

TEMA 7

REGRESIÓN CON SERIES TEMPORALES



A. Beyaert, M. Camacho, M. González, A. Quesada
Departamento de Métodos Cuantitativos para la Economía y la Empresa

UNIVERSIDAD DE
MURCIA

Lo que estudiaremos en este tema:

7.1. Características básicas de la estimación con series temporales

7.2. Tendencia en la media y tendencia en la varianza

7.3. Regresión espuria y solución

7.4. Estacionalidad



Bibliografía básica: Wooldridge, 2006, cap. 10 y parte del 18

7.1. Características básicas de la estimación con series temporales

Recordemos:

Datos de corte transversal: observaciones sobre distintas unidades (individuos , familias, ciudades,...) en un momento dado del tiempo

- Características:
 - **el orden** de los datos **no importa**
 - **muestreo aleatorio** → **no hay interrelación** entre los elementos de la muestra
- El **supuesto de observaciones independientes** es **esencial** para las buenas propiedades de los estimadores MCO.

7.1. Características básicas de la estimación con series temporales

En cambio:

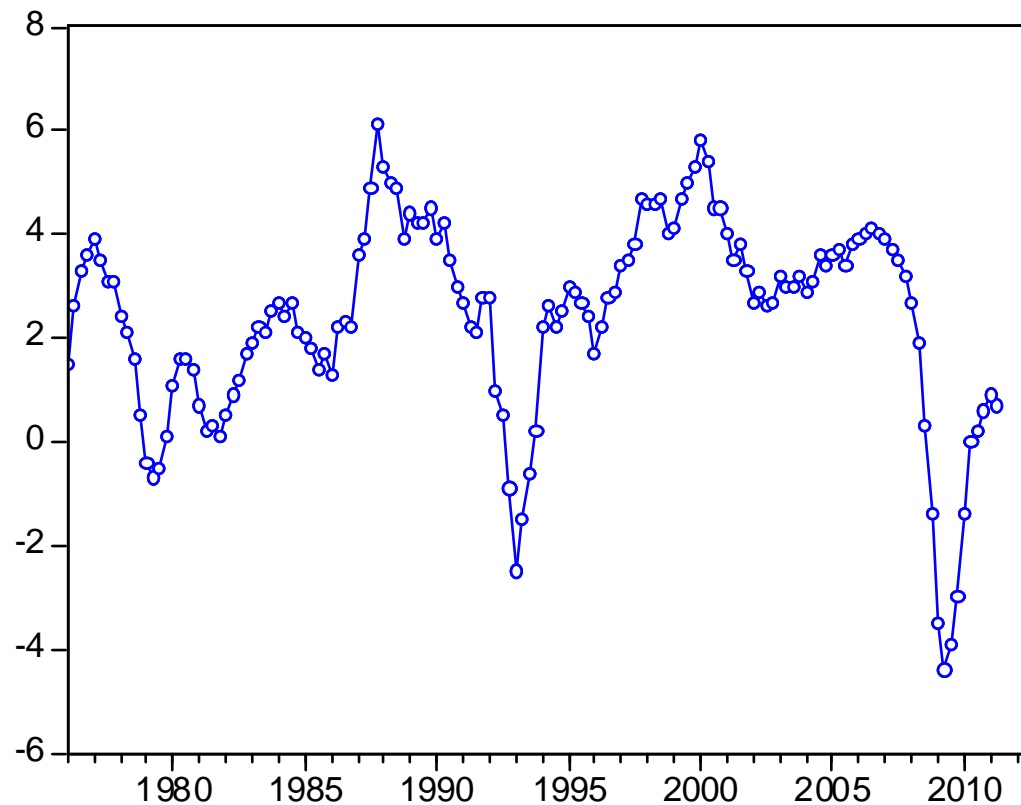
Datos de serie temporal: Observaciones sobre una o distintas variables a lo largo del tiempo

- Características muy diferentes:
 - **1. El orden de los datos sí importa: los datos son interdependientes**
 - **2. Puede haber tendencias en la media, en la varianza o en ambas**
 - **3. Puede haber estacionalidad**

7.1. Características básicas de la estimación con series temporales

- Ejemplos: Evolución a lo largo del tiempo de:
 - Tasa de crecimiento de la economía española (según contabilidad nacional trimestral)
 - Importaciones de bienes, a precios corrientes
 - Número de parados, registrados en el INEM
 - Índice de la bolsa de Paris (CAC-40)
 - ...
- Veamos, con estos ejemplos, estas características diferentes

7.1. Características básicas de la estimación con series temporales



—○— TASA DE CRECIMIENTO INTERANUAL, ESPAÑA 1976.T1-2011.T2

- Media constante
- Varianza estable
- Cierta **inercia**: influencia del pasado sobre el presente

7.1. Características básicas de la estimación con series temporales

- Una herramienta útil en series temporales: el **correlograma**

- **¿qué es?**

- Componente básico: “**correlacion de orden j**”

correlación muestral media entre datos distantes entre sí “j” períodos

- correlograma:**

representación gráfica de las correlaciones de orden $j=1,2,\dots,k$
en función del “retardo” j

- **¿qué información da?**

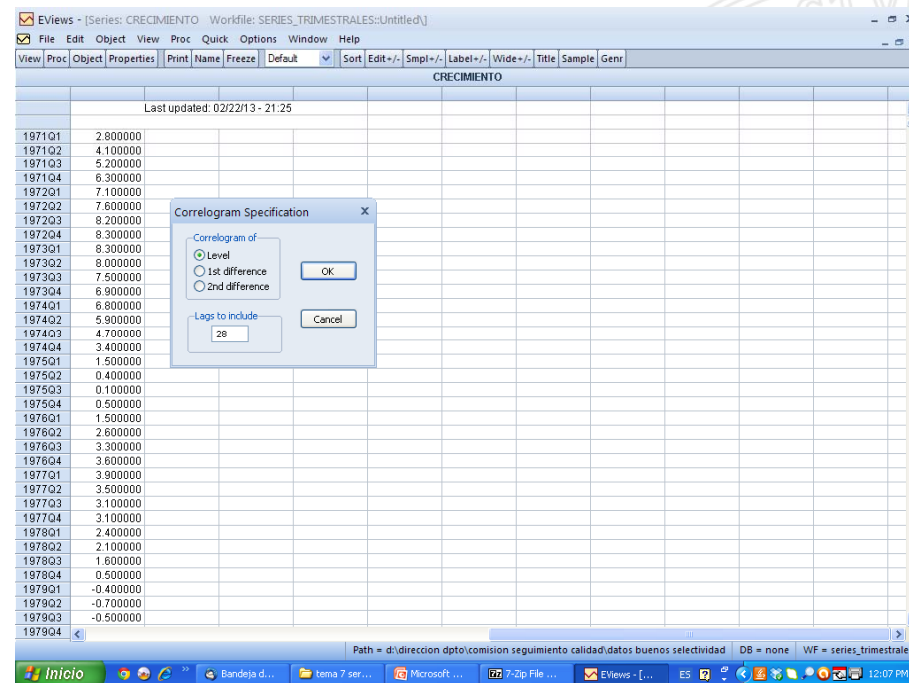
Resume el comportamiento dinámico de la serie,
informa sobre cómo depende de su propio pasado.

7.1. Características básicas de la estimación con series temporales

Para obtener el correlograma en EViews:

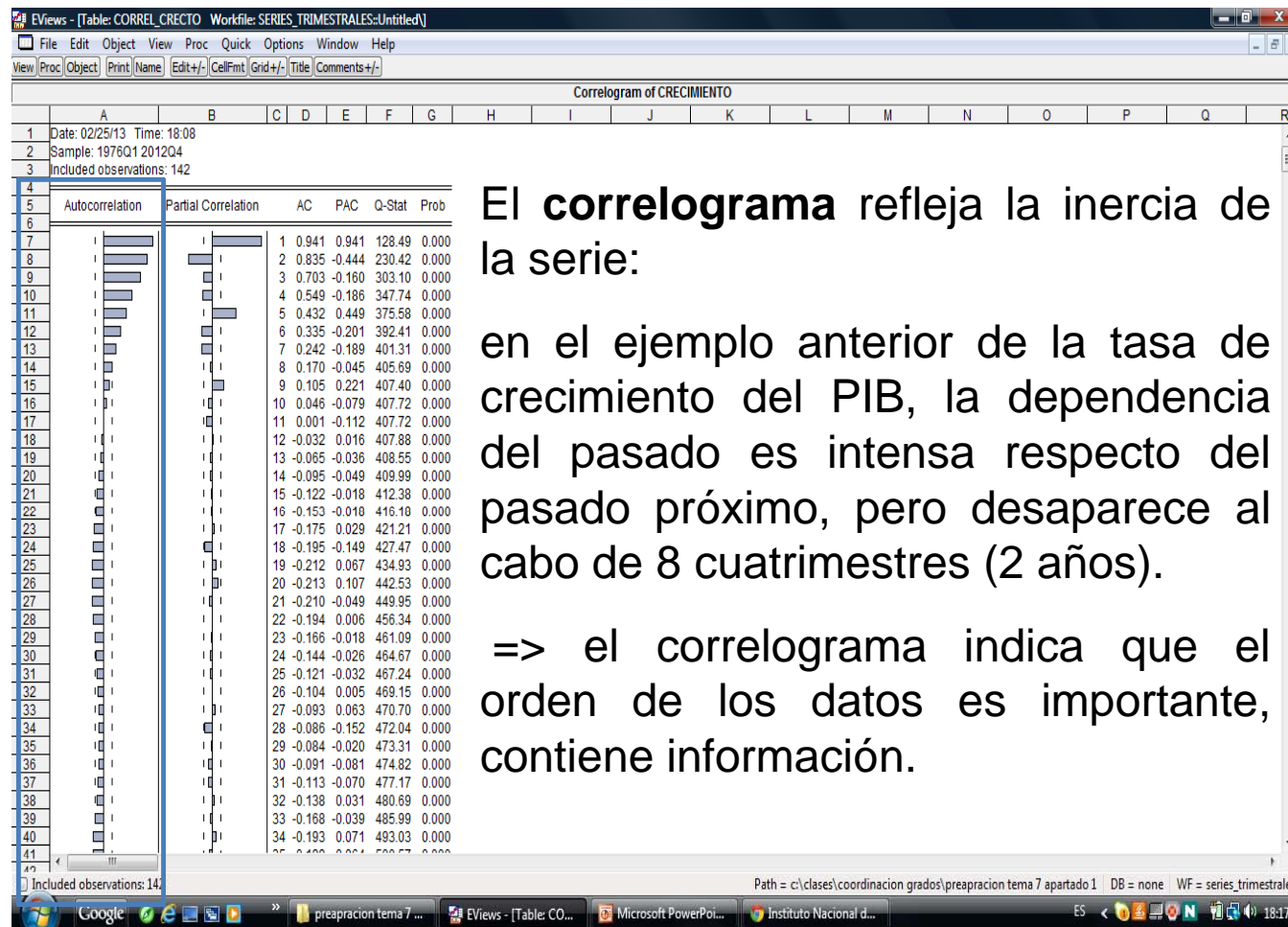
1) En la ventana de la serie , hacemos view/correlogram...

2) Aparece esta ventana:



3) Se selecciona “level” y damos a OK

7.1. Características básicas de la estimación con series temporales

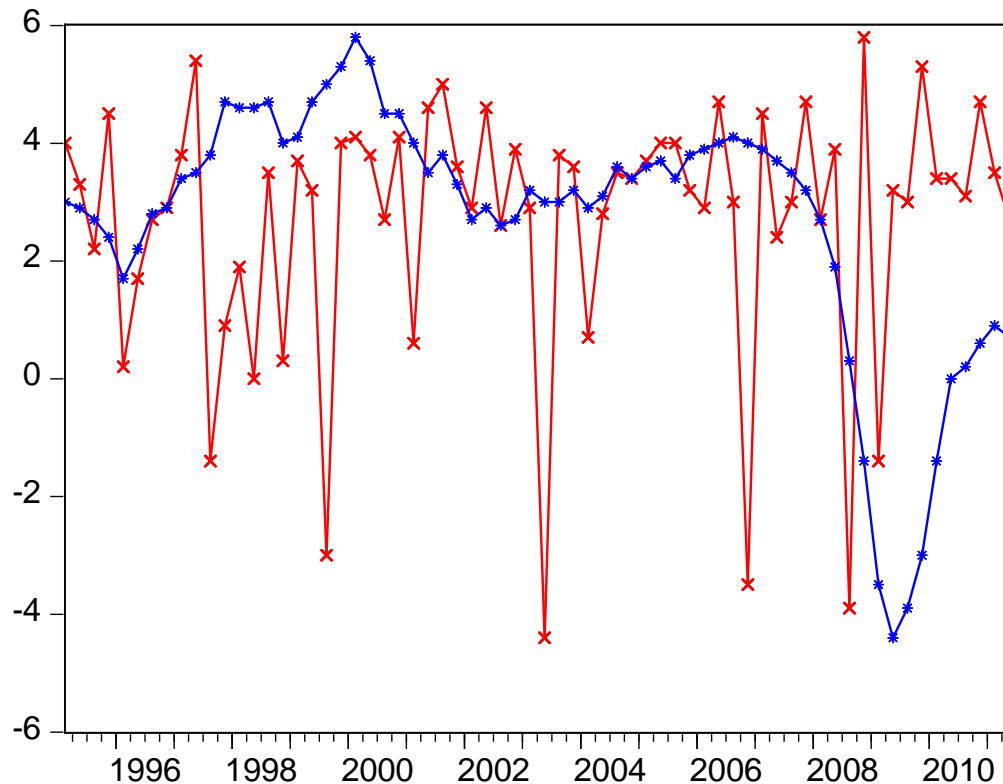


El correlograma refleja la inercia de la serie:

en el ejemplo anterior de la tasa de crecimiento del PIB, la dependencia del pasado es intensa respecto del pasado próximo, pero desaparece al cabo de 8 cuatrimestres (2 años).

=> el correlograma indica que el orden de los datos es importante, contiene información.

