

AUTOCORRELACIÓN

Eco. Douglas C. Ramírez V.

La autocorrelación surge cuando los términos de error del modelo no son independientes entre sí, es decir, cuando: $E(u_i u_j) \neq 0$, para todo $i \neq j$. Entonces los errores estarán vinculados entre sí. Los estimadores mínimos cuadrados ordinarios (MCO) obtenidos, bajo esta circunstancia, dejan de ser eficientes. La autocorrelación generalmente aparece en datos en serie de tiempo aunque también se presenta en el caso de una muestra de corte transversal. Aquí se estudiará el primer caso.

Como primera aproximación se asume que las observaciones se generan de la siguiente manera:

$$Y_t = \alpha + \beta X_t + u_t$$
$$u_t = \rho u_{t-1} + \varepsilon_t \quad ; \quad \text{para} \quad -1 < \rho < 1$$

Este modelo expresa un comportamiento autorregresivo de primer orden de los errores y se denota como AR(1). En este caso a "rho" (ρ) se le conoce como coeficiente de autocovarianza o de autocorrelación y el error (ε) es una perturbación estocástica que satisface los supuestos MCO tradicionales. El coeficiente de autocorrelación (rho) se mueve entre los valores de menos uno y uno ($-1 < \rho < 1$). Al encontrarse entre los valores extremos se dice que existe un proceso autorregresivo de las perturbaciones o disturbios de los errores

- ▶ Si $\rho = 0$ no existe autocorrelación
- ▶ Si $\rho = 1$ existe autocorrelación positiva perfecta
- ▶ Si $\rho = -1$ existe autocorrelación negativa perfecta

En términos matriciales, este problema supone que los elementos que se encuentran fuera de la diagonal principal de la matriz de varianza-covarianza ($E[UU']$) pueden ser no ceros.

1) AUTO CORRELACIÓN CAUSAS

El problema de la autocorrelación se presenta mucho en series históricas, en las cuales la “memoria” se transmite a través de los errores. Un “shock” —grande o pequeño— se mantiene en el tiempo.

Entre las principales causas que hacen que aparezca la autocorrelación en una muestra tenemos las siguientes:

- ▶ Inercia. Cuando existen tendencias marcadas que influyen en los valores futuros de la serie.
- ▶ Sesgos de especificación. Cuando se elige mal la forma funcional o cuando se omiten variables, lo cual genera un comportamiento sistemático en el término estocástico.
- ▶ Tiempo de ajuste. Implica el tiempo que los agentes económicos deben tomar para procesar información de un período dado. Así un fenómeno sucedido en un período determinado puede impactar en uno o varios posteriores.
- ▶ Preparación de datos. En datos de corte transversal al “ordenar” los datos con respecto alguna variable (consumo ordenado de mayor a menor por la variable ingreso) puede introducir un proceso “aparentemente” autocorrelacionado.

Las consecuencias inmediatas, producto de la autocorrelación, es que los estimadores son poco eficientes, ya que sus varianzas estarán sobre o subestimada lo cual imposibilita utilizar las pruebas de contrastes —“test”— estadístico usuales para verificar la validez de las estimaciones. Pero los estimadores siguen siendo lineales, insesgados y consistentes pero han perdido

(como consecuencia de autocorrelación) su propiedad de varianza mínima, pero la insesgadez será útil para resolver el problema.

2) AUTO CORRELACIÓN PROBLEMAS

Con la presencia de autocorrelación los estimadores obtenidos con mínimos cuadrados ordinarios dejan de ser MELI, en el sentido que dejan de ser aquellos con varianza mínima, aun cuando sean insesgados.

En consecuencia se tendría como principales problemas:

- ▶ Estimadores poco eficientes
- ▶ Invalidez de las pruebas de contraste usuales

Ignorar el problema de la autocorrelación lleva a que las pruebas t y F dejen de ser válidas ya que muy probablemente arrojen conclusiones erradas. Debido a que la matriz de varianzas y covarianzas estarán erradas.

No se debe dejar de considerar ciertos aspectos al analizar el problema:

- ▶ En la practica el ρ no se conoce por lo tanto hay que estimarlo y esto genera perdidas de grados de libertad, el problema, en muestras pequeñas, es relevante si el ρ estimado es mayor a 0,3. cuando el proceso es autorregresivo de primer orden.
- ▶ Si la estructura de autorregresión de los errores es más compleja no se debe ignorar el problema ni asumir supuestos heroicos.
- ▶ En series con datos mensuales (o trimestrales) pueden surgir procesos autorregresivos de duodécimo (o cuarto) orden y no detectar por tanto el proceso si se realizan pruebas de contraste de primer orden
- ▶ Es posible confundir una mala especificación con un proceso autorregresivo.

- ▶ También es posible inducir al surgimiento de un proceso autorregresivo al ordenar, los datos de corte transversal, los valores de la variable dependiente de mayor a menor o viceversa.

3) AUTOCORRELACIÓN IDENTIFICACIÓN

Existen diversos métodos formales y gráficos para identificar la existencia o no de procesos AR(1) de los errores, entre los formales, existen pruebas para detectar autocorrelación de primer orden y de orden superior (de orden p) entre los métodos gráficos tenemos el dibujo de los errores y el correlograma. En esta sección se verán los dos métodos más populares, el de Durbin-Watson y el de graficación de residuos.

3.1. Durbin Watson

Para detectar la presencia de autocorrelación en una serie de datos la prueba más utilizada y que es calculada en, prácticamente, todos los programas econométricos, es la de Durbin Watson. Para este fin se define el estadístico de la siguiente manera:

$$d = \frac{\sum_{t=2}^{t=T} (\epsilon_t \epsilon_{t-1})^2}{\sum_{t=1}^{t=T} \epsilon_t^2}$$

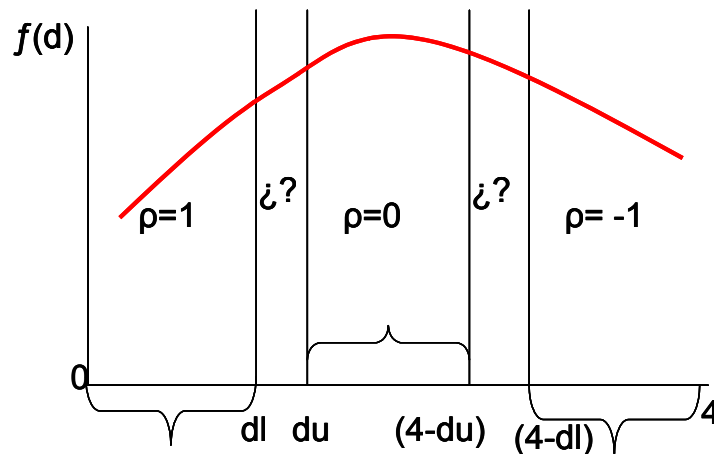
donde ϵ_t es el residuo estimado para el periodo t. Es posible escribir a d como

$$d = \frac{\sum_{t=2}^{t=T} \epsilon_t^2 + \sum_{t=2}^{t=T} \epsilon_{t-1}^2 + 2 \sum_{t=2}^{t=T} \epsilon_t \epsilon_{t-1}}{\sum_{t=1}^{t=T} \epsilon_t^2} \approx 2(1 - \rho)$$

