



# **Econometría II**

**Perturbaciones no esféricas**

**Miguel Jerez y Sonia Sotoca**

**Universidad Complutense de Madrid**

**Marzo 2010**

# Índice

- **Tratamiento general**
- **Autocorrelación**
- **Heterocedasticidad**
- **Soluciones: estimación en presencia de autocorrelación y/o heterocedasticidad**
- **Ejemplo de autocorrelación**
- **Ejemplo de heterocedasticidad**

# Tratamiento general (I): Planteamiento

En Econometría I se planteó el modelo:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} ; \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

En donde la expresión  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$  resulta de suponer que las perturbaciones son:

- **normales** y
- **esféricas**, supuesto que a su vez está formado por dos:
  - **homoscedasticidad**:  $\forall t : E(\varepsilon_t^2) = \sigma^2$
  - **ausencia de autocorrelación**:  $\forall t, k (t \neq k) : E(\varepsilon_t \varepsilon_k) = 0$

En este tema supondremos que las perturbaciones son no esféricas, esto es, que:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} ; \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{\Omega})$$

en donde  $\boldsymbol{\Omega}$  es una matriz  $n \times n$ , simétrica y definida positiva. Como veremos, cuando las perturbaciones no son esféricas:

- algunas propiedades de la estimación MCO se mantienen, mientras que
- otras se pierden y sólo pueden recuperarse usando estimadores alternativos.

## Tratamiento general (II): Propiedades de los MCO

Cuando las perturbaciones no son esféricas, se mantienen aquéllas propiedades que no dependen de los momentos de segundo orden de los estimadores:

- Insesgadez:  $E(\hat{\beta}) = \beta$
- Esperanza nula del error de previsión:  $E(\mathbf{y}^* - \mathbf{X}^* \hat{\beta}) = \mathbf{0}$

Con perturbaciones no esféricas, la matriz de covarianzas del estimador MCO no tiene la forma habitual:

$$\mathbf{cov}(\hat{\beta}) = E\left[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \varepsilon \varepsilon^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1}\right] = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \Omega \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1}$$

... y, por tanto, cuando las perturbaciones no son esféricas:

- se pierde la eficiencia de los MCO, ya que las covarianzas de las estimaciones no llegan a la cota de Cramer-Rao:

$$\mathbf{cov}(\hat{\beta}) = \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \Omega \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \neq \sigma^2 (\mathbf{X}^T \mathbf{X})^{-1}$$

- Asimismo, no son válidos los resultados habituales de contraste de hipótesis y
- ... de previsión por intervalo.

## Tratamiento general (III): Estimación por mínimos cuadrados generalizados (MCG)

Sabiendo que:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} ; \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{\Omega}) \quad (1)$$

$$\boldsymbol{\Omega}^{-1} = \mathbf{P}^T \mathbf{P}$$

Premultiplicando (1) por  $\mathbf{P}$  se obtiene el modelo transformado:

$$\bar{\mathbf{y}} = \bar{\mathbf{X}}\boldsymbol{\beta} + \bar{\boldsymbol{\varepsilon}} \quad (2)$$

en donde:  $\bar{\mathbf{y}} = \mathbf{P}\mathbf{y}$  ;  $\bar{\mathbf{X}} = \mathbf{P}\mathbf{X}$  y  $\bar{\boldsymbol{\varepsilon}} = \mathbf{P}\boldsymbol{\varepsilon}$  y las perturbaciones del modelo transformado (2) son esféricas, ya que:

$$E(\bar{\boldsymbol{\varepsilon}}) = \mathbf{P}E(\boldsymbol{\varepsilon}) = \mathbf{0}$$

$$\begin{aligned} \text{cov}(\bar{\boldsymbol{\varepsilon}}) &= E(\bar{\boldsymbol{\varepsilon}} \bar{\boldsymbol{\varepsilon}}^T) = E(\mathbf{P} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T \mathbf{P}^T) = \mathbf{P} E(\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T) \mathbf{P}^T = \sigma^2 \mathbf{P} (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \\ &= \sigma^2 \mathbf{P} (\mathbf{P})^{-1} (\mathbf{P}^T)^{-1} \mathbf{P}^T = \sigma^2 \mathbf{I} \end{aligned}$$

## Tratamiento general (IV): Estimación por mínimos cuadrados generalizados (MCG)

El resultado anterior puede leerse de dos maneras:

- **Como una transformación de los datos**, en función de la matriz  $\mathbf{P}$ , que permite transformar (1) en un modelo equivalente con perturbaciones esféricas (2).
- **Como un estimador alternativo**, al que llamaremos “de Mínimos Cuadrados Generalizados” o MCG, con las mismas propiedades que el MCO en el caso de perturbaciones esféricas:

$$\hat{\beta}_{MCG} = (\bar{\mathbf{X}}^T \bar{\mathbf{X}})^{-1} \bar{\mathbf{X}}^T \bar{\mathbf{y}} = (\mathbf{X}^T \mathbf{P}^T \mathbf{P} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{P}^T \mathbf{P} \mathbf{y} = (\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Omega^{-1} \mathbf{y}$$

$$\text{cov}(\hat{\beta}_{MCG}) = \sigma^2 (\bar{\mathbf{X}}^T \bar{\mathbf{X}})^{-1} = \sigma^2 (\mathbf{X}^T \mathbf{P}^T \mathbf{P} \mathbf{X})^{-1} = \sigma^2 (\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1}$$

$$\hat{\sigma}_{MCG}^2 = \frac{1}{n-k} \hat{\boldsymbol{\varepsilon}}^T \hat{\boldsymbol{\varepsilon}} = \frac{1}{n-k} (\bar{\mathbf{y}} - \bar{\mathbf{X}} \hat{\beta}_{MCG})^T (\bar{\mathbf{y}} - \bar{\mathbf{X}} \hat{\beta}_{MCG})$$

$$= \frac{1}{n-k} (\mathbf{y} - \mathbf{X} \hat{\beta}_{MCG})^T \mathbf{P}^T \mathbf{P} (\mathbf{y} - \mathbf{X} \hat{\beta}_{MCG}) = \frac{1}{n-k} (\mathbf{y} - \mathbf{X} \hat{\beta}_{MCG})^T \Omega^{-1} (\mathbf{y} - \mathbf{X} \hat{\beta}_{MCG})$$

# Autocorrelación (I): Causas

La autocorrelación es una característica específica de las series temporales, ya que requiere que los datos posean un orden natural.

