Econometría

Tema 6: Autocorrelación

2024-2025

1 Naturaleza, causas y consecuencias de la Autocorrelación 2

2 Procedimientos para la detección de la Autocorrelación 5

2.1 Métodos Gráficos 6
2.2 Contraste de Durbin-Watson 10
2.3 Contraste H de Durbin 14
2.4 Contraste de Ljung-Box 15

3 Estimación de modelos con perturbaciones autocorrelacionadas . . 20

4 Ejemplo 26

Naturaleza de la Autocorrelación

$$y = X\beta + u$$

con $\mathbb{E}[u] = 0$ y $\text{var}[u] = \mathbb{E}[uu^t] = \sigma^2 I_n$, lo que implica:

- $\mathbb{E}[u_t] = 0, \forall t \in 1, \ldots, n$
- $\mathbb{E}[u_t^2] = \sigma^2, \forall t \in 1, \ldots, n$ (Homocesdasticidad)
- $\mathbb{E}[u_i, u_j] = \text{Cov}[u_i, u_j] = 0, \forall i \neq j \in 1, \ldots, n$ (Incorrelación).

$$\text{Si Cov}[u_t, u_{t-k}] \neq 0 \ \forall k > 0 \rightarrow \textbf{Autocorrelación}.$$

Suponemos entonces que: $u_t = \rho u_{t-1} + \varepsilon_t$ donde ε_t verifica que es un ruido blanco gaussiano.

Causas de la Autocorrelación

- Error en la especificación del modelo, es decir, se han omitido en la construcción del modelo variables exógenas que son relevantes con el fin de estudiar el comportamiento de la variable endógena.
- Existencia de ciclos y tendencias. Si la variable endógena presenta ciclos, y éstos no están correctamente explicados por el conjunto de las variables exógenas, entonces el término de perturbación presentará autocorrelación.
- Existencia de relaciones no lineales.
- Existencia de relaciones dinámicas.

Consecuencias de la Autocorrelación

Los estimadores MCO son insesgados pero **no óptimos**.
El estimador por Mínimos Cuadrados Generalizados (MCG)

$$\widehat{\beta}_{MCG} = \left[X^t \Omega^{-1} X\right]^{-1} \left[X^t \Omega^{-1} y\right]$$

cuya matriz de varianzas-covarianzas viene dada por

$$\text{var}\left[\widehat{\beta}_{MCG}\right] = \sigma^2 \left[X^t \Omega^{-1} X\right]^{-1}$$

Sí es **óptimo**.

Detección de la Autocorrelación

- Gráficamente
  - Gráfico temporal de los residuos obtenidos por MCO. Si los residuos están incorrelados deben distribuirse de forma aleatoria alrededor de cero.
  - Gráfico de dispersión de los mismo frente a algún retardo suyo. Los residuos estarán autocorrelacionados cuando el gráfico presenta una tendencia creciente o decreciente.
- Analíticamente
  - Contraste de Durbin-Watson.
  - Estadístico H de Durbin.
  - Contraste de Ljung-Box

Métodos Gráficos

- Gráfico temporal de los residuos: si los residuos están incorrelados deben distribuirse de forma aleatoria alrededor de cero. Sin embargo, si están correlacionados:
  - observaremos rachas de residuos por debajo y por encima de la media (correlación positiva).
  - obervaremos una alternancia en el signo de los residuos (correlación negativa).
- Gráficos de dispersión: el gráfico de dispersión de los residuos, e_t, frente algún retardo suyo, e_{t-k} (normalmente se considera $k = 1$), puede mostrar la existencia de autocorrelación positiva (tendencia creciente en el gráfico) o negativa (tendencia decreciente).