7.1. Características básicas de la estimación con series temporales



—x— tasa de crecimiento interanual, orden aleatorio, 1995q1-2011q2
—*— tasa de crecimiento interanual, histórica, 1995q1-2011q2

En azul: la tasa de crecimiento, en orden cronológico

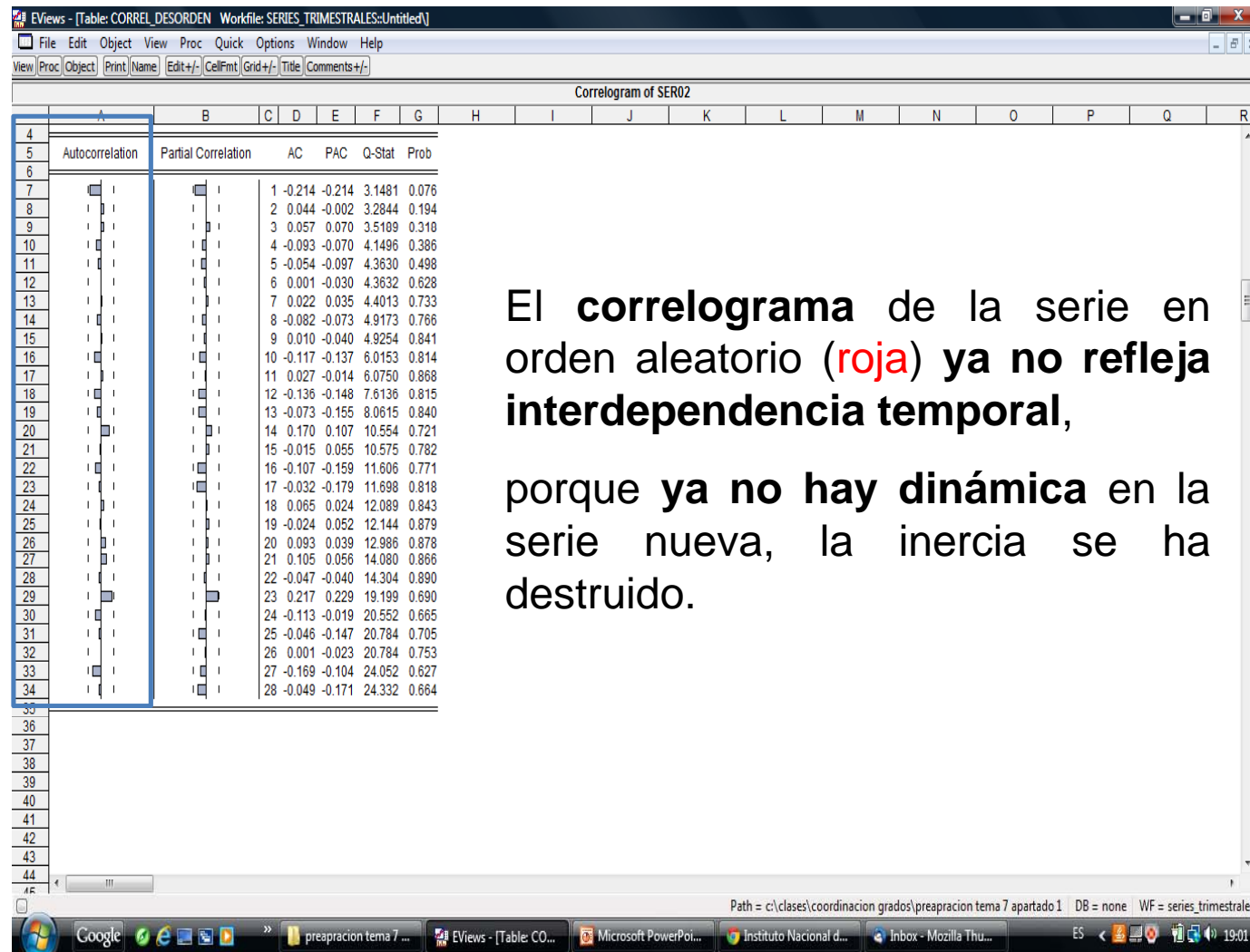
En rojo: los mismos datos, en orden aleatorio

⇒ hemos roto la dinámica

⇒ el orden importa: los datos cronológicos son interdependientes, los otros no.

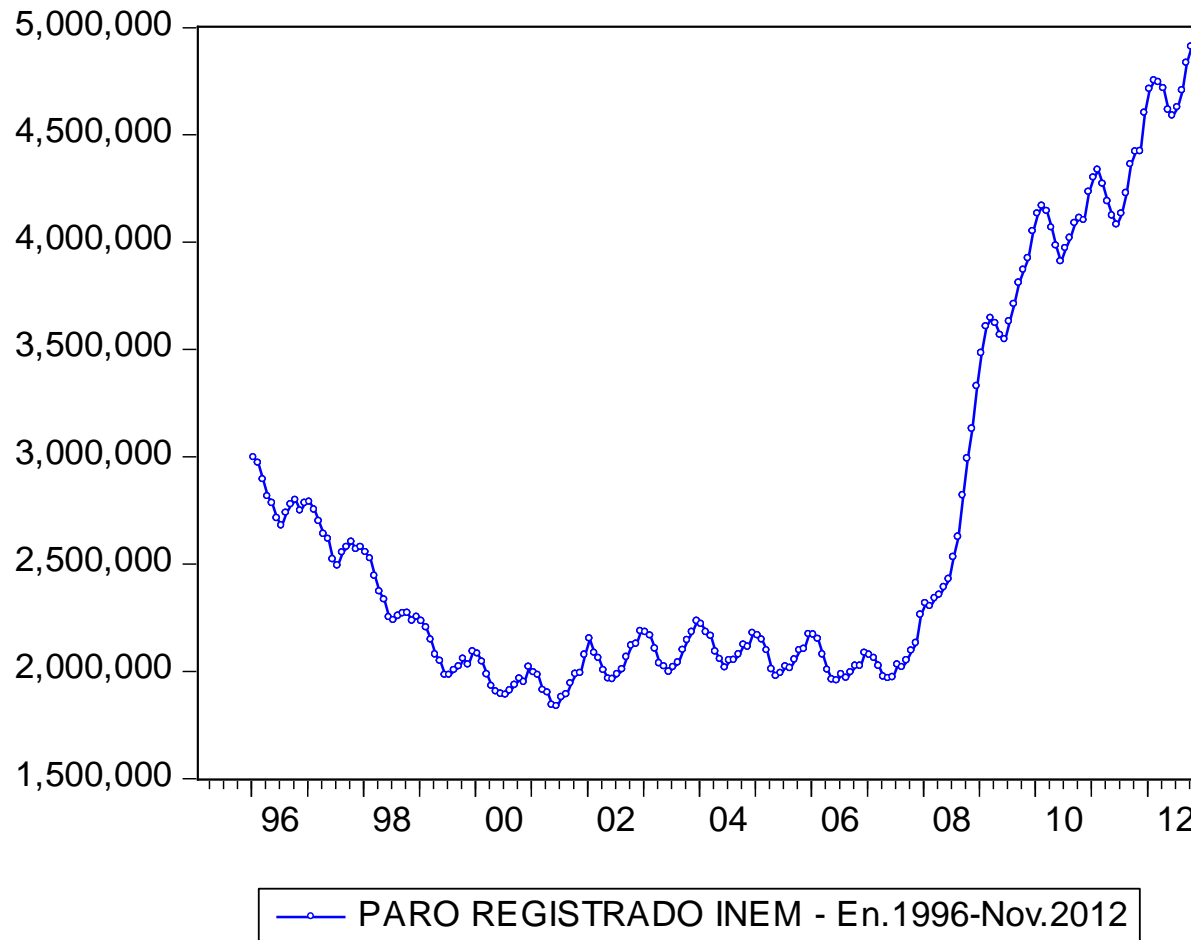
⇒ se notará en el correlograma

7.1. Características básicas de la estimación con series temporales



El correlograma de la serie en orden aleatorio (rojo) ya no refleja interdependencia temporal, porque ya no hay dinámica en la serie nueva, la inercia se ha destruido.

7.1. Características básicas de la estimación con series temporales



Media y varianza:

- ¿Media cambiante y varianza constante?

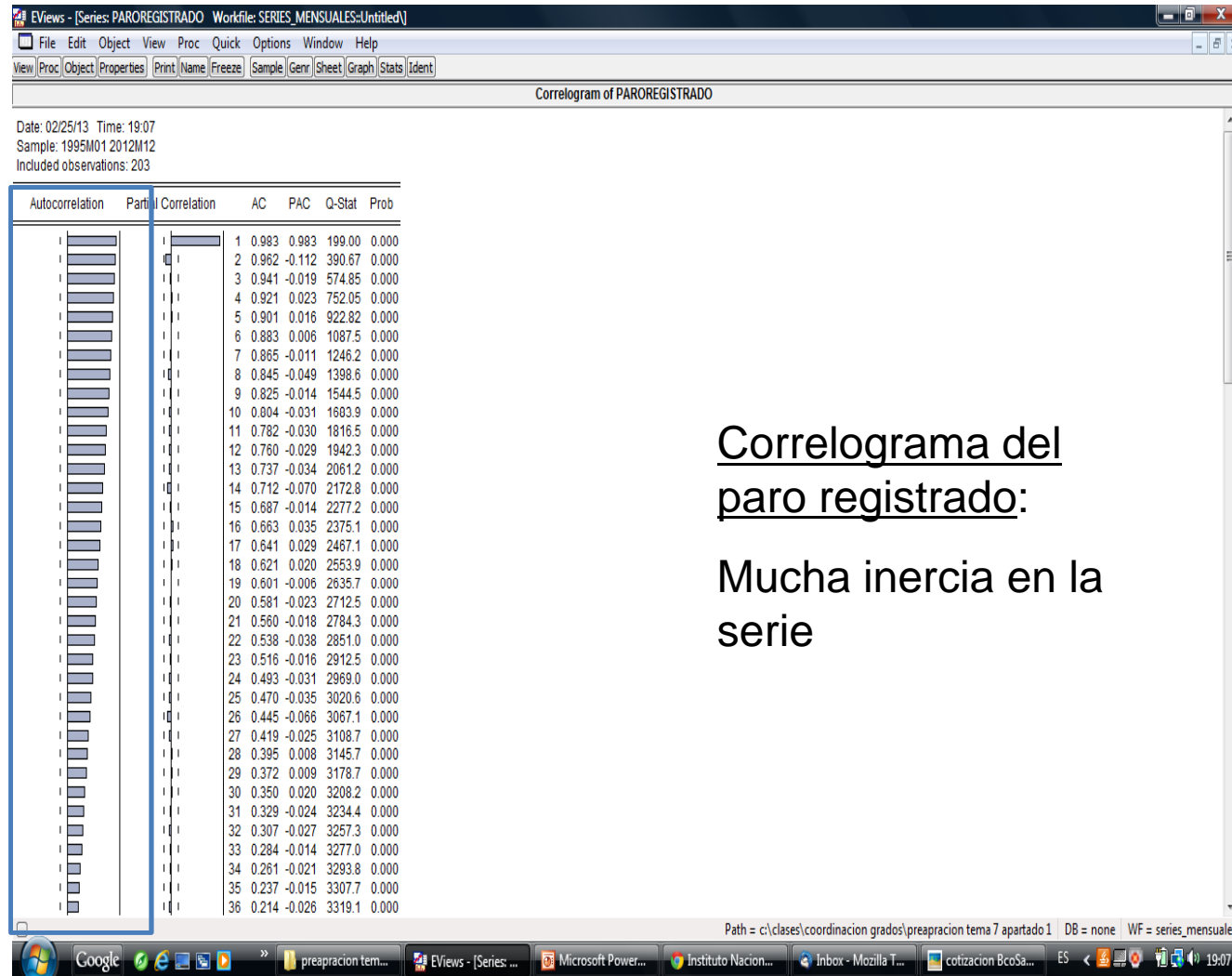
ó bien

- ¿Media constante y varianza creciente?

En cualquier caso:

- Algo de inercia: influencia del pasado sobre el presente

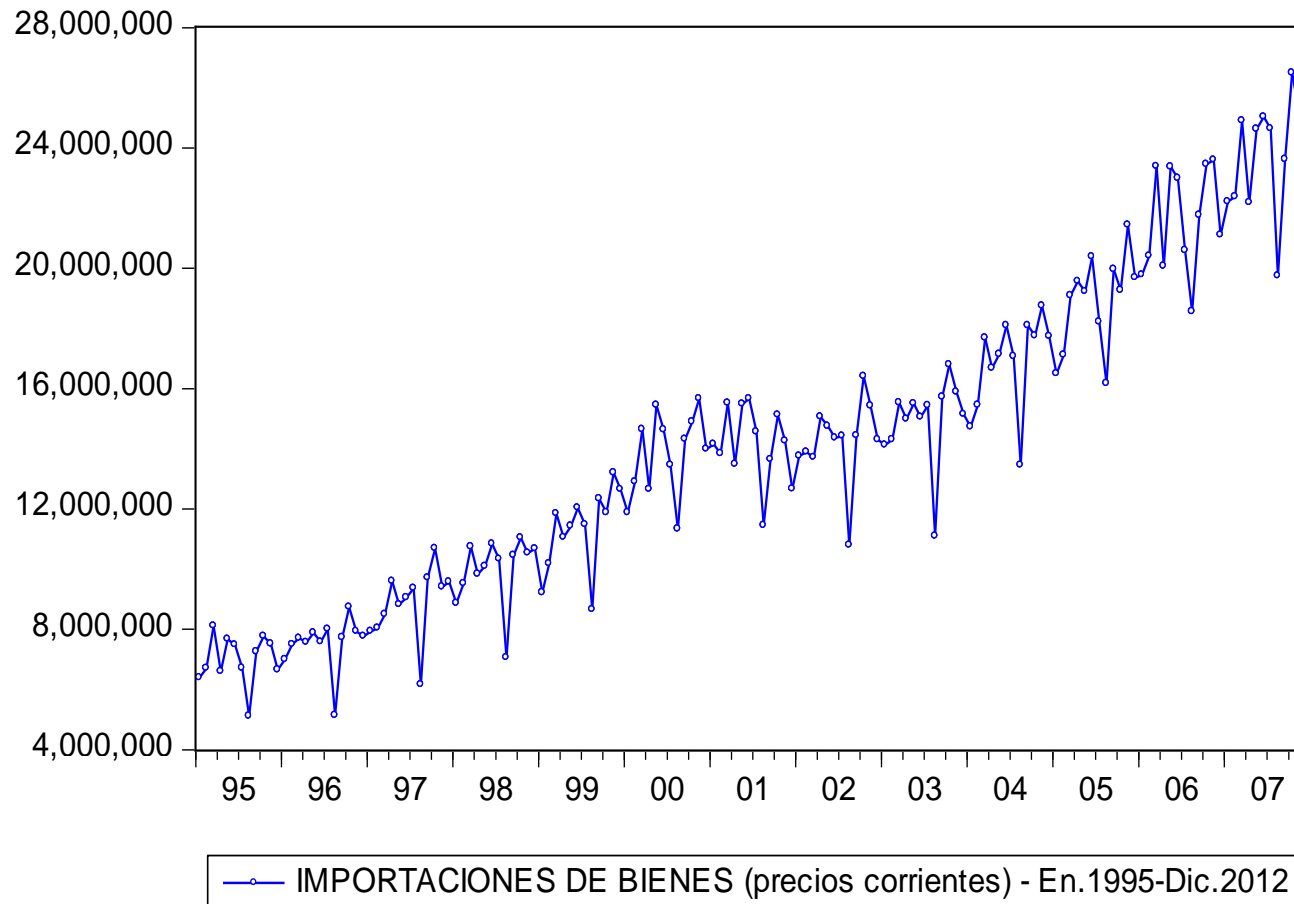
7.1. Características básicas de la estimación con series temporales