En general supondremos que la existencia de autocorrelación residual se debe a algún fallo en la especificación del modelo. Por ejemplo, a:

- Usar una especificación estática cuando debería utilizarse una dinámica.
- Omitir variables explicativas autocorreladas.
- Especificar un modelo lineal cuando la auténtica relación es no lineal.

En los años 50-60, la autocorrelación se contemplaba como un problema. Desde los 70 hasta la actualidad, se considera que:

- difícilmente un modelo econométrico podrá estar completo, por lo que
- las variables inevitablemente omitidas producirán alguna autocorrelación de los errores y esta autocorrelación, por tanto,
- es una característica más de los datos, que hay que modelizar.

## Autocorrelación (II): Detección

Los contrastes para detectar autocorrelación usan los resultados de la estimación MCO ya que, al ser insesgada, los residuos permiten inferir las propiedades de los errores subyacentes. Algunos de los instrumentos de detección disponibles son:

- Herramientas del análisis de series temporales (ACF y PACF residuales, p. ej.)
- Contraste de significación individual de los coeficientes de la ACF residual:

$$H_0 : \rho_k = 0 \quad t = \frac{\hat{\rho}_k}{\text{s.e.}(\hat{\rho}_k)} \underset{H_0}{\sim} t_{n-k} \quad \text{s.e.}(\hat{\rho}_k) \simeq n^{-1/2}$$

- Significación conjunta de  $K$  coeficientes de la ACF muestral de los residuos MCO:

$$H_0 : \rho_1 = \rho_2 = \rho_3 = \dots = \rho_K = 0 \quad Q(K) = n(n+2) \sum_{k=1}^K \frac{1}{n-k} \hat{\rho}_k^2 \underset{H_0}{\sim} \chi_{k-p}^2$$

Test de Breusch-Godfrey. Contrasta la hipótesis de ausencia de autocorrelación frente a la alternativa de autocorrelación de orden  $p$  de la siguiente manera:

- Estimar el modelo por MCO y calcular los residuos:  $\hat{\varepsilon}_t = y_t - \mathbf{x}_t^T \hat{\beta}_{MCO}$
- Estimar por MCO la regresión:  $\hat{\varepsilon}_t = \mathbf{x}_t^T \boldsymbol{\delta} + \varphi_1 \hat{\varepsilon}_{t-1} + \varphi_2 \hat{\varepsilon}_{t-2} + \dots + \varphi_p \hat{\varepsilon}_{t-p} + \mathbf{a}_t$
- Calcular el  $R^2$  de la regresión y el estadístico de contraste:  $BG = nR^2 \underset{H_0}{\sim} \chi_p^2$

# Autocorrelación (III): Detección

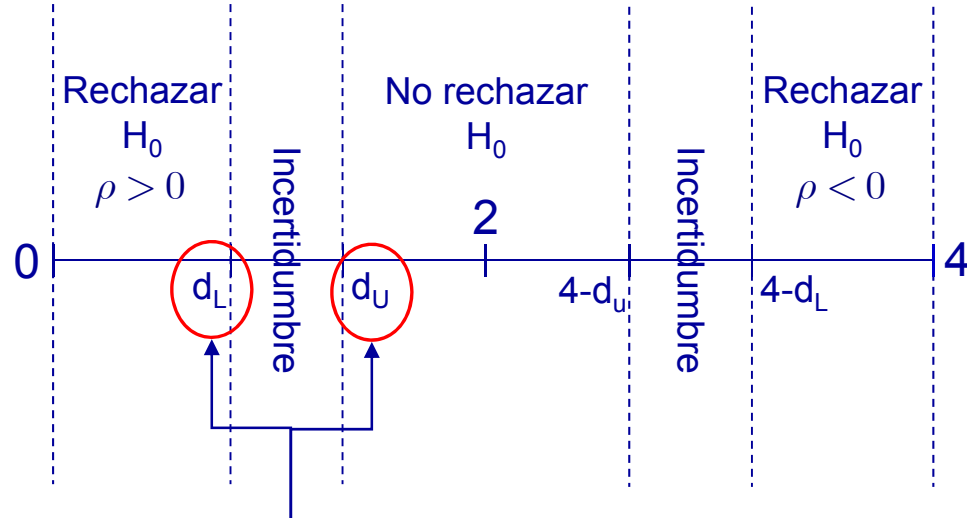
- Durbin-Watson:  $H_0 : \rho_1 = 0$

$$D = \frac{\sum_{t=2}^n (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^n \hat{\varepsilon}_t^2} \quad \hat{\varepsilon}_t = y_t - \mathbf{x}_t^T \hat{\beta}_{MCO}$$

$$D \simeq 2(1 - \hat{\rho}_1) \Rightarrow \rho_1 = 0 ; D \simeq 2$$

$$\rho_1 = 1 ; D \simeq 0$$

$$\rho_1 = -1 ; D \simeq 4$$



Estos límites, para diversos valores de  $k$  y  $n$ , figuran en las tablas del Test correspondientes al nivel de significación deseado

- Este test sólo detecta autocorrelación de orden 1. Para series trimestrales puede usarse el Test de Wallis:

$$D = \frac{\sum_{t=5}^n (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-4})^2}{\sum_{t=1}^n \hat{\varepsilon}_t^2} \simeq 2(1 - \hat{\rho}_4)$$

- Las tablas estándar no son válidas cuando hay regresores estocásticos. Para ese caso se dispone de tablas corregidas (Test Durbin-h).

