UNIVERSIDAD DE PANAMA

VICERRECTORIA DE INVESTIGACIONES Y POSTGRADO

PROGRAMA DE MAESTRIA EN MATEMATICA

ESTUDIO DE LA AUTOCORRELACION DE LAS PERTURBACIONES EN LOS MODELOS LINEALES

POR:

SAMUEL IBARRA MURILLO

Tesis presentada como uno de los requisitos para optar por el grado de Maestro en Ciencias con Especialización en Estadística Matemática.

FM

UNIVERSIDAD DE PANAMA

VICERRECTORIA DE INVESTIGACION Y POSTGRADO

12 ENE. 1987

Aprobado por:

Director de Tesis ______________________
DIMAS QUIEL, M.Sc.

Miembro del Jurado ______________________
JORGE POLTRONIERI, Ph.D.

Miembro del Jurado ______________________
PLUTARCO E. RAMOS, M.Sc.

Fecha 20 de Octubre de 1986

Obsequio del autor

121078

"Año 1986, Centenario del Natalicio del Dr. Harmodio Arias"
CIUDAD UNIVERSITARIA OCTAVIO MENDEZ PEREIRA
ESTAFETA UNIVERSITARIA
PANAMA, R. DE P.

DEDICATORIA

Dedico este humilde trabajo a todas aquellas personas que fueron el motor y el soporte que me impulsaron a continuar por el camino del estudio hasta obtener este logro, pues, sin ellos esto no hubiese sido posible. Por esto lo dedico:

A la mujer que me brindó sus embelesos, me envolvió en sus besos y me meció en la cuna; a mi madre LOLA, que es dulzura vertida en mi amargura y en éste momento de mi vida, estrella; porque sin ella no hubiese sido posible mi existencia.

A mis amados y preciosos hijos SAMUEL, SAMUEL IVAN, SHAYANNI y a quien en estos momentos viene en camino como una nueva estrella para iluminar mas el firmamento de mi vida; a todos ellos que se constituyen en el motor que impulsa mis deseos y mis esfuerzos de superación.

Con mucho amor a mi querida y adorada esposa MAYRA, quien supo soportar con valentía y estoicismo todos los malos momentos que le ocacionó éste esfuerzo y que supo comprenderme y alentarme en todo momento.

A mis queridos compañeros de estudios AURORA, GLADYS, JOSE, RAFAEL, ARGELIA, DILCIA y CEFERINO; quienes confiaron siempre en mi y me brindaron su ayuda en los estudios.

A mis amigos y colegas JULIO ALBERTO, JOSE RAFAEL, FERNANDO, RICARDO, MANUEL y a todos aquellos que escapan de mi mente y que de alguna manera me brindaron su colaboración de manera desinteresada para que este momento se hiciera una realidad.

Y muy especialmente lo dedico a mi DIOS TODO PODEROSO, pues sin su participación y anuencia nada es posible.

AGRADECIMIENTO

A todas las personas que hicieron posible el que yo haya podido culminar ésta etapa de mi vida, deseo expresarles mi más profundo y sincero agradecimiento. Por ello quiero expresarle mi agradecimiento :

A mi profesor DIMAS QUIEL, por sus sabias enseñanzas, por su gran paciencia y sus oportunos y certeros consejos.

A mi profesor JORGE POLTRONIERI, quien no solo me brindó sus grandes conocimientos atraves de sus sabias enseñanzas, sino que, tambien supo soportarme con infinita paciencia y quien ademas, se constituyó en el espejo en el cual debíamos de mirarnos.

Al Coordinador del Programa de Maestría, EGBERTO AGARD por su insistencia y preocupación en que culminaramos nuestros estudios.

Al Decano de la Facultad de Ciencias Naturales y Exactas, JUSTO MEDRANO, por todas las facilidades que nos brindó en todos los momentos de nuestros estudios, para poder estudiar con comodidad.

Al Rector de nuestra más alta casa de estudios, DR. CEFERINO SANCHEZ por todas las facilidades que nos brindó y por su confianza depositada en el programa y sus estudiantes.

A mis hermanos JULIO, VILMA, MARITZA, DAVID, FLORIDA y muy especialmente a LUCIANO quien siempre me brindo su apoyo cincero y desinteresado.

INDICE

INTRODUCCION

El trabajo que estamos presentando como uno de los requisitos para optar por el grado de Maestro en Ciencias con Especialización en Estadística Matemática, está inspirado en uno de los problemas fundamentales que se presentan en el estudio de los Modelos Lineales Generales.

A través del tiempo, el hombre a tratado de crear modelos matemáticos que le ayuden a interpretar la realidad que lo rodea con la intención de poder predecir, inferir o probar hipótesis de sucesos en estudio. Uno de los modelos matemáticos que ha tenido mayor utilidad y aplicabilidad ha sido el Modelo Lineal, ya que, este modelo permite que el análisis de muchos de los problemas concretos que nos presenta la vida y la naturaleza sea más sencillo, ademas, por el hecho de que muchos de los otros tipos de modelos matemáticos se pueden transformar en modelos del tipo lineal mediante ciertas transformaciones. Sin embargo, sabemos que las observaciones de hechos reales no siempre se ajustan exactamente a un modelo lineal ya que, no siempre la relación entre las variables observadas es lineal o pudiera ser que las observaciones no se ajusten al modelo lineal debido a, errores cometidos en la medición, por errores humanos o, de otra indole.

Existen diversos métodos para ajustar una serie de datos reales a un Modelo Lineal; uno de los métodos de mayor uso para efectuar dicho ajuste es el llamado " Método de los Mínimos Cuadrados ". Para poder aplicar este método, es necesario suponer básicamente que los términos de los errores o perturbaciones de los datos a ajustar en el modelo de regresión, son estadísticamente serialmente independientes.

En nuestro trabajo estudiaremos el Modelo Lineal General, en el cual supondremos que se cumplen algunas hipótesis con la intención de que las estimaciones obtenidas sean las mejores. Las hipótesis básicas que supondremos cumple nuestro modelo son las siguientes:

i .) Las variables X_i son no relacionadas entre sí.

ii .) Las Perturbaciones tienen Varianza constante.

iii .) Las variables X_i no son valores previos de las Y_i.

iv .) Las Perturbaciones u_i son serialmentes independientes.

De las hipótesis supuestas anteriormente, el no cumplimiento de la última de ellas y se conoce como el problema de la Autocorrelación de las Perturbaciones, es lo que nos ha inspirado a efectuar este trabajo.

En el Problema de la Autocorrelación de las Perturbaciones, nuestro estudio central se ha encaminado a analizar uno de los principales criterios que se usan para probar la hipótesis de Independencia entre los Errores llamado " Criterio de Durbin-Watson", el cual fué propuesto por Durbin y Watson ([3]), ([4]) y ([5]).

Para efectuar el estudio objeto de nuestra atención hemos desarrollado el presente trabajo en tres capítulos, así:

En el Capítulo I se efectua un estudio resumido del Modelo Lineal

General k Variante. Dentro del mismo se analizan los Estimadores de Mínimos Cuadrados y los de Máxima Verosimilitud de los parámetros desconocidos del modelo y sus principales propiedades, también se estudia el Estimador de la Varianza de las Perturbaciones, además, se presenta el concepto de lo que se conoce como Autocorrelación de los Errores y las pruebas que se pueden hacer al respecto de la misma.

En el Capítulo II se analiza en detalle el Criterio de Durbin-Watson para probar la Autocorrelación de las Perturbaciones. En el análisis se justifica la existencia del criterio, se estudia su distribución, se calculan los momentos de su numerador y denominador y además, se determina la Función Característica Conjunta del numerador y denominador del criterio. Por último, se hace un análisis de selección del criterio de la prueba para obtener una prueba que sea Uniformemente la más Potente entre las pruebas. También se revisan otros criterios para probar la Autocorrelación de los Errores en el supuesto de que el de Durbin-Watson nos conduzcan a inconclusiones.

En el Capítulo III analizamos dos métodos para la obtención de los límites de las Raíces Características de la Matríz Γ asociada a los errores, y que nos proporcionaran la Potencia de la Prueba del Estadístico de Durbin-Watson.

Por último, advertimos a los interesados en el tema que, el mismo exige del conocimiento general del Algebra Lineal, la Inferencia Estadística y los Modelos Lineales, los cuales se podrían obtener de los estudios de Franklin A. Graybill ([15]), Wonnacott ([16]), Boweman, O'Connell y Dicky ([17]) y de Courant and Hilbert ([18]).

MODELO LINEAL GENERAL

I EL MODELO LINEAL GENERAL

Sea $$Y = X\beta + U \qquad (1)$$

nuestro Modelo Lineal General k variante, donde,

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \qquad X = \begin{bmatrix} X_{11} & X_{21} & \cdots & X_{k1} \\ X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & & \vdots \\ X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \qquad U = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix} \qquad (2)$$

y donde supondremos que,

a.) $E[U] = 0$

b.) $E[UU'] = \sigma^2 \mathbf{I}_n$ (3)

c.) X es una matríz de rango completo $k \leq n$, cuyas columnas son las variables explicativas.

d.) β es el vector de parametros desconocidos.

e.) U vector de errores de componentes U_t , $t = 1,\ldots\ldots, n$.

Como $E[U] = 0 \Longrightarrow E[U_t] = 0$; $\forall\, t: 1 \leq t \leq n$

ademas, como

$$E[UU'] = \sigma^2 \mathbf{I} \Longrightarrow \begin{cases} E[U_t^2] = \sigma^2; \ \forall t: 1 \leq t \leq n \\ E[U_t U_{t+s}] = 0; \ \forall s \neq 0 \end{cases}$$

esto nos indica que las perturbaciones U_t son no correlacionadas a pares. La hipótesis (3c) nos indica que la única fuente de variación de Y es el vector U y que nuestro modelo está condicionado a X.