Dado que $\sum \epsilon_t^2$ y $\sum \epsilon_{t-1}^2$ son aproximadamente iguales para muestras grandes (y dado también que $\rho = [\sum \epsilon_t \epsilon_{t-1} / \sum \epsilon_t^2]$). También se ha demostrado que el valor esperado de d, cuando $\rho=0$, esta dado por la siguiente relación (Maddala, 1996).

$$E(d) \approx 2 + \frac{2(k-1)}{n-k}$$

La distribución del muestreo de la prueba y su contraste depende del número de observaciones, del número de parámetros, de la inclusión o no del intercepto y de la incorporación de variables rezagadas en el modelo, además del nivel de significancia. Los valores del estadístico Durbin-Watson poseen un rango que va de cero a cuatro.



<u>Valor del Durbin entre</u>	<u>Valor del ρ</u>	<u>Conclusión</u>
Limites 0— d_L	1	Autocorrelación Positiva
Limites d_L — d_U	$\zeta?$	Ausencia de Evidencia
Limites d_U — $4-d_U$	0	Ausencia de Autocorrelación
Limites $4-d_U$ — $4-d_L$	$\zeta?$	Ausencia de Evidencia
Limites $4-d_L$ —4	-1	Autocorrelación Negativa

Existen tablas para probar la hipótesis de autocorrelación cero ($\rho=0$) contra la hipótesis de autocorrelación positiva ($\rho>0$), que arrojan los límites: inferior (d_L) y superior (d_U), para la autocorrelación negativa se estima por diferencia con límites dado por la tabla que son $4-d_U$ y $4-d_L$.

--> REGRESS; LHS= GNP; RHS=ONE, MONEY, INTEREST;RES=E; AR1\$

```

+-----+
| Ordinary least squares regression Weighting variable = none |
| Dep. var. = GNP Mean= 2849.225000 , S.D.= 415.7382880 |
| Model size: Observations = 20, Parameters = 3, Deg.Fr.= 17 |
| Residuals: Sum of squares= 182804.7597 , Std.Dev.= 103.69774 |
| Fit: R-squared= .944334, Adjusted R-squared = .93778 |
| Model test: F[ 2, 17] = 144.20, Prob value = .00000 |
| Diagnostic: Log-L = -119.5832, Restricted(b=0) Log-L = -148.4670 |
| LogAmemiyaPrCrt.= 9.423, Akaike Info. Crt.= 12.258 |
| Autocorrel: Durbin-Watson Statistic = .46946, Rho = .76527 |
+-----+
+-----+-----+-----+-----+-----+-----+
|Variable | Coefficient | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+-----+-----+-----+-----+-----+-----+
Constant 1996.837680 69.531125 28.719 .0000
MONEY .5529854364 .49942732E-01 11.072 .0000 1247.0300
INTEREST 22.38233188 12.359266 1.811 .0879 7.2735000

```

En la figura anexa supra, se muestra la salida del programa LIMDEP donde se muestra los estadísticos. El valor del Durbin-Watson es de 0,469 y el rho es de 0,765 para 20 observaciones y dos variables al 1% los valores de la tabla son $d_L=0,773$ y $d_U=1,411$. Lo cual evidencia autocorrelación positiva.

```

+-----+
| AR(1) Model: e(t) = rho * e(t-1) + u(t) |
| Initial value of rho = .76527 |
| Iter= 3, SS= 78478.864, Log-L=-111.610745 |
| Final value of Rho = .78766 |
| Durbin-Watson: e(t) = .42467 |
| Std. Deviation: e(t) = 110.28000 |
| Std. Deviation: u(t) = 67.94412 |
| Durbin-Watson: u(t) = 1.18290 |
| Autocorrelation: u(t) = .40855 |
| N[0,1] used for significance levels |
+-----+
+-----+-----+-----+-----+-----+-----+
|Variable | Coefficient | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+-----+-----+-----+-----+-----+-----+
Constant 1938.445824 123.77443 15.661 .0000
MONEY .5809942446 .69991030E-01 8.301 .0000 1247.0300
INTEREST 22.19952554 10.925619 2.032 .0422 7.2735000
RHO .7876635894 .14134431 5.573 .0000

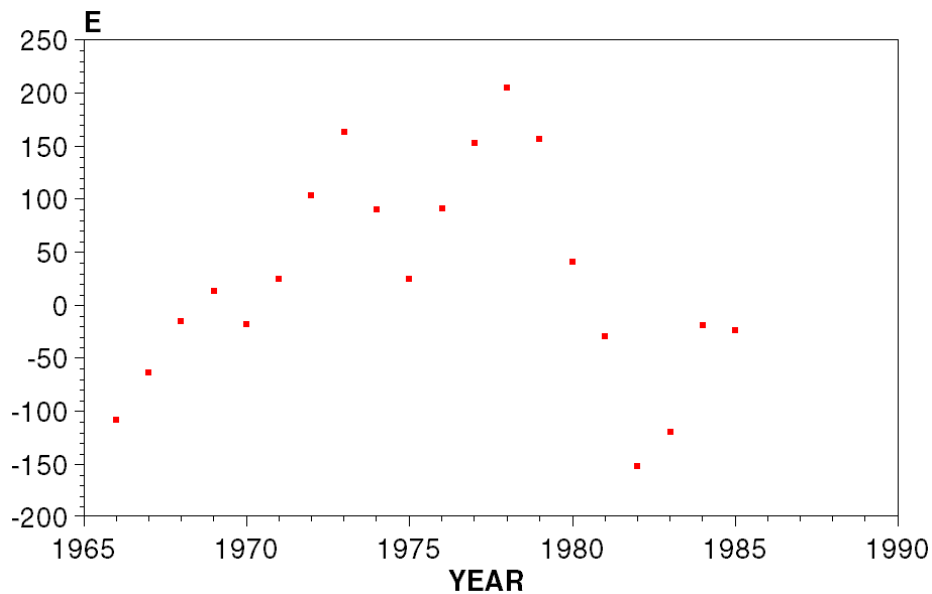
```

En la figura superior se ve la segunda parte del contraste de autocorrelación (utilizando la opción ;AR1) donde se valida la significancia del parámetro rho de autocorrelación.