# Heterocedasticidad (I): Causas

En un modelo de regresión, la heterocedasticidad supone que:

- la capacidad descriptiva del modelo cambia dentro de la muestra, y
- la varianza de la variable dependiente cambia, ya que:

$$y_i | \mathbf{x}_i \sim iid N(\mathbf{x}_i^T \boldsymbol{\beta}, \sigma_i^2)$$

La heterocedasticidad puede surgir tanto en series temporales como en datos de sección cruzada.

- En series temporales de baja frecuencia (mensual, trimestral, anual, ...) la heterocedasticidad suele manifestarse como una variabilidad de los datos que cambia con su nivel.
- En series de rendimientos de activos financieros observados con alta frecuencia (horaria, diaria, semanal, ...) la heterocedasticidad a menudo se muestra como períodos de alta y baja volatilidad en una serie estacionaria.
- Una causa frecuente de heterocedasticidad en datos de corte transversal es el distinto “tamaño” de las unidades que se comparan, sean éstas familias, empresas, países, etc.

## Heterocedasticidad (II): Detección

Igual que en el caso de autocorrelación, los contrastes para detectar heterocedasticidad utilizan los resultados de la estimación MCO, por ser ésta insesgada.

Los procedimientos de detección que veremos consideran distintas situaciones:

- La varianza de los datos depende de la media y esto causa heterocedasticidad (gráfico media-desviación típica)
- La varianza del error es distinta en dos partes de la muestra (Test de Goldfeld-Quandt)
- La varianza del error cambia de forma continua dentro de la muestra debido a:
  - un conjunto de variables conocidas (Test de Breusch-Pagan), o
  - una relación no lineal desconocida entre los errores y las variables exógenas del modelo de regresión (Test de White)

## Heterocedasticidad (III): Detección

**Test de Goldfeld-Quandt.** Contrasta la hipótesis nula de homoscedasticidad frente a la alternativa de que la varianza del error es distinta en dos partes de la muestra. Se calcula de la siguiente manera:

- Se seleccionan dos submuestras, de tamaños  $N_1$  y  $N_2$ , de manera que agrupen las observaciones con varianzas del error potencialmente distintas.
- Se estima por MCO el modelo para cada submuestra.
- Se calculan las estimaciones MCO de  $\hat{\sigma}_1^2$  y  $\hat{\sigma}_2^2$ . En estas condiciones y para muestras “grandes” puede demostrarse que:

$$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} \underset{H_0}{\sim} F_{N_1-K, N_2-K}$$

La mayor dificultad para aplicar este test se debe a que no resuelve cómo especificar las submuestras.

## Heterocedasticidad (IV): Detección

**Test de Breusch-Pagan.** Supone que la varianza de las perturbaciones es función de un vector de variables  $\mathbf{z}_t$ , que pueden estar incluidas o no en el modelo. La hipótesis nula es  $H_0 : \alpha = \mathbf{0}$  (homocedasticidad). El test se basa en estimar por MCO la regresión:

$$\frac{\hat{\varepsilon}_t^2}{\tilde{\sigma}^2} = \alpha_0 + \mathbf{z}_t^T \alpha + u_t$$

en donde  $\hat{\varepsilon}_t$  son los residuos MCO cuya posible heterocedasticidad se desea contrastar y  $\tilde{\sigma}^2$  es el estimador máximo-verosímil de la varianza del error. En estas condiciones, puede demostrarse que, para muestras “grandes”:

$$\frac{1}{2} SE \underset{H_0}{\sim} \chi_p^2$$

donde  $SE$  es la suma explicada en la regresión anterior y  $p$  es el número de variables en  $\mathbf{z}_t$ . Alternativamente, el contraste podría hacerse usando:

$$nR^2 \underset{H_0}{\sim} \chi_p^2$$

Este test tiene el inconveniente de que a menudo es difícil especificar qué variables causan la heterocedasticidad.

# Heterocedasticidad (V): Detección

**Test de White.** Contrasta la hipótesis nula de ausencia de heterocedasticidad sin suponer una forma funcional concreta para la varianza.

Se calcula estimando una regresión de los residuos MCO al cuadrado sobre: a) una constante, b) las variables explicativas, c) sus cuadrados y d) sus productos cruzados de segundo orden. Puede demostrarse que para muestras “grandes”:

$$nR^2 \underset{H_0}{\sim} \chi_p^2$$

en donde  $n$  es el número de residuos y  $p$  es el número de variables explicativas de la regresión de apoyo, exceptuando el término constante.

**Ejemplo.** Si el modelo de regresión es:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

entonces la regresión de apoyo del test de White tiene la forma:

$$\hat{\varepsilon}_i^2 = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \alpha_3 x_{i1}^2 + \alpha_4 x_{i2}^2 + \alpha_5 x_{i1} x_{i2} + u_i$$

y el valor de  $p$  sería 5.

# Soluciones: Estimación en presencia de autocorrelación y/o heterocedasticidad (I)

Si la muestra es grande y nuestro único problema es hacer inferencia a partir de las estimaciones MCO, puede ser adecuado calcular una “matriz de covarianzas robusta” a problemas de heterocedasticidad, como la matriz de covarianzas de White, o a problemas de autocorrelación y heterocedasticidad, como la matriz de Newey-West. En un ejemplo posterior se ilustra esta posibilidad.