Correlograma del
paro registrado:

Mucha inercia en la
serie

7.1. Características básicas de la estimación con series temporales



- Media creciente: **tendencia en media**
- Varianza estable
- Algo de inercia: influencia del pasado sobre el presente
- **Estacionalidad:** el mismo patrón dinámico se repite año tras año

7.1. Características básicas de la estimación con series temporales



Índice bursátil: CAC40 diario, al cierre

La serie “deambula”:

- Media no tendencial
- Varianza creciente
- Inercia: el pasado influye sobre el presente

7.1. Características básicas de la estimación con series temporales

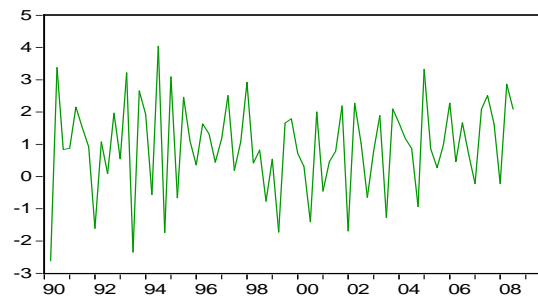
Característica 1:

- los datos de series temporales son interdependientes: *correlación intertemporal* o “*serial*”
- el orden de los datos importa
- el correlograma da información sobre la dinámica de la serie

7.1. Características básicas de la estimación con series temporales

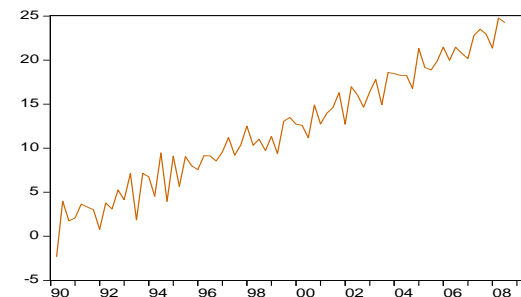
➔ **Característica 2:** hay series sin tendencia, otras con tendencia en la media y/o con tendencia en la varianza

Media constante
Y1

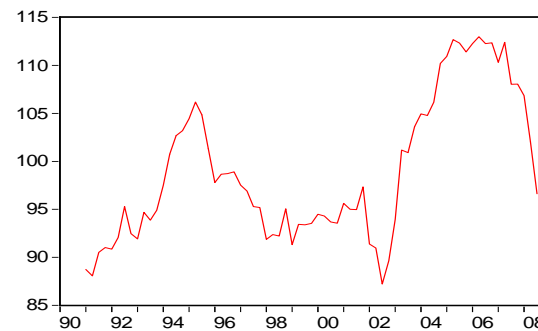


Varianza acotada

Media creciente
Y2

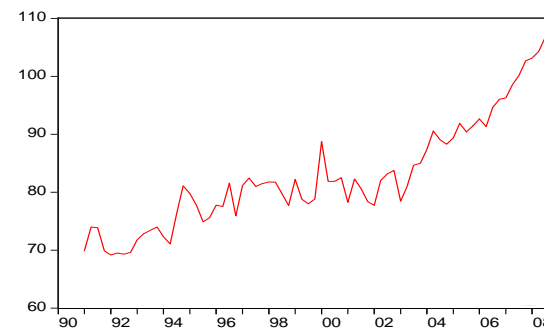


Y3



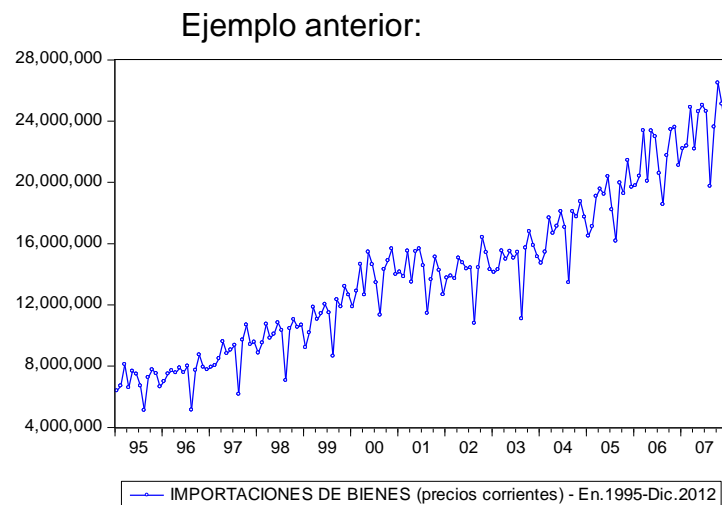
Varianza creciente

Y4

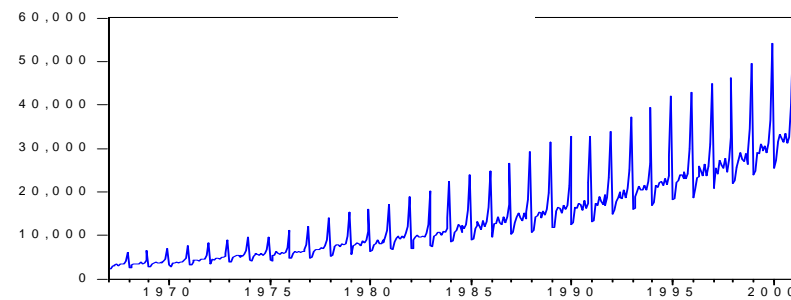


7.1. Características básicas de la estimación con series temporales

➔ Característica 3: a veces, hay estacionalidad

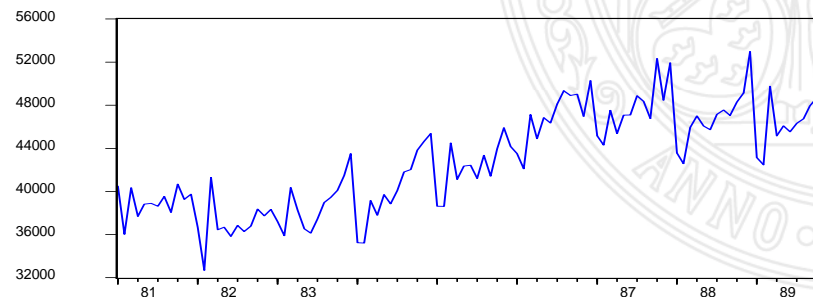


Estacionalidad muy intensa, dominante:



Ventas al por menor, USA, 1/1967-12/2000

Estacionalidad menos evidente:

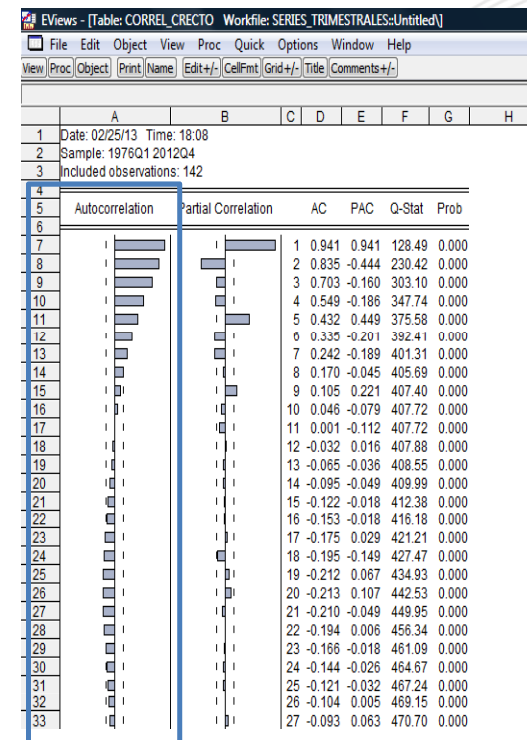
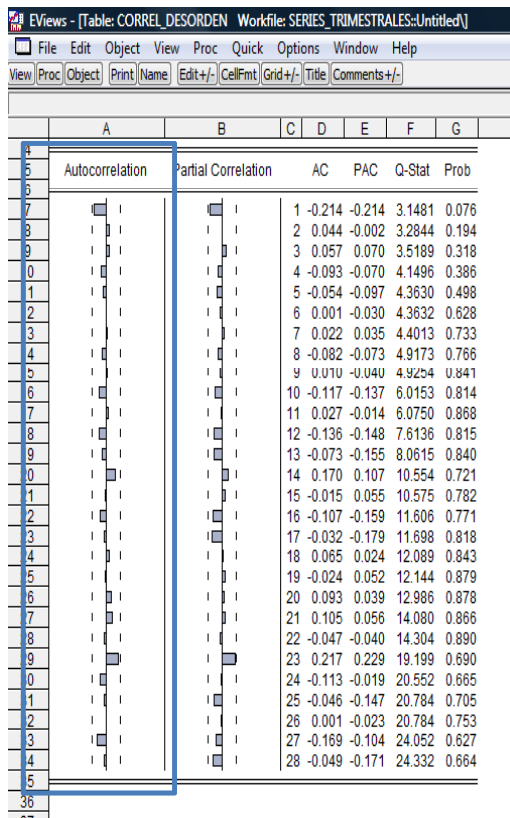


Numero accidentes de tráfico, California, 1/1981- 12/1989

7.1. Características básicas de la estimación con series temporales

El correlograma da información valiosa...

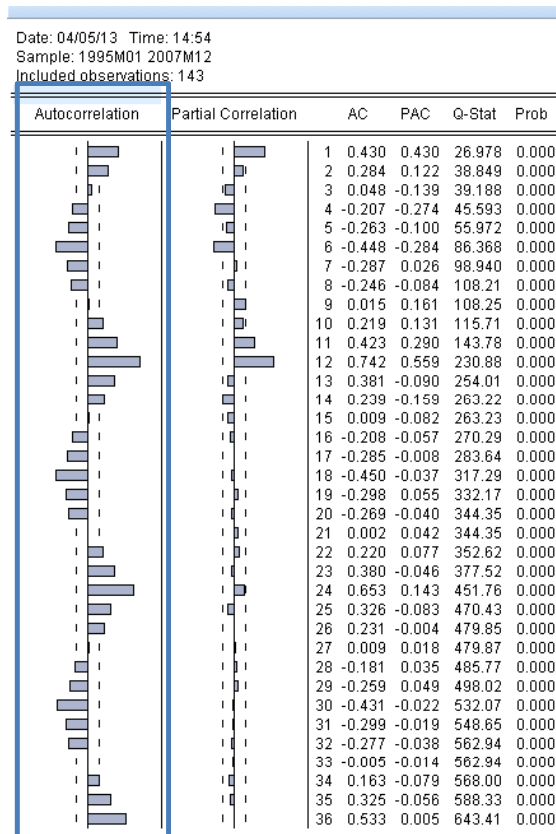
- 1) Serie **sin** dinámica, media y varianza estables:
- 2) Serie **con** dinámica, media y varianza estables:



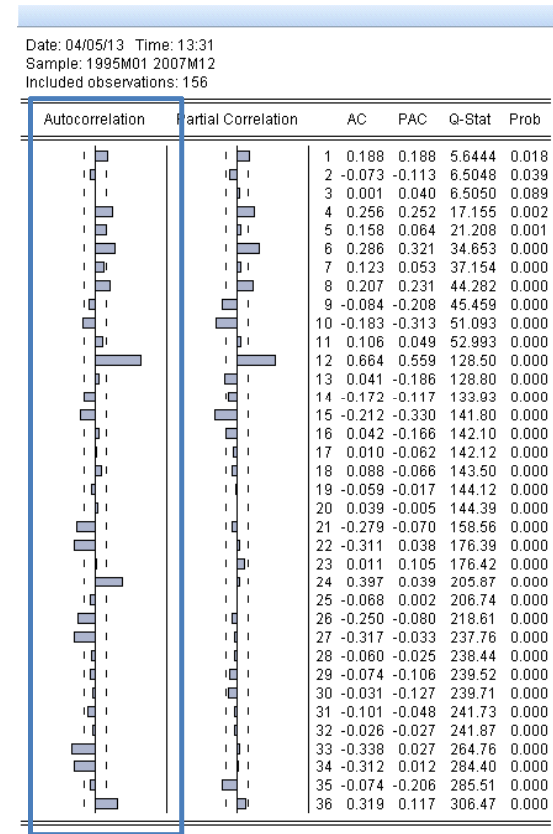
7.1. Características básicas de la estimación con series temporales

El correlograma da información valiosa...