3.2. Método Gráfico

Con el uso del programa se pueden construir y graficar los residuos de forma muy simple, se corre el modelo se estiman los residuos y se construyen los residuos rezagados para graficar. Las instrucciones correspondientes se listan a continuación.

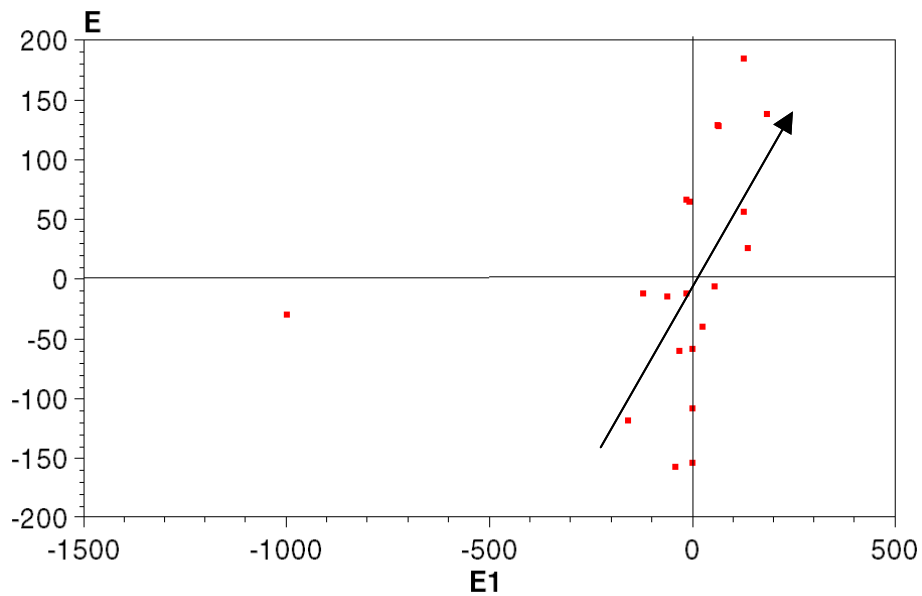
```
NAMELIST      ; X = ONE, MONEY, INTEREST; Y= GNP$
CREATE        ; E1 = 0 ; .E2 = 0 ; E3 = 0 $
REGRESS       ; Lhs = Y ; Rhs = X ; Res = E $
PERIOD        ; 1969 - 1985 $
CREATE        ; E1 = E[-1] ; E2 = E[-2]; E3 = E[-3] $
SAMPLE        ; All $
PLOT; RHS= E; LHS=YEAR$
```



Es fácil apreciar la existencia de un comportamiento sistemático de los errores graficados con respecto a los años, se puede observar, si unimos los puntos, como sube y bajan —parecido a una montaña rusa— los residuos con montañas y valles, lo cual hace pensar en un modelo o comportamiento subyacente que no ha sido capturado en la estimación.

Otro gráfico usual para detectar la presencia de procesos autorregresivos en los errores es graficar los errores contra ellos mismos pero rezagado, es decir se grafica el error ϵ_t versus ϵ_{t-1} y centrado en cero se ve si crecen o decrecen monótonamente. La instrucción que a continuación se lista muestra la existencia de un proceso autorregresivo positivo

PLOT; RHS= E; LHS=E1\$



Se puede observar claramente como los valores siguen una ruta que va desde el cuadro inferior izquierdo al cuadro superior derecho marcando un claro ascenso en diagonal. Este comportamiento es típico de un proceso autorregresivo positivo ($\rho=1$). Si el comportamiento hubiera sido desde el cuadro superior izquierdo al cuadro inferior derecho la autocorrelación sería ($\rho= -1$) negativa.

4) OTRAS PRUEBAS DE AUTOCORRELACIÓN

Existen otras pruebas para detectar autocorrelación para detectar autocorrelación de primer orden alternativas a las pruebas de Durbin-Watson cuando existen variables rezagadas o pruebas para muestras grandes.

4.1. Otras pruebas

Existen otras pruebas, para detectar autocorrelación de primer orden, que no son muy usuales como:

- La razón de Von Neumann.
- La prueba de Berenblut & Webb.

4.1.1. La razón de Von Neumann (RVN)

La razón de Von Neumann se define como:

$$RVN = \frac{\delta^2}{s^2} = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=2}^T (e_t - \bar{e})^2} \cdot (T-1)$$

Donde e representa los residuos.

La razón de Von Neumann sólo se puede utilizar con residuos independientes y homocedásticos condición que no se cumple con los residuos MCO. Para ello se sugiere construir la prueba de contraste con los “residuos recursivos¹” para muestras pequeñas. Estos residuos satisfacen las propiedades de independencia y varianza común y son muy utilizados para pruebas de diagnósticos, como las pruebas (Cusum y Cusum cuadrado) de estabilidad.

¹ Maddala, G. S. 1996. Introducción a la Econometría. PRENTICE-HALL HISPANOAMERICANA, S.A. Méjico.

Para muestras grandes , la RVN puede considerarse como normalmente distribuido con media y varianzas dadas por:

$$E\left(\frac{\delta^2}{s^2}\right) = \frac{2n}{n-1}$$

$$V\left(\frac{\delta^2}{s^2}\right) = \frac{4n^2(n-2)}{(n-1)(n-1)^3}$$

4.1.2. La prueba de Berenblut-Webb

La prueba de BW se basa en el estadístico:

$$g = \frac{\sum_{t=2}^T \hat{e}_t^2}{\sum_{t=2}^T \hat{u}_t^2}$$

Donde el residuo \hat{e} se obtienen de una regresión en primeras diferencias de la variable dependiente (Y) sobre las primeras diferencias de los regresores (X's), sin un termino constante, en caso de incluirlo es posible utilizar las tablas de DW con el estadístico g , y \hat{u} representa los errores de la regresión en niveles

4.2. Cuando el DW no es adecuado

En general la prueba DW se incluye en casi todos los programas y sus tablas aparecen en la mayoría de los textos de econometría y se construye con los residuos MCO. Las tablas son construidas para la autocorrelación positiva y en términos prácticos se prefiere utilizar el límite superior de la tabla (d_L) como punto de significancia verdadero (se incluye la región no concluyente como parte de la región de rechazo).