Por otra parte, si se conocen los valores en la matriz  $\Omega$ , los resultados de MCG permiten transformar un modelo con perturbaciones no esféricas en otro equivalente con perturbaciones esféricas y, por tanto, pueden aplicarse para:

- resolver problemas de contraste de hipótesis, y
- estimar los parámetros  $\beta$  del modelo por MCO.
- En general, los valores en  $\Omega$  son desconocidos, y es preciso estimarlos a la vez que  $\beta$ . El problema resultante es no lineal y, consecuentemente, su solución requiere el uso de métodos iterativos.
- A modo de ejemplo, veremos un método adecuado para esquemas sencillos de autocorrelación (método de Cochrane-Orcutt)

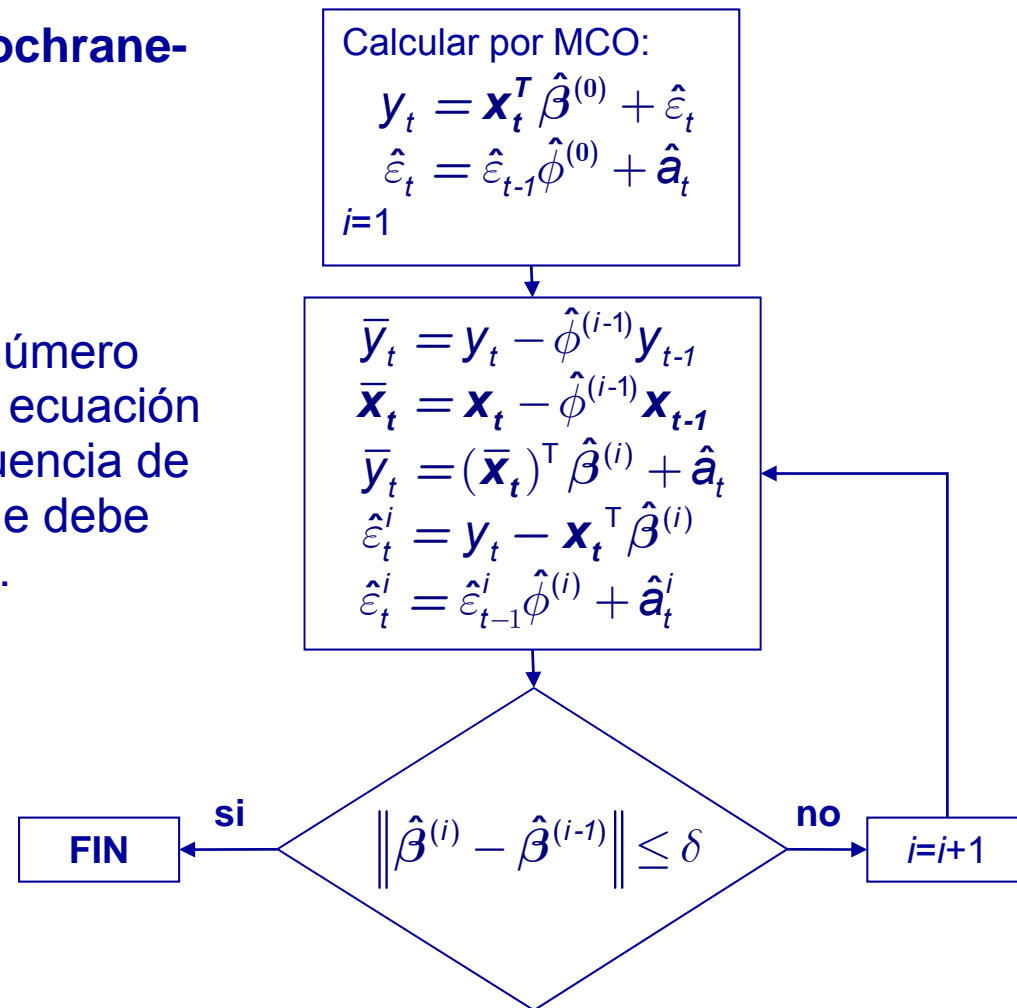
# Soluciones: Estimación en presencia de autocorrelación y/o heterocedasticidad (II)

**Organigrama del Método de Cochrane-Orcutt.** Suponemos que:

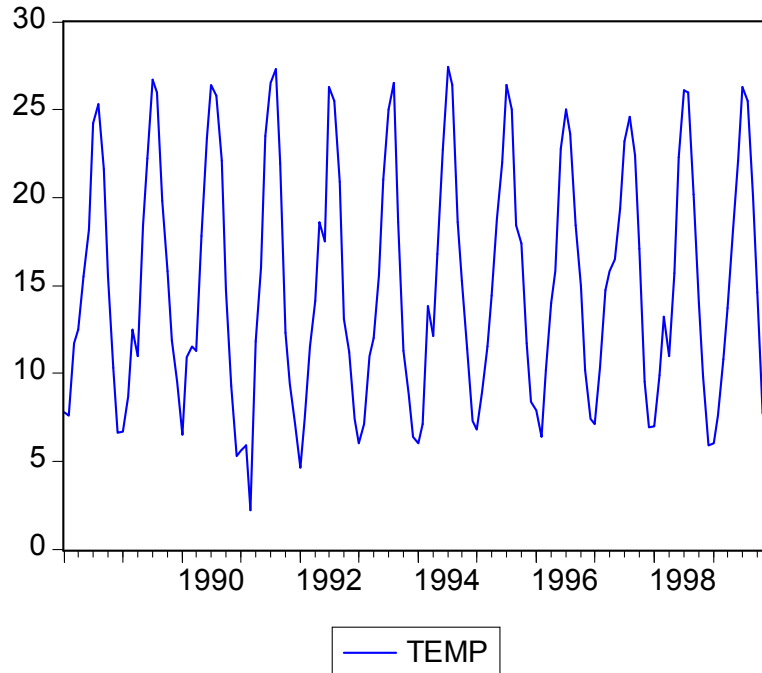
$$y_t = \mathbf{x}_t^T \boldsymbol{\beta} + \varepsilon_t ; \varepsilon_t = \phi \varepsilon_{t-1} + a_t$$

$$a_t \sim iidN(0; \sigma_a^2)$$

El método consiste en iterar un número indeterminado de veces sobre la ecuación de regresión, aplicando una secuencia de correcciones MCG mejoradas que debe converger a la corrección óptima.



## Ejemplo de autocorrelación (I)



Consideremos la serie histórica de temperaturas medias mensuales (en grados centígrados) registradas en El Retiro de Madrid, desde Enero de 1988 hasta Diciembre de 1999.

