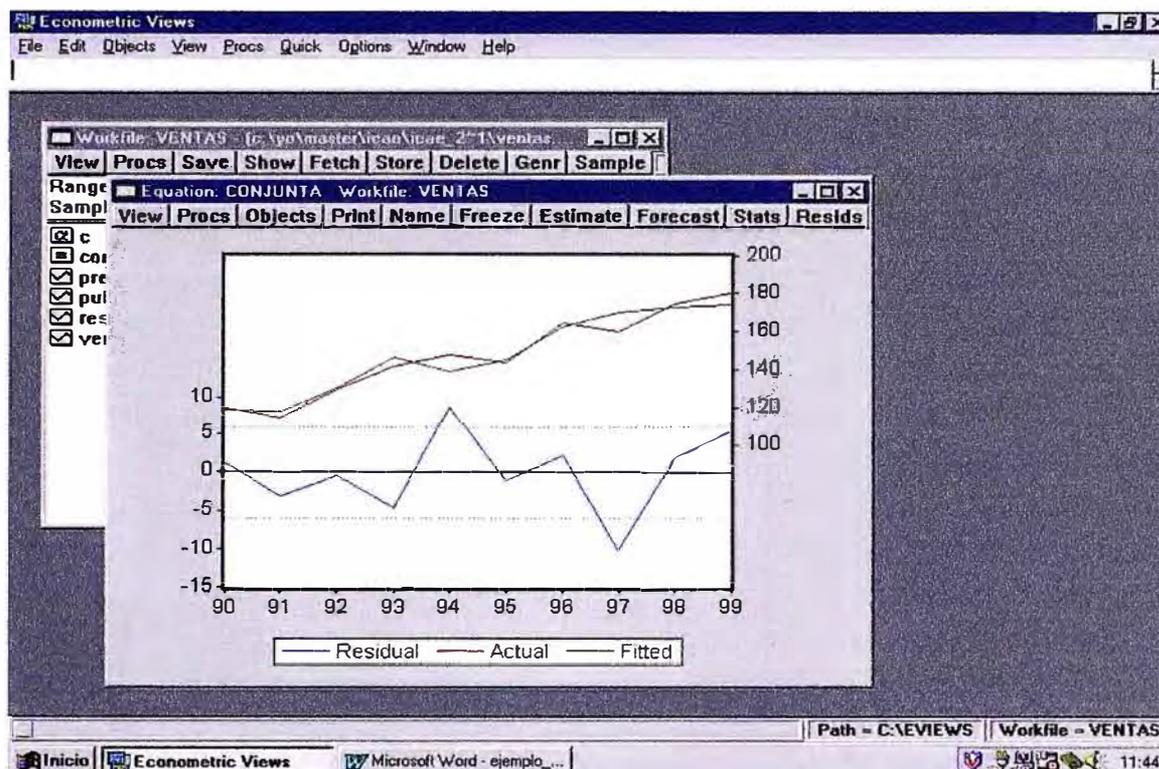
Ahora se tiene un nuevo objeto en el libro de trabajo: una ecuación con nombre "conjunta". La barra de herramientas que se encuentra en la ventana asociada a este nuevo objeto desplegará menús y submenús relacionados con este objeto (lo mismo ocurrirá si el objeto es una serie, o un grupo, o un gráfico). Por ejemplo, si se quiere un gráfico que compare la serie ajustada (*fitted*) con la serie original (*actual*) y, además, muestre los residuos resultantes, se han de dar los pasos descritos en la siguiente figura:

The screenshot shows the 'Econometric Views' application window with the 'Equation: CONJUNTA' selected. The 'Actual Fitted Residual' menu is open, showing options for 'Table' and 'Graph'. Below the menu, a table displays regression statistics for the 'CONJUNTA' equation.

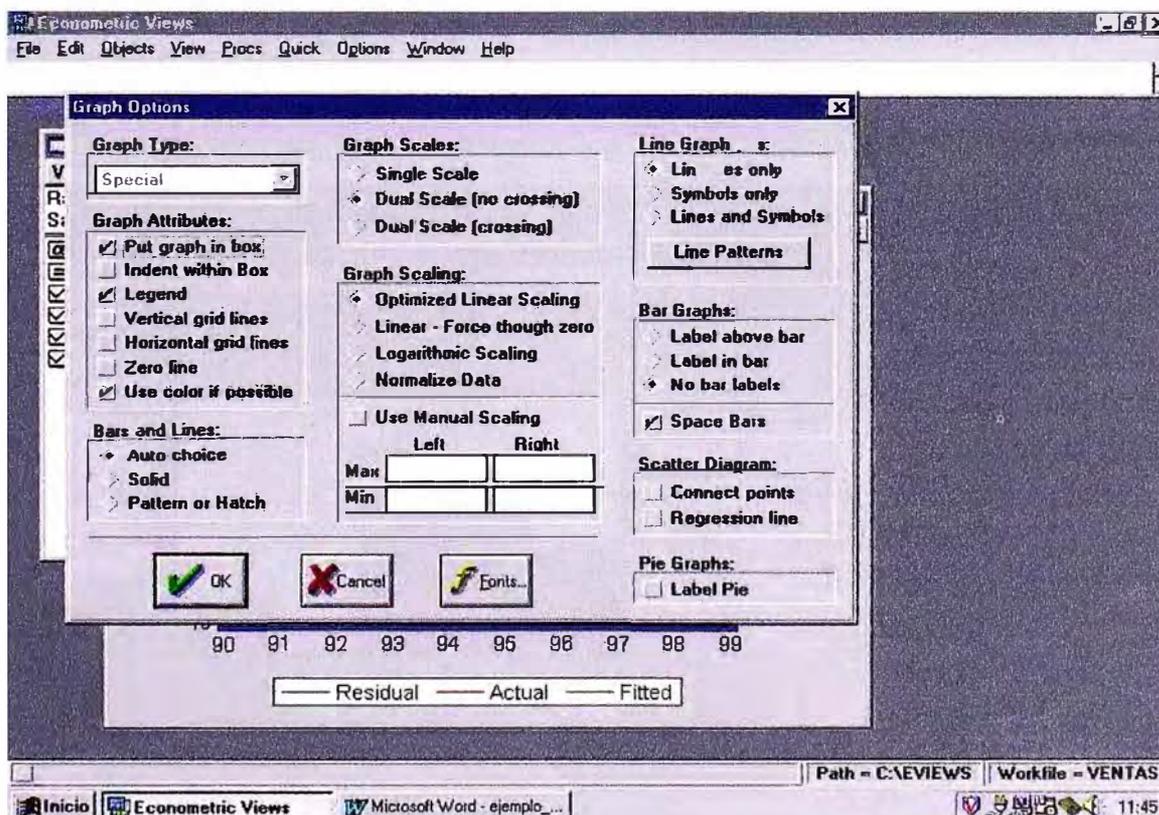
	Coefficient	Std. Error	t-Statistic	Prob.
Label	247.5675	67.35953	3.675315	0.0079
PRECIO	-1.464234	0.648685	-2.257233	0.0586
PUBLICIDAD	2.203809	0.545412	4.040634	0.0049

R-squared	0.943494	Mean dependent var	147.9000
Adjusted R-squared	0.927349	S.D. dependent var	22.19835
S.E. of regression	5.983315	Akaike info criterion	3.821275
Sum squared resid	250.6004	Schwarz criterion	3.912050
Log likelihood	-30.29576	F-statistic	58.43983
Durbin-Watson stat	2.549146	Prob(F-statistic)	0.000043

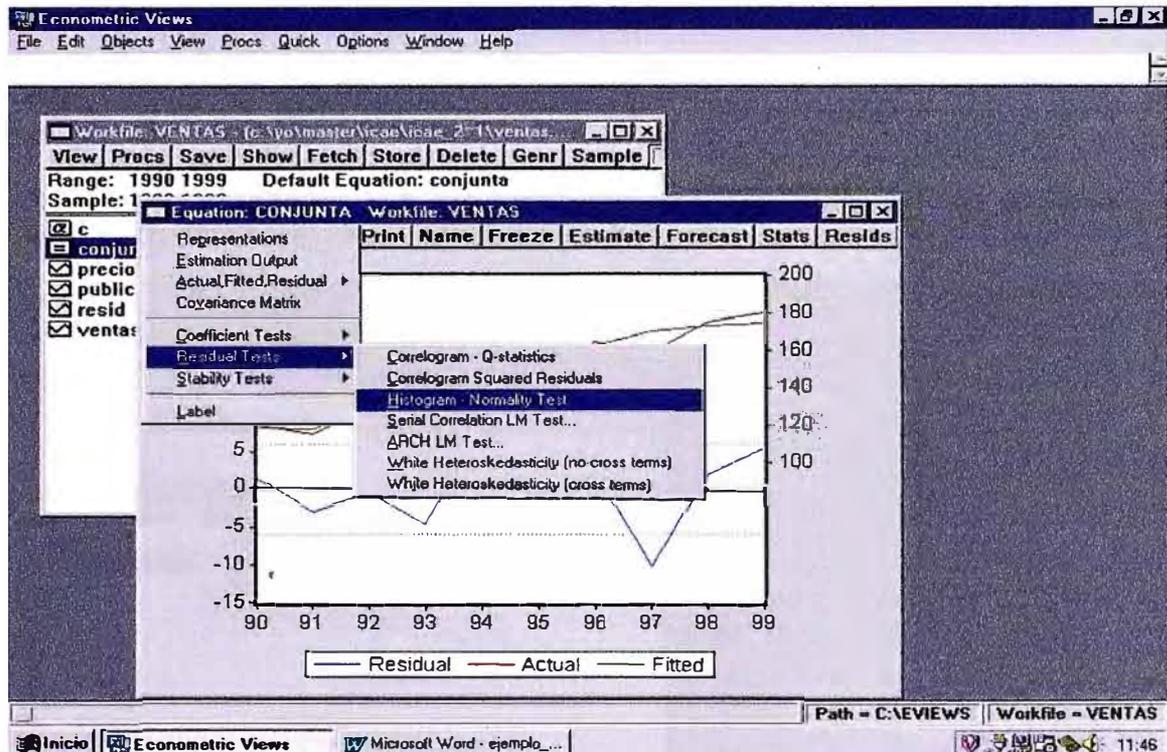
El resultado es el siguiente:



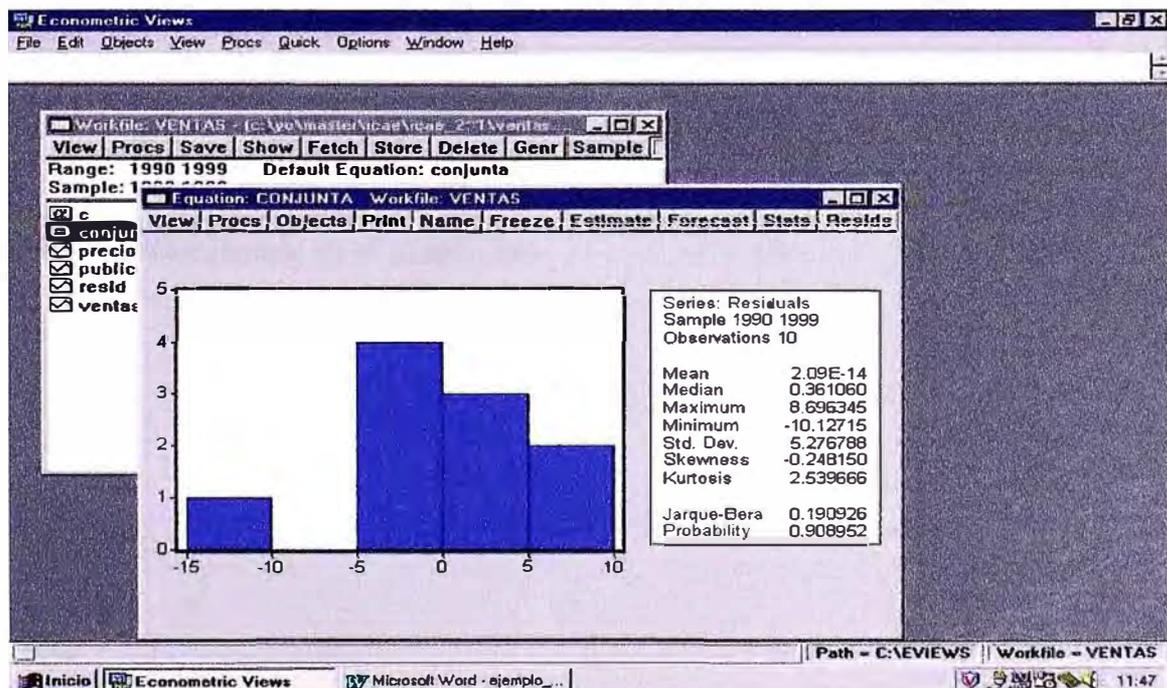
Haciendo doble click sobre el gráfico, EV desplegará una pantalla con opciones del gráfico (tipo de gráfico, doble escala, datos normalizados, añadir rejilla, añadir título, etc.). Esta pantalla será común para cualquier gráfico que se cree.



Por ejemplo, supongamos ahora que se quiere el *histograma* de los residuos para poder discutir si el supuesto de normalidad “parece” razonable o no. Para esto, en *Residual Test* seleccionamos la opción *Histogram Normality Test* :

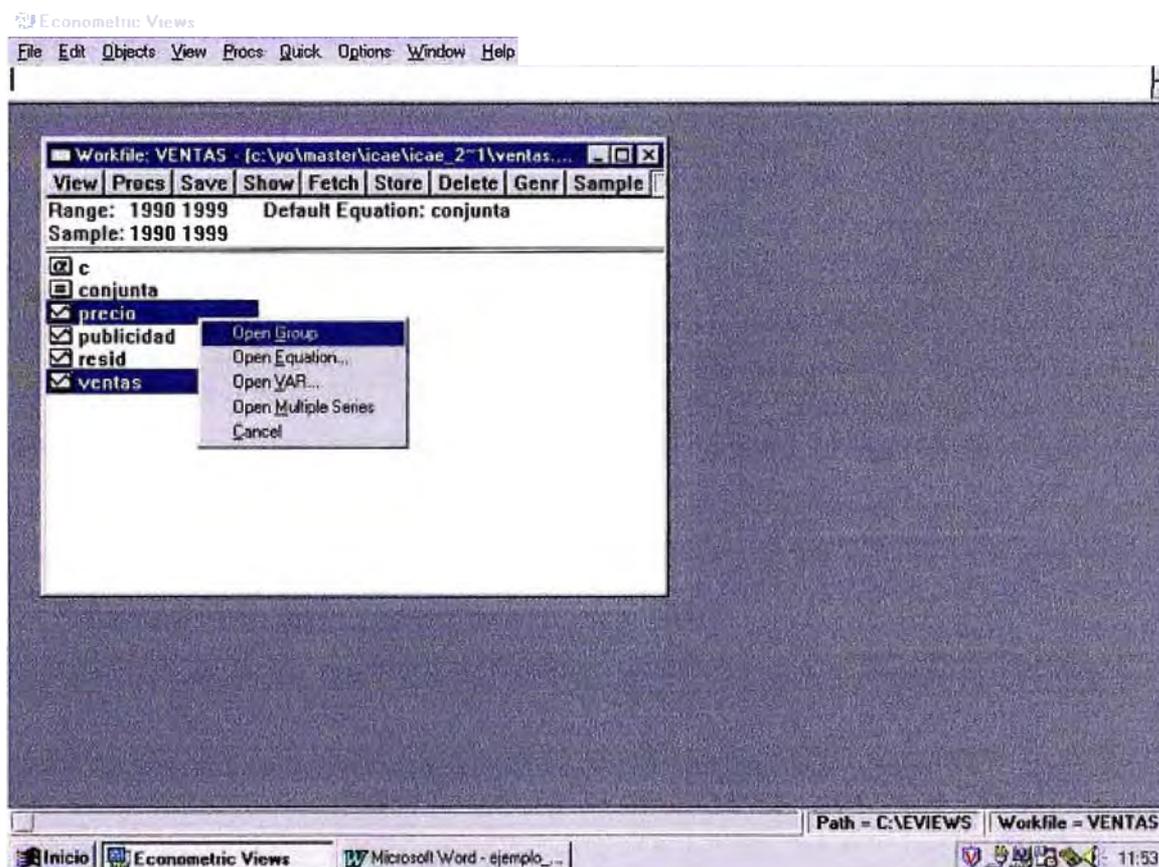


El resultado se muestra en la siguiente figura. Además de presentar el histograma de frecuencia, el programa ofrece un resumen de los estadísticos muestrales más representativos de la serie de residuos, y el resultado del contraste *Jarque-Bera* de Normalidad.



VI. CREACIÓN DE GRUPOS

Para crear un grupo de varias series, primero se han de seleccionar aquéllas que queramos que formen parte de dicho grupo (usar la tecla *Ctrl* del ordenador para marcar series no-consecutivas en el *Workfile*); a continuación, hacer doble click en alguna de las series marcadas y seleccionar la opción *Open Group* (un grupo también podría crearse a través de los submenús desplegados en *Objects* de la barra de herramientas principal):



Una vez creado el grupo, se tienen en el mismo objeto las series seleccionadas. Dándole el nombre que se quiera, se pueden dibujar conjuntamente las series, estimar una ecuación con las series del grupo (tomando como variable endógena la primera que haya seleccionada en el grupo), etc.

The screenshot shows the 'Econometric Views' application window. The main window displays a data table for the group 'VENTA_PRECIO' in the workfile 'VENTAS'. The table has columns for 'obs', 'VENTAS', and 'PRECIO'. The data rows are as follows:

obs	VENTAS	PRECIO
1990	120.0000	100.0000
1991	115.0000	102.0000
1992	130.0000	95.00000
1993	142.0000	90.00000
1994	148.0000	92.00000
1995	144.0000	94.00000
1996	165.0000	88.00000
1997	160.0000	86.00000
1998	175.0000	90.00000
1999	180.0000	86.00000

The taskbar at the bottom shows the system tray with the time 11:54 and the path 'C:\EVIEW5'.

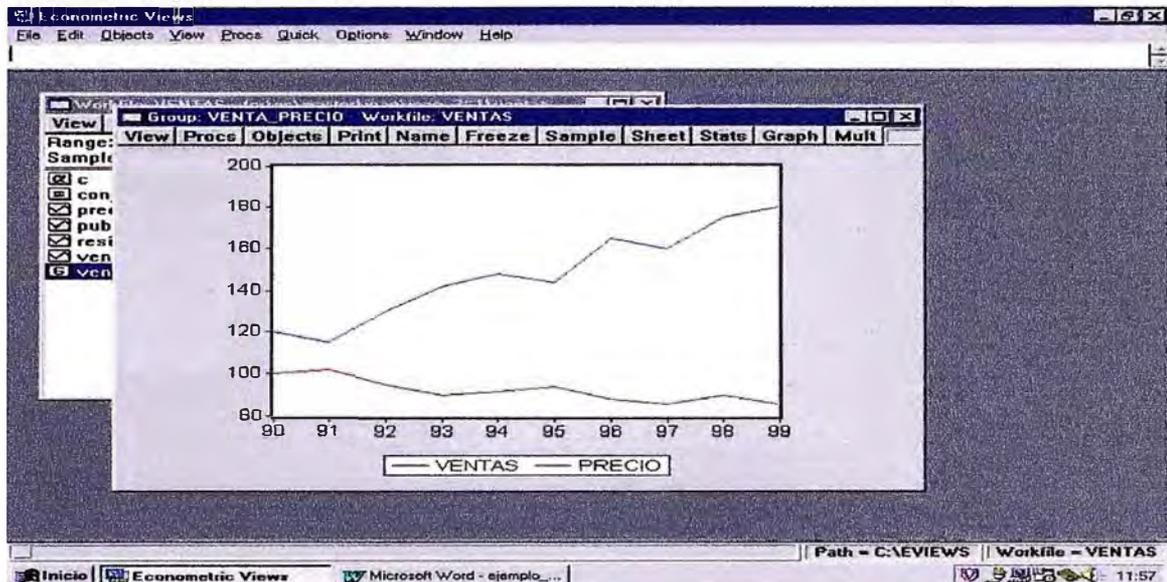
Por ejemplo, si se quieren dibujar conjuntamente las series del grupo, en *View* seleccionar la opción *Graph (Multiple Graphs* coloca todos los gráficos individuales en una misma ventana).