3) Estacionalidad fuerte:



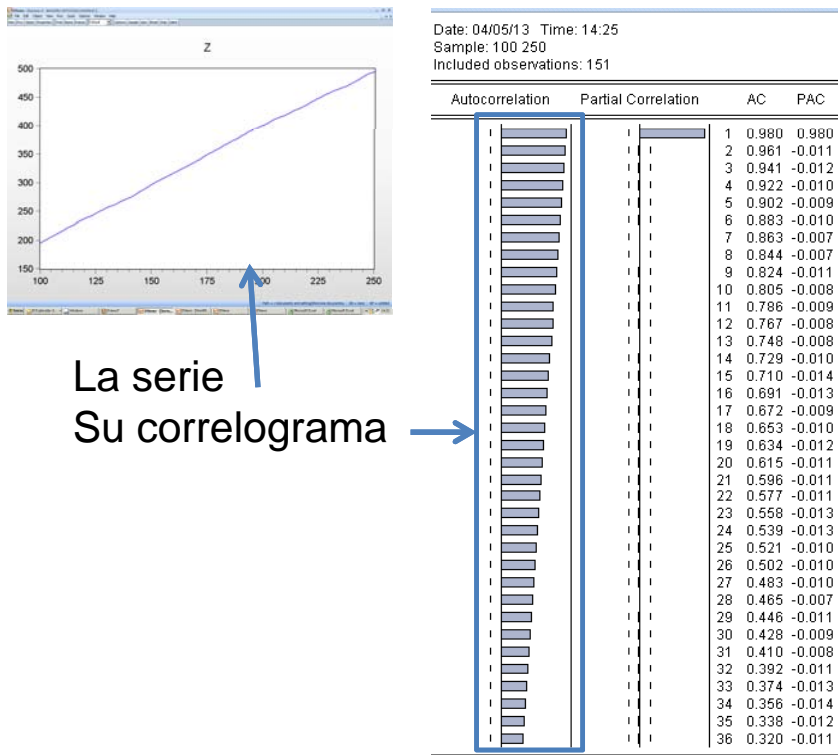
4) Algo de estacionalidad:



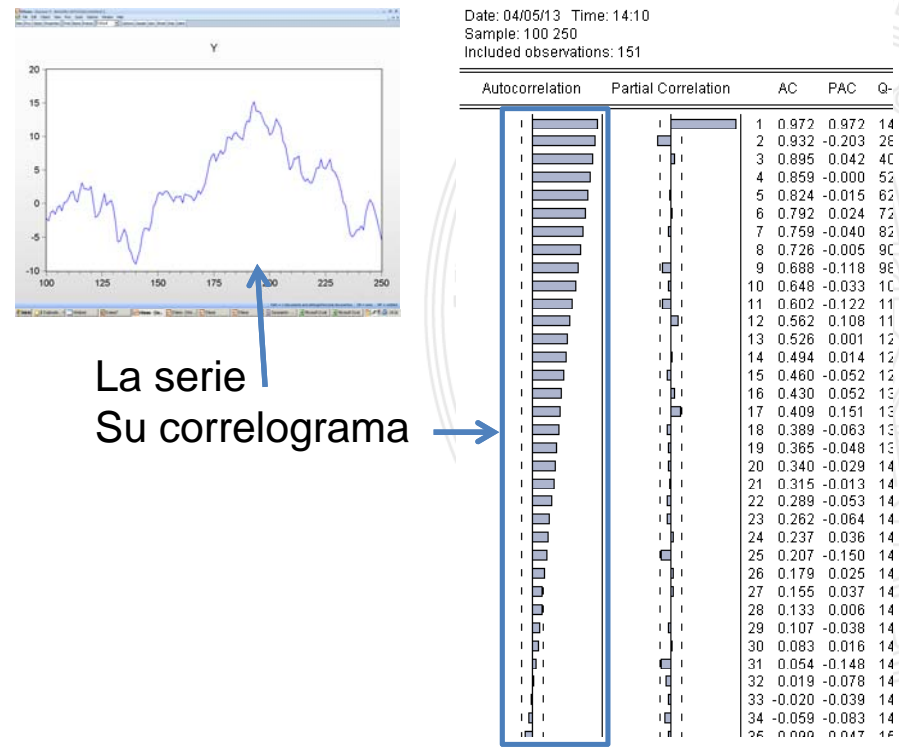
7.1. Características básicas de la estimación con series temporales

... el correlograma no discrimina entre tendencia en la media y tendencia en la varianza

5) Tendencia en la media:



6) Tendencia en la varianza:



7.2. Tendencia en la media y tendencia en la varianza

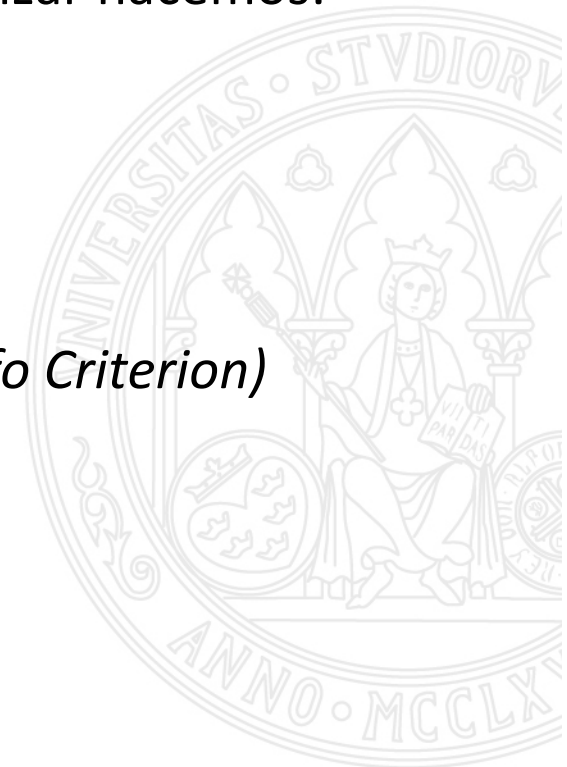
- Para distinguir entre series con tendencia en la media y/o con tendencia en la varianza se utilizan los “**contrastes de raíz unitaria**” (H_0 : hay raíz unitaria, H_1 : no hay raíz unitaria).
- Vamos a ver el **Contraste de Dickey-Fuller Aumentado (ADF)**
- Definiciones:
 - Una serie con *tendencia en la varianza* tiene raíz unitaria.
 - Una serie con *tendencia en la media* crece (decrece) en el tiempo, y además puede tener tendencia en la varianza.
 - Un serie que no tiene ni tendencia en la media ni tendencia en la varianza es ***estacionaria***.

7.2. Tendencia en la media y tendencia en la varianza

- En los **contrastes de raíz unitaria** vamos a distinguir **3 situaciones**:
 - Situación 1: sabemos que la serie crece (decrece) en media (ejemplo: PIB, consumo, inversión, etc.)
 - Situación 2: sabemos que la serie no crece (no decrece) en media (ejemplo: tipo de interés, tasa de inflación, tasa de desempleo, etc.)
 - Situación 3: no sabemos si crece o no (o si decrece o no)
- El procedimiento de contraste difiere de una situación a otra.

7.2. Tendencia en la media y tendencia en la varianza

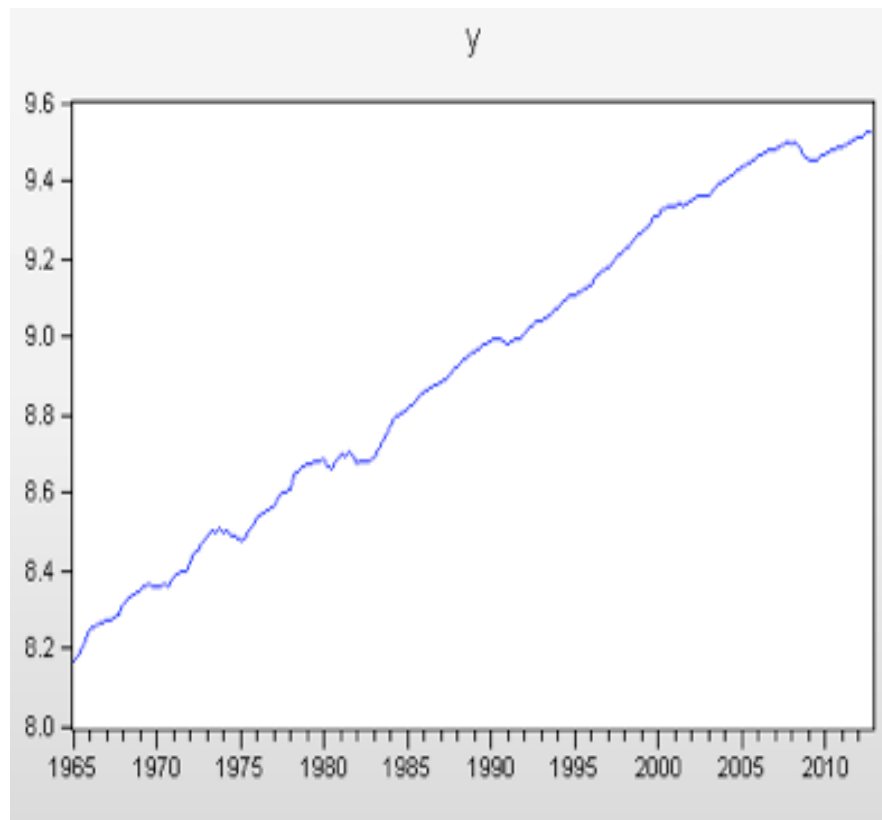
- En **Eviews**, el procedimiento de contraste se lleva a cabo de la siguiente manera:
 - 1) En la ventana de la serie que queremos analizar hacemos:
View/Unit Root Test...
 - 2) En la ventana que aparece se elige siempre:
Test Type: *Augmented Dickey-Fuller*
Test for unit root in: *Level*
Lag length: *Automatic Selection (Schwarz Info Criterion)*
 - 3) Dependiendo de la situación elegiremos:
 - Situación 1: Trend and intercept
 - Situación 2: Intercept
 - Situación 3: Trend and intercept



7.2. Tendencia en la media y tendencia en la varianza

➤ SITUACIÓN 1: sabemos que la serie (de)crece

- **Ejemplo 1**: Logaritmo del PIB trimestral de EE.UU. de 1965.1 a 2012.4



Correlogram of LPIB

Date: 04/18/13 Time: 13:04
Sample: 1965Q1 2012Q4
Included observations: 192

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.985	0.985	189.22	0.000	
2	0.970	-0.011	373.65	0.000	
3	0.955	-0.000	553.41	0.000	
4	0.940	0.002	728.66	0.000	
5	0.926	0.004	899.54	0.000	
6	0.912	-0.007	1066.1	0.000	
7	0.898	-0.009	1228.4	0.000	
8	0.883	-0.005	1386.4	0.000	
9	0.869	-0.008	1540.2	0.000	
10	0.855	-0.014	1689.7	0.000	
11	0.840	-0.007	1835.1	0.000	
12	0.826	-0.008	1976.3	0.000	
13	0.812	0.005	2113.5	0.000	
14	0.798	0.004	2246.8	0.000	
15	0.785	-0.008	2376.4	0.000	
16	0.771	-0.012	2502.2	0.000	
17	0.757	-0.012	2624.1	0.000	
18	0.742	-0.029	2742.1	0.000	
19	0.727	-0.019	2856.0	0.000	
20	0.712	-0.020	2965.8	0.000	

7.2. Tendencia en la media y tendencia en la varianza

- En el cuadro para hacer el contraste se selecciona:

Unit Root Test

Test type
Augmented Dickey-Fuller

Test for unit root in
 Level
 1st difference
 2nd difference

Include in test equation
 Intercept
 Trend and intercept
 None

Lag length
 Automatic selection:
Schwarz Info Criterion
Maximum lags: 14
 User specified: 4

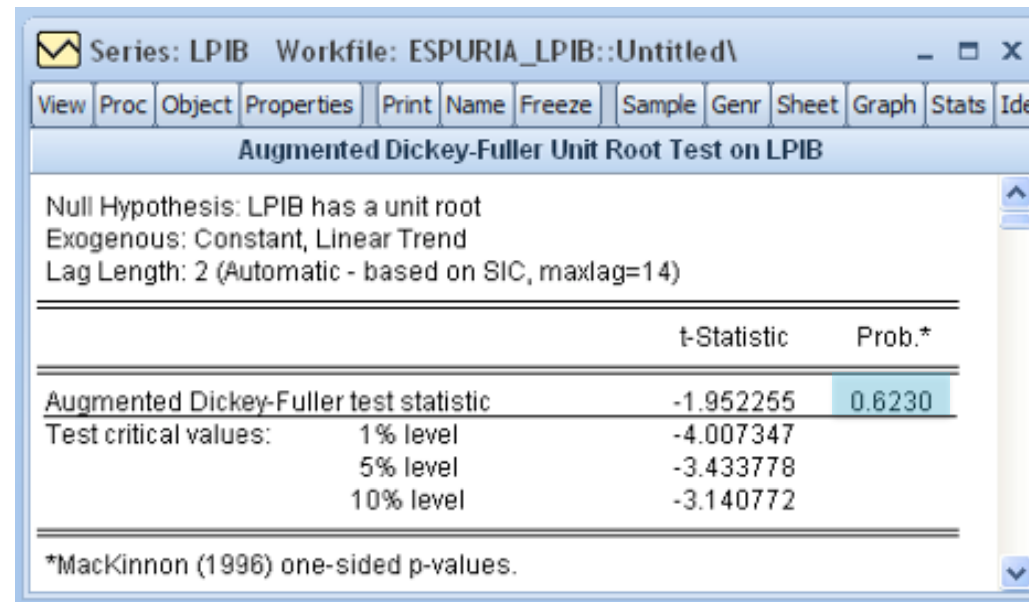
OK Cancel

H_0 : hay raíz unitaria

H_1 : no hay raíz unitaria

7.2. Tendencia en la media y tendencia en la varianza

- El resultado del contraste es:



Series: LPIB Workfile: ESPURIA_LPIB::Untitled\

View Proc Object Properties Print Name Freeze Sample Genr Sheet Graph Stats Ide

Augmented Dickey-Fuller Unit Root Test on LPIB

Null Hypothesis: LPIB has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic - based on SIC, maxlag=14)

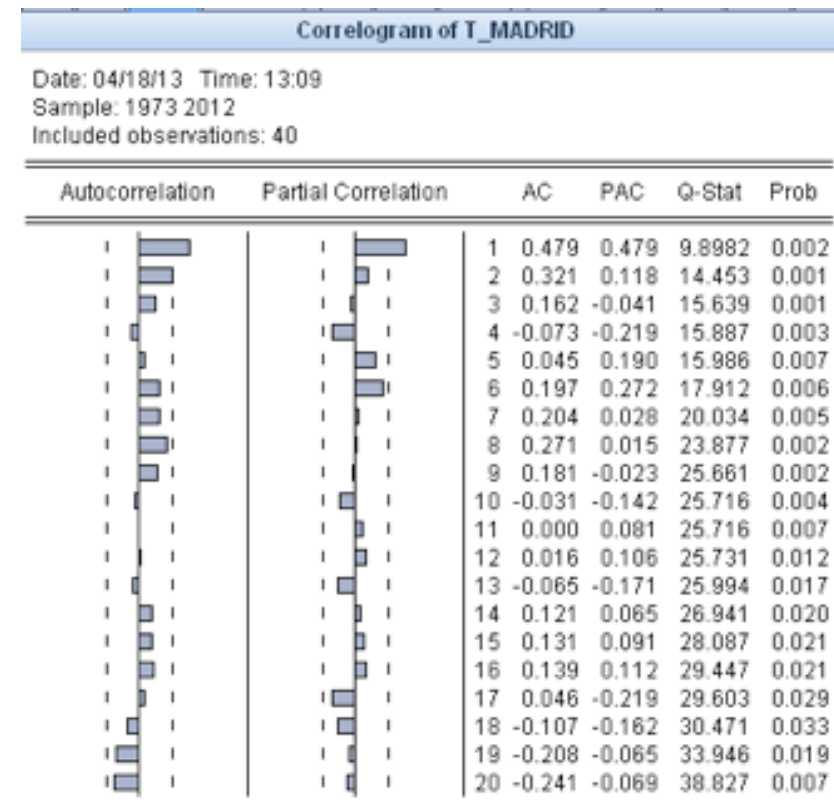
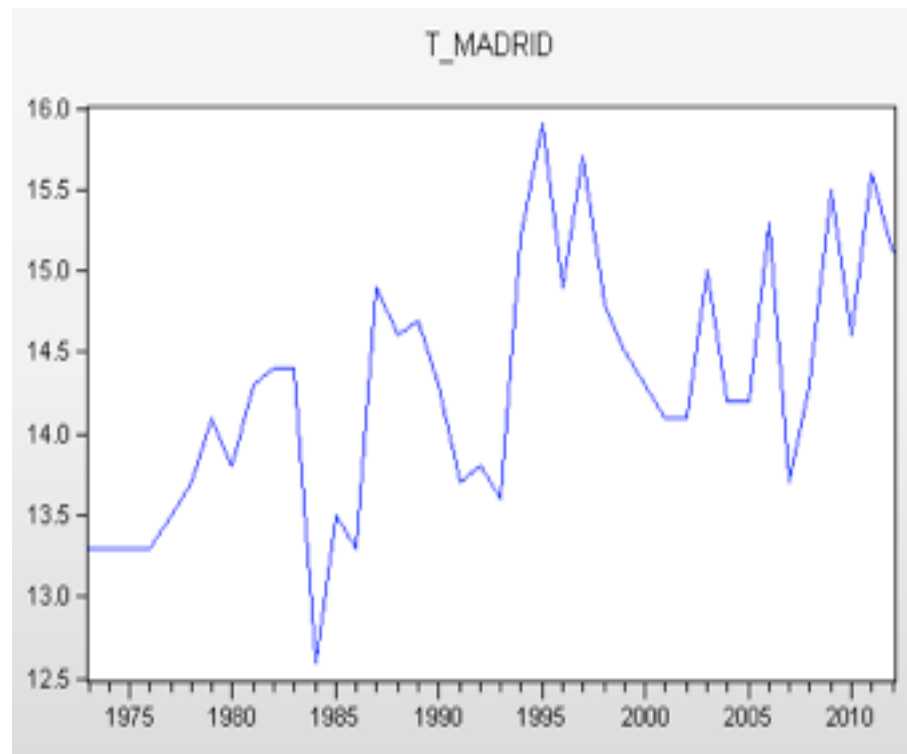
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.952255	0.6230
Test critical values:		
1% level	-4.007347	
5% level	-3.433778	
10% level	-3.140772	

*MacKinnon (1996) one-sided p-values.