Ejemplo

C: Consumo de Energía Eléctrica (miles de TEP)

$$C = \beta_0 + \beta_1 \text{PIB} + u$$

$$\Downarrow$$

$$\widehat{C}_t = -6234.54 + 0.0426873 \cdot PIB_t$$

$$R^2 = 0.992408$$

Año	C	PIB
1987	9427	355312
1988	9876	373412
1989	10410	391443
1990	10974	406252
1991	11372	416582
1992	11488	420462
1993	11569	416126
1994	11999	426041
1995	12462	437787
1996	12827	448457
1997	13331	466513
1998	14290	486785
1999	15364	507346
2000	16309	528714
2001	17282	543746
2002	17756	554852

Ejemplo

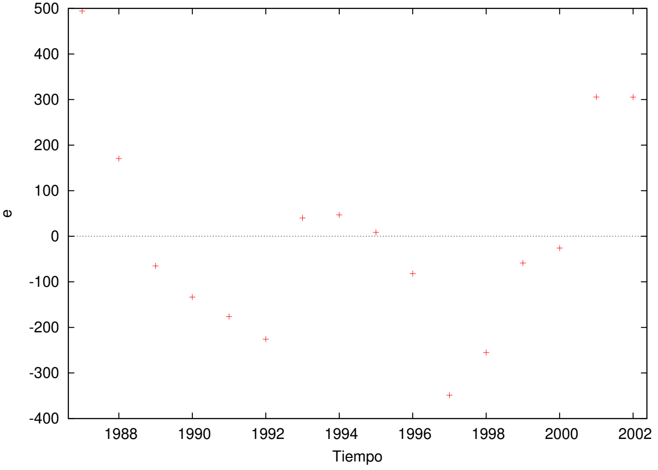


Figure. Gráfico temporal de los residuos

Ejemplo

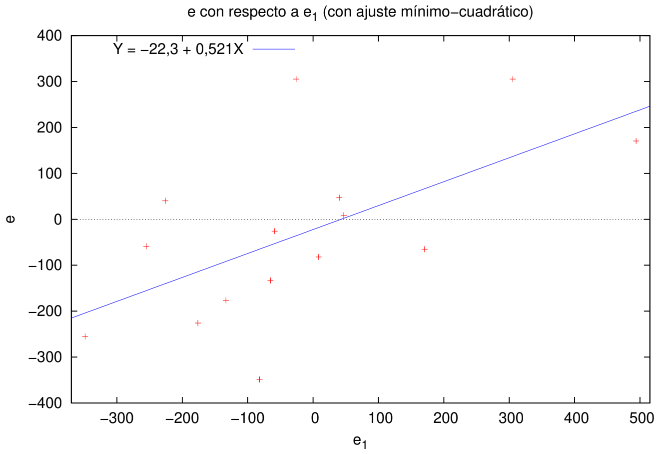


Figure. Gráfico de dispersión

Contraste de Durbin-Watson

El contraste de Durbin-Watson resulta adecuado si suponemos que la autocorrelación que presenta el término de perturbación pudiera ser descrito mediante un proceso autorregresivo de orden uno:

$$u_t = \rho u_{t-1} + \varepsilon_t$$

donde ε_t es ruido blanco.
Estudiamos si el coeficiente ρ es significativo o no mediante los contrastes:

$$\begin{cases} H_0 : \rho = 0 & \text{(Incorrelación)} \\ H_1 : \rho < 0 & \text{(Correlación negativa)} \end{cases} \qquad \begin{cases} H_0 : \rho = 0 & \text{(Incorrelación)} \\ H_1 : \rho > 0 & \text{(Correlación positiva)} \end{cases}$$

Estadístico de **Contraste de Durbin-Watson**:

$$DW = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} = \frac{\sum_{t=2}^{n}(e_t^2 - 2e_t e_{t-1} + e_{t-1}^2)}{\sum_{t=1}^{n} e_t^2}$$

siendo e_t los residuos MCO: $e_t = y_t - \widehat{y}_t = y_t - x\widehat{\beta}$.