The screenshot shows the 'Econometric Views' application window with the 'View' menu open. The menu options are listed in the table below:

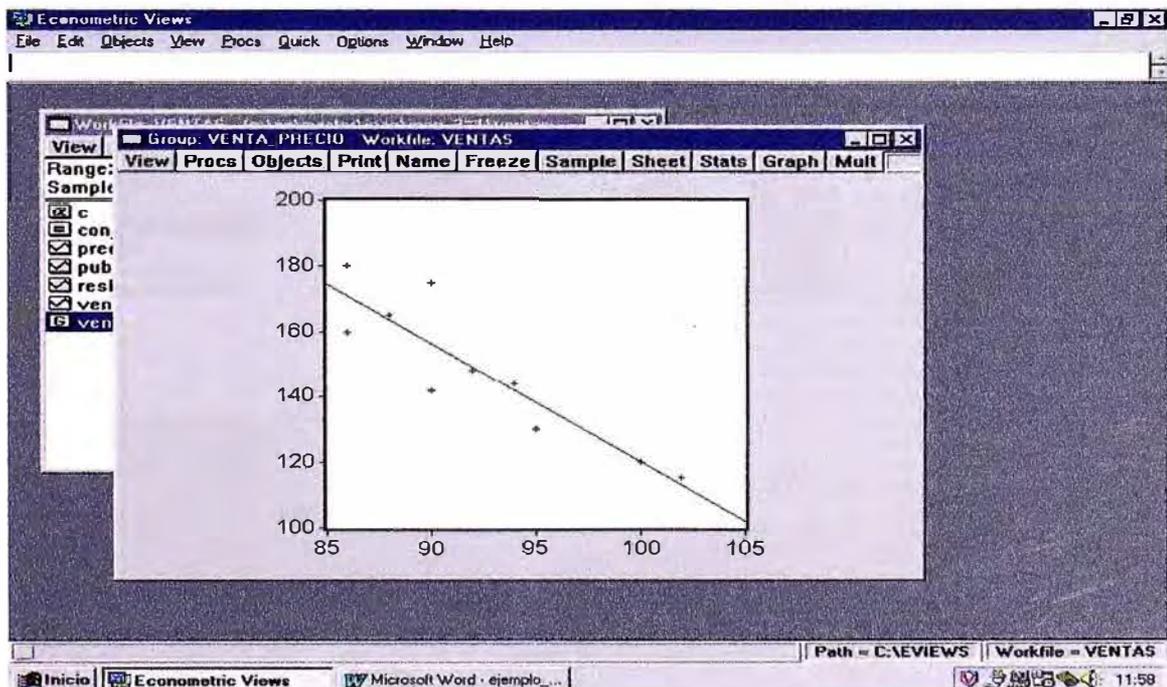
View	Print	Name	Freeze	Edit+/-	InsDel	Smpl+/-	Transpose	Tlt
SpreadSheet								
Graph		PRECIO						
Multiple Graphs		100.0000						
Descriptive Stats		102.0000						
Crosstab		95.00000						
Correlations		90.00000						
Covariances		92.00000						
Correlogram (1) ...		94.00000						
Cross Correlation (2) ...		88.00000						
Cointegration Test...		86.00000						
Granger Causality...		90.00000						
Label		86.0000						

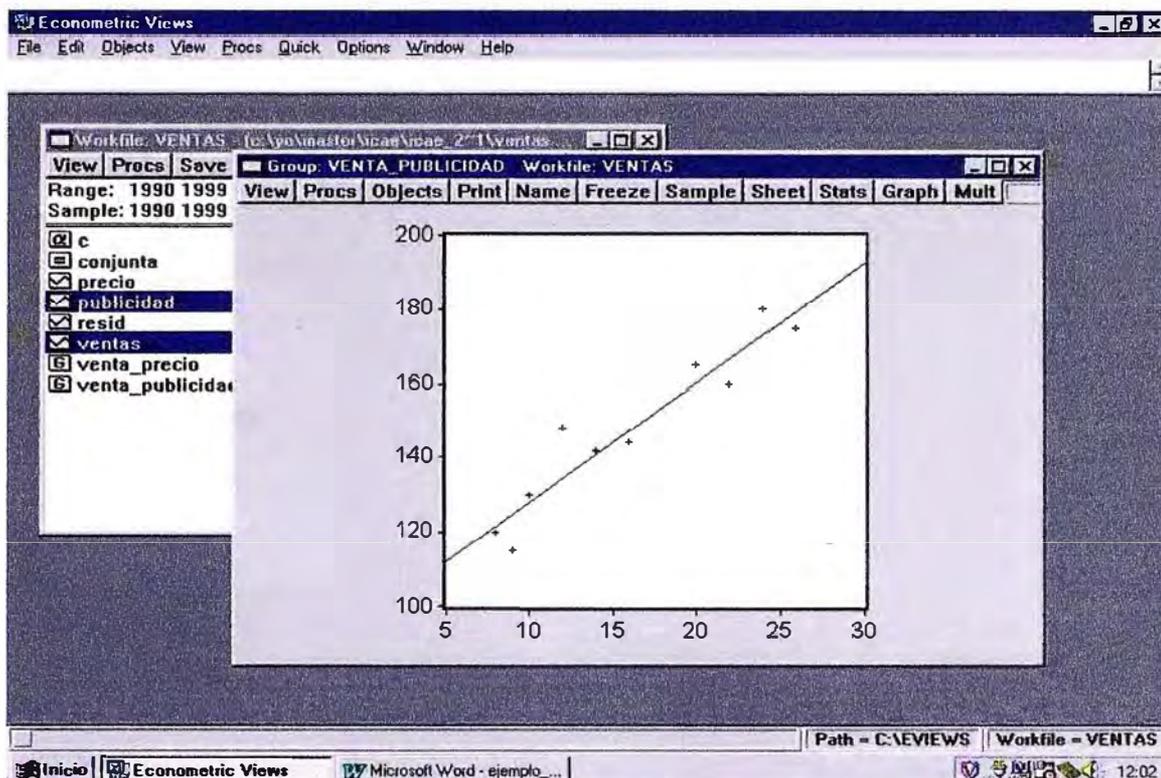
The taskbar at the bottom shows the system tray with the time 11:57 and the path 'C:\EVIEW5'.

Por defecto, aparece un gráfico conjunto. Hay que tener cuidado al construir estos gráficos conjuntos, ya que la escala de los datos es fundamental. En caso necesario, es mejor seleccionar la opción de doble escala o usar datos normalizados para poder apreciar sin engaños la evolución conjunta de las series.

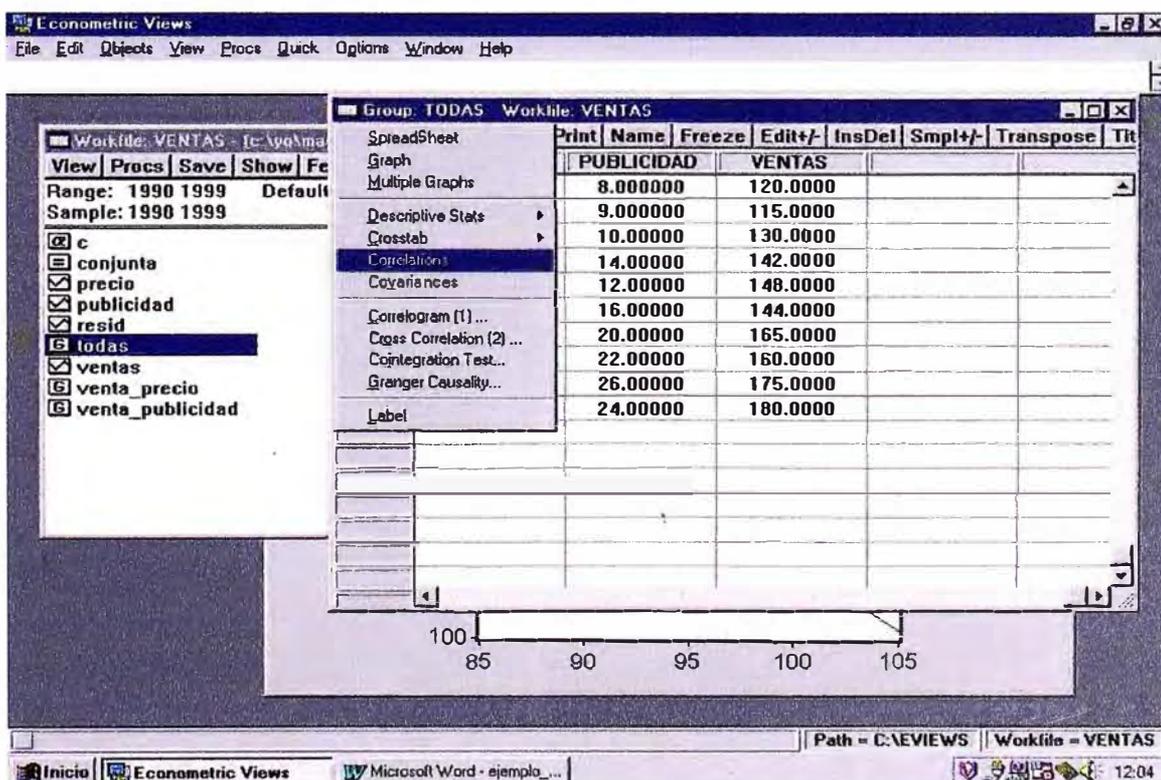


Con hacer doble click en el gráfico, se despliega una ventana con las propiedades del gráfico. Por ejemplo, en el caso que nos concierne, estaríamos interesados en un gráfico de dispersión X-Y (*Scatter Diagram*). Una vez seleccionado el gráfico de dispersión, podríamos activar la opción *Regression Line* para que, además, pintase la línea de regresión. A continuación, presentamos los gráficos X-Y del volumen de ventas con los gastos en publicidad y la ventas y el precio del producto. En el primer caso, se aprecia una relación claramente positiva, mientras que en el segundo resulta ser negativa.