- Como el estadístico de contraste de Dickey-Fuller Aumentado tiene un p-valor de $0.6230 > 0.05$, no RH_0 ➡ **Hay raíz unitaria**
- La serie tiene **tendencia en varianza** y como crece claramente también tendrá **tendencia en media**.

7.2. Tendencia en la media y tendencia en la varianza

- **Ejemplo 2:** Temperatura media anual en Madrid de 1973 a 2012



7.2. Tendencia en la media y tendencia en la varianza

- En el cuadro para hacer el contraste se selecciona:

Unit Root Test

Test type
Augmented Dickey-Fuller

Test for unit root in
 Level
 1st difference
 2nd difference

Include in test equation
 Intercept
 Trend and intercept
 None

Lag length
 Automatic selection:
Schwarz Info Criterion
Maximum lags: 14
 User specified: 4

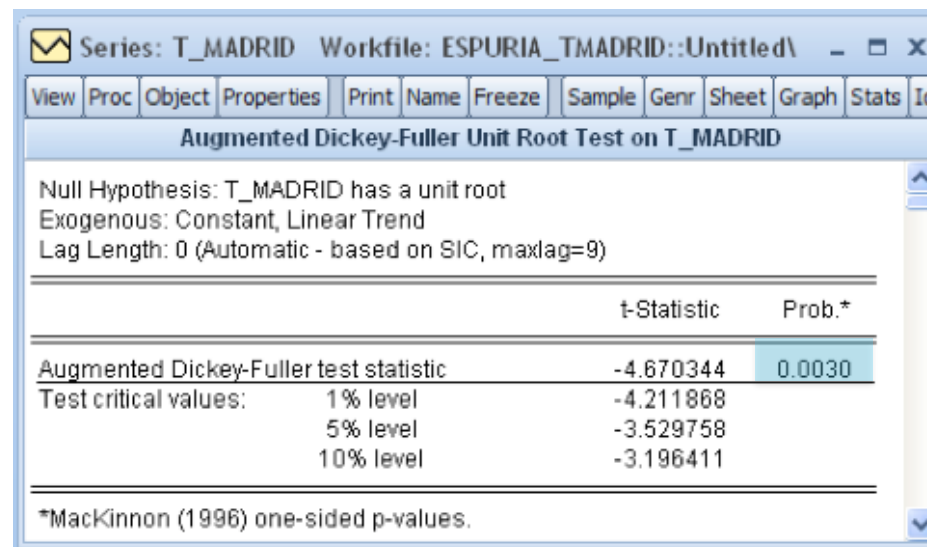
OK Cancel

H_0 : hay raíz unitaria

H_1 : no hay raíz unitaria

7.2. Tendencia en la media y tendencia en la varianza

- El resultado del contraste es:



Augmented Dickey-Fuller Unit Root Test on T_MADRID

Null Hypothesis: T_MADRID has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=9)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.670344	0.0030
Test critical values:		
1% level	-4.211868	
5% level	-3.529758	
10% level	-3.196411	

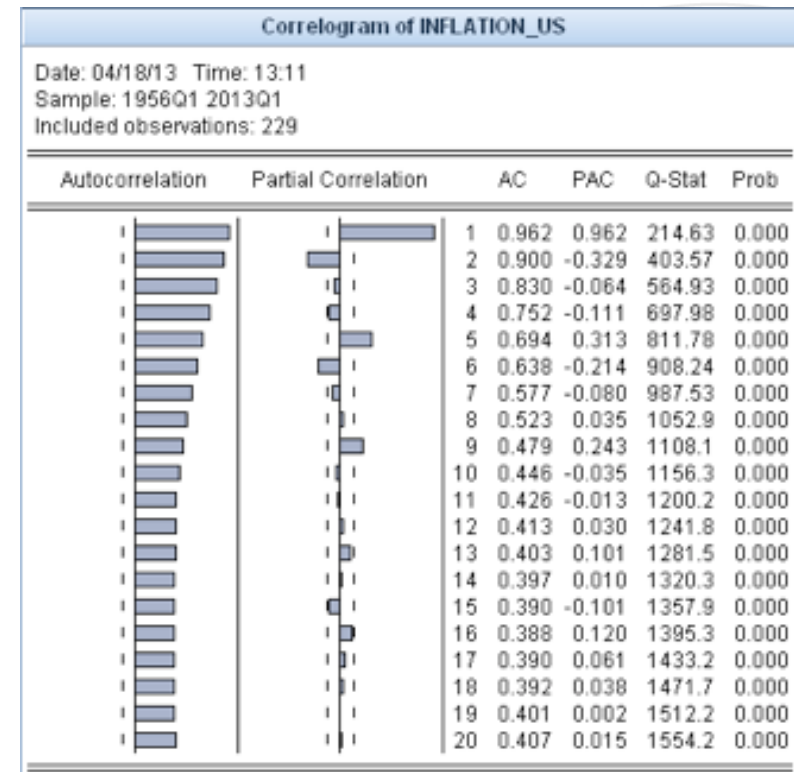
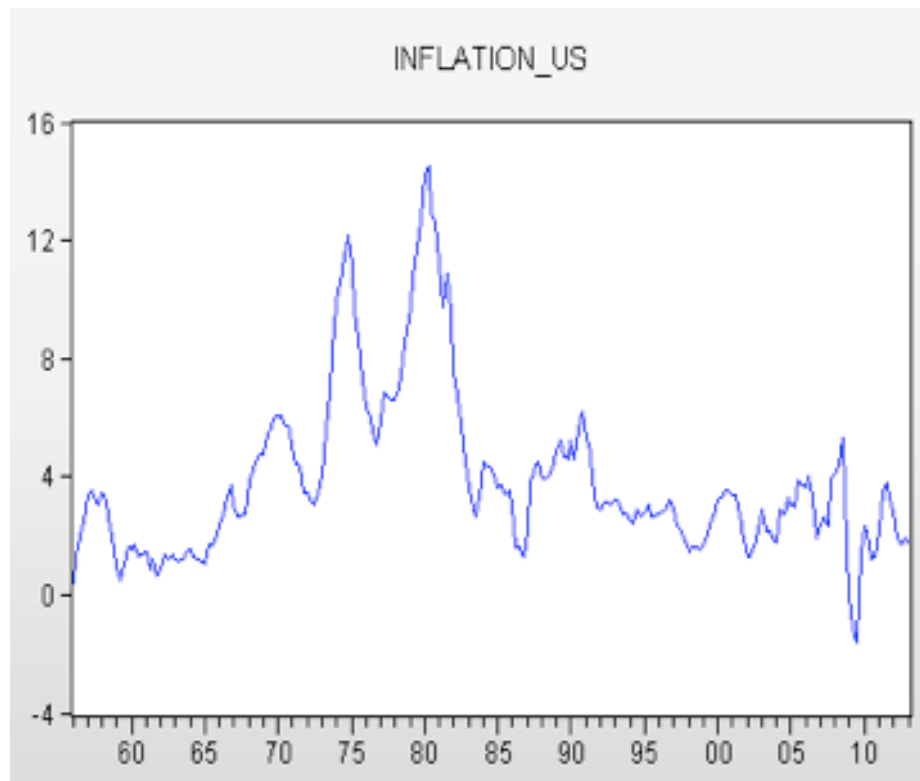
*MacKinnon (1996) one-sided p-values.

- Como el estadístico de contraste de Dickey-Fuller Aumentado tiene un p-valor de $0.0030 < 0.05$, se $RH_0 \rightarrow$ **No hay raíz unitaria**
- La serie **no** tiene tendencia en varianza, pero como crece claramente si que tendrá **tendencia en media**.

7.2. Tendencia en la media y tendencia en la varianza

➤ SITUACIÓN 2: sabemos que la serie no (de)crece

Ejemplo: Tasa de inflación de EE.UU. de 1956.1 a 2013.1



7.2. Tendencia en la media y tendencia en la varianza

- En el cuadro para hacer el contraste se selecciona:

Unit Root Test

Test type
Augmented Dickey-Fuller

Test for unit root in
 Level
 1st difference
 2nd difference

Include in test equation
 Intercept
 Trend and intercept
 None

Lag length
 Automatic selection:
Schwarz Info Criterion
Maximum lags: 14
 User specified: 4

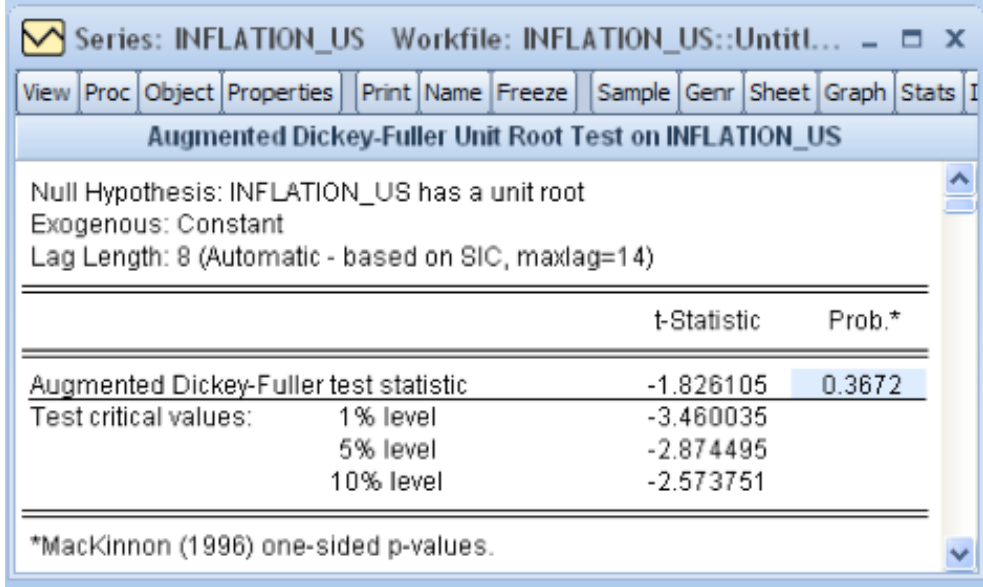
OK Cancel

H_0 : hay raíz unitaria

H_1 : no hay raíz unitaria

7.2. Tendencia en la media y tendencia en la varianza

- El resultado del contraste es:



Series: INFLATION_US Workfile: INFLATION_US::Untitl... - □ X

View Proc Object Properties Print Name Freeze Sample Genr Sheet Graph Stats I

Augmented Dickey-Fuller Unit Root Test on INFLATION_US

Null Hypothesis: INFLATION_US has a unit root
Exogenous: Constant
Lag Length: 8 (Automatic - based on SIC, maxlag=14)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.826105	0.3672
Test critical values:		
1% level	-3.460035	
5% level	-2.874495	
10% level	-2.573751	

*Mackinnon (1996) one-sided p-values.