II ALGUNOS ESTIMADORES Y SUS PROPIEDADES

A- LOS ESTIMADORES MINIMOS CUADRATICOS DE β

Sea $\hat{\beta}'=(\hat{\beta}_1,\hat{\beta}_2,\cdots\cdots\cdots,\hat{\beta}_k)$ el vector de las estimaciones de β, así tenemos que,

$$Y = X\hat{\beta} + \varepsilon \qquad (4)$$

donde ε es el vector de las estimaciones de los n residuos

$$\varepsilon = Y - X\hat{\beta}$$

$$\Longrightarrow \quad \varepsilon'\varepsilon = Y'Y - 2(X\hat{\beta})'Y + (X\hat{\beta})'(X\hat{\beta}) \qquad (5)$$

Si derivamos (5) con respecto a $\hat{\beta}$ e igualamos a cero el resultado, obtendremos que,

$$\hat{\beta} = (X'X)^{-1}X'Y \qquad (6)$$

el cual es el Estimador de Mínimos Cuadrados de β.

A 1. PROPIEDADES DEL ESTIMADOR $\hat{\beta}$.

1.) $\hat{\beta}$ es Función Lineal de las perturbaciones U, o sea que,

$$\hat{\beta} = \beta + (X'X)^{-1}X'U \qquad (7)$$

2.) $\hat{\beta}$ es un Estimador Insesgado de β o sea que.

$$E[\hat{\beta}] = \beta \qquad (8)$$

3.) $\hat{\beta}$ es de Varianza Mínima, esto significa que si B es otro Estimador Insesgado de β, tal que $B = AY$ donde A es una matríz de rxn; si hacemos $AX = C$, entonces tendremos que $var(B) = \sigma^2 AA'$. (9)

Para minimizar la varianza de nuestros estimadores debemos escoger los elementos de A tales que, $tr(A'A)$ sea mínima. Como la $var(\hat{\beta}) = \sigma^2(X'X)^{-1}$ y $var(B) = \sigma^2 C(X'X)^{-1}C'$ son mínimas, se tiene que si hacemos $C = I$ resultará que $\hat{\beta}$ es de varianza mínima y es el óptimo.

B-ESTIMADOR DE MAXIMA VEROSIMILITUD DE β:

Si en el modelo lineal (1), además de las hipótesis (3), suponemos que,

$$U \sim N(0,\sigma^2 I) \tag{10}$$

Se prueba a través de la Función de Verosimilitud de U que el Estimador de Máxima Verosimilitud de β está dado por,

$$\tilde{\beta} = (X'X)^{-1}X'Y \tag{11}$$

el cual, por ser igual al Estimador de Mínimos Cuadrados de β, es el Mejor Estimador Lineal Insesgado y además que,

$$\tilde{\beta} \sim N(\beta,\sigma^2(X'X)^{-1}) \tag{12}$$

C-ESTIMADOR DE LA VARIANZA DE LA PERTURBACION.

De (4) se tiene que,

$$\varepsilon = Y - X\hat{\beta} \tag{13}$$

y luego, de (1) y (6) tenemos que,

$$\varepsilon = \left[I - X(X'X)^{-1}X'\right]U \tag{14}$$

donde $\left[I - X(X'X)^{-1}X'\right] = A$, es una matríz semidefinida positiva, simétrica e idempotente, o sea que,

$$A = A' \qquad \wedge \qquad A.A = A \tag{15}$$

La suma de los cuadrados de los residuos es,

$$\varepsilon'\varepsilon = U'\left[I - X(X'X)^{-1}X'\right]U \tag{16}$$

$$\Longrightarrow \quad E\left[\varepsilon'\varepsilon\right] = E\left[U'(I - X(X'X)^{-1}X')U\right]$$

$$= \sigma_u^2(n - k) \tag{17}$$

Así, si designamos por $S^2 = \frac{\varepsilon'\varepsilon}{n - k_2}$ al estimador de la varianza de las perturbaciones, tendremos que S^2 es un estimador insesgado, tal que,

$$S^2 = \frac{Y'Y - Y'X(X'X)^{-1}X'Y}{n - k} \tag{18}$$

III AUTOCORRELACION DE ERRORES:

A. MINIMOS CUADRADOS GENERALIZADOS.

Sea

$$Y = X\beta + U \tag{19}$$

con Y, X, β y U como en (2), donde además, se supone que,

$$E[U] = 0 \qquad E[UU'] = V \tag{20}$$

Con V Matríz de Varianza-Covarianza del término de perturbación U, donde V es una matríz no singular de orden nxn.

Supongamos que U sigue un proceso estacionario de Markov de Primer Orden, o sea que,

$$U_t = \rho U_{t-1} + \varepsilon_t \tag{21}$$

en que $|\rho| < 1$ y donde ε_t satisface,

$$\begin{aligned} &\text{a.)} \quad E[\varepsilon_t] = 0 \\ &\text{b.)} \quad E[\varepsilon_t \varepsilon_{t+s}] = \begin{cases} \sigma^2 & \text{si } s = 0 \ ; \ \forall t \\ 0 & \text{si } s \neq 0 \ ; \ \forall t \end{cases} \end{aligned} \tag{22}$$

Las ecuaciones (21) y (22) definen la forma más simple de un esquema autorregresivo a fin de considerar la dependencia serial que pudiera haber en las perturbaciones.

De (21) se tiene que,

$$U_t = \rho^{\infty} U_{t-\infty} + \sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i} \tag{23}$$

Si hacemos $\rho^{\infty} = 0$ en (23), resulta que,

$$U_t = \sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i} \tag{24}$$

Como $\forall$ t se tiene que $E[\varepsilon_t] = 0$, entonces, $E[U_t] = 0$.

Por otro lado,

$$E[U_t^2] = E[(\sum_{i=0}^{\infty}\rho^i \varepsilon_{t-i})^2]$$

$$= [1 + \rho^2 + \rho^4 + \ldots\ldots.]\sigma_\varepsilon^2 \tag{25}$$

Como $1 + \rho^2 + \rho^4 + \ldots\ldots\ldots.$ es la suma de los términos de una Progresión Geométrica Infinita Decreciente con primer término $a = 1$ y razón $r = \rho^2$, entonces, se tiene que,

$$1 + \rho^2 + \rho^4 + \ldots\ldots = \frac{1}{1-\rho^2} \tag{26}$$

$$E[U_t^2] = \frac{\sigma_\varepsilon^2}{1-\rho^2} \qquad \forall\ t \tag{27}$$

Como $E[U_t^2] = E[(U_t - E[U_t])^2] = \sigma_u^2$; $\quad \forall t$

$$\Longrightarrow \quad \sigma_u^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2} \quad \forall\ t \tag{28}$$

Por otro lado tenemos que,

$$\begin{aligned}
Cov(U_t, U_{t-1}) &= E[U_t U_{t-1}] \\
&= E[(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \ldots.)(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \ldots\ldots.)] \\
&= E[\{\varepsilon_t + \rho(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \ldots..)\}(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \ldots\ldots..)] \\
&= E[\varepsilon_t(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \ldots..) + \rho(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \ldots.)^2] \\
&= E[\varepsilon_t(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \ldots..)] + \rho E[(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \ldots)^2]
\end{aligned}$$

Luego, por (22b) se tiene que,

$$E[\varepsilon_t(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \ldots\ldots..)] = 0$$

$$\Longrightarrow \quad Cov(U_t, U_{t-1}) = \rho E\left[(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \rho^2\varepsilon_{t-3} + \ldots\ldots) \right]$$

$$= E\left[U_{t-1}^2 \right]\rho$$

$$= \rho\sigma_u^2 \tag{29}$$

Si efectuamos el mismo procedimiento a U_t y U_{t-2}, se tiene que,

$$Cov(U_t, U_{t-2}) = E\left[U_t U_{t-2} \right] = \rho^2\sigma_u^2$$

Luego, generalizando este resultado $\forall\ s \neq 0$, se tiene que,

$$Cov(U_t, U_{t-s}) = \rho^s\sigma_u^2 \qquad \forall\ s \neq 0 \tag{30}$$

Por lo tanto, (19) no satisface la hipótesis de independencia serial entre las perturbaciones.

Sea

$$V = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix} \sigma_u^2 \tag{31}$$

y

$$\beta^* = AY \tag{32}$$

donde A es una matríz de kxn y β^* es un estimador lineal de β (lineal en los valores de Y).

Si hacemos $AX = I$, tendremos que:

1.) β^* es un Estimador Insesgado de β, o sea que,

$$E\left[\beta^* \right] = E\left[AY \right] = \beta \tag{33}$$

2.) $Var\left[\beta^* \right] = E\left[(\beta^* - \beta)\cdot(\beta^* - \beta)' \right] = AVA'$ (34)

Para determinar un estimador de β con mínima varianza, debemos encontrar una matríz A tal que tra(A'AV) sea mínima y cumpla con (33). Para ello introducimos k^2 multiplicadores de Lagrange ϕ_{ij} tal que:

$$\Psi = \text{tra}(A'AV) - \text{tra}\left[\,\Phi'(AX - I)\,\right] \tag{35}$$

donde

$$\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} \cdots\cdots\cdots & \phi_{1k} \\ \phi_{21} & \phi_{22} \cdots\cdots\cdots & \phi_{2k} \\ \vdots & \vdots & \vdots \\ \phi_{k1} & \phi_{k2} & \phi_{kk} \end{bmatrix} \tag{36}$$

Luego, derivando (35) parcialmente con respecto a A e igualando a cero tendremos que,

$$2AV = \Phi X' \tag{37}$$

Multiplicando (37) por $V^{-1}X$ y despejando se tiene que,

$$\Phi = 2(X'V^{-1}X)^{-1} \tag{38}$$

De (37) y (38) se tiene por sustitución que,

$$A = (X'V^{-1}X)^{-1}X'V^{-1} \tag{39}$$

Para demostrar que A, dada en (39) proporciona una varianza mínima, definimos un estimador lineal insesgado arbitrario de β de la siguiente forma,

$$B = (A + R)Y \tag{40}$$

donde R es una matríz de kxn no nula y donde, si se hace $RX = 0$ se tendrá que,

$$\text{Var}\left[\,B\,\right] = \text{Var}\left[\,\beta^*\,\right] + E\left[\,RUU'R'\,\right] \tag{41}$$

Como para cualquier matríz real R, la matríz dada por RUU'R' es definida positiva si rang(R) = k y es semidefinida positiva si rang(R) < k, entonces A proporciona una varianza mínima, con A definida por (39) y β^* definido por (32).