Si se quiere mayor precisión existen otras pruebas iterativas —que se verán más adelante— y pruebas que consideran hipótesis asintóticas sobre el comportamiento del rho y cuando el modelo incluye variables rezagadas que invalidan la prueba DW.

4.2.1. Prueba h de Durbin

Considere el siguiente modelo

$$Y_t = \alpha Y_{t-1} + \beta X_t + u_t$$

$$u_t = \rho u_{t-1} + \varepsilon_t \quad ; \quad \text{para} \quad -1 < \rho < 1$$

Donde ε_t son residuos $\sim N(0, \sigma^2)$.

Nótese que tanto u_t , como Y_{t-1} dependen de u_{t-1} , las dos variables estarán correlacionadas. El estimador MCO de α será inconsistente y si ρ es positivo, el parámetro será sesgado hacia arriba y el estimador de rho será sesgado hacia abajo, de forma tal que el DW estimado estará sesgado hacia 2 y no encontrará evidencia de autocorrelación serial.

Dado que la prueba DW no es aplicable a los modelos con rezago se sugiere una prueba alternativa llamada, prueba h de Durbin y esta se define como:

$$h = \hat{\rho} \sqrt{\frac{T}{1 - T * S_{\beta}^2}}$$

Esta prueba se contrasta como una variable normal estándar. Donde $\hat{\rho}$ es estimado por los residuos MCO y S_{β}^2 es la varianza del parámetro estimado y T es el tamaño de la muestra

4.2.2. Prueba Asintótica²

Esta es otra prueba de contraste que puede interesar porque no posee regiones inconclusa y aplicable a muestras grandes finitas bajo la hipótesis nula;

$$H_0: \rho=0$$

Versus la hipótesis en una u otra dirección —aun cuando la hipótesis corriente es de autocorrelación positiva—, es decir;

$$H_0: \rho \neq 0$$

Dado que asintóticamente $\sqrt{n} \hat{\rho} \sim N(0,1)$ por lo cual se utiliza las tablas de la normal estandarizada para la prueba de contraste.

La varianza estimada se obtiene por.

$$Var(\hat{\rho}) = \frac{\sqrt{(1 - \hat{\rho}^2)}}{T}$$

La normal estandarizada calculada (Z_c) se obtiene por.

$$Z_c = \frac{\hat{\rho} - \rho}{\sqrt{Var(\hat{\rho})}} \sim Z(0,1)$$

En las muestras pequeñas (<30) se puede utilizar esta prueba como aproximación.

5) PRUEBA DE AUTOCORRELACIÓN DE ORDEN P

Para probar la hipótesis de autocorrelación en los residuos el programa provee, inicialmente, el estadístico Durbin-Watson y el coeficiente de autocorrelación (ρ) para probar la ausencia o presencia de autocorrelación de primer orden en muestras pequeñas. Godfrey (1978, 1988)³ ha desarrollado una prueba más general para el caso de autocorrelaciones de orden superior para procesos

² Kamenta, Jan. 1977. Elementos de Econometría. Editorial Vicens-Vives Barcelona.

³ Godfrey, L.G., 1978. «Testing Against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables.» *Econometrica*: 46.

Godfrey, L.G., 1988. *Misspecification Tests in Econometrics*. Cambridge: Cambridge University Press.

autorregresivos y de promedios móviles (o una mezcla de ambos a la vez). Las formalidades de los procedimientos pueden ser vistas en sus artículos y textos.

3.2. El contraste de Godfrey-Breusch

El contraste de Godfrey (1978) & Breusch⁴ (1978) es un contraste de multiplicadores de Lagrange cuya hipótesis de trabajo son:

- ▶ Ho: ausencia de autocorrelación
- ▶ Ha: $\varepsilon_t = \text{AR}(p)$ o $\varepsilon_t = \text{MA}(p)$

Donde los errores (ε_t) siguen: un $\text{AR}(p)$, proceso autorregresivo del orden p ; o un $\text{MA}(p)$, proceso de promedios móviles de orden p .

Nótese que el contraste se puede realizar para cualquier estructura. La prueba se realiza regresando los residuos mínimos cuadráticos ε_t , versus las X 's , ε_{t-1} , ε_{t-2} ... ε_{t-p} , se obtiene el coeficiente de determinación (R^2) el cual se multiplica por el numero de observaciones (T) este estadístico ($T \cdot R^2$) sigue una distribución Chi cuadrado (χ^2_p) con p grados de libertad —que es igual al numero de rezagos—

Dado que $\mathbf{X}'\boldsymbol{\varepsilon} = \mathbf{0}$, la prueba equivale a obtener la parte no explicada de los residuos por X . Si se encuentra cualquier ajuste esta es debida a la correlación existente entre el residuo actual y los retardos. La prueba es un contraste conjunto de las primeras p autocorrelaciones de ε_t y no sólo de la primera.

En términos prácticos, la prueba puede ser obtenida a través de los residuos MCO de la regresión original para el conjunto de regresores y para los p rezagos de los valores residuales (ver Greene, 1999⁵ Cáp. 13 Sec. 13.5.2., y

⁴ Breusch, T. 1978. «Testing for Autocorrelation in Dynamic Linear Model.» *Australian Economic Papers*, 17

⁵ Greene, W.H., 1999. *Análisis Econométrico*. PRENTICE HALL IBERIA. Madrid. Tercera Edición

ejemplo 13.6). Obtener los residuos con LIMDEP, es trivial, sólo requiere añadir al comando *REGRESS*, simplemente *;RES=<nombre>*.

Al usar el comando *REGRESS* se estima la regresión por MCO, incluyendo las variables independientes o regresores originales y los residuos rezagados en Rhs y en el lado dependiente o Lhs se incluye el residuo estimado sin rezagar. Luego se obtiene el estadístico LM o contraste de Multiplicador de Lagrange

$$\chi^2_c = T \cdot R^2 \sim \chi^2_p.$$

El criterio de decisión es que si el $\chi^2_c > \chi^2_p$. Se rechaza la hipótesis nula de ausencia de autocorrelación. La principal crítica a esta prueba consiste en la dificultad de elegir el número de rezagos óptimos al igual que las pruebas de Box & Pierce (1970)⁶ y la Ljung & Box (1979)⁷