Contraste de Durbin-Watson

Si la muestra seleccionada es suficientemente grande, se puede considerar que $\sum_{t=2}^{n} e_t^2 \simeq \sum_{t=2}^{n} e_{t-1}^2$, ya que ambas sumatorias difieren únicamente en la primera de las observaciones muestrales:

$$DW = \frac{2\sum_{t=2}^{n} e_t^2 - 2\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2} = 2\left(\frac{\sum_{t=2}^{n} e_t^2}{\sum_{t=1}^{n} e_t^2} - \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2}\right) \simeq 2(1-\widehat{\rho})$$

donde $\widehat{\rho}$ es la estimación del parámetro ρ.

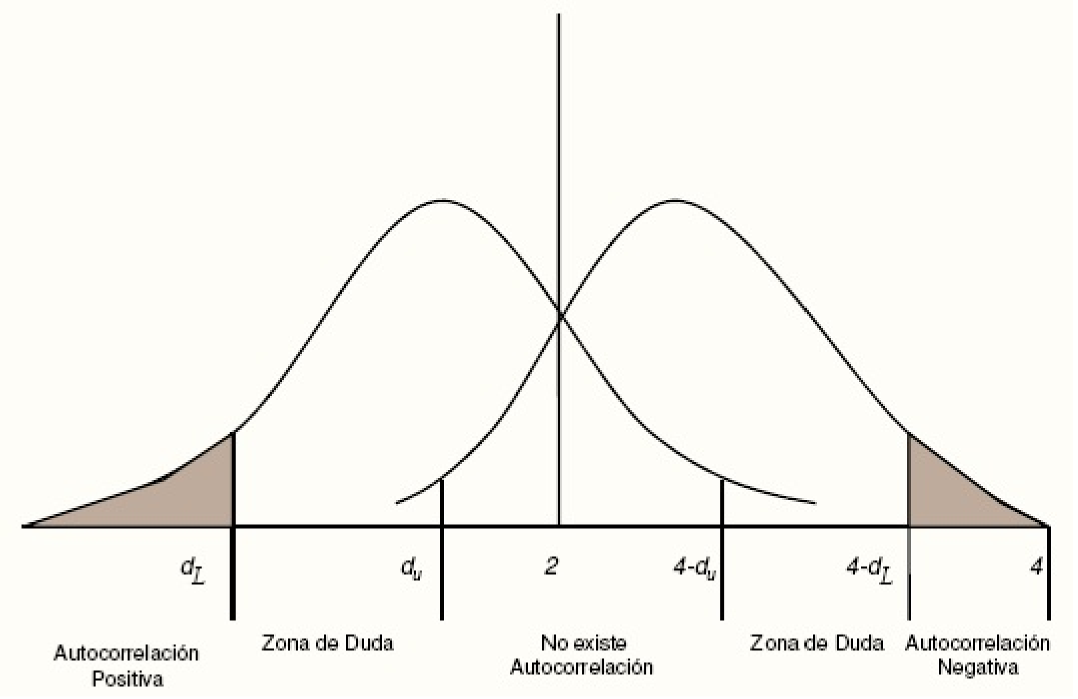
Como $|\rho| < 1$:

- $DW \simeq 0$ cuando $\widehat{\rho} \simeq 1$ (correlación positiva)
- $DW \simeq 2$ cuando $\widehat{\rho} \simeq 0$ (residuos incorrelacionados)
- $DW \simeq 4$ cuando $\widehat{\rho} \simeq -1$ (correlación es negativa)

```
from statsmodels.stats.stattools import durbin_watson
```
durbin_watson(mco.resid)

Contrastes de Durbin-Watson

Gráficamente se tiene:



Contrastes de Durbin-Watson

Las tablas estadísticas nos proporcionan una cota inferior, d_L, y otra cota superior, d_U. Si:

- $0 < DW < d_L$, se mantiene la hipótesis $\rho > 0$, autocorrelación positiva.
- $d_L < DW < d_U$, test no concluyente.
- $d_U < DW < 4 - d_U$, se mantiene que $\rho = 0$, residuos incorrelados.
- $4 - d_U < DW < 4 - d_L$, test no concluyente.
- $4 - d_L < DW < 4$, mantenemos que $\rho < 0$, autocorrelación negativa.