β^* es un estimador lineal insesgado de varianza mínima. (BLUE)

B. ESTIMADOR DE MAXIMA VEROSIMILITUD.

Si en el modelo (19), además de la hipótesis (20) se supone que,

$$U \sim N(0,V) \tag{43}$$

entonces, su Función de Verosimilitud estará dada por,

$$L = (2\pi)^{-1/2 n}|V|^{-1/2}\ \mathrm{Exp}\{-\tfrac{1}{2}(Y - X\beta)'V^{-1}(Y - X\beta)\} \tag{44}$$

Si elegimos como estimador de β aquel que minimixa la forma cuadrática,

$$(Y - X\beta)'V^{-1}(Y - X\beta) \tag{45}$$

llegaremos a obtener (39) y además, tendremos que,

1.) $\beta^* = AY = \beta + (X'V^{-1}X)^{-1}X'V^{-1}U$ (46)

2.) $E\left[(\beta^* - \beta)\cdot(\beta^* - \beta)'\right] = (X'V^{-1}X)^{-1}$ (47)

3.) β^* es de Varianza Mínima

Así, β^* es un Estimador Lineal Insesgado de Varianza Mínima, tal que,

$$\beta^* \sim N(\beta,\ (X'V^{-1}X)^{-1}) \tag{48}$$

C. IMPLICACIONES DE LOS ESTIMADORES MINIMOS CUADRATICOS GENERALIZADOS. CASO SIMPLE DE PERTURBACIONES AUTOCORRELACIONADAS.

Supongamos que se tiene un esquema de la forma,

$$Y_t = \alpha + \beta X_t + U_t \tag{49}$$

en donde U_t sigue un proceso autorregresivo como el descrito en (21) y Y_t es tal que cumple con la descripción de Y en (19). Entonces, tendremos que:

$$E\left[UU'\right] = V = \frac{\sigma_\varepsilon^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \cdots\cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots\cdots & \rho^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots\cdots & 1 \end{bmatrix} \qquad (50)$$

$$\Longrightarrow \quad V^{-1} = \frac{1}{\sigma_\varepsilon^2}\begin{bmatrix} 1 & -\rho & 0 & \cdots\cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots\cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots\cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots\cdots & -\rho & 1 \end{bmatrix} \qquad (51)$$

Se demuestra que el Estimador de Mínimos Cuadrados Generalizado es equivalente al estimador que se obtiene de un proceso de mínimos cuadrados en dos etápas, de la siguiente forma:

1.) Transformando las variables originales de acuerdo con la estructura autorregresiva del término de perturbación.

En efecto, sea $Y = X\beta + U$ y T una matríz de transformación tal que,

$$TY = TX\beta + TU \qquad (52)$$

2.) Aplicando Mínimos Cuadrados Ordinarios a las variables transformadas en (52).

El estimador de mínimos cuadrados ordinarios de β será:

$$\beta^* = \left[\,(TX)'(TX)\,\right]^{-1}(TX)'(TY)$$

$$= (X'T'TX)^{-1}X'T'TY \qquad (53)$$

Si hacemos $T'T = V^{-1}$ como en (51), resulta que,

$$\beta^* = (X'V^{-1}X)^{-1}X'V^{-1}Y \tag{54}$$

de donde, por (39) se tiene que, $\beta^* = AY$.

Si por razones de simplicidad hacemos $\sigma_\varepsilon^2 = 1$, entonces, una matríz T de (n-1)xn que satisface casi completamente a (54) es,

$$T = \begin{bmatrix} -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}_{(n-1)xn} \tag{55}$$

puesto que,

$$T'T = \begin{bmatrix} \rho^2 & -\rho & 0 & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{nxn} \tag{56}$$

De (51) y (56) se observa que la única diferencia entre V^{-1} y $T'T$ es el primer elemento de la primera fila y primera columna.

Luego, las variables transformadas indicadas por T son:

$$\begin{bmatrix} Y_2 - \rho Y_1 \\ Y_3 - \rho Y_2 \\ \vdots \\ Y_n - \rho Y_{n-1} \end{bmatrix} \quad \hat{} \quad \begin{bmatrix} X_2 - \rho X_1 \\ X_3 - \rho X_2 \\ \vdots \\ X_n - \rho X_{n-1} \end{bmatrix} \tag{57}$$

Con lo cual se tiene que:

$$E[\,(TU)(TU)'\,] = TVT' = \sigma_\varepsilon^2 \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{58}$$

D. PRUEBAS DE HIPOTESIS PARA LA AUTOCORRELACION DE LAS PERTURBACIONES SERIALES.

Dado un Modelo Lineal General $Y = X\beta + U$, como el descrito en la sección III-A, se quiere probar la siguiente hipótesis:

$H_0 : \rho > 0$
$H_1 : \rho = 0$

Esto equivale a querer probar que:

H_0 : Los errores están autocorrelacionados positivamente.
H_1 : Los errores no están autocorrelacionados.

Para probar ésta hipótesis, existen varios métodos, algunos de los cuales presentan muy pocas diferencias pero, casi todos están basados en el estudio de los residuos.

En el siguiente capítulo estudiaremos algunos de los criterios utilizados para probar estas hipótesis, entre los cuales podemos mencionar el Criterio del Cociente d, de Durbin - Watson. Este método lo estudiaremos con bastante profundidad debido a que es uno de los más utilizados y que tiene mayor aplicabilidad. Este método se basa en el criterio del cociente de verosimilitud de los errores, pues presupone que los errores se distribuyen como variables aleatorias normales.

Si suponemos que, $\varepsilon \sim N(0,\sigma^2 \mathbf{I})$
donde $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots\ldots\ldots, \varepsilon_n)'$, con ε_t definido como en (22), $\forall$ t; $1 \leq t \leq n$, entonces, se tiene que,

$$\sum_{t=1}^{n} \varepsilon_t^2 = \varepsilon'\varepsilon = U'\left[I - X(X'X)^{-1}X' \right]U \qquad (59)$$

donde $I - X(X'X)^{-1}X' = P$ es una matríz simétrica e idempotente y de rango n, por lo tanto se tiene que,

$$\varepsilon'\varepsilon \sim \chi^2_\nu \qquad \text{con} \quad \nu = n \quad \text{g.d.l} \qquad (60)$$

Por otro lado tenemos que,

$$\sum_{t=2}^{n} (\varepsilon_t - \varepsilon_{t-1})^2 = \varepsilon' A \varepsilon \qquad (61)$$

donde $A = T'T$ con $T'T$ definida como en (56), de aquí se deduce que

$$\sum_{t=2}^{n} (\varepsilon_t - \varepsilon_{t-1})^2 = U'\left[A - AX(X'X)^{-1}X' \right]U \qquad (62)$$

donde, $A - AX(X'X)^{-1}X' = Q$ no es una matríz idempotente ya que A no lo es. Por lo tanto $\varepsilon'A\varepsilon$ no se distribuye como una Chi-cuadrado; ademas, como

$$P \cdot Q = \left[I - X(X'X)^{-1}X' \right] \cdot \left[A - AX(X'X)^{-1}X' \right] \neq 0 \qquad (63)$$

entonces, se tiene que $\varepsilon'\varepsilon$ y $\varepsilon'A\varepsilon$ no se distribuyen en forma independiente, con lo cual el cociente formado por (62) y (59):

$$d = \frac{\sum_{t=2}^{n} (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=0}^{n} \varepsilon_t^2} = \frac{\varepsilon' A \varepsilon}{\varepsilon'\varepsilon} \qquad (64)$$

no se distribuye bajo ninguna de las formas comunes conocidas.

ALGUNOS CRITERIOS PARA LA PRUEBA DE LA AUTOCORRELACION DE LAS PERTURBACIONES

I CRITERIO DE DURBIN-WATSON.

A- INTRODUCCION AL TRATAMIENTO TEORICO.

Sabemos que una simple ecuación de un Modelo de Regresión puede ser escrita de la forma

$$Y = \beta_1 X_1 + \beta_2 X_2 + \cdots\cdots + \beta_k X_k + \varepsilon \qquad (1)$$

donde Y es la variable dependiente, X_i es la i-ésima variable independiente, β_i es el i-ésimo parámetro, el cual es desconocido y ε es un término de error, tambien desconocido. Para una muestra de n observaciones, nuestro modelo toma la forma,

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} X_{11} & X_{21} & \cdots & X_{k1} \\ X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & & \vdots \\ X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \qquad (2)$$

o sea que

$$Y = X\beta + \varepsilon \qquad (3)$$

de donde se tiene que $\hat{\beta}$ es el Estimador de Mínimos Cuadrados de β, el cual está dado por,

$$\hat{\beta} = (X'X)^{-1}X'Y \qquad (4)$$

además, $Z = Y - X\hat{\beta}$ es el Vector de Residuos de la Regresión, donde

$$Z = \{I_n - X(X'X)^{-1}X'\}Y = \{I_n - X(X'X)^{-1}X'\}\varepsilon = M\varepsilon \qquad (5)$$

donde M es una matríz simétrica semidefinida positiva e idempotente, o sea que,

$$M = (m_{ij}) = M' = (m_{ji}) = M^2 \qquad (6)$$

con estos elementos definimos el estadístico:

$$r = \frac{Z'AZ}{Z'Z} = \frac{\varepsilon'M'AM\varepsilon}{\varepsilon'M\varepsilon} \qquad (7)$$

el cual será el punto esencial de nuestro análisis.