Con el comando *CALC* se puede obtener el valor de tablas de Chi cuadrado para un nivel de significancia prefijado (de un 95 % a un 99% de confianza). El conjunto de comandos que pueden servir de ejemplo se lista a continuación.

```
/*Prueba de orden p de Multiplicadores de Lagrange*/
NAMELIST      ; X = ONE, MONEY, INTEREST; Y= GNP$
CREATE        ; E1 = 0 ; .E2 = 0 ; E3 = 0 $
REGRESS       ; Lhs = Y ; Rhs = X ; Res = E $
PERIOD        ; 1969 - 1985 $
CREATE        ; E1 = E[-1] ; E2 = E[-2]; E3 = E[-3] $
SAMPLE        ; All $
REGRESS       ; Lhs = E ; Rhs = X,E1,E2,E3 $
CALC          ; LIST; LMTEST = N * RSQRD
              ; Chi(LMTEST, (KREG-1)) $
CALC          ; list; Ctb(.99, (3)) $
```

⁶ Box, G., & D. Pierce, 1970. «Distribution of Residual Autocorrelations in Autoregressive Moving Average Time Series Models.» *Journal of American Statistical Association*, 65

⁷ Ljung, G. and G. Box, 1979. «On a Measure of Lack of Fit in Time Series Models.» *Biometrika*, 66, 265–270.

```

Limdep - [Output *]
File Edit Insert Project Model Run Tools Window Help
-----
Status Trace
Matrix LastOutp
[3,4]
--> PERIOD ; 1969 - 1985 $
--> CREATE ; E1 = E[-1] ; E2 = E[-2]; E3 = E[-3] $
--> SAMPLE ; All $
--> REGRESS ; Lhs = E ; Rhs = X,E1,E2,E3 $

-----+-----+-----+-----+-----+-----+-----+
| Ordinary least squares regression Weighting variable = none
| Dep. var. = E Mean= .2273736754E-13, S.D.= 98.08824161
| Model size: Observations = 20, Parameters = 6, Deg.Fr.= 14
| Residuals: Sum of squares= 103817.9451 , Std.Dev.= 86.11369
| Fit: R-squared= .432083, Adjusted R-squared = .22926
| Model test: F[ 5, 14] = 2.13, Prob value = .12174
| Diagnostic: Log-L = -113.9254, Restricted(b=0) Log-L = -119.5832
| LogAmemiyaPrCrt.= 9.174, Akaike Info. Crt.= 11.993
| Autocorrel: Durbin-Watson Statistic = 1.10486, Rho = .44757
|-----+-----+-----+-----+-----+-----+-----+
|Variable | Coefficient | Standard Error |t-ratio |P[|T|>t] | Mean of X|
|-----+-----+-----+-----+-----+-----+-----+
| Constant | 154.7261588 | 80.975608 | 1.911 | .0767 |
| MONEY | -.4328916224E-01 | .45186837E-01 | -.958 | .3543 | 1247.0300
| INTEREST | -8.529829514 | 10.662809 | -.800 | .4371 | 7.2735000
| E1 | .1953184487E-01 | .10860542 | .180 | .8599 | -33.109073
| E2 | .1746338747 | .60283973E-01 | 2.897 | .0117 | -232.24318
| E3 | -.1980647375E-01 | .69005793E-01 | -.287 | .7783 | -126.35508

Matrix LastOutp
[6,4]
--> CALC ; LIST; LMTEST = N * RSQRD
; Chi(LMTEST,(KREG-1)) $
LMTEST = .86416584225318230D+01
Result = .87576046167934460D+00
--> CALC ;list:Ctb(.99,(3))$
Result = .11344866676609990D+02
-----
Ready Ln 440/440 Idle 02:12

```

En la figura el Chi calculado es 8.64 y Chi de tabla con 99% de confianza y con 3 grados de libertad ($p=3$) es de 11.34 y se rechaza la hipótesis nula de ausencia de autocorrelación, si se disminuye a un 95% el nivel de confianza no se podría rechazar la hipótesis nula.

La prueba de multiplicadores de Lagrange requiere que los valores iniciales de los residuos rezagados sean fijados en cero ya que de otro modo el programa los inicializara en valores extraviados con el número de código -999.

5.2. Las pruebas de Box-Pierce & Ljung-Box

Finalmente, el estadístico de Box-Pierce Q-estadístico o contraste-Q puede ser usado para los mismos propósitos de probar la existencia de autocorrelaciones de orden superior y se lleva a cabo comparando

$$Q = T \sum_{j=1}^L r_j^2 \sim \chi^2_L, \text{ donde } \rightarrow r_j = \frac{\sum_{t=j+1}^T \varepsilon_t \varepsilon_{t-j}}{\sum_{t=1}^T \varepsilon_t^2}$$

Bajo la hipótesis nula de ausencia de autocorrelación la Q se contrasta con los valores críticos de la tabla de la Chi-cuadrado con L grados de libertad. Un refinamiento es sugerido por Ljung & Box (1979) el cual es:

$$Q_{LB} = T[T + 2] \sum_{j=1}^L \frac{r_j^2}{T - j} \sim \chi^2_L.$$

Estos contrastes se han revelado razonablemente potentes. Pero sufre la misma crítica que el contraste de Breusch-Godfrey, cuál es el criterio de elección del número de rezagos (L) adecuados. El problema de determinar la longitud de los rezagos a considerar, en estas pruebas de autocorrelación, no ha concluido aun.

En el programa para obtener estos últimos estadísticos se utiliza la estimación por MCO, se guardan los residuos de la regresión —como se describió anteriormente—. Luego se usa el comando *IDENTIFY* con el cual se obtiene el correlograma con los estadísticos asociados para cada rezago estimado, esto se realiza de la siguiente manera.

```
IDENTIFY ; Rhs = residuos ; Pds = L $
```



```

--> IDENTIFY ; Rhs = E ; Pds = 5$
Time series identification for E
Box-Pierce Statistic = 13.0280      Box-Ljung Statistic = 15.4176
Degrees of freedom   = 5            Degrees of freedom   = 5
Significance level   = .0231        Significance level   = .0087
* => |coefficient| > 2/sqrt(N) or > 95% significant.
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Lag | Autocorrelation Function |Box/Prc| Partial Autocorrelations X
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
  1 | .720* |          | 10.37* | .720* |          |          |          X
  2 | .307 |          | 12.26* | -.447 |          |          |          X
  3 | .010 |          | 12.26* | .096 |          |          |          X
  4 | -.121 |          | 12.55* | -.027 |          |          |          X
  5 | -.155 |          | 13.03* | -.167 |          |          |          X
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

```

Matrix: LastIden [5,2]

En la figura anexa el estadístico Box-Pierce es 13.028 y el Ljung-Box es de 15.417 los grados de libertad son (L=5) 5 para ambos. Aquí se rechaza la hipótesis nula de ausencia de autocorrelación, los residuos pueden seguir un proceso autorregresivo AR(1) a un AR(5). Aun cuando la evidencia apunta a un AR1.