Estos datos están disponibles en formatos .WF1 y .XLS en [www.ucm.es/info/ecocuan/mjm/ecoaplimj](http://www.ucm.es/info/ecocuan/mjm/ecoaplimj)

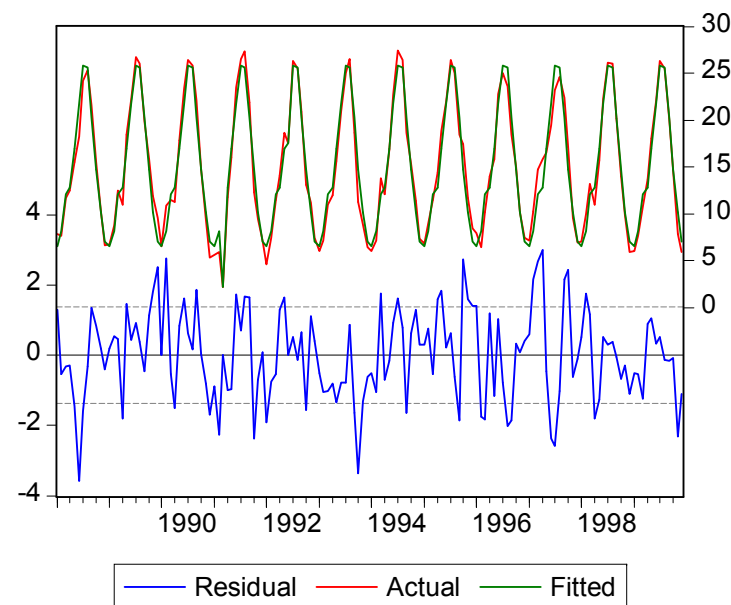
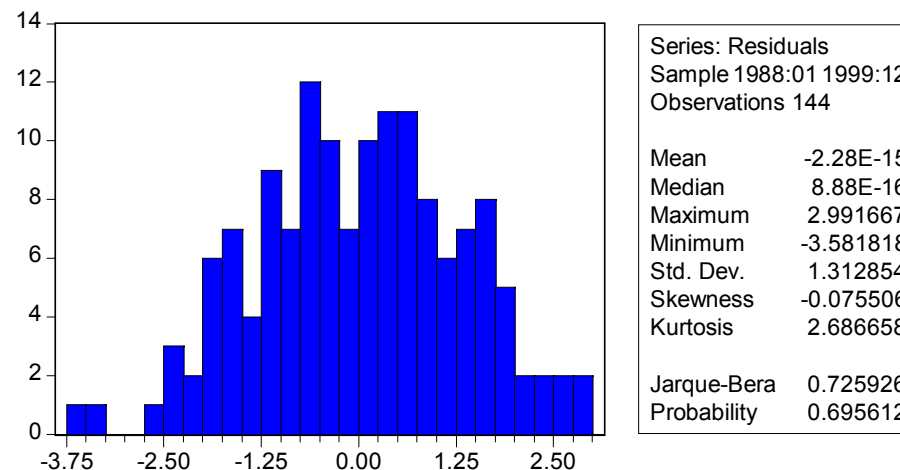
Según nuestra experiencia personal esta serie no puede ser estacionaria. ¿Por qué?

Como la temperatura media en el mismo mes de distintos años es aproximadamente constante, vamos a ajustar un modelo que la explica TEMP en función de:

- Doce variables *dummy*, caracterizando los doce meses del año
- Dos variables tipo impulso, para captar valores extremos en 3/91 y 6/92

# Ejemplo de autocorrelación (II)

Dependent Variable: TEMP				
Method: Least Squares				
Date: 02/24/03 Time: 12:53				
Sample: 1988:01 1999:12				
Included observations: 144				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
ENERO	6.500000	0.397486	16.35277	0.0000
FEBRERO	8.150000	0.397486	20.50386	0.0000
MARZO	12.03636	0.415161	28.99205	0.0000
ABRIL	12.80833	0.397486	32.22334	0.0000
MAYO	16.95833	0.397486	42.66395	0.0000
JUNIO	21.78182	0.415161	52.46598	0.0000
JULIO	25.79167	0.397486	64.88695	0.0000
AGOSTO	25.62500	0.397486	64.46764	0.0000
SEPTIEMBRE	20.25000	0.397486	50.94516	0.0000
OCTUBRE	14.67500	0.397486	36.91952	0.0000
NOVIEMBRE	10.10833	0.397486	25.43065	0.0000
DICIEMBRE	7.008333	0.397486	17.63164	0.0000
PULSE0391	-9.836364	1.438159	-6.839551	0.0000
PULSE0692	-4.281818	1.438159	-2.977291	0.0035
R-squared	0.963160	Mean dependent var	15.04306	
Adjusted R-squared	0.959476	S.D. dependent var	6.839996	
S.E. of regression	1.376933	Akaike info criterion	3.569759	
Sum squared resid	246.4727	Schwarz criterion	3.858491	
Log likelihood	-243.0227	Durbin-Watson stat	1.357921	