Dadas las condiciones anteriores, enunciaremos un teorema, el cual demostraremos y a partir del cual se desprende un corolario que nos será de gran utilidad en nuestro trabajo.

TEOREMA 1:

Si Z y ε son vectores nx1 tales que $Z = M\varepsilon$, donde $M = I_n - X(X'X)^{-1}X'$ y si $r = \frac{Z'AZ}{Z'Z}$, donde A es una matríz real simétrica, entonces:

a.) Existe una transformación ortogonal $\varepsilon = M\zeta$, tal que:

$$r = \frac{\sum_{i=1}^{n-k} \nu_i \zeta_i^2}{\sum_{i=1}^{n-k} \zeta_i^2} \qquad (8)$$

donde $\nu_1, \nu_2, \ldots\ldots\ldots\ldots, \nu_{n-k}$ son las raíces propias de $M \cdot A$, menos k ceros.

b.) Si s de las columnas de X son combinaciones lineales de s vectores propios de A, y si las raíces de A asociadas con los restantes vectores propios n-s de A son reenumerados tal que:

$$\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots\cdots\cdots \leq \lambda_{n-s}$$

entonces,

$$\lambda_i \leq \nu_i \leq \lambda_{i+k-s} \qquad (i = 1,2,\ldots,n-k) \qquad (9)$$

DEMOSTRACION:

a.) Demostraremos que existe una transformación ortogonal $\varepsilon = H\zeta$ tal que:

$$r = \frac{\sum_{i=1}^{n-k} \nu_i \zeta_i^2}{\sum_{i=1}^{n-k} \zeta_i^2}$$

En efecto, sabemos que existe una matríz ortogonal L tal que:

$$L'M\,L = \begin{bmatrix} I_{n-k} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 0 \end{bmatrix}$$

donde I_{n-k} es la matríz de unidades de orden n-k y 0 representa una matríz cero con apropiados números de filas y columnas.

Así,

$$\begin{aligned} L'M\,A\,M\,L &= (L'M\,L)(L'A\,L)(L'M\,L) \\ &= \begin{bmatrix} I_{n-k} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 0 \end{bmatrix} \begin{bmatrix} B_1 & \vdots & B_3 \\ \cdots & \cdots & \cdots \\ B_2 & \vdots & B_4 \end{bmatrix} \begin{bmatrix} I_{n-k} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 0 \end{bmatrix} \\ &= \begin{bmatrix} B_1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 0 \end{bmatrix} \end{aligned} \qquad (10)$$

donde $\begin{bmatrix} B_1 & \vdots & B_3 \\ \cdots & \cdots & \cdots \\ B_2 & \vdots & B_4 \end{bmatrix}$ es la partición apropiada de la matríz real simétrica L'A L.

Sea N_1 la matríz ortogonal que diagonaliza a B_1 o sea que,

$$N_1' B_1 N_1 = \begin{bmatrix} \nu_1 & & & \\ & \nu_2 & & 0 \\ & & \ddots & \\ 0 & & & \nu_{n-k} \end{bmatrix} \qquad (11)$$

entonces, $N = \begin{bmatrix} N_1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & I_k \end{bmatrix}$ es ortogonal, de modo que $H = L\cdot N$ es ortogonal.

Por consiguiente,

$$H'MH = N'L'MLN = N'\begin{bmatrix} I_{n-k} & \vdots & 0 \\ \cdots & & \cdots \\ 0 & \vdots & 0 \end{bmatrix} N$$

$$= \begin{bmatrix} I_{n-k} & \vdots & 0 \\ \cdots & & \cdots \\ 0 & \vdots & 0 \end{bmatrix} \tag{12}$$

Por lo tanto,

$$H'M\,A\,M\,H = (H'M\,H)(H'A\,H)(H'M\,H)$$

$$= \begin{bmatrix} I_{n-k} & \vdots & 0 \\ \cdots & & \cdots \\ 0 & \vdots & 0 \end{bmatrix} \begin{bmatrix} N_1'B_1N_1 & \vdots & 0 \\ \cdots & & \cdots \\ 0 & \vdots & 0 \end{bmatrix} \begin{bmatrix} I_{n-k} & \vdots & 0 \\ \cdots & & \cdots \\ 0 & \vdots & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \nu_1 & & & & & \vdots & \\ & \nu_2 & & 0 & & \vdots & 0 \\ & & \nu_3 & & & \vdots & \\ & 0 & & \ddots & & \vdots & \\ & & & & \nu_{n-k} & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots & & \cdots \\ & 0 & & & & \vdots & 0 \end{bmatrix} \tag{13}$$

Si hacemos $\varepsilon = H\zeta$, se tiene que,

$$r = \frac{\sum_{i=1}^{n-k} \nu_i \zeta_i^2}{\sum_{i=1}^{n-k} \zeta_i^2}$$

Cuando $\varepsilon' M\,A\,M\varepsilon$ = constante y $\varepsilon' M\varepsilon$ = constante, se tiene què son hipercilindros con generadores paralelos.

Sabemos que $\nu_1, \nu_2, \nu_3, \ldots\ldots\ldots, \nu_{n-k}$ son las raíces propias de $M\,A\,M$ menos k ceros, o sea que son las raíces propias de $M^2 A$, pero, como $M^2 = M$, entonces, $M^2A = M\,A$; por lo tanto $\nu_1, \nu_2, \nu_3, \ldots\ldots\ldots\ldots, \nu_{n-k}$ son las raíces propias de $M\,A$ menos

κ ceros.

Si efectuamos la transformación real no singular de las $\{X_i\}$, $P = XG$, se tiene que $M = I_n - P(P'P)^{-1}P'$, es decir, M es invariante bajo este tipo de transformación; si escogemos G de modo que los vectores columnas $P_1, P_2, P_3, \ldots\ldots\ldots, P_k$; de P son ortogonales.

$$\text{i.e.} \qquad P_i'P_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \tag{14}$$

tenemos que $P'P = I_n$.

Esto nos indica que podemos cambiar las variables independientes originales por un conjunto ortogonal sin afectar los residuos. Por lo tanto,

$$\begin{aligned} M &= I_n - (P_1P_1' + P_2P_2' + \cdots\cdots + P_kP_k') \\ &= (I_n - P_1P_1')(I_n - P_2P_2')\cdots(I_n - P_kP_k') \\ &= M_1M_2M_3\cdots\cdots M_k \end{aligned} \tag{15}$$

donde cada M_i tiene la misma forma de M; es claro que los M_i conmutan, o sea que en términos algebráicos podemos ajustar regresiones sobre variables ortogonales separadamente y en cualquier orden sin afectar el resultado final.

Dada la forma de las matrices M_i tenemos que, cualquier resultado que establezcamos acerca de las raíces de M_1A en términos de las de A, será cierto para las raíces de M_2M_1A en término de las de M_1A y así sucesivamente. Esto sugiere un método para desarrollar un conocimiento de las raíces de $M_kM_{k-1}\cdots\cdots M_2M_1A$ en etápas partiendo de las raíces de A, las cuales se asumen conocidas.

Por lo anteriormente dicho, estudiaremos las raíces propias de M_1A; $\theta_1, \theta_2, \theta_3, \ldots\ldots, \theta_{n-1}, 0$; las cuales son raices de la ecuación determinante,

$$|I_n - M_1A| = |I_n - (I_n - P_1P_1')A| = 0 \tag{16}$$

Sea T una matríz ortogonal que diagonaliza a A así:

$$T'AT = \Lambda = \begin{bmatrix} \lambda_1 & & & & 0 \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ 0 & & & & \lambda_n \end{bmatrix} \tag{17}$$

donde $\lambda_1, \lambda_2, \lambda_3, \ldots\ldots\ldots\ldots, \lambda_n$ son las raíces propias de A.

Si multiplicamos por $|T'|$ y $|T|$ a (16) se tiene que:

$$|T'||I_n\theta - (I_n - P_1P_1')A||T| = |I_n\theta - (I_n - 1_11_1')\Lambda| = 0 \tag{18}$$

donde $1_1 = (1_{11}, 1_{12}, 1_{13}, \ldots\ldots\ldots\ldots, 1_{1n})$ con $1_1 = TP_1$ es el vector de dirección coseno de P_1 referente a los vectores de A como ejes. Si eliminamos el subíndice 1 en (18), tendremos que:

$$|I_n\theta - (I_n - 11')\Lambda| = \begin{vmatrix} \theta-\lambda_1+1_1^2\lambda_1 & 1_11_2\lambda_2 & 1_11_n\lambda_n \\ 1_21_1\lambda_1 & \theta-\lambda_2+1_2^2\lambda_2 & 1_21_n\lambda_n \\ \vdots & \vdots & \vdots \\ 1_n1_1\lambda_1 & 1_n1_2\lambda_2 & \theta-\lambda_n+1_n^2\lambda_n \end{vmatrix} \tag{19}$$

Si restamos $1_2/1_1$ veces la primera fila de la segunda, $1_3/1_1$ veces la primera fila de la tercera y así sucesivamente, se tiene que:

$$\begin{vmatrix} \theta-\lambda_1+1_1^2\lambda_1 & 1_1 1_2\lambda_2 & 1_1 1_3\lambda_3 & \cdots & 1_1 1_n\lambda_n \\ -(\theta-\lambda_1)1_2/1_1 & (\theta-\lambda_2) & 0 & \cdots & 0 \\ -(\theta-\lambda_1)1_3/1_1 & 0 & (\theta-\lambda_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -(\theta-\lambda_1)1_n/1_1 & 0 & 0 & \cdots & (\theta-\lambda_n) \end{vmatrix} = 0 \qquad (20)$$