5.3. Correlograma

El comando “*Identify*” despliega el grafico de la función de autocorrelación y la función de autocorrelación parcial, esta se especifica para un orden de rezago. Esta función permite caracterizar el patrón temporal de dependencia de la serie de temporal como si fuera una serie univariante y permite detectar autocorrelaciones de orden superior a uno.

En el correlograma, en la primera columna se reporta el estadístico Q de Box-Pierce y en la segunda columna se reporta el estadístico Ljung-Box Q-statistics para L valores de rezagos, estas prueba se basa en la hipótesis nula de ausencia de autocorrelación de orden L

El estadístico Ljung-Box Q-statistics y el estadístico Q-Box-Pierce están asintóticamente distribuidos como una Chi con grados de libertad igual al número de auto correlaciones.

La línea punteada en el grafico de autocorrelación parcial son construidas para dos desviaciones estándar y son calculadas como la raíz cuadrada del número de observaciones $[\pm \sqrt{T}]$ y están construidas para un nivel de 5% de error tipo I (nivel de significancia) o nivel de confianza del 95% (aproximadamente).

6) PROCEDIMIENTOS ITERATIVOS.

Consideremos el siguiente modelo matricial:

$$Y = X\beta + U; \text{ con } E(U) = 0 \text{ y } E(UU') = \sigma^2 \Omega$$

Donde Y es un vector de $T \times 1$; X es una matriz de $T \times K$, β es un vector de $K \times 1$ y U es un vector de $T \times 1$. y $\sigma^2 \Omega$ es una matriz de varianza –covarianza desconocida.

La estimación por Mínimos Cuadrados Generalizados (MCG) es posible realizarse si se encuentra una matriz de transformación W de parámetros conocidos tal que $w'w = \Omega$. Luego estimar el modelo transformado por MCO tal que se puedan obtener estimadores MELI del modelo:

$$w Y = w X \beta + w U$$

6.1. Procesos Alternativos

Suponga que se sigue un proceso AR1 como. La matriz Ω para este modelo esta dada por:

$$E(UU') = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \dots & 1 \end{bmatrix}$$

Su inversa por.

$$\Omega^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho & 0 & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Considérese la matriz de transformación de orden T-1 x T:

$$\omega = \begin{bmatrix} -\rho & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}_{T-1,T}$$

Aplicando esta matriz de transformación el modelo transformado será

$$\begin{bmatrix} Y_2 - \rho Y_1 \\ Y_3 - \rho Y_2 \\ \vdots \\ Y_T - \rho Y_{T-1} \end{bmatrix} = \begin{bmatrix} 1 & X_2 - \rho X_1 \\ 1 & X_3 - \rho X_2 \\ \dots & \dots \\ 1 & X_T - \rho X_{T-1} \end{bmatrix} \begin{bmatrix} \alpha(1-\rho) \\ \beta \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{T-1} \end{bmatrix}$$

El proceso iterativo de Cochrane-Orcutt utiliza este tipo de transformación para estimar el valor del rho el cual sería utilizado para la corrección del modelo.

Otra alternativa de transformación a utilizar es la matriz de transformación de orden TxT, como:

$$\omega = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \dots & 0 & 0 \\ -\rho & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}_{T,T}$$

El modelo transformado a estimar sería:

$$\begin{bmatrix} \sqrt{1-\rho^2} Y_1 \\ Y_2 - \rho Y_1 \\ \vdots \\ Y_T - \rho Y_{T-1} \end{bmatrix} = \begin{bmatrix} \sqrt{1-\rho^2} & \sqrt{1-\rho^2} X_1 \\ 1 & X_2 - \rho X_1 \\ \dots & \dots \\ 1 & X_T - \rho X_{T-1} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \sqrt{1-\rho^2} u_1 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_{T-1} \end{bmatrix}$$

Nótese que el factor $\sqrt{1-\rho^2}$ se necesita para que las perturbaciones transformadas sean homoscedásticas ya que; $\sigma_u^2 = \sigma_\varepsilon^2 (1-\rho^2) I_T$.

La diferencia entre una y otra matriz de transformación es que la segunda da un tratamiento especial al primer termino. Este proceso se conoce como método Prais-Winsten. Con dicha modificación es posible mejorar la eficiencia de la estimación, en especial para muestras pequeñas.

Por último existe un tercer procedimiento por Máximo Verosimilitud (MV) con información completa para perturbaciones AR1. El modelo considerado es:

$$Y = X\beta + U;$$

con: $u_t = \rho u_{t-1} + \varepsilon_t$; $E(\varepsilon_t) = 0$; $E(\varepsilon\varepsilon') = \sigma_\varepsilon^2 I_T$

La función de verosimilitud es:

$$L(\varepsilon) = \frac{1}{(2\pi\sigma_\varepsilon^2)^{T/2}} \exp\left[-\frac{1}{2\sigma_\varepsilon^2}(\varepsilon'\varepsilon)\right]$$

Haciendo uso de la matriz de transformación ω definida se tiene que: $\omega U = \varepsilon$; en donde U y ε son vectores de $T \times 1$. Un cambio de variable en la función de verosimilitud da:

$$L(U) = L(\varepsilon) \left| \frac{\partial \varepsilon}{\partial U} \right|,$$

Donde $|\partial \varepsilon / \partial U|$ es el valor del determinante de la matriz de derivadas parciales de los ε con respecto a U . Luego entonces se tiene que:

$$\left| \frac{\partial \varepsilon}{\partial U} \right| = |\det T| = \sqrt{1 - \rho^2}$$

por lo tanto:

$$\ln L(U) \propto -\frac{T}{2} \ln \sigma_\varepsilon^2 + (1 - \rho^2) - \frac{1}{2\sigma_\varepsilon^2} (\varepsilon' \varepsilon)$$

Donde \propto significa que es "proporcional a".

Las discusiones sobre los diversos procedimientos iterativos en Johnston⁸ (1989) muestran que el método iterativo de la transformación Prais-Winsten con ρ estimada a partir de los residuos MCO resultan ser el mejor de los métodos realizables y es tan eficiente como el MV con información completa. Para muestras grandes (>60) no existe gran diferencia entre los métodos utilizados ya que producen resultados similares.