Estadístico de Durbin-Watson - Puntos críticos de d_L y d_u al nivel de significación del 5 %
k^* corresponde al número de regresores del modelo excluido el término independiente (es decir, $k^* = k - 1$)

	$k^*=1$		$k^*=2$		$k^*=3$		$k^*=4$		$k^*=5$		$k^*=6$	
n	d_L	d_u	d_L	d_u	d_L	d_u	d_L	d_u	d_L	d_u	d_L	d_u
6	0.610	1.400										
7	0.700	1.356	0.467	1.896								
8	0.763	1.332	0.559	1.777	0.368	2.287						
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588				
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822		
11	0.927	1.324	0.658	1.604	0.595	1.928	0.444	2.283	0.316	2.645	0.203	3.005
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.379	2.506	0.268	2.832
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390	0.328	2.692
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.472
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318

H de Durbin

El contraste de Durbin-Watson es válido cuando la autocorrelación de los errores es autorregresiva de orden 1 y cuando la regresión no incluye entre las variables explicativas algún retardo de la variable dependiente.

En este segundo caso se recurre al contraste h de Durbin. Esto es, en modelos del tipo

$$Y_t = \alpha Y_{t-1} + \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + u_t,$$

donde $u_t = \rho u_{t-1} + v_t$, para contrastar

$$\begin{aligned} H_0 &: \rho = 0 \text{ (incorrelación)} \\ H_1 &: \rho \neq 0 \text{ (correlación)} \end{aligned}$$

utilizaremos el estadístico $h = \rho \cdot \sqrt{\frac{n}{1-n \cdot \widehat{Var(\widehat{\alpha})}}} \sim N(0,1)$, de forma que se rechazará la hipótesis nula si $|h| > Z_{1-\frac{\alpha}{2}}$.

Contraste de Ljung-Box

Si los residuos son independientes, sus primeras m autocorrelaciones son cero, para cualquier valor de m. El test de Ljung-Box contrasta la hipótesis nula de que las primeras m autocorrelaciones, ρ_m, son cero. Esto es:

$$\left.\begin{array}{l} H_0 : \rho_1 = \rho_2 = \cdots = \rho_m = 0 \\ H_1 : \exists\, i \in \{1, 2, \ldots, m\} \text{ tal que } \rho_i \neq 0 \end{array}\right\}$$

Contraste de Ljung-Box

Se rechaza la hipótesis nula (de incorrelación) si $Q_{LB} = n(n+2)\sum_{s=1}^{m}\frac{r(s)^2}{n-s} > \chi^2_{m-1}(1-\alpha)$ donde

$$r(s) = \frac{\sum_{t=s+1}^{n} e_t e_{t-s}}{\sum_{t=1}^{n} e_t^2},$$

es el coeficiente de autocorrelación muestral de orden k.

Si las observaciones son independientes (incorrelación), los coeficientes $r(s)$ serán próximos a cero, por lo que no se rechazaría la hipótesis nula.

```
from statsmodels.stats.diagnostic import acorr_ljungbox
```

acorr_ljungbox(mco.resid, lags=3)

Ejemplo

Consideremos el modelo que ajusta el consumo energético en función del PIB para aplicar el anterior contraste para $m = 3$.