Este determinante es de la forma:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & 0 & 0 & \cdots & 0 \\ a_{31} & 0 & a_{33} & 0 & \cdots & 0 \\ a_{41} & 0 & 0 & a_{44} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \\ a_{n1} & 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix} = 0 \qquad (21)$$

Donde resolviendo por menores se tiene que este determinante es igual a:

$$|\mathbb{I}_n\theta - (\mathbb{I}_n - 11')\Lambda| = \prod_{j=1}^{n} a_{jj} - \sum_{j=2}^{n} \prod_{\substack{j=2 \\ j\neq i}}^{n} a_{jj}a_{1i}a_{j1} \qquad (22)$$

Así, reemplazando los a_{jj} por los valores correspondientes dados en el determinante, es decir:

$$\begin{aligned} &a_{11} = (\theta - \lambda_1) + 1_1^2\lambda_1 \quad \wedge \quad a_{jj} = (\theta - \lambda_j); \;\; \forall\, j: j \neq 1 \\ &a_{j1} = -(\theta - \lambda_1)1_j/1_1 \qquad\quad a_{1i} = 1_1 1_i\lambda_i \;\; ; \;\; \forall\, i: i \neq 1 \end{aligned} \qquad (23)$$

se tiene que:

$$\prod_{j=1}^{n} a_{jj} = a_{11} \prod_{j=2}^{n} a_{jj} = \{(\theta - \lambda_1) + 1_1^2\lambda_1\} \prod_{j=2}^{n}(\theta - \lambda_j)$$

$$= \prod_{j=1}^{n}(\theta - \lambda_j) + 1_1^2\lambda_1 \prod_{j=2}^{n}(\theta - \lambda_j) \qquad (24)$$

$$a_{j1} \cdot a_{1i} = -(\theta - \lambda_i)1_i/1_1 \cdot 1_1 1_i \lambda_i = -(\theta - \lambda_i)1_i^2\lambda_i$$

Luego, sustituyendo estos valores en la ecuación (22) se tiene que:

$$|I_n\theta - (I_n - 11')\Lambda| = \prod_{j=1}^{n}(\theta-\lambda_j) + 1_1^2\lambda_1 \prod_{i=2}^{n}(\theta-\lambda_i) + \sum_{i=2}^{n} 1_i^2\lambda_i \prod_{\substack{j=1 \\ j\neq i}}^{n}(\theta-\lambda_j)$$

$$= \prod_{j=1}^{n}(\theta - \lambda_j) + \sum_{i=1}^{n} 1_i^2\lambda_i \prod_{\substack{j=1 \\ j\neq i}}^{n}(\theta - \lambda_j) = 0 \qquad (25)$$

Si reducimos (25) y sacamos un factor θ correspondiente a la raiz cero conocida de $M_1 \cdot A$, obtendremos:

$$\sum_{i=1}^{n} 1_i^2 \prod_{j\neq i}^{n}(\theta - \lambda_j) = 0 \qquad (26)$$

donde $\theta_1, \theta_2, \theta_3, \ldots\ldots\ldots\ldots, \theta_{n-1}$ son las raices de ésta ecuación.

Demostraremos que (26) se deduce de (25), para ello usaremos el Método de Inducción.

1.) Sea n = 2: $\prod_{i=1}^{2}(\theta-\lambda_j) + \sum_{i=1}^{2} 1_i^2\lambda_i \prod_{j\neq i}^{2}(\theta-\lambda_j) = 0$ con $\sum_{i=1}^{2} 1_i^2 = 1$

$$\Longrightarrow \prod_{i=1}^{2}(\theta-\lambda_j) + \sum_{i=1}^{2} 1_i^2\lambda_i \prod_{j\neq i}^{2}(\theta-\lambda_j) = (\theta-\lambda_1)(\theta-\lambda_2) + 1_1^2\lambda_1(\theta-\lambda_2) + 1_2^2\lambda_2(\theta-\lambda_1)$$

$$= \theta^2 - \theta\lambda_1 - \theta\lambda_2 + \lambda_1\lambda_2 + 1_1^2\lambda_1\theta - 1_1^2\lambda_1\lambda_2 + 1_2^2\lambda_2\theta - 1_2^2\lambda_1\lambda_2$$

$$= \theta(\theta - \lambda_1 - \lambda_2 + 1_1^2\lambda_1 + 1_2^2\lambda_2) + \lambda_1\lambda_2(1 - 1_1^2 - 1_2^2)$$

$$= \theta(\theta - \lambda_1 - \lambda_2 + 1_1^2\lambda_1 + 1_2^2\lambda_2) + \lambda_1\lambda_2 \cdot 0$$

$$\prod_{j=1}^{2}(\theta-\lambda_j)+\sum_{i=1}^{2}l_i^2\lambda_i\prod_{j\neq i}^{2}(\theta-\lambda_j) = \theta(\theta-\lambda_1-\lambda_2+l_1^2\lambda_1+l_2^2\lambda_2+l_2^2\lambda_1-l_2^2\lambda_1+$$

$$+l_1^2\lambda_2-l_1^2\lambda_2)$$

$$= \theta(\theta-\lambda_1\{1-(l_1^2+l_2^2)\}-\lambda_2\{1-(l_1^2+l_2^2)\}-l_1^2\lambda_2-l_2^2\lambda_1)$$

$$= \theta(\theta-\lambda_1 0-\lambda_2 0-l_1^2\lambda_2-l_2^2\lambda_1)$$

$$= \theta(\theta\{l_1^2+l_2^2\}-l_1^2\lambda_2-l_2^2\lambda_1)$$

$$= \theta(l_1^2\{\theta-\lambda_2\}+l_2^2\{\theta-\lambda_1\})$$

$$= \theta\left[\sum_{i=1}^{2}l_i^2\prod_{j\neq i}^{2}(\theta-\lambda_j)\right]$$

$$\sum_{i=1}^{2}l_i^2\prod_{j\neq i}^{2}(\theta-\lambda_j) = 0$$

2.) Hipotesis de Inducción:

Sea $n = k$: $\prod_{j=1}^{k}(\theta-\lambda_j) + \sum_{i=1}^{k}l_i^2\lambda_i\prod_{j\neq i}^{k}(\theta-\lambda_j)$ con $\sum_{i=1}^{k}l_i^2 = 1$

$$\Longrightarrow \prod_{j=1}^{k}(\theta-\lambda_j) + \sum_{i=1}^{k}l_i^2\lambda_i\prod_{j\neq i}^{k}(\theta-\lambda_j) = \theta\left[\sum_{i=1}^{k}l_i^2\prod_{j\neq i}^{k}(\theta-\lambda_j)\right]$$

Por demostrar:

Para $n = k + 1$, si $\sum_{i=1}^{k+1}l_i^2 = 1$, entonces,

$$\prod_{j=1}^{k+1}(\theta-\lambda_j) + \sum_{i=1}^{k+1}l_i^2\lambda_i\prod_{j\neq i}^{k+1}(\theta-\lambda_j) = \theta\left[\sum_{i=1}^{k+1}l_i^2\prod_{j\neq i}^{k+1}(\theta-\lambda_j)\right]$$

En efecto,

$$\prod_{j=1}^{k+1}(\theta-\lambda_j)+\sum_{i=1}^{k+1}l_i^2\lambda_i\prod_{j\neq i}^{k+1}(\theta-\lambda_j) = (\theta-\lambda_{k+1})\prod_{j=1}^{k}(\theta-\lambda_j) +$$

$$(\theta-\lambda_{k+1})\sum_{i=1}^{k}l_i^2\lambda_i\prod_{j\neq i}^{k}(\theta-\lambda_j) +$$

$$l_{k+1}^2\lambda_{k+1}\prod_{i=1}^{k}(\theta-\lambda_j)$$

$$\Rightarrow \prod_{j=1}^{k+1}(\theta-\lambda_j)+\sum_{i=1}^{k+1} l_i^2\lambda_i \prod_{j\neq i}^{k+1}(\theta-\lambda_j) = (\theta-\lambda_{k+1})\prod_{j=1}^{k}(\theta-\lambda_j)+(\theta-\lambda_{k+1})\sum_{i=1}^{k} l_i^2\lambda_i\cdot$$

$$\prod_{j\neq i}^{k}(\theta-\lambda_j)+l_{k+1}^2\lambda_k \prod_{j\neq k}^{k}(\theta-\lambda_j)(\theta-\lambda_{k+1})-l_{k+1}^2\lambda_k \prod_{j\neq k}^{k}(\theta-\lambda_j)\cdot$$

$$\cdot(\theta-\lambda_{k+1})+l_{k+1}^2\lambda_{k+1} \prod_{j=1}^{k}(\theta-\lambda_j)$$