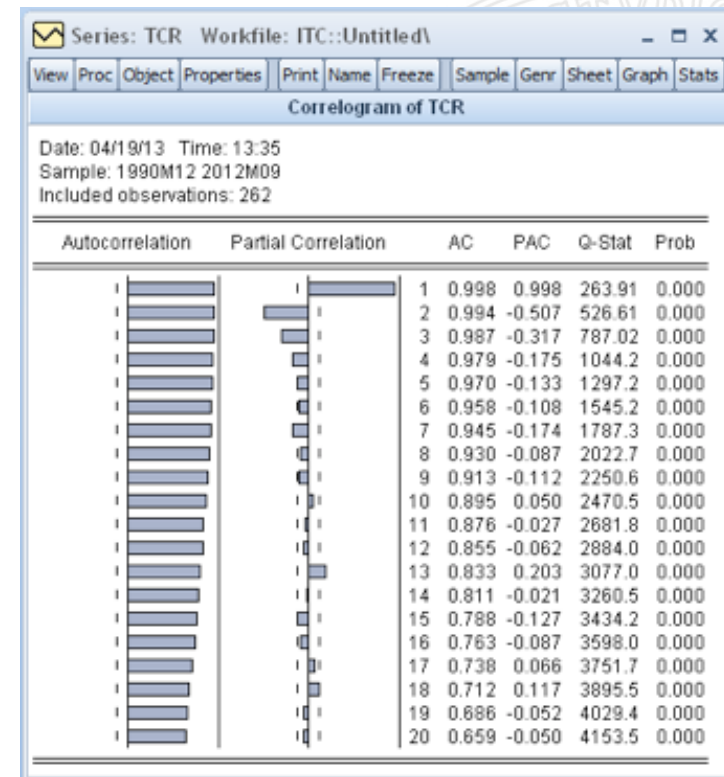
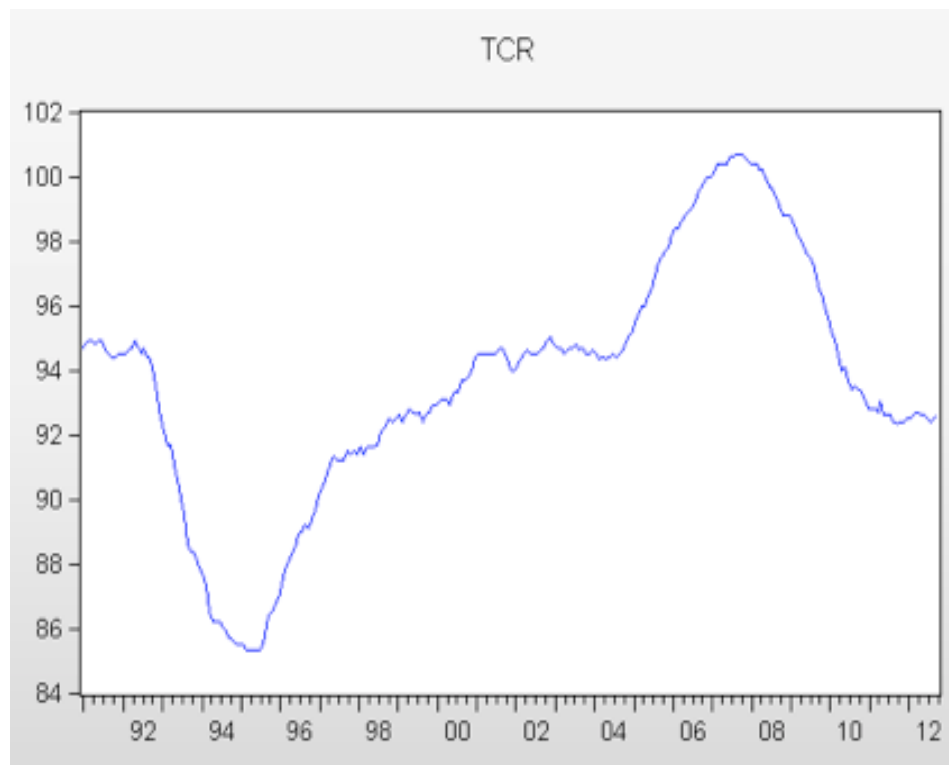
- Como el estadístico de contraste de Dickey-Fuller Aumentado tiene un p-valor de $0.3672 > 0.05$, no RH_0 ➡ **Hay raíz unitaria**
- La serie tiene **tendencia en varianza**, pero como no crece, no tendrá tendencia en media.

7.2. Tendencia en la media y tendencia en la varianza

➤ SITUACIÓN 3: no sabemos si la serie (de)crece o no

En este caso, trataremos a la serie igual que en la situación 1.

- **Ejemplo:** Tipo de cambio real efectivo de España frente a UM de 1990.12 a 2012.09



7.2. Tendencia en la media y tendencia en la varianza

- En el cuadro para hacer el contraste se selecciona:

Unit Root Test

Test type
Augmented Dickey-Fuller

Test for unit root in
 Level
 1st difference
 2nd difference

Include in test equation
 Intercept
 Trend and intercept
 None

Lag length
 Automatic selection:
Schwarz Info Criterion
Maximum lags: 15
 User specified: 4

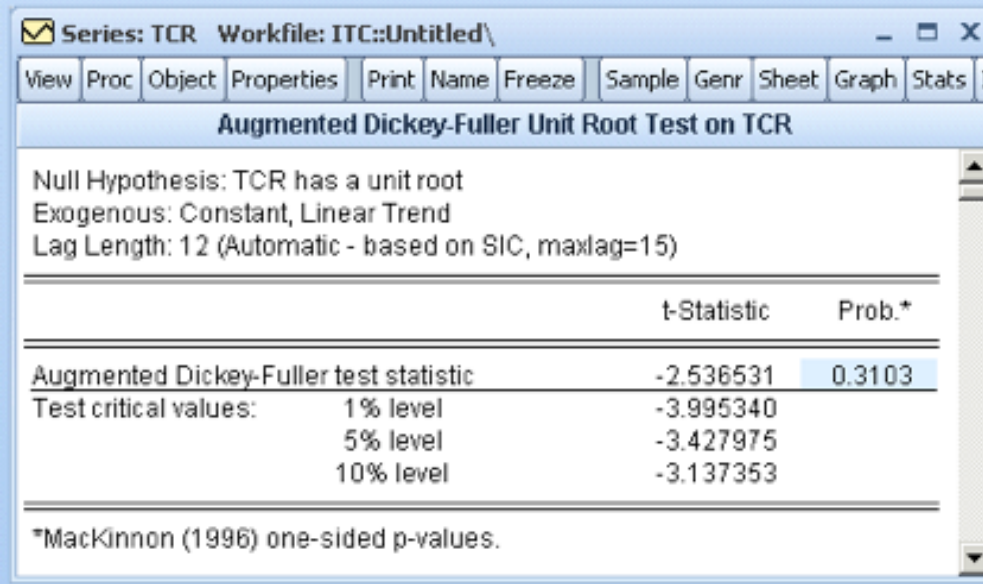
OK Cancel

H_0 : hay raíz unitaria

H_1 : no hay raíz unitaria

7.2. Tendencia en la media y tendencia en la varianza

- El resultado del contraste es:



Series: TCR Workfile: ITC::Untitled\

View Proc Object Properties Print Name Freeze Sample Genr Sheet Graph Stats I

Augmented Dickey-Fuller Unit Root Test on TCR

Null Hypothesis: TCR has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 12 (Automatic - based on SIC, maxlag=15)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.536531	0.3103
Test critical values:		
1% level	-3.995340	
5% level	-3.427975	
10% level	-3.137353	

*MacKinnon (1996) one-sided p-values.

- Como el estadístico de contraste de Dickey-Fuller Aumentado tiene un p-valor de $0.3103 > 0.05$, no RH_0 ➡ **Hay raíz unitaria**

7.3. Regresión espuria y solución

- Al regresar una serie temporal y_t sobre otra(s) x_t

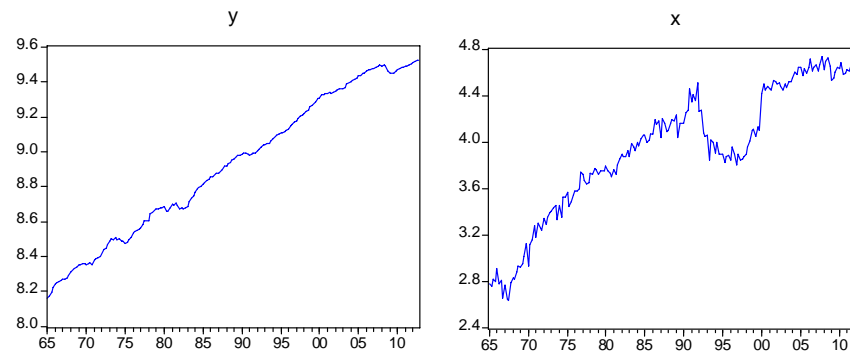
$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

estamos interesados en examinar si x_t tiene algún efecto sobre y_t

- En datos de corte transversal nos fijábamos en la bondad del ajuste del modelo (R^2) y en la significatividad de $\hat{\beta}_1$.
- Cuando esta regresión se realiza entre series temporales **no estacionarias** podría ocurrir que
 - El R^2 sea alto
 - El parámetro β_1 sea significativoaunque no exista ninguna relación económica entre ambas series
- Este problema se le conoce como **regresión espuria**

7.3. Regresión espuria y solución

- **Ejemplo 1:** regresión entre dos series con tendencia en varianza (en este ejemplo también hay tendencia en media)
 - Sean las series temporales y_t el logaritmo del PIB trimestral de EE.UU. y x_t el logaritmo de los coches registrados en Luxemburgo de 1965.1 a 2012.4



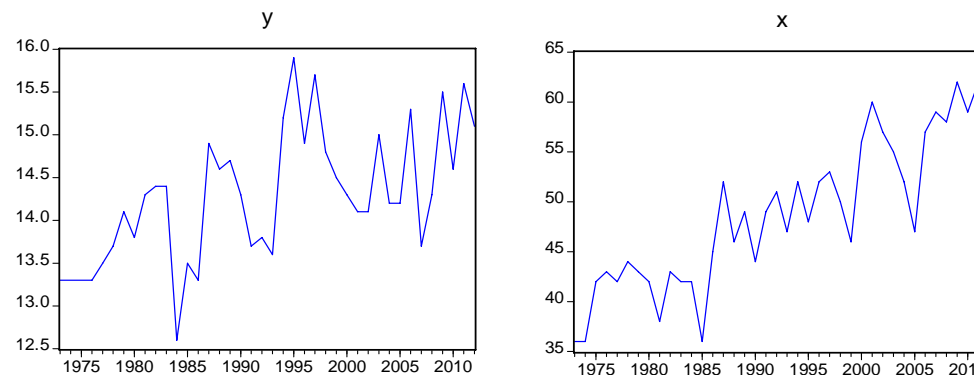
- Esperaríamos que la regresión de y_t sobre x_t nos diera un R^2 bajo y un coeficiente no significativo. Sin embargo:

$$\hat{y}_t = 6.24 + 0.67 x_t \quad R^2 = 0.86$$

(0.67) (0.02)

7.3. Regresión espuria y solución

- **Ejemplo 2:** regresión entre dos series con tendencia sólo en media
 - Sean las series temporales y_t la temperatura media anual en Madrid y x_t los partidos jugados en cada temporada por el FC Barcelona de 1973 a 2012



- Esperaríamos que la regresión de y_t sobre x_t nos diera un R^2 bajo y un coeficiente no significativo. Sin embargo:

$$\hat{y}_t = 11.38 + 0.06x_t \quad R^2 = 0.34$$

(0.65) (0.01)

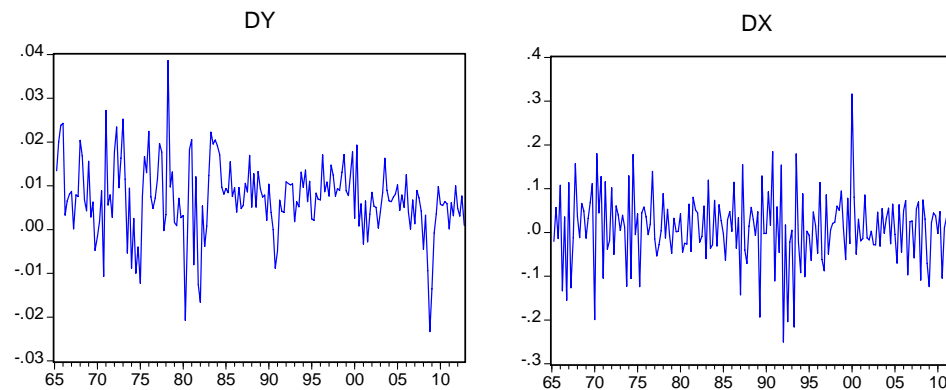
7.3. Regresión espuria y solución

- ¿Por qué parece buena la regresión entre las dos series?
 - R^2 puede ser alto sólo porque la varianza de y_t sea grande
 - Cuando hay tendencia en media el coeficiente es significativo porque la explicada tiene tendencia propia y la tendencia de la explicativa actúa de “variable proxy”
 - Cuando hay tendencia en la varianza (haya o no tendencia lineal en media) los estadísticos no tienen la distribución habitual y los contrastes de significatividad habituales no son válidos
- ¿Cómo debemos plantear la regresión para evitar este riesgo?
 - Tendencia en varianza: diferenciamos las series y las regresamos en diferencias (este método haría desaparecer la tendencia en varianza y también la tendencia en media si la hubiera)
 - Tendencia sólo en media: cualquiera de estos dos métodos equivalentes
 - Quitamos la tendencia lineal de cada serie y las regresamos sin ella
 - Añadimos una tendencia lineal a la regresión original

7.3. Regresión espuria y solución

- **Solución en ejemplo 1:**

- Ambas series tienen raíz unitaria (p-valores de ADF mayores que 0.05)
- Llamemos dy_t y dx_t a las diferencias de ambas series: $dy_t = y_t - y_{t-1}$ $dx_t = x_t - x_{t-1}$
- Para generar en Eviews las primeras diferencias de una serie hay dos opciones:
 - Utilizando el comando `d()`: “`genr dy=d(y)`”
 - Con la fórmula de las primeras diferencias: “`genr dy=y-y(-1)`”



- La regresión de las series en diferencias ya no muestra problemas de regresión espuria

$$d\hat{y}_t = 0.0071 - 0.00156 dx_t \quad R^2 = 0.0002$$

(-0.0006) (0.0078)

7.3. Regresión espuria y solución

- **Solución en ejemplo 2:**

- Ninguna de las series tiene raíz unitaria (p-valores de ADF menores que 0.05)
- **Solución 1:** quitamos la tendencia lineal y las regresamos sin ella

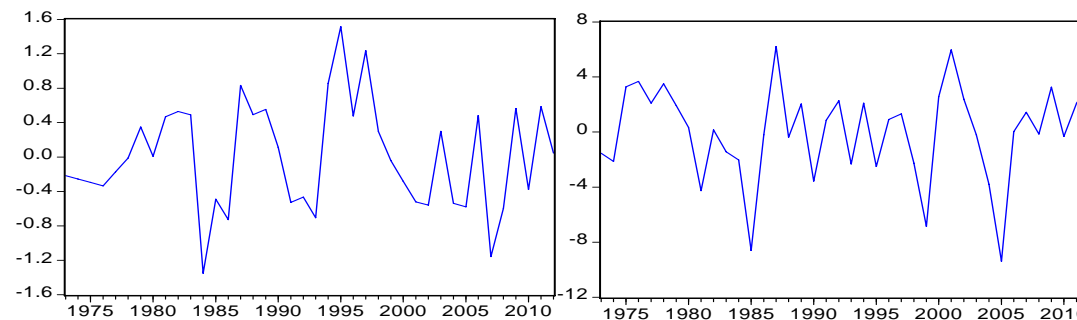
- Para ello, primero regresamos las series sobre una constante y una tendencia

$$y_t = \beta_{y0} + \beta_{y1}t + \varepsilon_{1t} \quad x_t = \beta_{x0} + \beta_{x1}t + \varepsilon_{2t}$$

- Llamamos a los residuos (variable menos tendencia estimada) $e1_t$ y $e2_t$

$$e1_t = y_t - (\hat{\beta}_{y0} + \hat{\beta}_{y1}t) \quad e2_t = x_t - (\hat{\beta}_{x0} + \hat{\beta}_{x1}t)$$

E1 E2



- Por último, regresamos $e1_t$ sobre $e2_t \Rightarrow$ no hay síntomas de regresión espuria

$$\hat{e1}_t = 0.00 + 0.03e2_t \quad R^2 = 0.024$$

(0.09) (0.03)

7.3. Regresión espuria y solución

- **Solución en ejemplo 2:**

- **Solución 2:** Añadimos una tendencia lineal a la regresión

$$\hat{y}_t = 12.462 + 0.022t + 0.028x_t \quad R^2 = 0.371$$

(1.116) (0.019) (0.029)

- El R^2 puede ser engañosamente alto debido a que y_t tiene una varianza grande al ser una variable con tendencia en media



- **Frecuencia de los datos:** diaria, mensual, trimestral, anual....