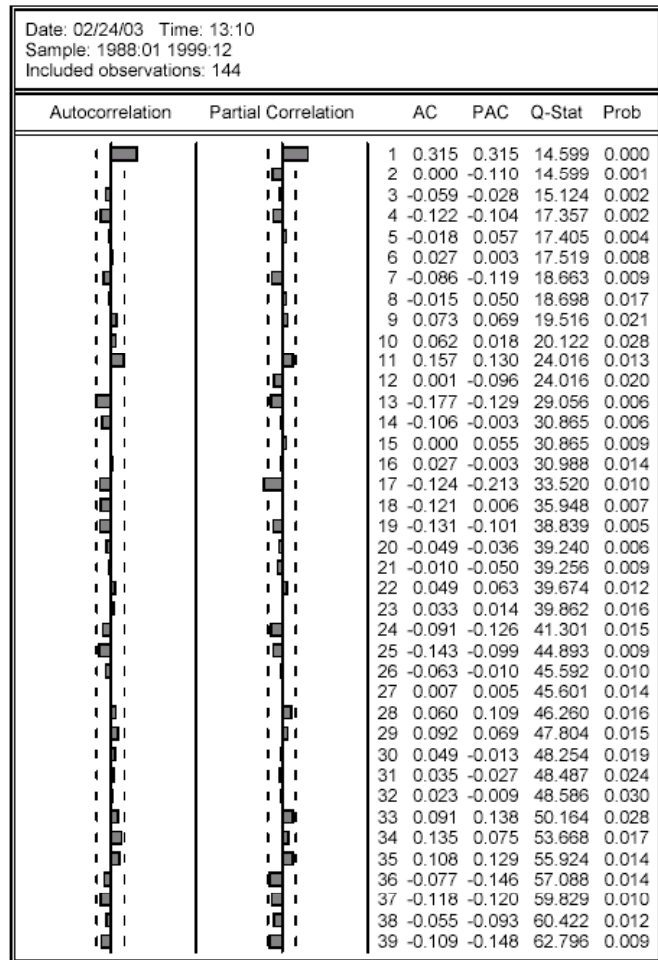


La estimación MCO parece adecuada en general. ¿Cuál es la interpretación de

- los parámetros estimados y
- de la desviación típica residual?

# Ejemplo de autocorrelación (III)

Sin embargo los residuos muestran una débil autocorrelación positiva de primer orden.  
¿Cómo podemos interpretarla?



Si ampliamos el modelo añadiendo un término AR(1), el ajuste mejora ligeramente. Asimismo, las estimaciones de los parámetros apenas cambian. ¿por qué?

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ENERO	6.347987	0.416176	15.25313	0.0000
FEBRERO	8.098959	0.402380	20.12765	0.0000
MARZO	11.92984	0.415334	28.72348	0.0000
ABRIL	12.80258	0.400310	31.98165	0.0000
MAYO	16.95640	0.400268	42.36263	0.0000
JUNIO	21.84015	0.414610	52.67638	0.0000
JULIO	25.79145	0.400260	64.43669	0.0000
AGOSTO	25.62493	0.400260	64.02068	0.0000
SEPTIEMBRE	20.24998	0.400260	50.59204	0.0000
OCTUBRE	14.67499	0.400260	36.66364	0.0000
NOVIEMBRE	10.10833	0.400260	25.25440	0.0000
DICIEMBRE	7.008332	0.400260	17.50945	0.0000
PULSE0391	-8.763786	1.302826	-6.726752	0.0000
PULSE0692	-4.989555	1.298489	-3.842585	0.0002
AR(1)	0.335767	0.084122	3.991440	0.0001
R-squared	0.967106	Mean dependent var	15.09371	
Adjusted R-squared	0.963508	S.D. dependent var	6.836884	
S.E. of regression	1.306043	Akaike info criterion	3.470858	
Sum squared resid	218.3359	Schwarz criterion	3.781645	
Log likelihood	-233.1663	Durbin-Watson stat	1.909043	
Inverted AR Roots	.34			

## Ejemplo de heterocedasticidad (I)

Caso	País	Gasto sanitario Púb. (MMill. \$)	GDP (MMill. \$)	Población 99 (Mill. Hab.)
1	Alemania	169.90	2150.6	82.09
2	Australia	22.07	374.0	18.97
3	Austria	12.25	211.2	8.09
4	Bélgica	19.77	250.2	10.22
5	Canadá	39.49	598.3	30.49
6	Corea	7.29	317.1	46.86
7	Dinamarca	11.87	174.5	5.32
8	EE.UU.	532.46	8728.8	272.99
9	España	31.64	586.0	39.42
10	Finlandia	6.84	129.0	5.17
11	Francia	104.54	1452.0	60.15
12	Grecia	5.71	121.4	10.53
13	Holanda	23.48	391.3	15.81
14	Hungría	2.44	47.0	10.06
15	Irlanda	4.14	86.3	3.74
16	Islandia	0.57	8.1	0.28
17	Italia	67.00	1196.5	57.65
18	Japón	236.43	3940.5	126.69
19	Luxemburgo	0.99	18.3	0.43
20	Méjico	11.79	421.0	97.44
21	Noruega	10.49	147.8	4.46
22	Nueva Zelanda	3.28	52.9	3.81
23	Polonia	6.69	159.3	38.67
24	Portugal	5.80	111.6	9.98
25	Reino Unido	78.98	1410.4	59.50
26	República Checa	3.74	56.7	10.28
27	Suecia	16.78	239.7	8.86
28	Suiza	20.24	262.8	7.14
29	Turquía	5.81	200.3	65.82

Consideremos los datos de Gasto Sanitario Público (GSAN\_PUB), PIB (GDP) y Población (POB) en los países de la OCDE (menos la República Eslovaca). El objetivo es decidir si el gasto en España está de acuerdo con los “estándares” de nuestro entorno socioeconómico.

Es habitual abordar este análisis comparando ratios de gasto sobre PIB ...