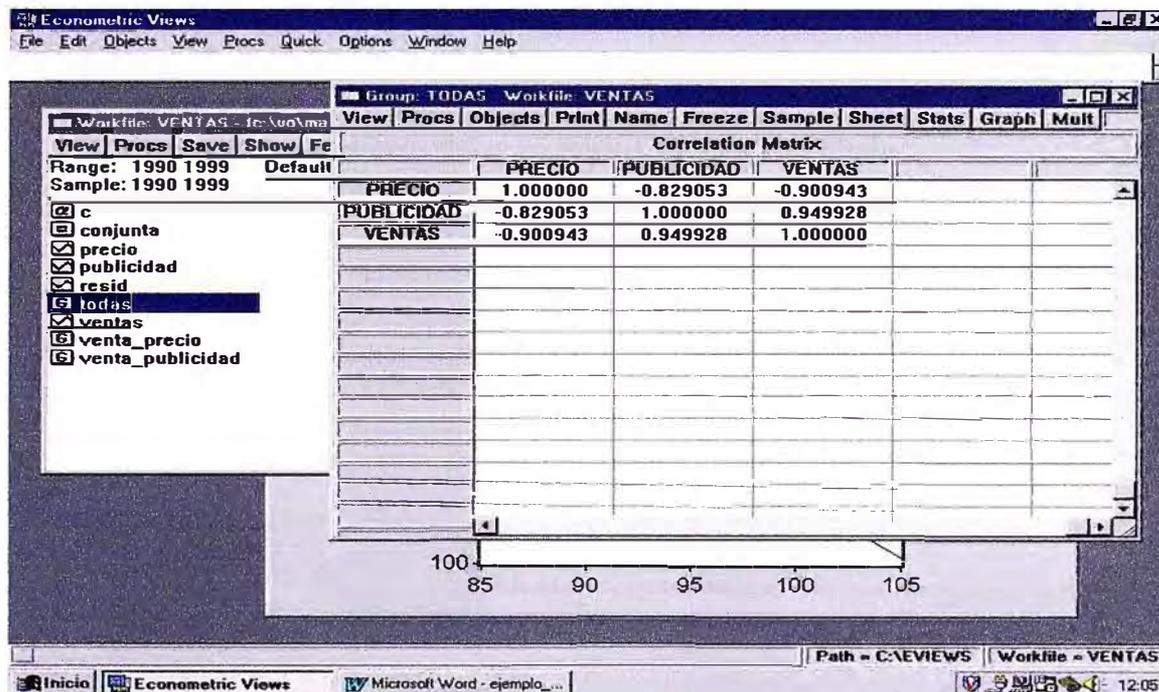




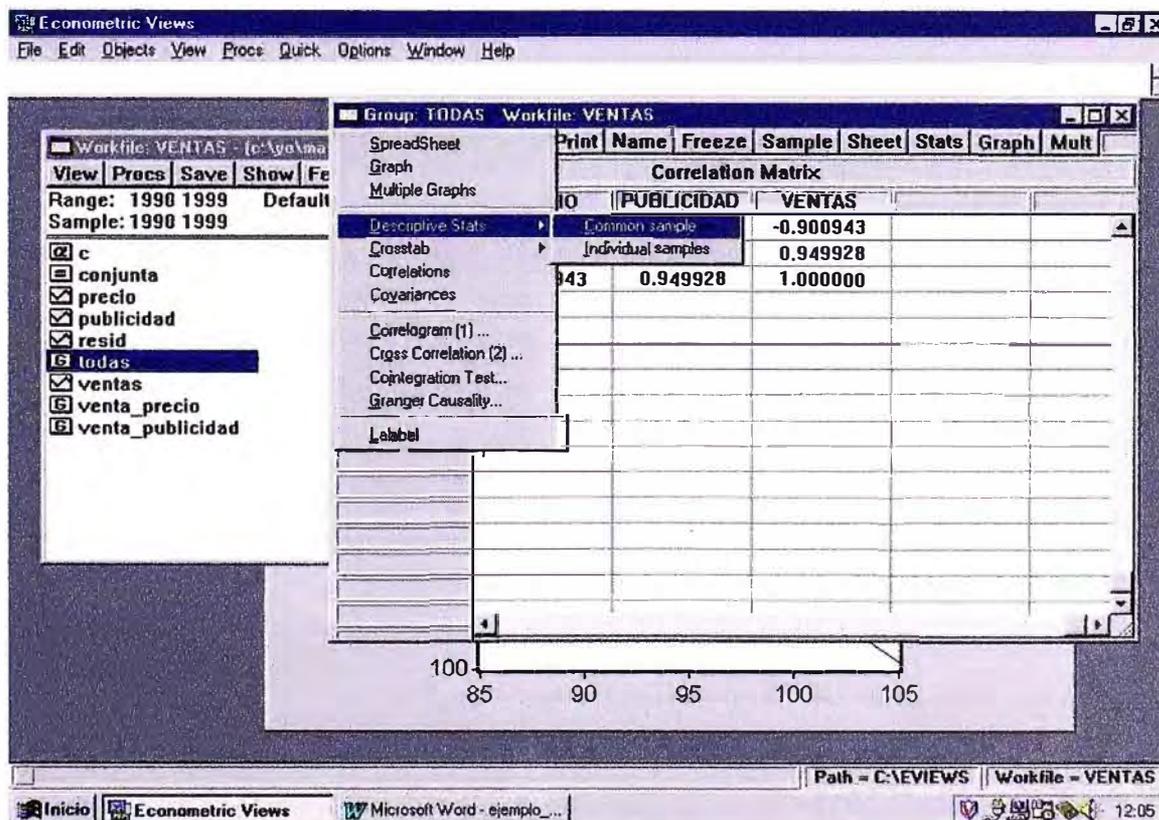
Si se define un grupo con las tres series del análisis (y por ejemplo, lo llamamos "todas"), podemos hacer que EV muestre una serie de estadísticos que resuman el grado de relación (lineal) entre las variables. Lo más común es la matriz de correlaciones entre las tres series. Para esto, en *View* se selecciona la opción *Correlations* :



La diagonal principal de dicha matriz de correlaciones es, obviamente, un vector de unos. Por ejemplo, la correlación entre ventas y publicidad es 0,9499 y la de ventas y precio es negativa e igual a $-0,9009$.



Si siguiendo los mismos pasos, se puede marcar del menú desplegado por *View* la opción *Descriptive Stats* y seleccionar *Common Sample*.



Se obtiene una tabla resumen de los principales estadísticos descriptivos de las tres series del grupo, usando muestras comunes a todas ellas.

Group: TODAS Workfile: VENTAS

	PRECIO	PUBLICIDAD	VENTAS
Mean	92.30000	16.10000	147.9000
Median	91.00000	15.00000	146.0000
Maximum	102.0000	26.00000	180.0000
Minimum	86.00000	8.000000	115.0000
Std. Dev.	5.498485	6.539623	22.19835
Skewness	0.527919	0.211958	-0.036350
Kurtosis	1.942284	1.426199	1.642596
Jarque-Bera	0.930650	1.106898	0.769929
Probability	0.627931	0.574963	0.680475
Observations	10	10	10

Tan sólo a modo de ejemplo, una vez que se tiene el grupo "todas" abierto, es muy sencillo llevar a cabo una regresión entre sus componentes. En *Procs*, seleccionar *Make Equation* y aparecerá una pantalla como las descritas anteriormente.

Group: TODAS Workfile: VENTAS

View	Make Equation	Freeze	Edit+/-	InsDel	Smpl+/-	Transpose	Tit
obs	Make Vector Autoregression..	AD					
1990	100.0000	8.000000					
1991	102.0000	9.000000					
1992	95.00000	10.00000					
1993	90.00000	14.00000					
1994	92.00000	12.00000					
1995	94.00000	16.00000					
1996	88.00000	20.00000					
1997	86.00000	22.00000					
1998	90.00000	26.00000					
1999	86.00000	24.00000					

VII. SELECCIONAR UNA SOLA SERIE

Hacer doble click en cualquier serie del libro de trabajo. Por ejemplo, una vez que se tiene la serie seleccionada, en *View - Histogram and Stats*, aparece el histograma de frecuencia de la serie y un resumen de sus estadísticos muestrales.

The screenshot shows the 'Econometric Views' application window. The 'Series: VENTAS' window is open, displaying a menu with 'Histogram and Stats' selected. The main data table is titled 'VENTAS' and shows the following data:

	1	2	3	4
Last updated: 10/22/00 - 11:18				
	115.0000	130.0000	142.0000	148.0000
	165.0000	160.0000	175.0000	180.0000

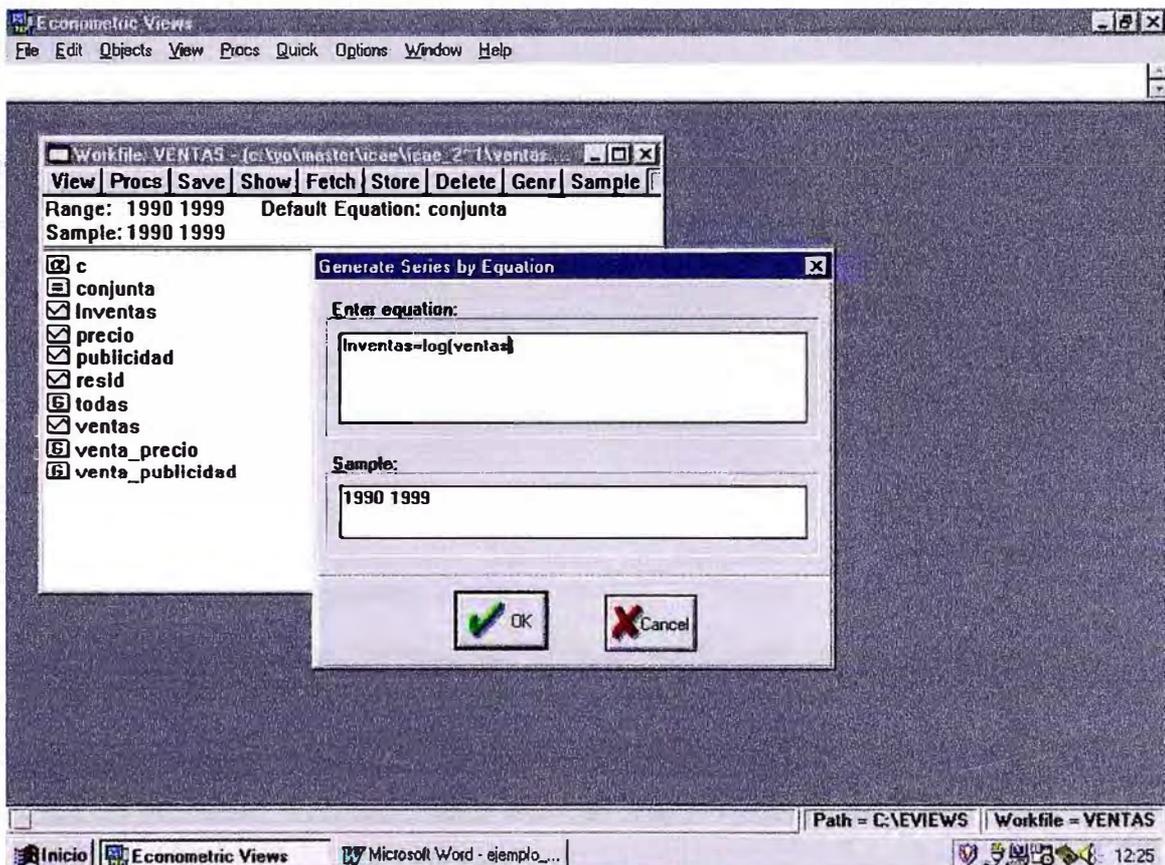
Para poder editar la tabla de datos de una serie (o de un grupo), pulsar pestaña *Edit +/-* en la barra de herramientas asociada a la serie (o grupo).