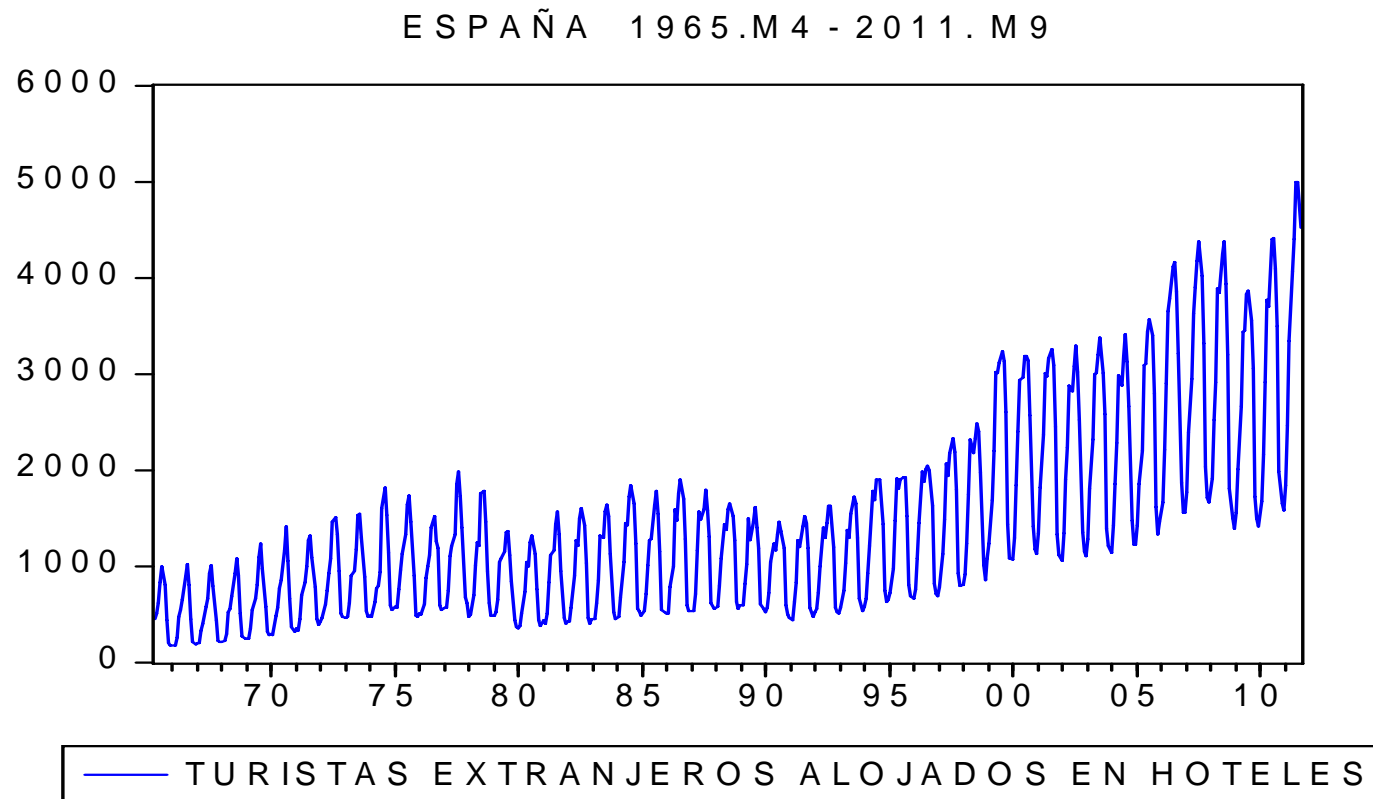
¿Cuál es la frecuencia de estas series?

- PIB de España
- Tasa de inflación de la UE
- Tasa de paro de España
- Número de ocupados en la Región de Murcia
- Precio de las acciones del IBEX 35

- **Estacionalidad:** Comportamiento cíclico de duración anual o menor, en series de frecuencia superior a la anual (frecuencia diaria, mensual, trimestral...)

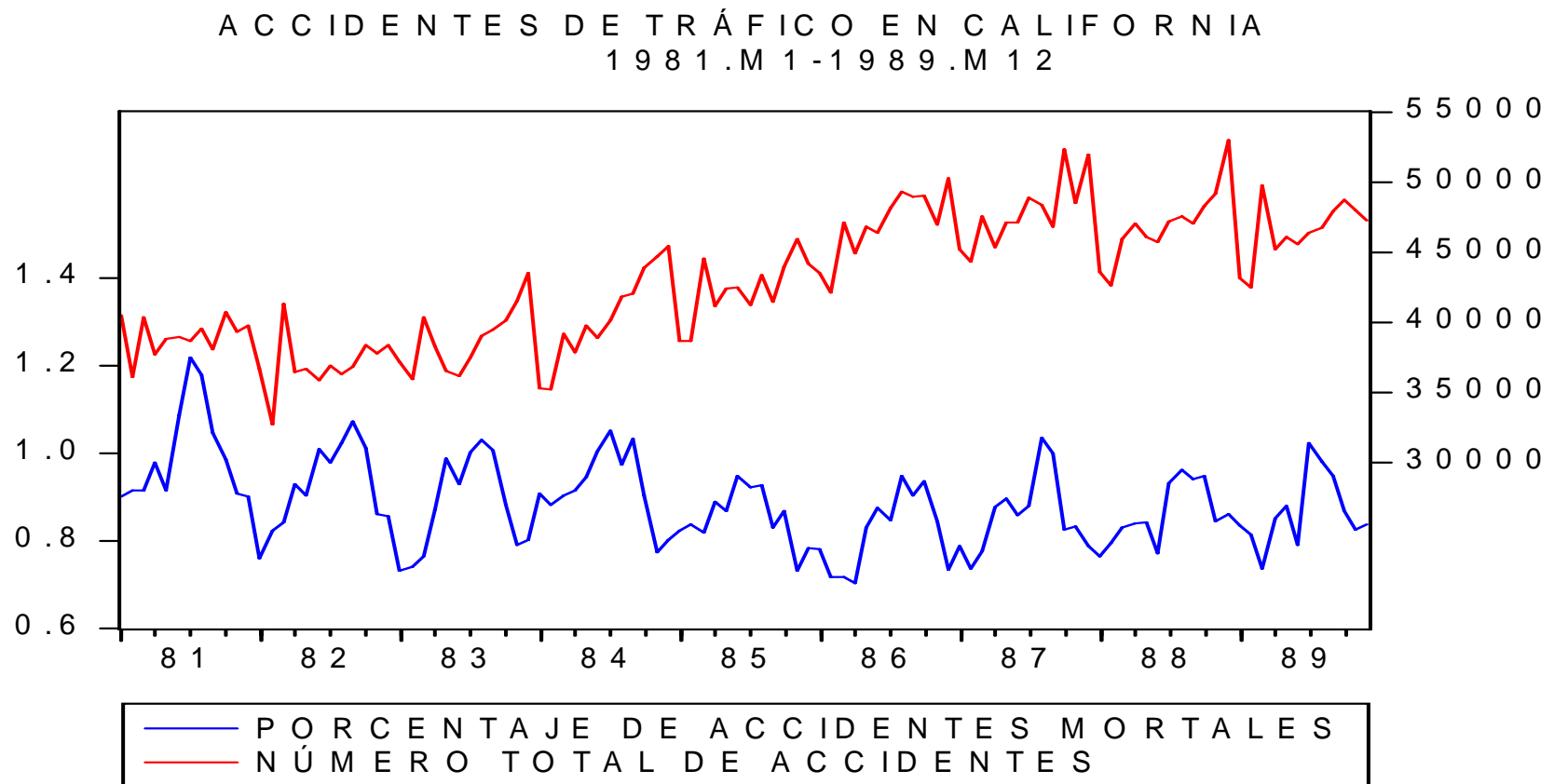
7.4. Estacionalidad

- Ejemplos: estacionalidad muy intensa



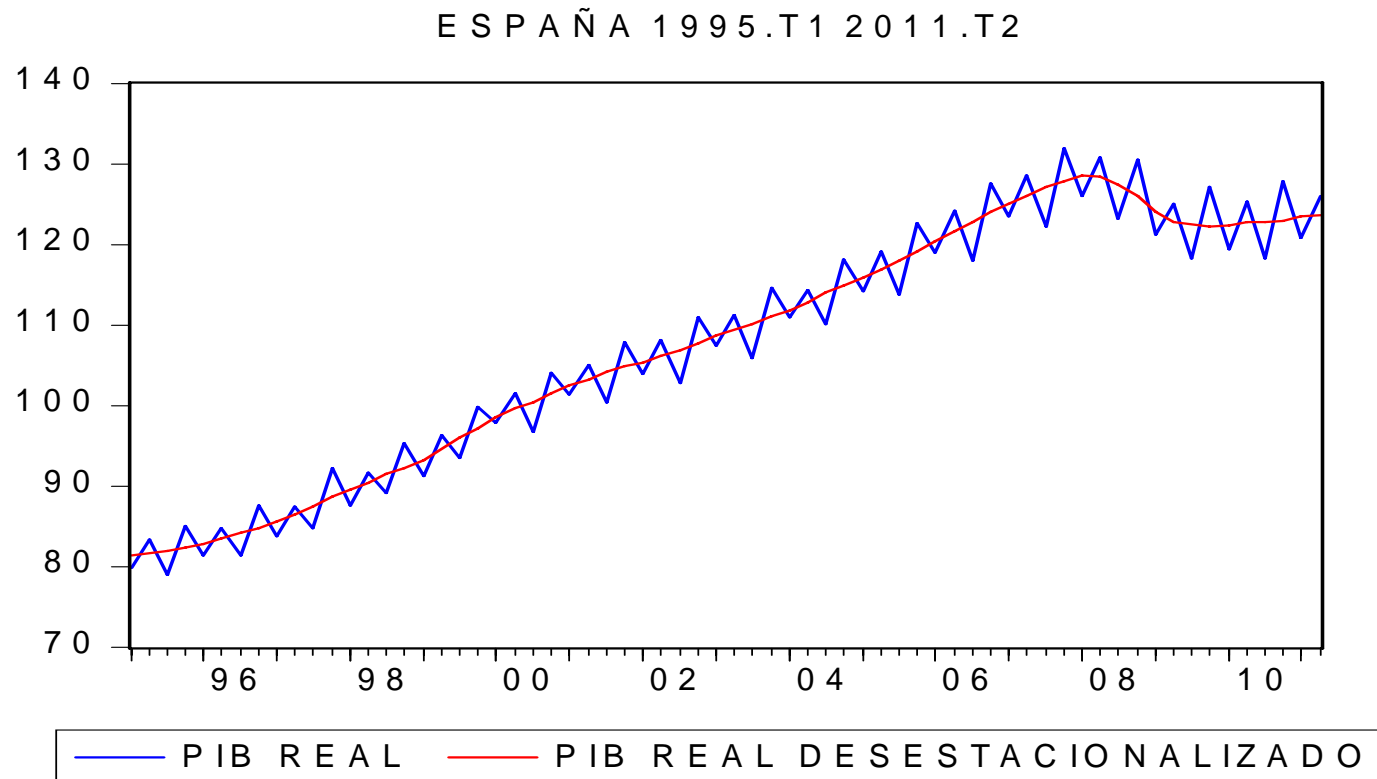
7.4. Estacionalidad

- Ejemplos: estacionalidad menos evidente



7.4. Estacionalidad

- Algunas series también se publican desestacionalizadas
- Por ejemplo, la serie trimestral del PIB (INE):



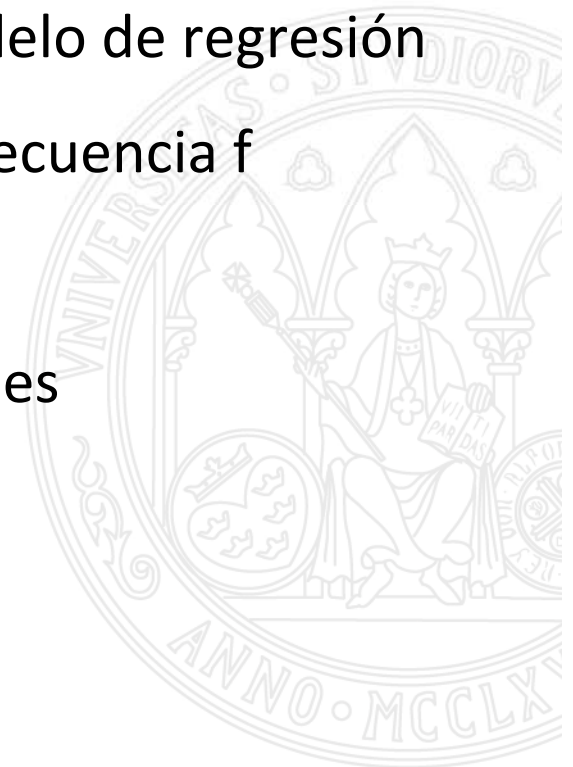
¿Qué hacer si los datos presentan comportamiento estacional?

- ❑ Introducir **ficticias estacionales** en el modelo de regresión

Se introducen $(f-1)$ ficticias en datos de frecuencia f

- ❑ **Desestacionalizar** previamente las variables

Existen varios métodos.



7.4. Estacionalidad

- **Ejemplo:** Sea el modelo de regresión con **datos trimestrales**

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t$$

Si hay estacionalidad \longrightarrow incluir **ficticias estacionales**

$$y_t = \beta_0 + \delta_1 T_{1t} + \delta_2 T_{2t} + \delta_3 T_{3t} + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t$$

$$T_{1t} = \begin{cases} 1 & t \in \text{trimestre 1} \\ 0 & \text{resto} \end{cases} \quad T_{2t} = \begin{cases} 1 & t \in \text{trimestre 2} \\ 0 & \text{resto} \end{cases} \quad T_{3t} = \begin{cases} 1 & t \in \text{trimestre 3} \\ 0 & \text{resto} \end{cases}$$

4 trimestres por año \longrightarrow introducir 3 ficticias

El cuarto trimestre es el período de referencia.

- ¿Qué hacer si no tenemos claro si hay comportamiento estacional?

□ Contrastar estacionalidad

Podemos contrastar estacionalidad en el modelo de regresión con ficticias estacionales.