6.2. Corrección

Cuando es conocida la estructura se puede estimar el coeficiente de autocorrelación serial (ρ) por algunos de los procedimientos referidos anteriormente. Si sigue un proceso autorregresivo de primer orden. Las variables se transforman y se corre la regresión para el conjunto de datos transformados. Para ejemplificar el proceso se muestra el proceso con un modelo de dos variables.

$$(1) \quad Y_t = \alpha + \beta X_t + u_t$$

Si multiplicamos (1) por $(1 - \rho L)$ donde L es un operador de rezagos, tal que si se multiplica una variable en el tiempo la rezaga en un periodo es decir: $LY_t = Y_{t-1}$ y si el operador de rezagos es; $L^2 Y_t = Y_{t-2}$ y si $L^k Y_t = Y_{t-k}$. Por tanto:

$$(2) \quad (1 - \rho L) Y_t = (1 - \rho L) [\alpha + \beta X_t + u_t]$$

⁸ Johnston, J. 1989. *Métodos Econometría*. Editorial Vicens-Vives. Barcelona

$$(3) \quad (Y_t - \rho Y_{t-1}) = [\alpha(1-\rho) + \beta (X_t - \rho X_{t-1}) + (u_t - \rho u_{t-1})]$$

$$(4) \quad Y_t^* = \alpha^* + \beta X_t^* + \varepsilon_t$$

Donde $\alpha^* = \alpha(1-\rho)$, $Y_t^* = (Y_t - \rho Y_{t-1})$, $X_t^* = (X_t - \rho X_{t-1})$ y $\varepsilon_t = (u_t - \rho u_{t-1})$. Nótese que con la transformación $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$

Puesto que ε_{t-1} cumple los supuestos de normalidad, al aplicar MCO en las variables transformadas se obtienen los estimadores y las estimaciones óptimas. La regresión realizada por la ecuación (4) es una estimación por mínimos cuadrados generalizados (MCG). La primera observación de las variables: dependiente (Y) y de los regresores (X's) se le aplica la transformación de Prais-Winsten de la siguiente forma $Y_1 \sqrt{1-\rho^2}$ y $X_1 \sqrt{1-\rho^2}$.

Las instrucciones que realizan dicho proceso en el programa son las siguientes:

```
REGRESS; LHS= GNPR; RHS=ONE, MONEYSR, INR;AR1; RES=E$
CREATE; GNPR=GNP-RHO*GNP[-1]; MONEYSR=MONEY -RHO*MONEY[-1]
; INR=INTEREST-RHO*INTEREST[-1]$
REGRESS; LHS= GNPR; RHS=ONE, MONEYSR, INR;AR1; RES=E1$
```

Nótese que la base del proceso de estimación iterativo consiste en:

- (a) Estimar el modelo inicial por MCO. Que es la primera salida que se tiene a continuación, donde se obtiene una estimación inicial del ρ , valor que será utilizado en el segundo paso.
- (b) Con el ρ estimado se inicia un proceso de búsqueda hasta que el modelo converja, el tiempo y rapidez y exactitud depende de los diversos algoritmos.
- (c) Luego el ρ estimado se utiliza para transformar las variables.
- (d) Se prueba de nuevo la estimación con los datos transformados. Si no se ha resuelto el problema se vuelve a repetir el proceso.

La primera salida se tiene a continuación donde el Durbin-Watson muestra claras evidencias de autocorrelación.

```

+-----+
| Ordinary least squares regression Weighting variable = none
| Dep. var. = GNP Mean= 2849.225000 , S.D.= 415.7382880
| Model size: Observations = 20, Parameters = 3, Deg.Fr.= 17
| Residuals: Sum of squares= 182804.7597 , Std.Dev.= 103.69774
| Fit: R-squared= .944334, Adjusted R-squared = .93778
| Model test: F[ 2, 17] = 144.20, Prob value = .00000
| Diagnostic: Log-L = -119.5832, Restricted(b=0) Log-L = -148.4670
| LogAmemiyaPrCrt.= 9.423, Akaike Info. Crt.= 12.258
| Autocorrel: Durbin-Watson Statistic = .46946, Rho = .76527
+-----+

```

Variable	Coefficient	Standard Error	t-ratio	P[T >t]	Mean of X
Constant	1996.837680	69.531125	28.719	.0000	
MONEY	.5529854364	.49942732E-01	11.072	.0000	1247.0300
INTEREST	22.38233188	12.359266	1.811	.0879	7.2735000

El ρ estimado

La salida siguiente con el proceso Prais-Winsten parte del ρ inicial cuyo valor es de 0.76527 y finaliza luego de tres iteraciones (Iter=3) con un valor del ρ estimado de 0.78766.

```

+-----+
| AR(1) Model: e(t) = rho * e(t-1) + u(t)
| Initial value of rho = .76527
| Iter= 3, SS= 78478.864, Log-L=-111.610745
| Final value of Rho = .78766
| Durbin-Watson: e(t) = .42467
| Std. Deviation: e(t) = 110.28000
| Std. Deviation: u(t) = 67.94412
| Durbin-Watson: u(t) = 1.18290
| Autocorrelation: u(t) = .40855
| N[0,1] used for significance levels
+-----+

```

El ρ estimado inicial

El ρ estimado final

El t calculado del ρ

Variable	Coefficient	Standard Error	b/St.Er.	P[Z >z]	Mean of X
Constant	1938.445824	123.77443	15.661	.0000	
MONEY	.5809942446	.69991030E-01	8.301	.0000	1247.0300
INTEREST	22.19952554	10.925619	2.032	.0422	7.2735000
RHO	.7876635894	.14134431	5.573	.0000	

Además la prueba de contraste muestra que el ρ estimado es significativo en la prueba t-Student calculado se tiene para ρ un valor de 5.573

El resultado que se muestra en la figura anexa es el modelo final corregido del problema de autocorrelación

```

+-----+
| Ordinary least squares regression      Weighting variable = none |
| Dep. var. = GNPR      Mean= 585.3650014 , S.D.= 388.5751178 |
| Model size: Observations = 20, Parameters = 3, Deg.Fr.= 17 |
| Residuals: Sum of squares= 78385.14199 , Std.Dev.= 67.90354 |
| Fit: R-squared= .972677, Adjusted R-squared = .96946 |
| Model test: F[ 2, 17] = 302.59, Prob value = .00000 |
| Diagnostic: Log-L = -111.1153, Restricted(b=0) Log-L = -147.1156 |
|              LogAmemiyaPrCrt.= 8.576, Akaike Info. Crt.= 11.412 |
| Autocorrel: Durbin-Watson Statistic = 1.65619, Rho = .17191 |
+-----+
+-----+-----+-----+-----+-----+-----+
| Variable | Coefficient | Standard Error | t-ratio | P[|T|>t] | Mean of X |
+-----+-----+-----+-----+-----+-----+
| Constant | 488.1639391 | 35.231626 | 13.856 | .0000 | |
| MONEYR | .4973837878 | .88448104E-01 | 5.623 | .0000 | 291.80259 |
| INR | .9919115204 | .13881475 | 7.146 | .0000 | -48.327714 |
+-----+-----+-----+-----+-----+-----+