Si hacemos $l_k^2 + l_{k+1}^2 = l_k^{2*}$ y además $l_i^{2*} = l_i^2$ con $1\leq i\leq k-1$ (*)

se tiene que:

$$\prod_{j=1}^{k+1}(\theta-\lambda_j)+\sum_{i=1}^{k+1} l_i^2\lambda_i \prod_{j\neq i}^{k+1}(\theta-\lambda_j) = (\theta-\lambda_{k+1})\prod_{j=1}^{k}(\theta-\lambda_j)+(\theta-\lambda_{k+1})\cdot$$

$$\sum_{i=1}^{k} l_i^{2*}\lambda_i \prod_{j\neq i}^{k}(\theta-\lambda_j)-l_{k+1}^2\lambda_k \prod_{j\neq k}^{k}(\theta-\lambda_j)(\theta-\lambda_{k+1}) +$$

$$+ l_{k+1}^2\lambda_{k+1} \prod_{j=1}^{k}(\theta-\lambda_j)$$

$$= (\theta-\lambda_{k+1})\Big[\prod_{j=1}^{k}(\theta-\lambda_j)+\sum_{i=1}^{k} l_i^{2*}\lambda_i \prod_{j\neq i}^{k}(\theta-\lambda_j)\Big]-$$

$$- l_{k+1}^2\lambda_k \prod_{j\neq k}^{k+1}(\theta-\lambda_j)(\theta-\lambda_{k+1})+l_{k+1}^2\lambda_{k+1} \prod_{j=1}^{k}(\theta-\lambda_j)$$

$$= (\theta-\lambda_{k+1})\theta\Big[\sum_{i=1}^{k} l_i^{2*}\prod_{j\neq i}^{k}(\theta-\lambda_j)\Big]+l_{k+1}^2 \prod_{j=1}^{k-1}(\theta-\lambda_j)\Big[\lambda_{k+1}(\theta-\lambda_k)-$$

$$-\lambda_k(\theta-\lambda_{k+1})\Big]$$

$$= \theta\Big[\sum_{i=1}^{k} l_i^{2*}\prod_{j\neq i}^{k+1}(\theta-\lambda_j)\Big] + l_{k+1}^2 \prod_{j=1}^{k-1}(\theta-\lambda_j)\theta(\lambda_{k+1}-\lambda_k)$$

$$= \theta\Big[\sum_{i=1}^{k} l_i^2 \prod_{j\neq i}^{k+1}(\theta-\lambda_j)+l_{k+1}^2 \prod_{j\neq k}^{k+1}(\theta-\lambda_j)+l_{k+1}^2 \prod_{j=1}^{k-1}(\theta-\lambda_j)(\lambda_{k+1}-\lambda_k)\Big]$$

Esto último se obtuvo haciendo uso de (*)

Así, tenemos que:

$$\prod_{j=1}^{k+1}(\theta-\lambda_j)+\sum_{i=1}^{k+1} l_i^2\lambda_i \prod_{j\neq i}^{k+1}(\theta-\lambda_j) = \theta\Big[\sum_{i=1}^{k} l_i^2 \prod_{j\neq i}^{k+1}(\theta-\lambda_j)+l_{k+1}^2 \prod_{j=1}^{k-1}(\theta-\lambda_j)\cdot$$

$$\cdot\{(\theta-\lambda_{k+1})+(\lambda_{k+1}-\lambda_k)\}\Big]$$

$$= \theta\Big[\sum_{i=1}^{k} l_i^2 \prod_{j\neq i}^{k+1}(\theta-\lambda_j) + l_{k+1}^2 \prod_{j=1}^{k-1}(\theta-\lambda_j)(\theta-\lambda_k)\Big]$$

$$= \theta\Big[\sum_{i=1}^{k} l_i^2 \prod_{j\neq i}^{k+1}(\theta-\lambda_j) + l_{k+1}^2 \prod_{j=1}^{k}(\theta-\lambda_j)\Big]$$

$$= \theta\Big[\sum_{i=1}^{k+1} l_i^2 \prod_{j\neq i}^{k+1}(\theta-\lambda_j)\Big].$$

Para $n = k+1$, si se cumple que $\sum_{i=1}^{k+1} l_i^2 = 1$, se tiene que:

$$\prod_{j=1}^{k+1}(\theta-\lambda_j) + \sum_{i=1}^{k+1} l_i^2\lambda_i \prod_{j\neq i}^{k+1}(\theta-\lambda_j) = \theta\Big[\sum_{i=1}^{k+1} l_i^2 \prod_{j\neq i}^{k+1}(\theta-\lambda_j)\Big]$$

Luego, de 1.) y 2.) se tiene que (26) se deduce de (25).

Notamos que cuando $l_r = 0$; $\theta - \lambda_r$ es un factor de (26), de modo que $\theta = \lambda_r$ es una solución.

En efecto,

$$\sum_{i=1}^{n} l_i^2 \prod_{j\neq i}^{n}(\theta-\lambda_j) = 0 \quad \text{y si } l_r = 0$$

$$\Longrightarrow \sum_{i=1}^{n} l_i^2 \prod_{j\neq i}^{n}(\theta-\lambda_j) = \sum_{i=1}^{r-1} l_i^2 \prod_{j\neq i}^{n}(\theta-\lambda_j) + \cancelto{0}{l_r^2 \prod_{j\neq r}^{n}(\theta-\lambda_j)} +$$

$$+ \sum_{i=r+1}^{n} l_i^2 \prod_{j\neq i}^{n}(\theta-\lambda_j)$$

$$\Longrightarrow \sum_{i=1}^{n} l_i^2 \prod_{j\neq i}^{n}(\theta-\lambda_j) = \sum_{i=1}^{r-1} l_i^2 \prod_{j\neq i}^{n}(\theta-\lambda_j) + \sum_{i=r+1}^{n} l_i^2 \prod_{j\neq i}^{n}(\theta-\lambda_j)$$

$$= (\theta-\lambda_r)\Big[\sum_{i=1}^{r-1} l_i^2 \prod_{\substack{j\neq i\\ j\neq r}}^{n}(\theta-\lambda_j)+\sum_{i=r+1}^{n} l_i^2 \prod_{\substack{j\neq i\\ j\neq r}}^{n}(\theta-\lambda_j)\Big]$$

$$= (\theta-\lambda_r)\Big[\sum_{i\neq r}^{n} l_i^2 \prod_{\substack{j\neq i\\ j\neq r}}^{n}(\theta-\lambda_j)\Big]$$

Como $\sum_{i=1}^{n} l_i^2 \prod_{j\neq i}^{n}(\theta-\lambda_j) = 0$

$$\Longrightarrow \theta - \lambda_r = 0 \qquad \vee \qquad \sum_{i\neq r}^{n} l_i^2 \prod_{\substack{j\neq i\\ j\neq r}}^{n}(\theta-\lambda_j) = 0$$

$$\Longrightarrow \theta = \lambda_r$$

$\therefore \quad \theta = \lambda_r$ es una solución de (26).

Esto nos indica que cuando P_1 coincide con un vector propio de A; $\theta_1, \theta_2, \ldots\ldots, \theta_{n-1}$ son iguales a las raíces propias asociadas con los restantes n-1 vectores propios de A. De igual forma tenemos que, si P_2 coincide con un vector propio de A, las raíces de M_2M_1A distintas de cero son iguales a las raíces propias asociadas con los restantes n-2 vectores propios de A. En términos generales tenemos que, si los k vectores de regresión P_k coinciden con k de los vectores propios de A, las raíces de $M_k \cdots M_2M_1A$ ó sean $\nu_1, \nu_2, \ldots\ldots, \nu_{n-k}$ son iguales a las raíces propias asociadas con los restantes n - k vectores propios de A.

En efecto;

1.) Supongamos que P_1 coincide con un vector propio de A asociado al valor propio λ_1.

$\Longrightarrow \quad AP_1 = \lambda_1 P_1 \qquad (a)$

Esto implica que P_1 asocia $\theta = 0$ en M_1A; lo que es equivalente a que: $\qquad (M_1A)(P_1) = 0 \qquad (b)$

En efecto,

$$(M_1A)(P_1) = (I - P_1P_1')AP_1 = AP_1 - P_1P_1'AP_1$$

$$= \lambda_1P_1 - P_1P_1'\lambda_1P_1 \qquad \text{por (a)}$$

$$= \lambda_1P_1 - \lambda_1P_1P_1'P_1$$

$$= \lambda_1P_1 - \lambda_1P_1 \qquad \text{pues } P_1'P_1 = 1$$

$$= 0$$

$\Longrightarrow P_1 \in \operatorname{Ker}(M_1A) \iff \theta = 0$

Sea ahora, β un vector propio de A asociado al valor propio $\lambda_1 \Longrightarrow \lambda_1$ es valor propio de M_1A.

i.e. $A\beta = \lambda_1\beta \Longrightarrow \exists\, \gamma : (M_1A)(\gamma) = \lambda_1\gamma$

En efecto,

$$\text{Sea } \gamma = M_1\beta \Longrightarrow (M_1A)(\gamma) = (M_1A)(M_1\beta)$$

$$\Longrightarrow (M_1A)\gamma = (M_1A)(I - P_1P_1')\beta = M_1A\beta - M_1AP_1P_1'\beta$$

$$= M_1\lambda_1\beta - 0P_1'\beta \qquad \text{por (b)}$$

$$= \lambda_1M_1\beta = \lambda_1\gamma$$

2.) Supongamos ahora que, P_2 es un vector propio de A asociado al valor propio $\lambda_2 \Longrightarrow AP_2 = \lambda_2 P_2$

$$\Longrightarrow \begin{cases} (M_2M_1A)P_1 = 0 \\ (M_2M_1A)P_2 = 0 \end{cases} \qquad \text{(c)}$$

En efecto,

$$(M_2M_1A)(P_1) = M_2(M_1AP_1) = M_2 0 = 0$$

$\Longrightarrow \quad P_1 \in Ker(M_2M_1A)$

Ademas,

$$\begin{aligned}(M_2M_1A)(P_2) &= (I - P_2P_2' - P_1P_1')AP_2 \\ &= AP_2 - P_2P_2'AP_2 - P_1P_1'AP_2 \\ &= \lambda_2P_2 - P_2P_2'\lambda_2P_2 - P_1P_1'\lambda_2P_2 \\ &= \lambda_2P_2 - \lambda_2P_2P_2'P_2 - \lambda_2P_1P_1'P_2 \\ &= \lambda_2P_2 - \lambda_2P_2 \qquad \text{pues } P_2'P_2 = 1 \quad P_1'P_2 = 0 \\ &= 0\end{aligned}$$