De este ejemplo ya sabemos que $\sum_{t=1}^{n} e_t^2 = 765312'5$. Además, a partir de los residuos podemos construir la siguiente tabla:

Año	e_t	e_{t-1}	e_{t-2}	e_{t-3}	$e_t \cdot e_{t-1}$	$e_t \cdot e_{t-2}$	$e_t \cdot e_{t-3}$
1987	494'1584						
1988	170'5190	494'1584			84263'40		
1989	-65'1749	170'5190	494'1584		-11113'56	-32206'724	
1990	-133'3305	-65'1749	170'5190	494'1584	8689'80	-22735'383	-65886'386
1991	-176'2899	-133'3305	-65'1749	170'5190	23504'82	11489'676	-30060'777
1992	-225'9164	-176'2899	-133'3305	-65'1749	39826'78	30121'546	14724'078
1993	40'1755	-225'9164	-176'2899	-133'3305	-9076'31	-7082'534	-5356'619
1994	46'9314	40'1755	-225'9164	-176'2899	1885'49	-10602'572	-8273'531
1995	8'5268	46'9314	40'1755	-225'9164	400'18	342'568	-1926'343
1996	-81'9462	8'5268	46'9314	40'1755	-698'74	-3845'849	-3292'229
1997	-348'7073	-81'9462	8'5268	46'9314	28575'23	-2973'357	-16365'321
1998	-255'0634	-348'7073	-81'9462	8'5268	88942'46	20901'476	-2174'874
1999	-58'7561	-255'0634	-348'7073	-81'9462	14986'52	20488'681	4814'839
2000	-25'8974	-58'7561	-255'0634	-348'7073	1521'63	6605'478	9030'612
2001	305'4278	-25'8974	-58'7561	-255'0634	-7909'78	-17945'746	-77903'453
2002	305'3431	305'4278	-25'8974	-58'7561	93260'26	-7907'592	-17940'769
Total					357073'5	-15350'3338	-200610'77

Ejemplo

A partir de la cual podemos obtener:

$$r(1) = \frac{\sum_{t=2}^{n} e_t \cdot e_{t-1}}{\sum_{t=1}^{n} e_t^2} = \frac{357073'5}{765312'5} = 0'4666,$$

$$r(2) = \frac{\sum_{t=3}^{n} e_t \cdot e_{t-2}}{\sum_{t=1}^{n} e_t^2} = \frac{-15350'3338}{765312'5} = -0'0201,$$

$$r(3) = \frac{\sum_{t=4}^{n} e_t \cdot e_{t-3}}{\sum_{t=1}^{n} e_t^2} = \frac{-200610'77}{765312'5} = -0'2622.$$

Ejemplo

De forma que:

$$
\begin{aligned}
Q_{LB} &= n(n+2)\sum_{s=1}^{m}\frac{r(s)^2}{n-s} = 16\cdot 18\cdot\left(\frac{r(1)^2}{15}+\frac{r(2)^2}{14}+\frac{r(3)^2}{13}\right)\\
&= 288\cdot\left(\frac{0'21771556}{15}+\frac{0'00040401}{14}+\frac{0'06874884}{13}\right)\\
&= 288\cdot(0'01452+0'000028+0'0053) = 288\cdot 0'01983\\
&= 5'7115.
\end{aligned}
$$

Como $Q_{LB} = 5'7115 \not> 5'991 = \chi^2_2(0'95)$, no rechazo la hipótesis nula de incorrelación.

Adviértase que para $m = 2$, se tiene que $Q_{LB} = 4'1885 > 3'841 = \chi^2_1(0'95)$. Por tanto, en tal caso si se rechazaría la hipótesis nula de incorrelación.

Estimación bajo Autocorrelación

$$y = X\beta + u$$

con:

- $\mathbb{E}[u] = 0$,
- $\mathbb{E}[u_i u_j] \neq 0$ ($i \neq j$ con $i, j \in \{1, \ldots, n\}$) y
- $\text{var}[u] = \sigma^2 \Omega$.

(no se verifican los supuestos de MCO)
Por el método de Mínimos Cuadrados Generalizado: necesitamos calcular P tal que $\Omega^{-1} = P^t P$.
Suponemos que u_t viene determinada por un proceso autorregresivo de orden 1, $AR(1)$: $u_t = \rho u_{t-1} + \varepsilon_t$ (ε_t un ruido blanco) y ρ el coeficiente de autocorrelación con $-1 < \rho < 1$.