The screenshot shows the 'Econometric Views' application window. The 'Series: VENTAS' window is open, displaying a menu with 'Edit +/-' selected. The main data table is titled 'VENTAS' and shows the following data:

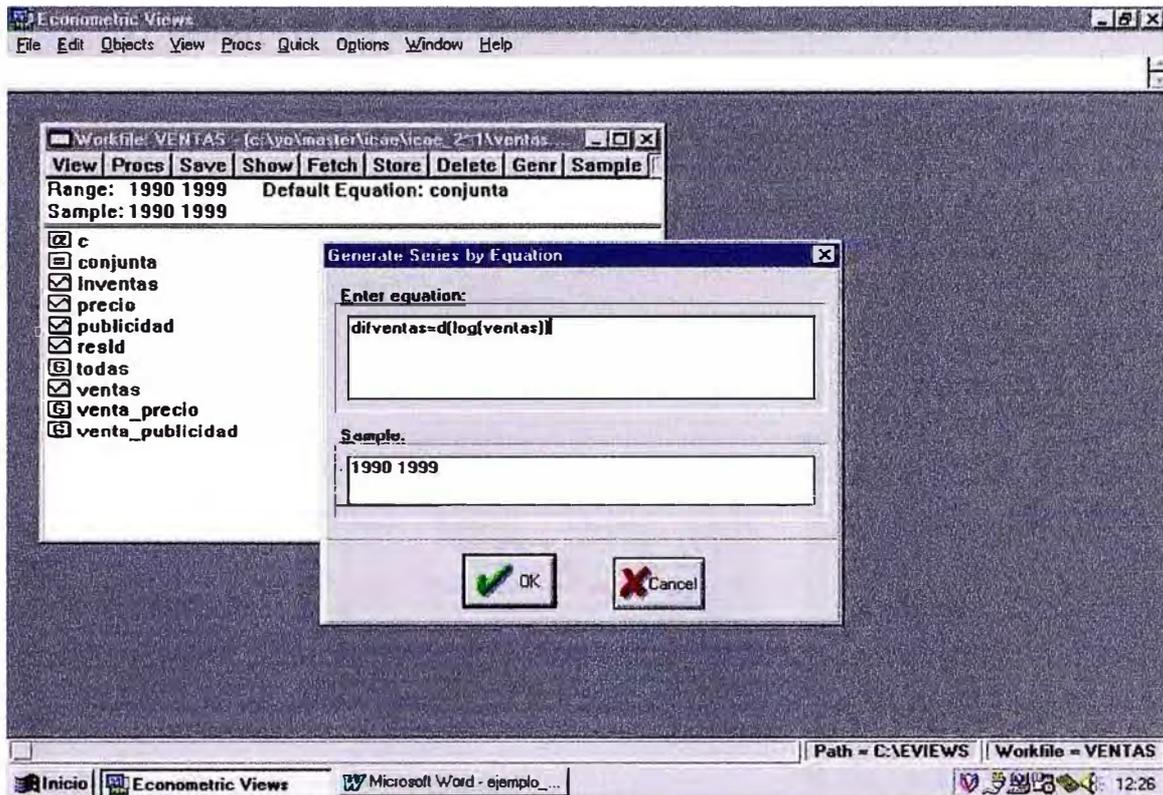
	0	1	2	3	4
Last updated: 10/22/00 - 11:18					
1990	120.0000	115.0000	130.0000	142.0000	148.0000
1995	144.0000	165.0000	160.0000	175.0000	180.0000

VIII. TRANSFORMACION DE DATOS

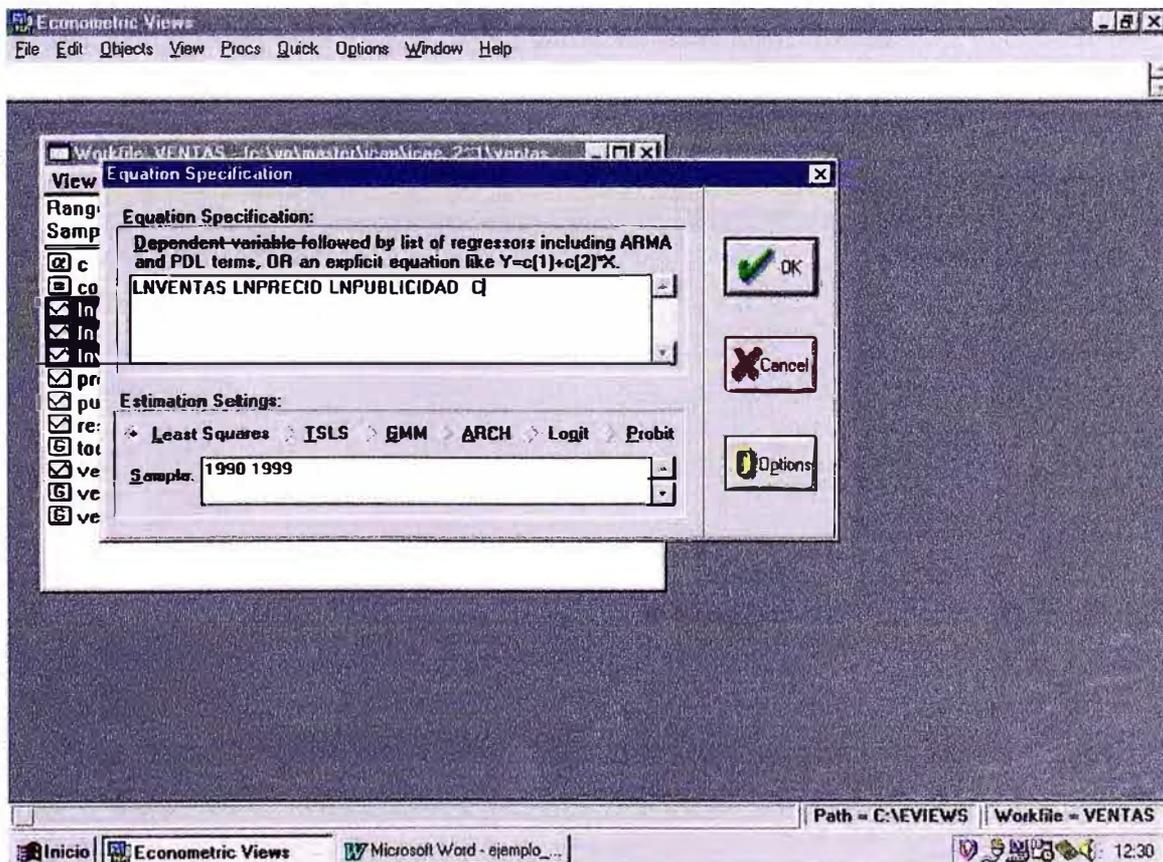
Se pueden generar transformaciones de las series cargadas en el *Workfile*. Por ejemplo, supongamos que se quiere estimar un modelo de regresión *log-lineal*. Para esto, una alternativa es transformar primero las variables originales tomando logaritmos neperianos y proceder a estimar una ecuación lineal con las nuevas variables. Pulsando la pestaña *Genr* en la barra de herramientas del *Workfile*, se desplegará una ventana como la siguiente:



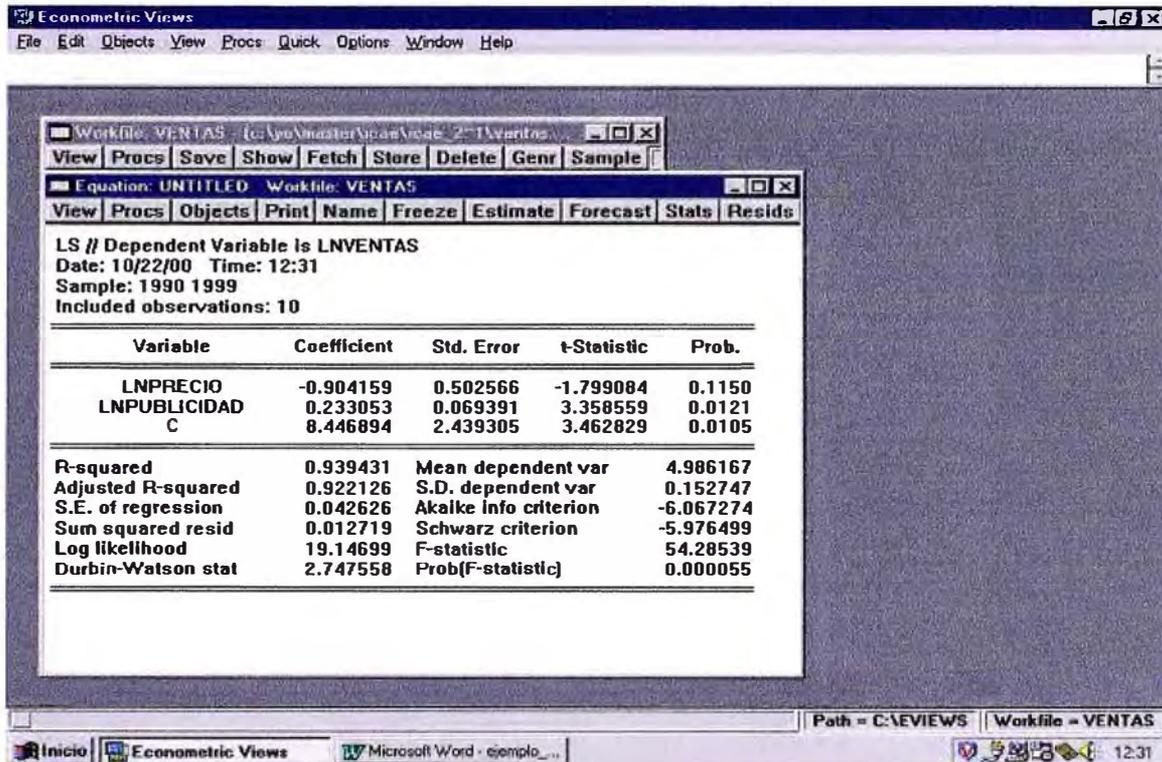
En el cuadro blanco se tiene que introducir una instrucción con la siguiente estructura: “*nombre de la variable transformada = función (variable a transformar)*”. Por ejemplo, para tomar el logaritmo neperiano de ventas: $Inventas = \log(ventas)$; para hallar la tasa de crecimiento logarítmico: $difventas = d(\log(ventas))$, donde el comando “ $d(.)$ ” denota el operador primeras diferencias ($d(y) = y(t) - y(t-1)$).