$\Longrightarrow \quad P_2 \in Ker(M_2M_1A)$

Sea ahora, β_1 vector propio de A asociado al valor propio λ_2

o sea que $A\beta_1 = \lambda_2\beta_1$

$\Longrightarrow \lambda_2$ es valor propio de M_2M_1A

i.e. $\exists\ \beta_1' : (M_2M_1A)(\beta_1') = \lambda_2\beta_1'$

En efecto, sea $\beta_1' = M_2M_1\beta_1$

$\Longrightarrow \quad (M_2M_1A)(\beta_1') = (M_2M_1A)(M_2M_1\beta_1) = (M_2M_1A)(I-P_2P_2'-P_1P_1')\beta_1$

$$\Longrightarrow \quad (M_2M_1A)(\beta_1') = M_2M_1A\beta_1-(M_2M_1AP_2)P_2'\beta_1-(M_2M_1AP_1)P_1'\beta_1$$

$$= M_2M_1\lambda_2\beta_1 - 0P_2'\beta_1 - 0P_1'\beta_1 \qquad \text{por (c)}$$

$$= \lambda_2M_2M_1\beta_1$$

$$= \lambda_2\beta_1'$$

Así, en general si los k vectores de regresión coinciden con k vectores de A, las restantes raíces $\nu_1,\nu_2,\nu_3,\ldots\ldots\ldots,\nu_{n-k}$ coinciden con las restantes raíces de A.

b.) Supongamos que s vectores de regresión coinciden con los vectores propios de A o son combinaciones lineales de ellos. Si reenumeramos las raíces de modo que:

$$\lambda_1 \leq \lambda_2 \leq \cdots\cdots\cdots\cdots \leq \lambda_{n-s}$$

Procederemos a demostrar que,

$$\lambda_i \leq \nu_i \leq \lambda_{i+k-s} \qquad i = 1,\ 2,\ldots\ldots,\ n-k$$

Establecemos primero un resultado análogo para el conjunto completo de los ν's y λ's y ordenamos los ν's, λ's y θ's de modo que,

$$\nu_1 \leq \nu_2 \leq \cdots\cdots\cdots\cdots \leq \nu_{n-k}$$

$$\lambda_1 \leq \lambda_2 \leq \cdots\cdots\cdots\cdots \leq \lambda_n$$

$$\theta_1 \leq \theta_2 \leq \cdots\cdots\cdots\cdots \leq \theta_{n-1}$$

Hemos visto que si $l_r = 0$, entonces λ_r es la raíz de (26), además, si λ_r y λ_{r+1} son iguales, tenemos que $\lambda_r = \lambda_{r+1}$ es una raíz de

(26). Estos son los dos únicos casos en los cuales los λ_i son raíces de (26).

Para las restantes raíces, consideremos la función:

$$f(\theta) = \sum_{i=1}^{n} l_i^2 \prod_{j\neq i}^{n} (\theta - \lambda_j)$$

$$\Longrightarrow \qquad f(\lambda_r) = l_r^2 \prod_{j\neq r}^{n} (\lambda_r - \lambda_j) \qquad (27)$$

además,

$$\text{si} \quad f(\lambda_r) > 0 \Longrightarrow f(\lambda_{r+1}) \leq 0 \qquad (28)$$

$$\wedge$$

$$\text{si} \quad f(\lambda_r) < 0 \Longrightarrow f(\lambda_{r+1}) \geq 0 \qquad (29)$$

La implicación (27) es evidente, demostraremos la condición (28) y de manera análoga se demuestra la condición (29).

En efecto, sabemos que,

$$\lambda_1 \leq \lambda_2 \leq \cdots\cdots\cdots \leq \lambda_{r-1} \leq \lambda_r \leq \lambda_{r+1} \leq \cdots\cdots\cdots \leq \lambda_n \qquad (*)$$

Si $f(\lambda_r) = l_r^2 \prod_{j\neq r}^{n} (\lambda_r - \lambda_j) > 0$ se tiene que:

$$f(\lambda_{r+1}) = l_{r+1}^2 \prod_{j\neq r+1}^{n} (\lambda_{r+1} - \lambda_j)$$

Vamos a verificar que: $f(\lambda_{r+1}) \leq 0$

Sabemos que $\forall$ j: $1 \leq j \leq k-1$; $(\lambda_r - \lambda_j) > 0$ por (*) y además, que $f(\lambda_r) = l_r^2 \prod_{j=1}^{r-1} (\lambda_r - \lambda_j) \prod_{j=r+1}^{n} (\lambda_r - \lambda_j) > 0$

$$\Longrightarrow \quad \prod_{j=1}^{r-1} (\lambda_r - \lambda_j) > 0$$

además, como $l_r^2 \geq 0$ y $f(\lambda_r) > 0$ tenemos que:

$$\prod_{j=r+1}^{n} (\lambda_r - \lambda_j) > 0$$

Pero, como $\forall$ j: $r + 1 \leq j \leq n$; $(\lambda_r - \lambda_j) < 0$, entonces se tiene que $\prod_{j=r+1}^{n}(\lambda_r - \lambda_j)$ es el producto de un número par de factores negativos, es decir, $n - r = n - [(r + 1) - 1]$ **es par.**

Por otro lado,

$$f(\lambda_{r+1}) = 1^2_{r+1} \prod_{j=1}^{r}(\lambda_{r+1} - \lambda_j)\prod_{j=r+2}^{n}(\lambda_{r+1} - \lambda_j)$$

donde $\forall$ j: $1 \leq j \leq r$; $(\lambda_{r+1} - \lambda_j) > 0$ por (*)

$$\Longrightarrow \prod_{j=1}^{r}(\lambda_{r+1} - \lambda_j) > 0 \qquad \text{(a)}$$

pero, como $\forall$ j: $r + 2 \leq j \leq n$; $(\lambda_{r+1} - \lambda_j) < 0$ por (*) y como $n - [(r + 1) - 1]$ es par, entonces , se tiene que, $n - (r + 1) = n - [(r + 2) - 1]$ es impar y por lo tanto,

$$\prod_{j=r+2}^{n}(\lambda_{r+1} - \lambda_j) \leq 0 \qquad \text{y} \quad 1^2_{r+1} \geq 0 \qquad \text{(b)}$$

$$\Longrightarrow f(\lambda_{r+1}) = 1^2_{r+1} \prod_{j=1}^{r}(\lambda_{r+1}-\lambda_j)\prod_{j=r+2}^{n}(\lambda_{r+1}-\lambda_j) \leq 0$$

$$\therefore f(\lambda_{r+1}) \leq 0$$

Puesto que $f(\theta)$ es contínua, exíste una raíz en cada intervalo

$$\lambda_r \leq \theta \leq \lambda_{r+1}$$

Así: $\lambda_i \leq \theta_i \leq \lambda_{i+1} \quad \forall$ i: i = 1, 2, ..., n-1 (30)

Para extender este resultado, observemos que M_1A tiene una raíz cero en adición a $\theta_1, \theta_2, \ldots\ldots\ldots\ldots, \theta_{n-1}$.

Supongamos que: $\theta_t \leq 0 \leq \theta_{t+1}$

entonces, las raíces de M_1A pueden ser ordenadas en la siguiente forma:

$$\theta_1 \le \theta_2 \le \cdots\cdots \le \theta_t \le 0 \le \theta_{t+1} \le \cdots\cdots \le \theta_{n-1}$$

Sean ahora $\phi_1, \phi_2, \ldots\ldots\ldots, \phi_{n-1}$ las raíces de $M_2(M_1A)$ junto con una raíz cero,

$$\Longrightarrow \quad \theta_1 \le \phi_1 \le \theta_2 \le \phi_2 \le \cdots\cdots \le \theta_t \le \phi_t \le 0 \le \phi_{t+1} \le \cdots\cdots\cdots \le \phi_{n-1} \le \theta_{n-1}$$

pero, como M_2M_1A tiene dos raíces cero y tiene rango a lo sumo de n-2, entonces, uno de los dos ϕ_t v ϕ_{t+1} debe ser cero. Rechazando una de ellas y reenumerando, se tiene que,

$$\lambda_i \le \phi_i \le \lambda_{i+2} \quad \forall\ i:\ i = 1, 2, \ldots\ldots, n-2$$

Si aplicamos el mismo argumento de manera sucesiva, tendremos que,

$$\lambda_i \le \nu_i \le \lambda_{i+k} \quad \forall\ i:\ i = 1, 2, \ldots\ldots\ldots, n-k$$

Considerando el caso en que s vectores de regresión coinciden con los vectores propios de A o combinaciones lineales de ellos tendremos, $\lambda_i \le \nu_i \le \lambda_{i+k-s} \quad \forall\ i:\ i = 1, 2, \ldots\ldots\ldots, n-k$ (31)

COROLARIO 1:

$r_L \le r \le r_U$ donde

$$r_L = \frac{\sum_{i=1}^{n-k} \lambda_i \zeta_i^2}{\sum_{i=1}^{n-k} \zeta_i^2} \quad \text{y} \quad r_U = \frac{\sum_{i=1}^{n-k} \lambda_{i+k-s} \zeta_i^2}{\sum_{i=1}^{n-k} \zeta_i^2} \tag{32}$$

PRUEBA:

Sabemos que $\lambda_i \le \nu_i \le \lambda_{i+k-s} \quad \forall\ i:\ i = 1, 2, \ldots\ldots\ldots, n-k$

Así, $\lambda_i \zeta_i^2 \le \nu_i \zeta_i^2 \le \lambda_{i+k-s} \zeta_i^2 \quad \forall\ i:\ i = 1, 2, \ldots\ldots, n-k$

$$\Longrightarrow \sum_{i=1}^{n-k} \lambda_i \zeta_i^2 \le \sum_{i=1}^{n-k} \nu_i \zeta_i^2 \le \sum_{i=1}^{n-k} \lambda_{i+k-s} \zeta_i^2$$

$$\Longrightarrow \frac{\sum_{i=1}^{n-k} \lambda_i \zeta_i^2}{\sum_{i=1}^{n-k} \zeta_i^2} \leq \frac{\sum_{i=1}^{n-k} \nu_i \zeta_i^2}{\sum_{i=1}^{n-k} \zeta_i^2} \leq \frac{\sum_{i=1}^{n-k} \lambda_{i+k-s} \zeta_i^2}{\sum_{i=1}^{n-k} \zeta_i^2}$$

es decir; $r_L \leq r \leq r_U$

La importancia de este resultado es que éste conjunto de límites de r no depende del conjunto de vectores de regresión

B- DISTRIBUCION DE r

Hemos señalado que cuando los errores están distribuidos independientemente de las variables explicativas, estas variables pueden considerarse como fijas. Sin embargo hay un caso especial en el cual es más conveniente considerar las $\{X_i\}$ como variables aleatorias.