Los datos están disponibles en formatos .WF1 y .XLS en: [www.ucm.es/info/ecocuan/mjm/ecoaplimj](http://www.ucm.es/info/ecocuan/mjm/ecoaplimj)

Estadísticas sanitarias, económicas y demográficas OCDE (menos República Eslovaca)

Fuente: OCDE (<http://www.oecd.org>)

Todos los datos están referidos a 1998 excepto cuando se indica

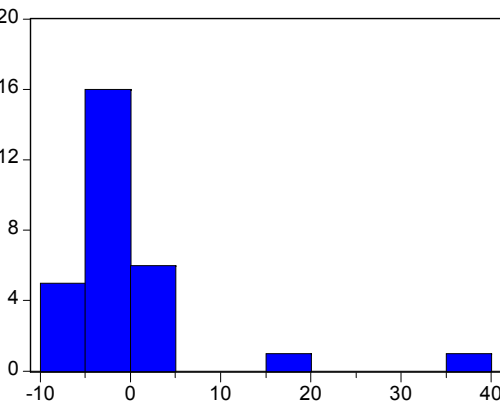
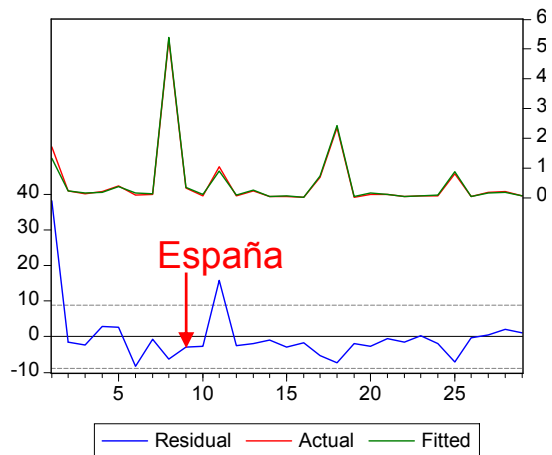
## Ejemplo de heterocedasticidad (II)

Rkg	País	Gasto sanitario Púb. (MMill. \$)	GDP (MMill. \$)	Gasto Púb/GDP %	Rkg	País	Gasto sanitario Púb. (MMill. \$)	Población 99 (Mill. Hab.)	Gasto Púb/POB Miles \$/Hab
1	Alemania	169.9	2150.6	7.9%	1	Suíza	20.2	7.1	2.8
2	Bélgica	19.8	250.2	7.9%	2	Noruega	10.5	4.5	2.4
3	Suíza	20.2	262.8	7.7%	3	Luxemburgo	1.0	0.4	2.3
4	Francia	104.5	1452.0	7.2%	4	Dinamarca	11.9	5.3	2.2
5	Noruega	10.5	147.8	7.1%	5	Alemania	169.9	82.1	2.1
6	Suecia	16.8	239.7	7.0%	6	Islandia	0.6	0.3	2.1
7	Islandia	0.6	8.1	7.0%	7	EE.UU.	532.5	273.0	2.0
8	Dinamarca	11.9	174.5	6.8%	8	Bélgica	19.8	10.2	1.9
9	Canadá	39.5	598.3	6.6%	9	Suecia	16.8	8.9	1.9
10	República Checa	3.7	56.7	6.6%	10	Japón	236.4	126.7	1.9
11	Nueva Zelanda	3.3	52.9	6.2%	11	Francia	104.5	60.2	1.7
12	EE.UU.	532.5	8728.8	6.1%	12	Austria	12.2	8.1	1.5
13	Japón	236.4	3940.5	6.0%	13	Holanda	23.5	15.8	1.5
14	Holanda	23.5	391.3	6.0%	14	Reino Unido	79.0	59.5	1.3
15	Australia	22.1	374.0	5.9%	15	Finlandia	6.8	5.2	1.3
16	Austria	12.2	211.2	5.8%	16	Canadá	39.5	30.5	1.3
17	Italia	67.0	1196.5	5.6%	17	Australia	22.1	19.0	1.2
18	Reino Unido	79.0	1410.4	5.6%	18	Italia	67.0	57.6	1.2
19	<b>España</b>	<b>31.6</b>	<b>586.0</b>	<b>5.4%</b>	19	Irlanda	4.1	3.7	1.1
20	Luxemburgo	1.0	18.3	5.4%	20	Nueva Zelanda	3.3	3.8	0.9
21	Finlandia	6.8	129.0	5.3%	21	<b>España</b>	<b>31.6</b>	<b>39.4</b>	<b>0.8</b>
22	Hungría	2.4	47.0	5.2%	22	Portugal	5.8	10.0	0.6
23	Portugal	5.8	111.6	5.2%	23	Grecia	5.7	10.5	0.5
24	Irlanda	4.1	86.3	4.8%	24	República Checa	3.7	10.3	0.4
25	Grecia	5.7	121.4	4.7%	25	Hungría	2.4	10.1	0.2
26	Polonia	6.7	159.3	4.2%	26	Polonia	6.7	38.7	0.2
27	Turquía	5.8	200.3	2.9%	27	Corea	7.3	46.9	0.2
28	Méjico	11.8	421.0	2.8%	28	Méjico	11.8	97.4	0.1
29	Corea	7.3	317.1	2.3%	29	Turquía	5.8	65.8	0.1
<b>Media</b>				<b>5.8%</b>	<b>Media</b>				<b>1.3</b>
<b>Déficit español sobre media</b>				<b>0.4%</b>	<b>Déficit español sobre media</b>				<b>0.5</b>
<b>Gasto adicional requerido (Mmill. \$)</b>				<b>2.14</b>	<b>Gasto adicional requerido (Mmill. \$)</b>				<b>19.3</b>

# Ejemplo de heterocedasticidad (III)

Dependent Variable: GSAN_PUB Method: Least Squares Date: 03/04/03 Time: 11:41 Sample: 1 29 Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.843333	2.162583	0.852375	0.4018
GDP	0.066420	0.002775	23.93551	0.0000
POB	-0.157221	0.086365	-1.820420	0.0802
R-squared	0.993709	Mean dependent var	50.43034	
Adjusted R-squared	0.993225	S.D. dependent var	107.1874	
S.E. of regression	8.822777	Akaike info criterion	7.290248	
Sum squared resid	2023.876	Schwarz criterion	7.431692	
Log likelihood	-102.7086	F-statistic	2053.354	
Durbin-Watson stat	1.306955	Prob(F-statistic)	0.000000	

White Heteroskedasticity Test:				
F-statistic	4.133801	Probability	0.007970	
Obs*R-squared	13.72601	Probability	0.017447	
Test Equation: Dependent Variable: RESID^2 Method: Least Squares Date: 03/04/03 Time: 13:25 Sample: 1 29 Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.81744	66.31333	0.178206	0.8601
GDP	0.027544	0.246734	0.111633	0.9121
GDP^2	-0.000985	0.000320	-3.080540	0.0053
GDP*POB	0.034140	0.011911	2.866164	0.0087
POB	-5.092047	7.097719	-0.717420	0.4803
POB^2	-0.068532	0.087048	-0.787284	0.4392



Series: Residuals	
Sample 1 29	
Observations 29	
Mean	-1.67E-14
Median	-1.767310
Maximum	38.12086
Minimum	-8.247623
Std. Dev.	8.501840
Skewness	3.399792
Kurtosis	15.47356
Jarque-Bera	243.8706
Probability	0.000000

Se observan indicios claros de no normalidad y heterocedasticidad residual, ¿cuál es la causa?