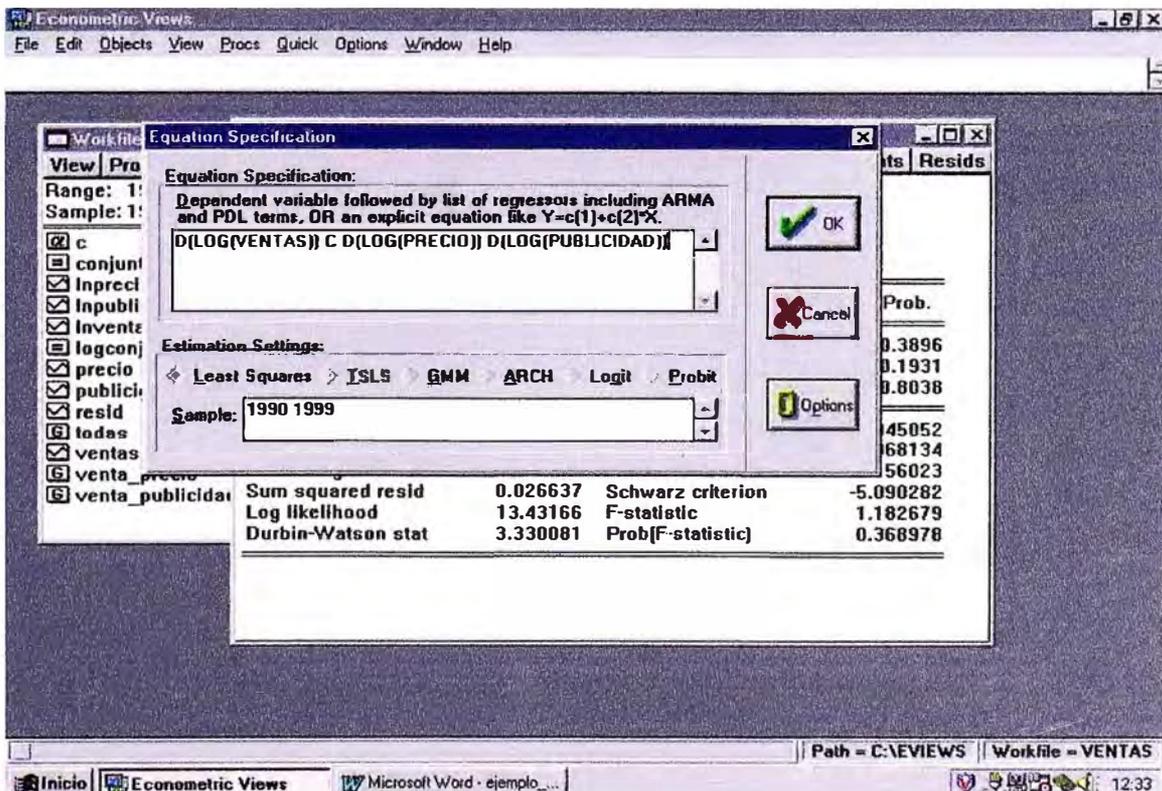
Por tanto, para estimar el modelo log-lineal, una vez generadas las nuevas series en logaritmos (*lnventas*, *lnprecios* y *lnpublicidad*), especificaremos una ecuación de la siguiente manera:



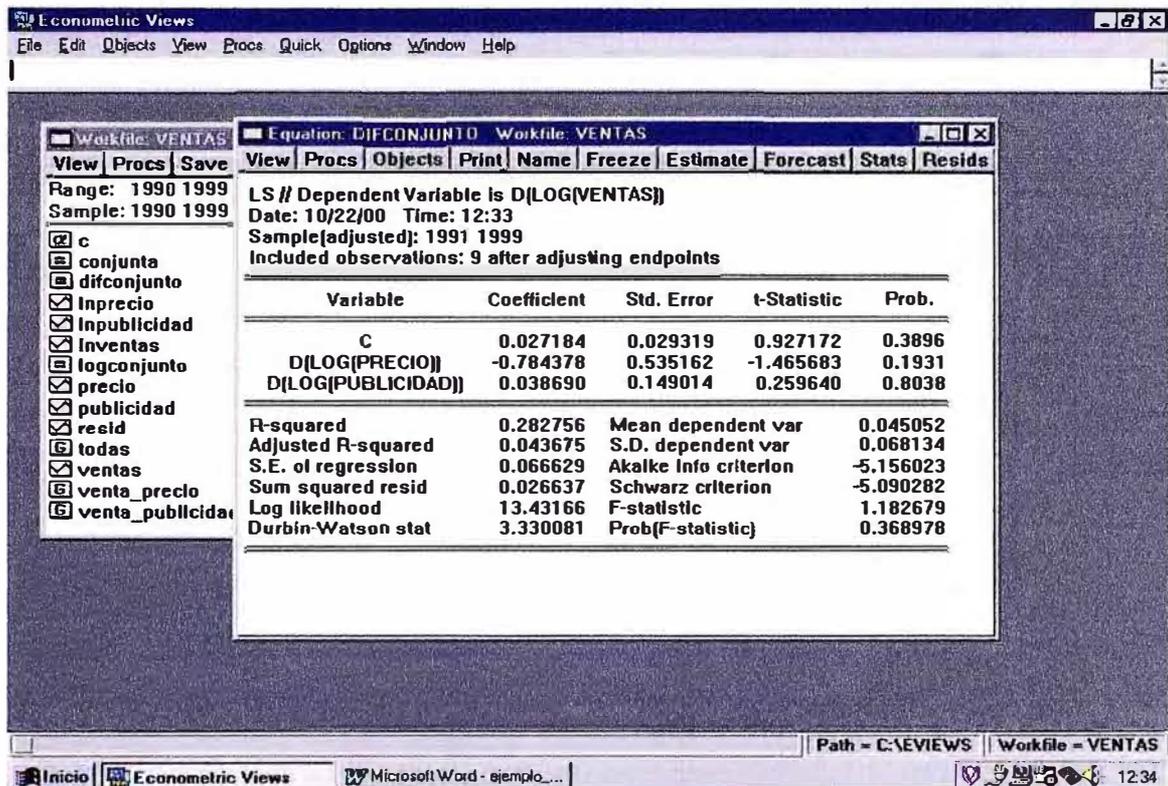
El resultado de la estimación del modelo log-lineal queda reflejado en la siguiente figura:



Para estimar una ecuación con variables transformadas, se podría indicar tal transformación en el propio recuadro de la ecuación. Por ejemplo, para estimar de forma directa el modelo usando las variables en tasas logarítmicas:



El resultado de la estimación del modelo en tasas logarítmicas se presenta en la siguiente figura:



IX. CONTRASTES SOBRE LOS COEFICIENTES

Para contrastar cualquier restricción lineal sobre los parámetros del modelo, se selecciona la opción *View* en la pantalla de la ecuación correspondiente. Entonces, EV desplegará la siguiente caja de diálogos. Seleccionar la opción *Coefficient Test* y posteriormente *Wald Coefficient Restrictions*.

Label	Name	Freeze	Estimate	Forecast	Stats	Resids
			247.5675	67.35953	3.675315	0.0079
PRECIO			-1.464234	0.648685	-2.257233	0.0586
PUBLICIDAD			2.203809	0.545412	4.040634	0.0049
R-squared			0.943494	Mean dependent var		147.9000
Adjusted R-squared			0.927349	S.D. dependent var		22.19835
S.E. of regression			5.983315	Akaike info criterion		3.821275
Sum squared resid			250.6004	Schwarz criterion		3.912050
Log likelihood			-30.29576	F-statistic		58.43983
Durbin-Watson stat			2.549146	Prob(F-statistic)		0.000043

Al hacerlo, EV desplegará la siguiente pantalla:

Label	Name	Freeze	Estimate	Forecast	Stats	Resids
			247.5675	67.35953	3.675315	0.0079
PRECIO			-1.464234	0.648685	-2.257233	0.0586
PUBLICIDAD			2.203809	0.545412	4.040634	0.0049
R-squared			0.943494	Mean dependent var		147.9000
Adjusted R-squared			0.927349	S.D. dependent var		22.19835
S.E. of regression			5.983315	Akaike info criterion		3.821275
Sum squared resid			250.6004	Schwarz criterion		3.912050
Log likelihood			-30.29576	F-statistic		58.43983
Durbin-Watson stat			2.549146	Prob(F-statistic)		0.000043

En el recuadro blanco se introducen las restricciones (a contrastar) separadas por comas. Por ejemplo, si uno está interesado en contrastar la existencia de rendimientos constantes a escala en una determinada función de producción, se debe introducir en el recuadro blanco la siguiente instrucción: $c(2)+c(3)=1$; si se quiere contrastar la significación conjunta de las pendientes del modelo: $c(2)=c(3)=0$ (o $c(2)=0, c(3)=0$).

Posteriormente, EV desplegará la siguiente pantalla, donde aparece el valor de los estadísticos "F" y el asintótico de "Wald" (en el ejemplo, hemos contrastado la significación conjunta de las pendientes del modelo).