Estimación bajo Autocorrelación

$$u_t \cdot u_{t-1} = (\rho u_{t-1} + \varepsilon_t) \cdot u_{t-1} = \rho\, u_{t-1}^2 + \varepsilon_t \cdot u_{t-1} \Rightarrow \mathbb{E}[u_t u_{t-1}] = \rho\sigma^2$$

De la misma forma:

$$\mathbb{E}[u_t u_{t-2}] = \mathbb{E}[(\rho u_{t-1} + \varepsilon_t)\, u_{t-2}] = \rho\mathbb{E}[u_{t-1}u_{t-2}] = \rho(\rho\sigma^2) = \rho^2\sigma^2$$
$$\mathbb{E}[u_t u_{t-3}] = \mathbb{E}[(\rho u_{t-1} + \varepsilon_t)\, u_{t-3}] = \rho\mathbb{E}[u_{t-1}u_{t-3}] = \rho(\rho^2\sigma^2) = \rho^3\sigma^2$$

$$\mathbb{E}[u_t u_{t-k}] = \rho^k\sigma^2$$

$$\Omega = \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \rho^2 & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \rho & \cdots & \rho^{n-3} \\ & & \ddots & & & \\ \rho^{n-1} & \rho^{n-2} & \cdots & \rho^2 & \rho & 1 \end{pmatrix}$$

Estimación bajo Autocorrelación: MCG

$$\Rightarrow \Omega^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho & 0 & 0 & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \cdots & 0 \\ & & \ddots & & 1+\rho^2 & -\rho \\ 0 & 0 & \cdots & 0 & -\rho & 1 \end{pmatrix}$$

Estimación bajo Autocorrelación

Suponemos;

$$y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + u_t$$

y

$$y_{t-1} = \beta_1 + \beta_2 X_{2t-1} + \cdots + \beta_k X_{kt-1} + u_{t-1}$$

Luego:

$$y_t - \rho y_{t-1} = \beta_1(1-\rho) + \beta_2(X_{2t} - \rho X_{2t-1}) + \cdots + \beta_k(X_{kt} - \rho X_{kt-1}) + u_t - \rho u_{t-1}$$

Estimación bajo Autocorrelación

Llamando $y_t^* = y_t - \rho y_{t-1}$ y $X_{it}^* = X_{it} - \rho X_{it-1}$:

$$y_t^* = \beta_1 + \beta_2 X_{1t}^* + \cdots + \beta_k X_{kt}^* + \varepsilon_t$$

donde ahora ε cumple las hipótesis para aplicar MCO.
<u>**Nota:**</u> Se perdería la primera observación, para ello Prais-Wistem proponen $y_1^* = \sqrt{1-\rho^2} y_1$, $X_{i1}^* = \sqrt{1-\rho^2} X_{i1}$...

Modificación Prais-Winsten

1. Estimando por MCO y se obtienen los residuos e_t.
2. Se estima $u_t = \rho u_{t-1} + \varepsilon_t$ el valor de ρ empleando los e_t:

$$\widehat{\rho} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{{}_{t=2}^{n} e_t^2}$$

3. Se estima por MCO $y^* = X^* \beta + u^*$: e_t^*.
4. Estimamos $\widehat{\rho}$ de la regresión $e_t^* = \widehat{\rho} e_{t-1}^* + \varepsilon_t^*$: $\widehat{\widehat{\rho}} = \frac{\sum_{t=2}^{n} e_t^* e_{t-1}^*}{{}_{t=2}^{n} (e_t^*)^2}$.
5. Repetimos... La estimación de ρ es el valor que estabiliza la secuencia $\widehat{\rho}$, $\widehat{\widehat{\rho}}$, $\widehat{\widehat{\widehat{\rho}}}$, ... (hasta que la diferencia sea $< 10^{-3}$).