```

En la prueba de contraste por el método de iteración de Prais-Winsten se tiene que el ρ de la ecuación con las variables transformadas ya nos es significativo estadísticamente.

```

+-----+-----+-----+-----+-----+-----+
| AR(1) Model: e(t) = rho * e(t-1) + u(t) |
| Initial value of rho = .17191 |
| Iter= 2, SS= 76059.576, Log-L=-110.829569 |
| Final value of Rho = .17418 |
| Durbin-Watson: e(t) = 1.65163 |
| Std. Deviation: e(t) = 67.92704 |
| Std. Deviation: u(t) = 66.88866 |
| Durbin-Watson: u(t) = 1.86620 |
| Autocorrelation: u(t) = .06690 |
| N[0,1] used for significance levels |
+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+-----+-----+
| Variable | Coefficient | Standard Error | |b/St.Er.| P[|Z|>z] | Mean of X |
+-----+-----+-----+-----+-----+-----+
| Constant | 487.1090280 | 40.908210 | 11.907 | .0000 | |
| MONEYR | .4991237464 | .10188615 | 4.899 | .0000 | 291.80259 |
| INR | .9841524515 | .15112582 | 6.512 | .0000 | -48.327714 |
| RHO | .1741830568 | .22590872 | .771 | .4407 | |
+-----+-----+-----+-----+-----+-----+

```

Como se dijo antes existen otros métodos iterativos como el método de Cochrane-Orcutt y el de MV completo. A continuación se listan las instrucciones que se requieren para las misma. Dado que los resultados no son muy distintos

del proceso aquí descrito no se ejemplificaran pero se recomienda realizar su propia verificación.

```
/* Metodo de Cochrane-Orcutt */

?Estimación inicial
  REGRESS; LHS= GNP; RHS=ONE, MONEY, INTEREST;
  AR1; ALG=CORC; ; RES=EC$
  PLOT; RHS= EC; LHS=YEAR$
  IDENTIFY ; Rhs = EC ; Pds = 5$

?Transformación
  Create ; GNPR=GNP-RHO*GNP[-1]
  ; MONEYR=MONEY -RHO*MONEY[-1]
  ; INR=INTEREST-RHO*INTEREST[-1]$

?Estimación Corregida
  REGRESS; LHS= GNPR; RHS=ONE, MONEYR, INR;
  AR1; ALG=CORC; RES=E2$
  PLOT; RHS= E2; LHS=YEAR$
  IDENTIFY ; Rhs = E2 ; Pds = 5$

/* Metodo de Maximoverosimilitud */

?Estimación inicial
  REGRESS; LHS= GNP; RHS=ONE, MONEY, INTEREST;
  AR1; ALG=MLE; RES=EML$
  PLOT; RHS= EML; LHS=YEAR$
  IDENTIFY ; Rhs = EML; PDS=5$

?Transformación
  Create; GNPR=GNP-RHO*GNP[-1];
  MONEYR=MONEY -RHO*MONEY[-1];
  INR=INTEREST-RHO*INTEREST[-1]$

?Estimación Corregida
  REGRESS; LHS= GNPR; RHS=ONE, MONEYR, INR;
  AR1; ALG=MLE; RES=E3$
  PLOT; RHS= E3; LHS=YEAR$
  IDENTIFY ; Rhs = E3 ; Pds = 5$
```

7) DATOS AUTOCORRELACIÓN

A continuación se listan los datos utilizados en el desarrollo de este documento. Se espera que repliquen los resultados mostrados aquí.

Reset\$

? GNPM1.DAT

Read ; Nobs=20 ; Nvar=4 ; Names=

Year, Interest, Money, GNP \$

1966	4.5	480.0	2208.3
1967	4.19	524.3	2271.4
1968	5.16	566.3	2365.6
1969	5.87	589.5	2423.3
1970	5.95	628.2	2416.2
1971	4.88	712.8	2484.8
1972	4.50	805.2	2608.5
1973	6.44	861.0	2744.1
1974	7.83	908.4	2729.3
1975	6.25	1023.1	2695.0
1976	5.50	1163.6	2826.7
1977	5.46	1286.6	2958.6
1978	7.46	1388.9	3115.2
1979	10.28	1497.9	3192.4
1980	11.77	1631.4	3187.1
1981	13.42	1794.4	3248.8
1982	11.02	1954.9	3166.0
1983	8.50	2188.8	3277.7
1984	8.80	2371.7	3492.0
1985	7.69	2563.6	3573.5

DATES ; 1966 \$

PERIOD ; 1966 - 1985 \$

8) **Bibliografía**

- ▶ Box, G., & D. Pierce, 1970. «Distribution of Residual Autocorrelations in Autoregressive Moving Average Time Series Models.» *Journal of American Statistical Association*, 65
- ▶ Breusch, T. 1978. «Testing for Autocorrelation in Dynamic Linear Model.» *Australian Economic Papers*, 17
- ▶ Godfrey, L.G., 1978. «Testing Against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables.» *Econometrica*: 46.
- ▶ Godfrey, L.G., 1988. *Misspecification Tests in Econometrics*. Cambridge: Cambridge University Press.
- ▶ Greene, W.H., 1999. *Análisis Econométrico*. PRENTICE HALL IBERIA. Madrid. Tercera Edición
- ▶ Johnston, J. 1989. *Métodos Econometría*. Editorial Vicens-Vives. Barcelona
- ▶ Kamenta, Jan. 1977. *Elementos de Econometría*. Editorial Vicens-Vives Barcelona.
- ▶ Ljung, G. and G. Box, 1979. «On a Measure of Lack of Fit in Time Series Models.» *Biometrika*, 66, 265–270.
- ▶ Maddala, G. S. 1996. *Introducción a la Econometría*. PRENTICE-HALL HISPANOAMERICANA, S.A. Méjico.
- ▶ LIMDEP. 1998. *User's Manual Ver. 7.0*
- ▶ Davidson, R., and MacKinnon, J., 1993. *Estimation and Inference in Econometrics*, Oxford University Press.
- ▶ Judge, G. et. al., 1985. *The Theory and Practice of Econometrics and Introduction to the Theory and Practice of Econometrics*, John Wiley and Sons.
- ▶ Fomby, T., et. al., 1984. *Advanced Econometric Methods*, Springer Verlag.
- ▶ Berndt, E., 1991. *The Practice of Econometrics*, Addison Wesley.