The screenshot shows the 'Econometric Views' window with the 'Equation. CONJUNTA' dialog box open. The dialog box displays the Wald Test results for the equation 'CONJUNTA'. The null hypothesis is $C(2)=0$ and $C(3)=0$. The F-statistic is 58.43983 with a probability of 0.000043, and the Chi-square is 116.8797 with a probability of 0.000000.

Statistic	Value	Probability
F-statistic	58.43983	0.000043
Chi-square	116.8797	0.000000

En el siguiente ejemplo, se contrasta si los coeficientes asociados a la variable "publicidad" y "precio" son iguales en magnitud (es decir, si suman cero).

The screenshot shows the Econometric Views interface with a window titled 'Equation: CONJUNTA'. The main window displays the results of a Wald Test for the equation 'CONJUNTA'. The null hypothesis is $C(2)+C(3)=0$. The test results are as follows:

Statistic	Value	Probability
F-statistic	0.419165	0.538000
Chi-square	0.419165	0.517354

The interface also shows a list of objects on the left, including 'conjunta', 'difconj', 'inprec', 'inpubl', 'invent', 'logconj', 'precio', 'publicidad', 'resid', 'todas', 'ventas', 'venta', and 'venta_'. The status bar at the bottom indicates the path is 'C:\EVIEWWS' and the workfile is 'VENTAS'.

X. COEFICIENTES SIMPLES Y PARCIALES DE REGRESIÓN Y DETERMINACIÓN

Los coeficientes simples de regresión y determinación se obtienen al regresar la variable endógena (Ventas) sobre cada una de las potencialmente explicativas (Publicidad y Precio). El primero es la pendiente del modelo y el segundo el coeficiente de determinación (el R2) asociado a la ecuación. Los resultados son los siguientes:

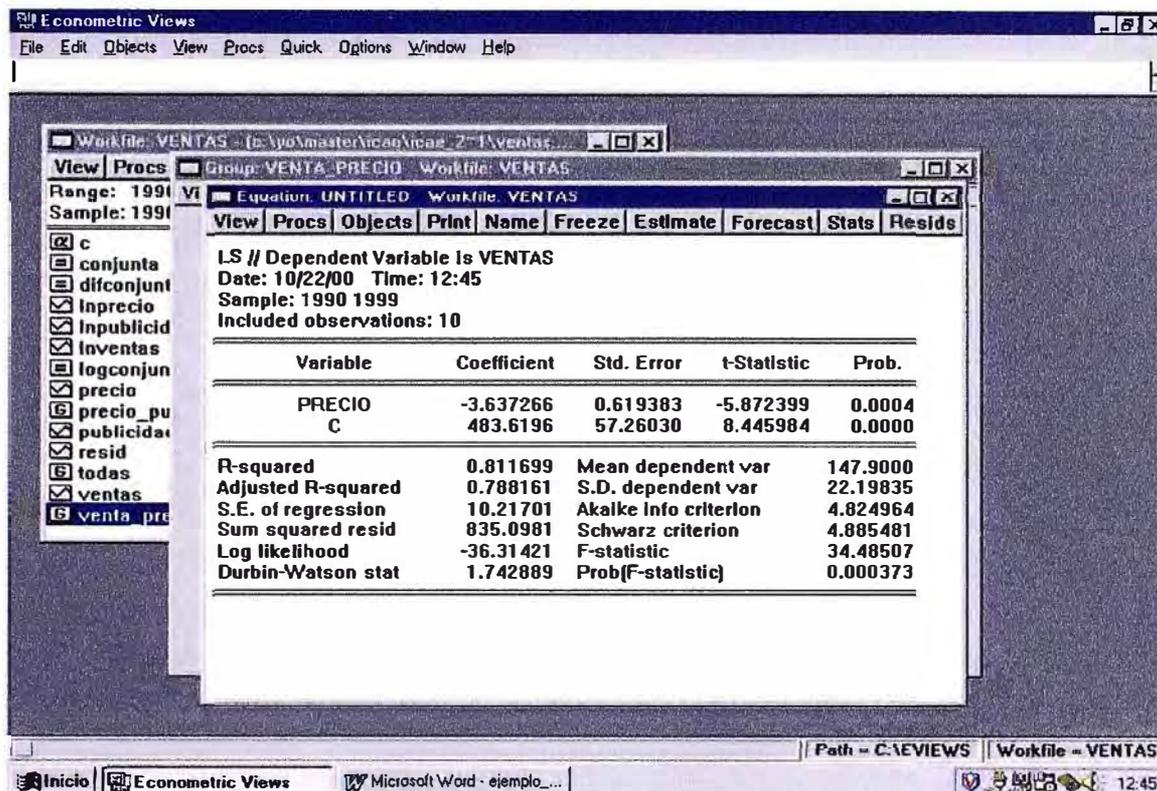
The screenshot shows the Econometric Views interface with a window titled 'Equation EQ_SIMPLE_PUBLI'. The main window displays the results of a simple regression for the dependent variable 'VENTAS'. The results are as follows:

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PUBLICIDAD	3.224474	0.374997	8.598667	0.0000
C	95.98597	6.470192	14.83510	0.0000

Additional statistics shown include:

- R-squared: 0.902364
- Adjusted R-squared: 0.890160
- S.E. of regression: 7.357016
- Sum squared resid: 433.0055
- Log likelihood: -33.03021
- Durbin-Watson stat: 2.262896
- Mean dependent var: 147.9000
- S.D. dependent var: 22.19835
- Akaike info criterion: 4.168165
- Schwarz criterion: 4.228682
- F-statistic: 73.93707
- Prob(F-statistic): 0.000026

The interface also shows a list of objects on the left, including 'conjunta', 'difconjunto', 'eq_simple_precio', 'eq_simple_publ', 'inprecio', 'inpublicidad', 'inventas', 'logconjunto', 'precio', 'precio_publicidad', 'publicidad', 'resid', and 'resid_ven_pre'. The status bar at the bottom indicates the path is 'C:\EVIEWWS' and the workfile is 'VENTAS'.



Nótese que los coeficientes de regresión simple son distintos a los coeficientes resultantes cuando se estimaba el modelo conjuntamente. La razón de esta diferencia está en que las dos variables exógenas son significativas para explicar la evolución de la endógena y, además, están correlacionadas entre sí.

Los coeficientes de regresión parciales coinciden con las pendientes obtenidas en la regresión conjunta estimada anteriormente. Sin embargo, los coeficientes de determinación parcial se tienen que obtener de forma indirecta, ya que el que se obtuvo en la regresión conjunta es lo que llamamos coeficiente de determinación global.

El cálculo de los coeficientes parciales se obtiene regresando variables ortogonales al resto de variables explicativas. Por ejemplo, si quiere calcular el coeficiente parcial entre ventas y publicidad, se han de regresar primero ambas variables sobre la variable precios. Los residuos resultantes son ortogonales a esta variable, ya que son el resultado de descontar de ambas variables el efecto de la variable precio. Finalmente, al regresar los residuos resultantes de la ecuación de ventas sobre la de publicidad, obtendremos los coeficientes de determinación (el R² resultante) y el coeficiente de regresión parcial (la pendiente del modelo).

De este modo, un paso previo es guardar la serie de residuos de las regresiones simples entre ventas y precios, y publicidad y precios. Para esto, en *Procs*, seleccionamos la opción *Make Residual Series*.

Workfile: VENTAS - [c:\yvo\master\iccae\iccae_2\1\ventas...]

View Procs Group: VENTA_PRECIO Workfile: VENTAS

Range: 1991 VI Equation: EQ_SIMPLE_PRECIO Workfile: VENTAS

Sample: 1991

View Specify/Estimate... Name Freeze Estimate Forecast Stats Resids

LS #1 Forecast... VENTAS

Date: Make Regressor Group

Sample: Make Residual Series

Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PRECIO	-3.637266	0.619383	-5.872399	0.0004
C	483.6196	57.26030	8.445984	0.0000

R-squared 0.811699 Mean dependent var 147.9000

Adjusted R-squared 0.788161 S.D. dependent var 22.19835

S.E. of regression 10.21701 Akaike info criterion 4.824964

Sum squared resid 835.0981 Schwarz criterion 4.885481

Log likelihood -36.31421 F-statistic 34.48507

Durbin-Watson stat 1.742889 Prob(F-statistic) 0.000373

Path = C:\REVIEWS Workfile = VENTAS

Inicio Econometric Views Microsoft Word - ejemplo... 12:46

Con esto, se genera una nueva serie que podremos renombrar (por ejemplo, *resid_ven_pre*, es decir, los residuos generados al regresar ventas sobre precios).

Workfile: VENTAS - [c:\yvo\master\iccae\iccae_2\1\ventas...]

View Procs Group: VENTA_PRECIO Workfile: VENTAS

Range: 1991 VI Equation: EQ_SIMPLE_PRECIO Workfile: VENTAS

Sample: 1991

View Procs Objects Print Name Freeze Edit+/- InsDel Smp+/- Label+/- Title

Series: RESID1

Object Name 10/22/00 - 12:47

Enter name of 16 or fewer characters:

RESID_VEN_PRE

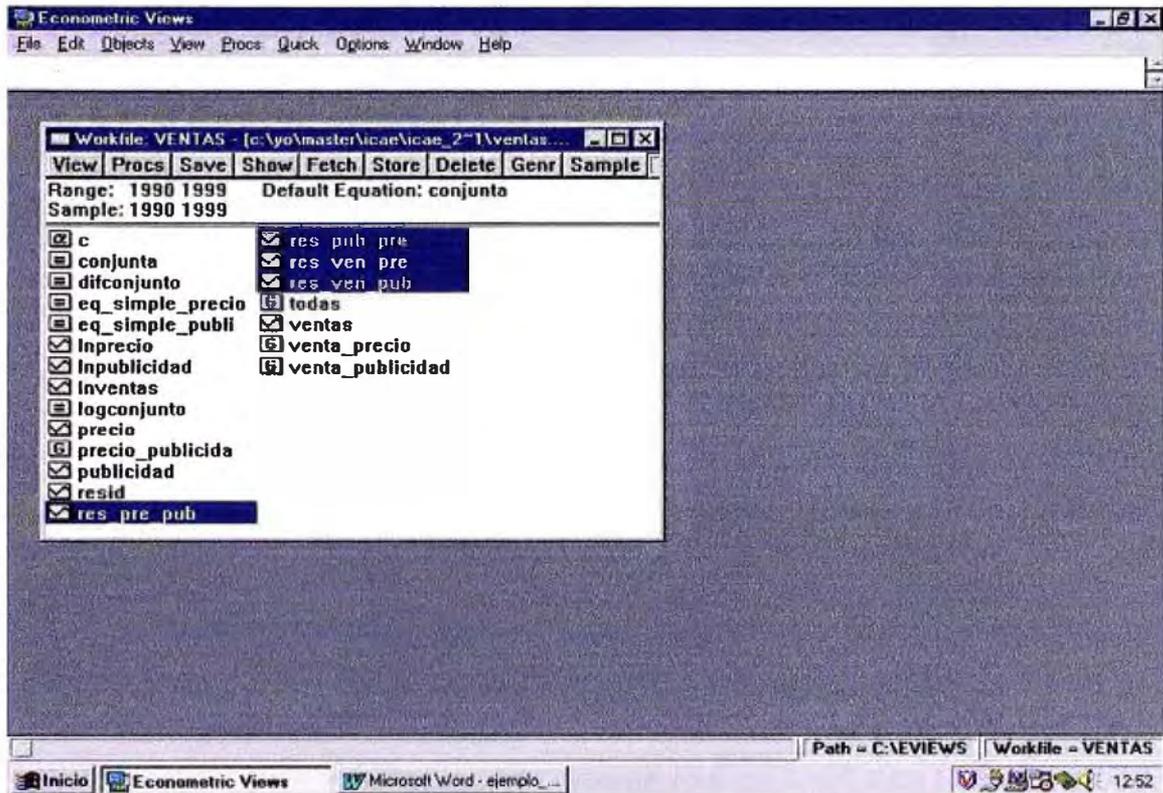
OK Cancel

	0	1	2	3	4
1990	0.		79383	-14.26571	-0.991180
1995	2.		81477	18.73429	9.185226

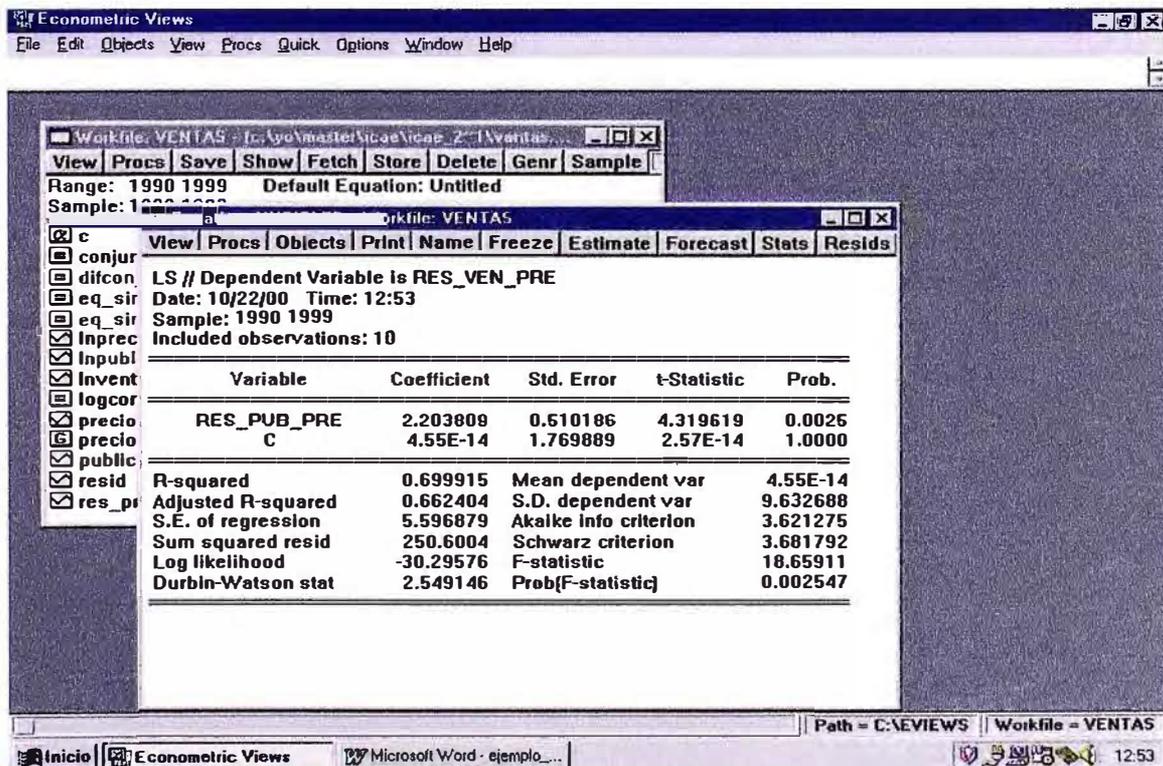
Path = C:\REVIEWS Workfile = VENTAS

Inicio Econometric Views Microsoft Word - ejemplo... 12:47

Una vez generadas las series de residuos resultantes de las regresiones simples comentadas anteriormente,



se realiza el siguiente conjunto de regresiones para calcular los coeficientes de regresión parcial:



Econometric Views

Equation: UNTITLED Workfile: VENTAS

View Procs Objects Print Name Freeze Estimate Forecast Stats Resids

LS // Dependent Variable is RES_VEN_PUB
Date: 10/22/00 Time: 12:55
Sample: 1990 1999
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RES_PRE_PUB	-1.464234	0.606790	-2.413084	0.0423
C	-3.40E-15	1.769889	-1.92E-15	1.0000

R-squared 0.421253 Mean dependent var 2.84E-15
Adjusted R-squared 0.348910 S.D. dependent var 6.936261
S.E. of regression 5.596879 Akaike info criterion 3.621275
Sum squared resid 250.6004 Schwarz criterion 3.681792
Log likelihood -30.29576 F-statistic 5.822975
Durbin-Watson stat 2.549146 Prob(F-statistic) 0.042304

Path = C:\EVIEWSW Workfile = VENTAS

XI. REGRESIONES DE VARIABLES CON TENDENCIA

En ocasiones, es interesante regresar variables de las que se ha extraído una tendencia temporal. La instrucción para que, como regresor, incluyamos una tendencia determinista en el modelo es la función @TREND(1990), donde 1990 está indicando el comienzo de la tendencia.

Econometric Views

Equation: TENDENCIA Workfile: VENTAS

Equation Specification

Equation Specification:
Dependent variable followed by list of regressors including ARMA and PDL terms. OR an explicit equation like Y=c[1]+c[2]*X.

VENTAS C PRECIO PUBLICIDAD @TREND(1990)

Estimation Settings:
Least Squares ISLS GMM ARCH Logit Probit

Sample: 1990 1999

			Prob.
			0.0174
			0.1754
			0.5900
			0.1602
			47.9000
			2.19835
			665116
Sum squared resid	175.5110	Schwarz criterion	3.786150
Log likelihood	-28.51497	F-statistic	48.53701
Durbin-Watson stat	3.142792	Prob(F-statistic)	0.000134

Path = C:\EVIEWSW Workfile = VENTAS

El resultado de la estimación del modelo con tendencia es:

Workfile: VENTAS Equation: TENDENCIA Workfile: VENTAS

View Procs S View Procs Objects Print Name Freeze Estimate Forecast Stats Resids

Range: 1990 1999
Sample: 1990 1999

LS // Dependent Variable is VENTAS
Date: 10/22/00 Time: 12:59
Sample: 1990 1999
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	211.3555	64.94777	3.254238	0.0174
PRECIO	-1.003311	0.653136	-1.536143	0.1754
PUBLICIDAD	0.626464	1.101042	0.568974	0.5900
@TREND(1990)	4.236439	2.644161	1.602186	0.1602

R-squared 0.960425 Mean dependent var 147.9000
Adjusted R-squared 0.940638 S.D. dependent var 22.19835
S.E. of regression 5.408496 Akaike info criterion 3.665116
Sum squared resid 175.5110 Schwarz criterion 3.786150
Log likelihood -28.51497 F-statistic 48.53701
Durbin-Watson stat 3.142792 Prob(F-statistic) 0.000134

Path = C:\EVIEW5 Workfile = VENTAS

Otra posibilidad hubiese sido crear primero una tendencia temporal, guardarla en el libro de trabajo y posteriormente incluirla en el modelo de regresión a estimar. Para generar esta tendencia lineal, en *Genr*, se escribe la siguiente instrucción:

Econometric Views

File Edit Objects View Procs Quick Options Window Help

Workfile: VENTAS In: Workmaster\Users\user_2\1\ventas...

View Generate Series by Equation

Range: Enter equation:
Sample: TEND=@TREND(1990)

Sample: 1990 1999

OK Cancel

Path = C:\EVIEW5 Workfile = VENTAS

BIBLIOGRAFÍA

1. Prof. Rafael de Arce, Prof. Ramón Mahía, “Modelos Arima”, Departamento Economía Aplicada, Universidad Autónoma de Madrid – España.
http://www.uam.es/personal_pdi/economicas/rarce/pdf/Box-Jenkins.PDF
2. Prof. Antonio Pulido, Prof. Ana López, Prof. Jorge Rodríguez “Curso Combinado de Predicción y Simulación”, Universidad Autónoma de Madrid – España Edición 2004.
http://www.uam.es/docencia/predysim/principal_combinado.htm
3. Ana Cecilia Kikut Valverde, Evelyn Muñoz Salas, Juan Carlos Quirós Solano “Aspectos Conceptuales Sobre Series de Tiempo - DIE/02-2002-NT”, Dpto. de Investigaciones Económicas del Banco Central de Costa Rica – Costa Rica 2002.
4. Miguel Jerez, “Econometría II” Universidad Complutense de Madrid – España 2002.
5. Federico López Fagúndez “Estadística & Econometría” Universidad del CEMA – Argentina 2005.
6. Josefa Mula Bru, Raúl Poler Escoto, David Peidro Payá, Angel Ortiz Bas, “Enfoques de Programación Matemática para la Previsión de la Demanda mediante descomposición de series temporales”, Centro de Investigación Gestión e Ingeniería de Producción Universidad Politécnica de Valencia – España 2005.
7. José. Jiménez, Juan Gázquez, Raquel Sánchez “La Capacidad Predictiva en los Metodos Box-Jenkins y Holt Winters”, Universidad de Almería – España 2005.
8. Dr. Julian Moral Caicedo “Análisis del Ciclo Económico: Descomposición De Series Temporales”, Universidad Autónoma de Madrid – España 2002.