Proceso iterativo de Cochrane-Orcutt: Obviamos la primera observación en las transformaciones.

```
sm.GLSAR(y, sm.add_constant(x), rho=rho).iterative_fit(maxiter
```

Ejemplo

El número de pequeños accidentes, Y, ocurridos en las calles de una ciudad y el número de coches que han sido matriculados, X, en la misma durante 10 años están recogidos en la siguiente tabla:

Y	X
25	510
27	520
28	528
32	540
33	590
36	650
38	700
40	760
41	800
45	870

Dado el modelo $Y_t = \beta_1 + \beta_2 X_t + u_t$:

- Contrastar la hipótesis de no autocorrelación por medio de Durbin-Watson.
- En caso de detectar problemas de autocorrelación, obtenga una estimación aplicando MCG.

EJEMPLO

Se estima el modelo empleando MCO.

$$\widehat{Y}_t = 2.56755 + 0.0493699 X_t \quad (R^2 = 0.942095)$$

Y	$\widehat{Y}$	e_t	e_{t-1}	$(e_t - e_{t-1})^2$	e_t^2
25	27.7462	-2.7462			7.5416
27	28.2399	-1.2399	-2.7462	2.2689	1.5374
28	28.6349	-0.6349	-1.2399	0.3661	0.4030
32	29.2273	2.7727	-0.6349	11.6115	7.6879
33	31.6958	1.3042	2.7727	2.1565	1.7010
36	34.6580	1.3420	1.3042	0.0014	1.8010
38	37.1265	0.8735	1.3420	0.2195	0.7630
40	40.0887	-0.0887	0.8735	0.9258	0.0079
41	42.0635	-1.0635	-0.0887	0.9502	1.1310
45	45.5194	-0.5194	-1.0635	0.2961	0.2697
				18.7960	22.8435

$$DW = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} = \frac{18.7960}{22.8435} = 0.8228 \underset{\alpha=5\%}{\overset{\text{k=1,n=10}}{\Rightarrow}} \begin{cases} d_L = 0.879 \\ d_U = 1.320 \end{cases} \overset{DW<dL}{\Rightarrow} \begin{cases} \text{Autocorrelación} \\ \text{Positiva} \end{cases}$$

Ejemplo

Estima la regresión $e_t = \rho e_{t-1} + \varepsilon_t$:

$$\widehat{e}_t = 0.4226 e_{t-1} \Rightarrow \widehat{\rho} = 0.4226$$

Como el n es pequeño usamos la Modificación Prais-Winsten (transformación):

$$y_t^* = \begin{cases} \sqrt{1-(0.4226)^2}\, y_t & t = 1 \\ y_t - 0.4226 y_{t-1} & t > 1 \end{cases}$$

$$X_t^* = \begin{cases} \sqrt{1-(0.4226)^2}\, X_t & t = 1 \\ X_t - 0.4226 X_{t-1} & t > 1 \end{cases}$$

$$cte_t^* = \begin{cases} \sqrt{1-(0.4226)^2}\, cte_t & t = 1 \\ cte_t - 0.4226\, cte_{t-1} & t > 1 \end{cases}$$

Ejemplo

Y^*	cte^*	X^*
22.61101004	0.904440402	461.2646048
16.335	0.5734	302.434
16.4818	0.5734	306.168
20.0552	0.5734	314.7552
19.3488	0.5734	359.636
21.9222	0.5734	398.306
22.6424	0.5734	422.71
23.7892	0.5734	461.38
23.936	0.5734	475.784
27.5094	0.5734	528.72

$Y_t^* = \beta_3 + \beta_4 X_t^* + u_t^* \Rightarrow \widehat{Y}^* = 2.1071 + 0.04975X^*.$

Nota: Como el valor de ρ ha sido estimado deberíamos realizar el proceso iterativo. Otra alternativa es calcular el estadístico de D-W y comprobar si se ha eliminado el problema de autocorrelación. En este caso $DW = 1.5305$, indicando pues que la autocorrelación ha desaparecido.