## Ejemplo de heterocedasticidad (IV)

Una primera solución consiste en mantener el método de estimación y calcular errores estándar “robustos” a problemas de heterocedasticidad

Dependent Variable: GSAN_PUB				
Method: Least Squares				
Date: 05/05/03 Time: 09:42				
Sample: 1 29				
Included observations: 29				
White Heteroskedasticity-Consistent Standard Errors & Covariance				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.843333	0.993556	1.855288	0.0749
GDP	0.066420	0.001242	53.47846	0.0000
POB	-0.157221	0.033995	-4.624868	0.0001
R-squared	0.993709	Mean dependent var	50.43034	
Adjusted R-squared	0.993225	S.D. dependent var	107.1874	
S.E. of regression	8.822777	Akaike info criterion	7.290248	
Sum squared resid	2023.876	Schwarz criterion	7.431692	
Log likelihood	-102.7086	F-statistic	2053.354	
Durbin-Watson stat	1.306955	Prob(F-statistic)	0.000000	

En el listado anterior:

- la constante era claramente no significativa
- la variable población no era significativa al 95% de confianza (sí al 90%)

Utilizando errores estándar robustos:

- la constante no es significativa al 95% de confianza (sí al 90%)
- la variable población resulta muy significativa

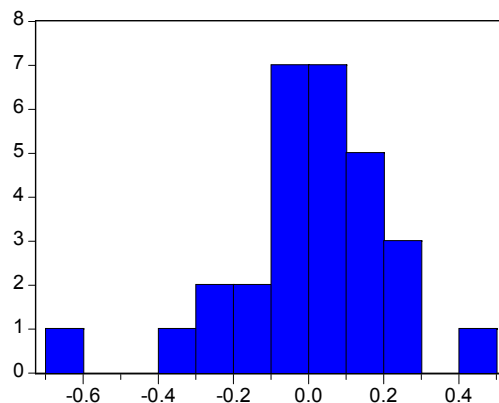
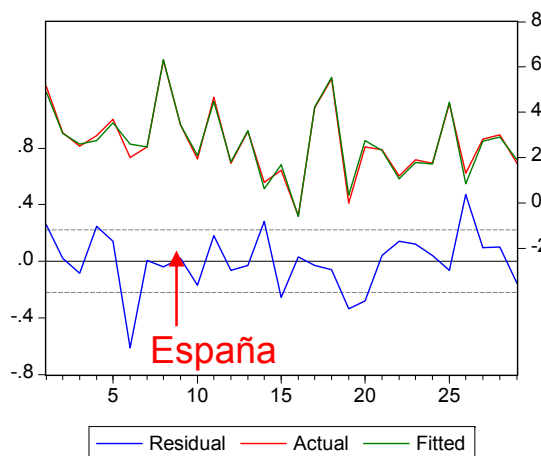
Esta solución es muy limitada ya que sólo resuelve el problema de inferencia y en este caso, en donde sólo hay 29 observaciones, mantener la eficiencia resulta importante. Para ello podemos:

- Aplicar MCG, por ejemplo, usando como ponderaciones los valores ajustados de la regresión del test de White.
- Transformar logarítmicamente las variables. Esto suele servir cuando el problema se debe a la escala de los datos.

# Ejemplo de heterocedasticidad (V)

Dependent Variable: LOG(GSAN_PUB)				
Method: Least Squares				
Date: 02/21/02 Time: 16:37				
Sample: 1 29				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.627720	0.203319	-17.84252	0.0000
LOG(GDP)	1.275227	0.057762	22.07741	0.0000
LOG(POB)	-0.290029	0.056148	-5.165409	0.0000
R-squared	0.981916	Mean dependent var	2.669273	
Adjusted R-squared	0.980525	S.D. dependent var	1.570579	
S.E. of regression	0.219177	Akaike info criterion	-0.100180	
Sum squared resid	1.248999	Schwarz criterion	0.041265	
Log likelihood	4.452605	F-statistic	705.8844	
Durbin-Watson stat	2.036269	Prob(F-statistic)	0.000000	

White Heteroskedasticity Test:				
F-statistic	0.979457	Probability	0.451281	
Obs*R-squared	5.090862	Probability	0.404892	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 03/04/03 Time: 13:31				
Sample: 1 29				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.015339	0.406544	-0.037729	0.9702
LOG(GDP)	0.098136	0.253220	0.387553	0.7019
(LOG(GDP))^2	-0.024340	0.042522	-0.572405	0.5726
(LOG(GDP))*LOG(POB)	0.041005	0.076145	0.538511	0.5954
LOG(POB)	-0.089954	0.216468	-0.415554	0.6816
(LOG(POB))^2	-0.014338	0.034613	-0.414249	0.6825



Series: Residuals	
Sample 1 29	
Observations 29	
Mean	1.91E-17
Median	0.020357
Maximum	0.473552
Minimum	-0.614297
Std. Dev.	0.211204
Skewness	-0.554134
Kurtosis	4.339692
Jarque-Bera	3.652829
Probability	0.160990

La transformación log mejora los resultados, ¿por qué?  
¿debemos gastar más en sanidad pública?