En el ejemplo anterior (datos trimestrales), el contraste de estacionalidad es:

$$y_t = \beta_0 + \delta_1 T_{1t} + \delta_2 T_{2t} + \delta_3 T_{3t} + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t$$

$$H_0 : \delta_1 = \delta_2 = \delta_3 = 0$$

$$H_1 : \text{no } H_0$$

Si se RH_0  comportamiento estacional

7.4. Estacionalidad

Ejemplo: Disponemos de observaciones mensuales (enero 1981-diciembre 1989), sobre accidentes automovilísticos y desempleo correspondientes al estado de California.

Objetivo:

Averiguar si hay relación entre los accidentes de tráfico y el desempleo, teniendo en cuenta la posible estacionalidad (datos mensuales).

Las series son:

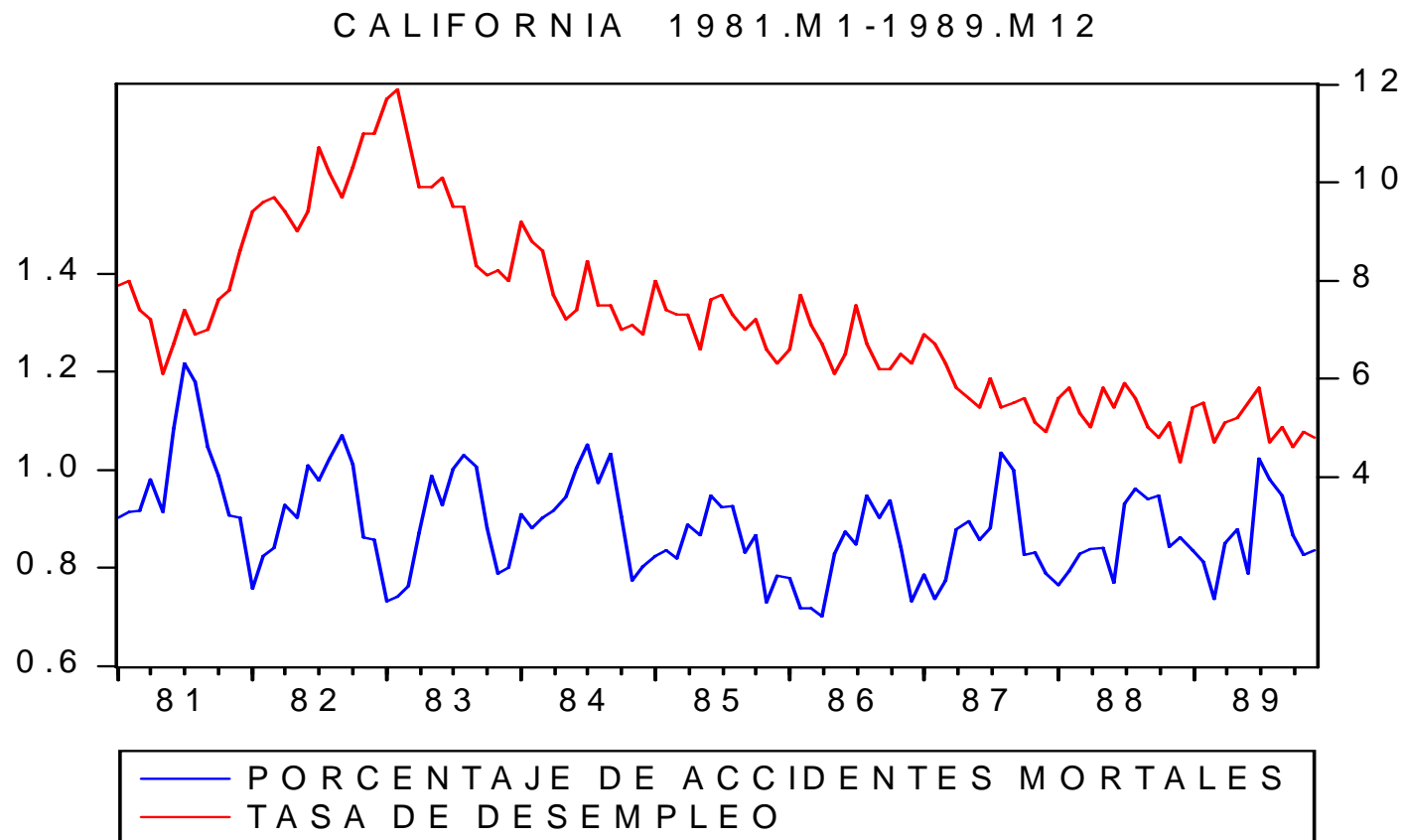
PRCFAT: porcentaje de accidentes con resultado de muerte.

DESEM: tasa de desempleo.

Ninguna de las dos series tiene tendencia en varianza (o raíz unitaria)

7.4. Estacionalidad

- Un gráfico de las series:



7.4. Estacionalidad

- Ambas variables presentan:
 - tendencia en media decreciente (más evidente en DESEM)
 - estacionalidad
- El modelo de regresión ha de tomar en cuenta la tendencia y estacionalidad de los datos

$$\text{PRCFAT}_t = \beta_0 + \beta_1 t + \delta_1 M_{1t} + \delta_2 M_{2t} + \delta_3 M_{3t} + \delta_4 M_{4t} + \delta_5 M_{5t} + \delta_6 M_{6t} \\ + \delta_7 M_{7t} + \delta_8 M_{8t} + \delta_9 M_{9t} + \delta_{10} M_{10t} + \delta_{11} M_{11t} + \beta_2 \text{DESEM}_t + \varepsilon_t$$

$$M_{1t} = \begin{cases} 1 & \text{si } t \in \text{mes 1} \\ 0 & \text{resto} \end{cases} \quad M_{2t} = \begin{cases} 1 & \text{si } t \in \text{mes 2} \\ 0 & \text{resto} \end{cases} \quad \dots$$

- Diciembre es el mes de referencia

7.4. Estacionalidad

- ¿Cómo estimar con ficticias estacionales en Eviews?

Con la función “**@seas(i)**”, donde i es el período del año para el cual queremos que la ficticia valga 1

Ejemplo 1 (para datos trimestrales):

“ls Y c X @seas(1) @seas(2) @seas(3)”

Ejemplo 2 (para datos mensuales):

“ls Y c X @seas(1) @seas(2) @seas(3) @seas(4) @seas(5) @seas(6)
@seas(7) @seas(8) @seas(9) @seas(10) @seas(11)”

7.4. Estacionalidad

- Modelo estimado:

La tasa de desempleo es significativa y con signo negativo.

La tendencia lineal es significativa y con signo negativo.

¿Hay estacionalidad?

Sí, porque algunas ficticias son significativas.

Equation: EQ01 Workfile: ESTACIONALIDAD

View Procs Objects Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: PRCFAT
Method: Least Squares
Date: 04/16/13 Time: 11:39
Sample: 1981:01 1989:12
Included observations: 108

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.053195	0.056331	18.69659	0.0000
@TREND	-0.001966	0.000306	-6.414027	0.0000
@SEAS(1)	-0.011081	0.028580	-0.387729	0.6991
@SEAS(2)	-0.011363	0.028649	-0.396649	0.6925
@SEAS(3)	-0.012701	0.028424	-0.446835	0.6560
@SEAS(4)	0.044242	0.028370	1.559493	0.1222
@SEAS(5)	0.064563	0.028398	2.273535	0.0253
@SEAS(6)	0.094035	0.028354	3.316503	0.0013
@SEAS(7)	0.170791	0.028564	5.979264	0.0000
@SEAS(8)	0.185284	0.028348	6.536115	0.0000
@SEAS(9)	0.151371	0.028343	5.340612	0.0000
@SEAS(10)	0.092586	0.028334	3.267671	0.0015
@SEAS(11)	0.005031	0.028335	0.177560	0.8595
DESEM	-0.017609	0.005471	-3.218373	0.0018

R-squared	0.681300	Mean dependent var	0.885636
Adjusted R-squared	0.637224	S.D. dependent var	0.099778
S.E. of regression	0.060097	Akaike info criterion	-2.665291
Sum squared resid	0.339495	Schwarz criterion	-2.317608
Log likelihood	157.9257	F-statistic	15.45754
Durbin-Watson stat	1.257141	Prob(F-statistic)	0.000000

7.4. Estacionalidad

- ¿Hay comportamiento estacional?
- Recordar: el **contraste de estacionalidad** es un contraste de significatividad conjunta de todas las ficticias estacionales

$$\text{PRCFAT}_t = \beta_0 + \beta_1 t + \delta_1 M_{1t} + \delta_2 M_{2t} + \delta_3 M_{3t} + \delta_4 M_{4t} + \delta_5 M_{5t} + \delta_6 M_{6t} \\ + \delta_7 M_{7t} + \delta_8 M_{8t} + \delta_9 M_{9t} + \delta_{10} M_{10t} + \delta_{11} M_{11t} + \beta_2 \text{DESEM}_t + \varepsilon_t$$

$$H_0 : \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = \delta_6 = \delta_7 = \delta_8 = \delta_9 = \delta_{10} = \delta_{11} = 0$$

$$H_1 : \text{no } H_0$$

- Estadístico de contraste: F

7.4. Estacionalidad

- Con Eviews el contraste de estacionalidad se puede hacer automáticamente.

En la ventana de ecuación:

“View/ Coefficient Tests/ Wald - Coefficient Restrictions...”

The screenshot shows the EViews software interface. The main window is titled "Equation: EQ01 Workfile: ESTACIONALIDAD". It displays the following information:

- Dependent Variable: PRCFAT
- Method: Least Squares
- Date: 04/16/13 Time: 11:47
- Sample: 1981:01 1989:12
- Included observations: 108

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.053195	0.056331	18.69659	0.0000
@TREND	-0.001966	0.000306	-6.414027	0.0000
@SEAS(1)	-0.011081	0.028580	-0.387729	0.6991
@SEAS(2)	-0.011363	0.028649	-0.396649	0.6925
@SEAS(3)				
@SEAS(4)				
@SEAS(5)				
@SEAS(6)				
@SEAS(7)				
@SEAS(8)				
@SEAS(9)				
@SEAS(10)				
@SEAS(11)				
DESEM				

Below the table, the following statistics are displayed:

- R-squared
- Adjusted R-squared
- S.E. of regression
- Sum squared resid
- Log likelihood
- Durbin-Watson stat
- Akaike info criterion
- Schwarz criterion
- F-statistic
- Prob(F-statistic)

The "Wald Test" dialog box is open, showing the following coefficient restrictions:

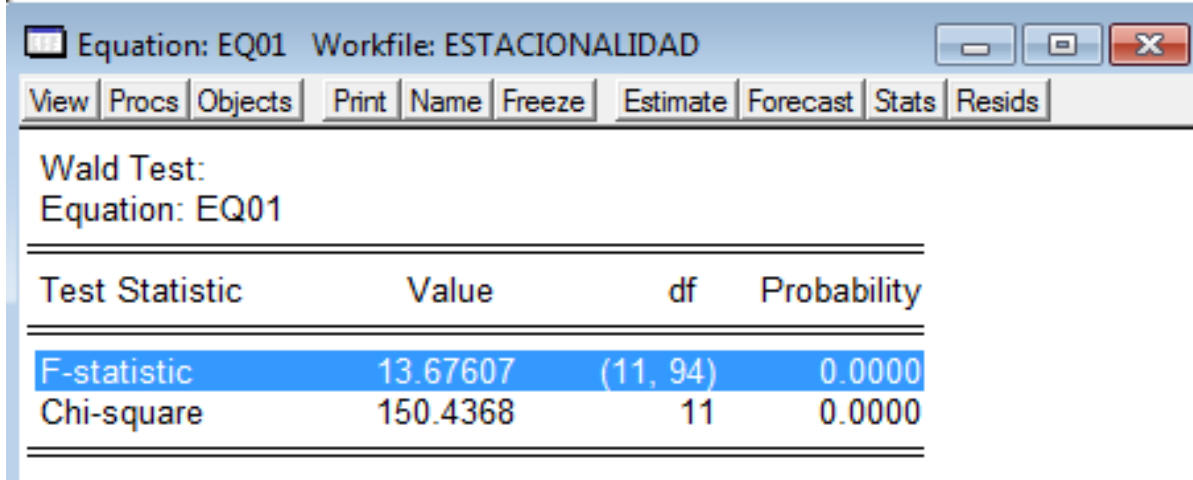
Coefficient restrictions separated by commas:
c(3)=c(4)=c(5)=c(6)=c(7)=c(8)=c(9)=c(10)=c(11)
=c(12)=c(13)=0

Examples:
C(1)=0, C(3)=2*C(4)

Buttons: OK, Cancel

7.4. Estacionalidad

- Contraste de estacionalidad. Resultado:



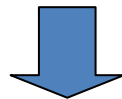
Test Statistic	Value	df	Probability
F-statistic	13.67607	(11, 94)	0.0000
Chi-square	150.4368	11	0.0000

Se $RH_0 \rightarrow$ hay comportamiento estacional

7.4. Estacionalidad

- Modelo definitivo: después de eliminar (una a una) las variables no significativas:

Las ficticias de enero, febrero, marzo y noviembre no son significativas.



Durante esos meses la evolución de PRCFAT es similar a la del mes de referencia (diciembre).

Equation: EQ01 Workfile: ESTACIONALIDAD

View Procs Objects Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: PRCFAT
Method: Least Squares
Date: 04/16/13 Time: 12:16
Sample: 1981:01 1989:12
Included observations: 108

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.051798	0.053431	19.68507	0.0000
@TREND	-0.001976	0.000300	-6.584959	0.0000
@SEAS(4)	0.050096	0.021647	2.314172	0.0227
@SEAS(5)	0.070269	0.021796	3.223978	0.0017
@SEAS(6)	0.099905	0.021613	4.622514	0.0000
@SEAS(7)	0.176969	0.021628	8.182241	0.0000
@SEAS(8)	0.191163	0.021595	8.852172	0.0000
@SEAS(9)	0.157103	0.021704	7.238512	0.0000
@SEAS(10)	0.098342	0.021673	4.537511	0.0000
DESEM	-0.018156	0.005305	-3.422605	0.0009

R-squared	0.679180	Mean dependent var	0.885636
Adjusted R-squared	0.649717	S.D. dependent var	0.099778
S.E. of regression	0.059053	Akaike info criterion	-2.732736
Sum squared resid	0.341753	Schwarz criterion	-2.484391
Log likelihood	157.5678	F-statistic	23.05191
Durbin-Watson stat	1.253180	Prob(F-statistic)	0.000000

7.4. Estacionalidad

Conclusiones:

- en promedio, un incremento en un punto porcentual de DESEM provoca una reducción de 0.018 puntos porcentuales en el porcentaje de accidentes (PRCFAT)
- hay una ligera tendencia decreciente
- los coeficientes positivos de las ficticias incluidas indican que, en promedio, el menor porcentaje de accidentes corresponde a los meses de invierno (noviembre a marzo).