1.) REGRESION SOBRE VARIABLES ALEATORIAS:

Consideraremos el caso de un sistema normal multivariante. En tal sistema las regresiones son lineales y los errores están distribuidos independientemente de las variables independientes.

Se puede demostrar que si Y, X_1, X_2,.........., X_n están distribuidas conjuntamente como normales tal que la regresión de Y sobre las $\{X_i\}$ pase por el origen y si observaciones sucesivas son independientes, entonces, r se distribuye como si los residuos $Z_1,\ldots,Z_n$ fuesen variables normales independientes o sea que el efecto de regresión desaparece del problema. De igual forma, cuando la regresión no pasa por el origen, r se distribuye como si las $\{Z_i\}$ fuesen residuos de la media muestral de n observaciones normales independientes.

Como en la práctica casi nunca se prueba la hipotesis de correlación serial de los errores cuando se sabe que observaciones sucesivas de las $\{X_i\}$ son independientes, es por ello que el caso planteado anteriormente no es el más importante.

M. S., Bartlett en 1934 demostró que si las $\{X_i\}$ no se consideran fijas, Z estará dirigida aleatoriamente en el n-espacio, si y solo si, las $\{X_i\}$ están distribuidamente como normales y sus observaciones sucesivas son independientes. En este caso, la dirección de Z está distribuida como si $Z_1, Z_2, \ldots\ldots\ldots, Z_n$ fuesen normales e independientes con la misma varianza. De igual forma se puede demostrar que si ajustamos una regresión incluyendo un término constante, r estará distribuida como si las $\{Z_i\}$ fuesen residuos de una media muestral de variables normales e independientes, aunque la regresión de la población pase o no pase por el origen.

2.) REGRESION SOBRE VARIABLES FIJAS:

Para examinar la distribución de r cuando las variables explicativas son fijas, asumiremos que los errores $\varepsilon_1, \varepsilon_2, \ldots\ldots\ldots\ldots, \varepsilon_n$ son normales independientes con varianza constante.
Transformando r como se hizo en I-A se tiene que:

$$r = \frac{\sum_{i=1}^{n-k} \nu_i \zeta_i^2}{\sum_{i=1}^{n-k} \zeta_i^2}$$

Como la transformación que se hizo es ortogonal, $\zeta_1, \zeta_2, \ldots\ldots, \zeta_{n-k}$ son variables independientes $N(0, \sigma^2)$ donde la variación de r está limitada al rango (ν_1, ν_{n-k}).

Bajo el supuesto de que las $\{\nu_i\}$ son desconocidas, R. L., Anderson [9] en 1942, dió la distribución exacta de r para dos casos especiales que son:

1.) Para n - k par, las $\{\nu_i\}$ siendo iguales a pares.
2.) Para n - k impar, siendo las $\{\nu_i\}$ iguales a pares con un valor mayor o menor que todos los otros.

Las expresiones para la Función de Distribución de r encontradas por Anderson fueron las siguientes:

$$P(r > r') = \sum_{i=1}^{m} \frac{(\tau_i - r')^{1/2(n-k)-1}}{\alpha_i} \quad (\tau_{m+1} \leq r' \leq \tau_m) \qquad (33)$$

donde:

a.) Para n - k par; Las $\{\nu_i\}$ forman 1/2(n-k) pares distintos denotados por $\tau_1 > \tau_2 > \cdots\cdots > \tau_{1/2(n-k)}$ y

$$\alpha_i = \prod_{i \neq j}^{1/2(n-k)} (\tau_i - \tau_j) \qquad (34)$$

b.) Para n - k impar; las $\{\nu_i\}$ forman 1/2(n-k-1) pares distintos simultaneamente como anteriormente con una raíz aislada τ menor que todas las otras y

$$\alpha_i = \prod_{j \neq i}^{1/2(n-k-1)} (\tau_i - \tau_j)\sqrt{(\tau_i - \tau_j)} \qquad (35)$$

Aquí en (b) $\tau > \tau_1$ se obtiene escribiendo $-r$ por r.

Anderson también encontró las fórmulas para la Función de Densidad.

Para el caso en el cual las $\{\nu_i\}$ son todas diferentes y n-k es par, la [1/2(n-k)-1]-ésima derivada de la Función de Densidad fué encontrada por Von Neumann en 1942, obteniendo: |10|

$$\frac{d^{1/2(n-k)-1}}{d_r^{1/2(n-k)-1}} f(r) = 0 \qquad \text{si m es par}$$

$$(36)$$

$$\text{y} \qquad \frac{d^{1/2(n-k)-1}}{d_r^{1/2(n-k)-1}} f(r) = \frac{(-1)^{1/2(n-k-m-1)}\left[(n-k)/2 - 1\right]!}{\pi\sqrt{-\prod_{i=1}^{n-k}(r - \nu_i)}}$$

si m es impar; con $\nu_m < r < \nu_{m+1}$: m = 1, 2, ..., n-k-1

Estos resultados pueden ser usados en casos particulares si las $\{\nu_i\}$ son conocidas. Lo que significa que los vectores de regresión deben ser vectores propios de A y las raíces asociadas con los restantes n-k vectores propios de A deben cumplir las condiciones de Anderson o de Von Neumann.

Los anteriores resultados pueden ser aplicados a la distribución de r_L y r_U, aportando las apropiadas $\{\lambda_i\}$.

Si F_L y F_U son las Funciones de Distribución de r_L y r_U respectivamente, se puede probar que:

$$F_L(r) \geq F(r) \geq F_U(r) \tag{37}$$

lo cual se obtiene notando que r_L y r son correspondencias en (30) y que $r_L \leq r$ siempre.

C- MOMENTOS DE r

El análisis anterior muestra que r es independiente de la escala de los $\{\zeta_i\}$ luego, podemos tomar a σ^2 igual a la unidad y r podemos expresarla como:

$$r = \frac{u}{v} \quad \text{donde} \quad u = \sum_{i=1}^{n-k} \nu_i \zeta_i^2 \quad \text{y} \quad v = \sum_{i=1}^{n-k} \zeta_i^2 \tag{38}$$

donde $\forall$ i: $i = 1, 2, \ldots\ldots, n-k$; $\zeta_i \sim N(0,1)$ independientes.

Probaremos que r y v se distribuyen independientemente; para ello primero partiremos de un teorema, el cual no será demostrado puesaparece demostrado en |1|.

TEOREMA 2:

Sea $X_m \sim N_m(0,I)$ y $r = ||X|| = (X'X)^{1/2}$, $T(X) = \dfrac{X}{||X||}$

entonces, r y $T(X)$ se distribuyen independientemente.

COROLARIO 2:

Sea $\zeta = (\zeta_i)_{1\leq i\leq n-k} \sim N_{n-k}(0,I)$, sean ademas, $r = \dfrac{u}{v}$

con $u = \sum_{i=1}^{n-k} \nu_i \zeta_i^2 = \zeta' D\zeta$ y $v = \sum_{i=1}^{n-k} \zeta_i^2 = ||\zeta||^2$

entonces, r y v son independientes.

DEMOSTRACION:

Sustituyendo ζ por X en el teorema 2, tendremos que $r = ||\zeta||$ y ademas que $T(X) = T(\zeta) = \zeta/||\zeta||$, entonces, r y $T(\zeta)$ son independientes.

Sea $f(r) = ||\zeta||$ y $g(T(\zeta)) = \zeta'D\zeta$ funciones de r y $T(\zeta)$ respectivamente, entonces tendremos que $f(r)$ y $g(T(\zeta))$ son independientes. i.e. r y v son independientes.

Así, se tiene que:

$$E(u^s) = E(r^s v^s) = E(r^s)E(v^s)$$

$$\Longrightarrow \quad E(r^s) = \frac{E(u^s)}{E(v^s)} \tag{39}$$

Como $\forall$ i: i = 1, 2,........., n-k; $\zeta_i \sim N(0,1)$ independientes, entonces, $\forall$ i: i = 1, 2,..........., n-k; $\zeta_i^2 \sim \chi_1^2$ independientes, entonces,

$$v = \sum_{i=1}^{n-k} \zeta_i^2 \sim \chi^2_{n-k} \Longrightarrow \begin{cases} E(v) = n - k \\ E(v^2) = (n - k)(n - k + 2) \\ \text{etc.} \end{cases} \tag{40}$$

Por otro lado como $\forall$ i: i = 1, 2,..........., n-k; ν_i es una constante y además, $\zeta_i^2 \sim \chi_1^2$ independientes, entonces, $\nu_i\zeta_i^2$ son variables independientes **y** como,

$$u = \sum_{i=1}^{n-k} \nu_i\zeta_i^2$$

se tiene que el n-ésimo cumulante de $\nu_i\zeta_i^2$ está dado por:

$$K_s(\nu_i\zeta_i^2) = 2^{s-1}\ (s - 1)!\,\nu_i^s \tag{41}$$

$$K\left(\sum_{i=1}^{n-k}\nu_i\zeta_i^2\right) = \sum_{i=1}^{n-k}\left[2^{s-1}(s-1)!\,\nu_i^s\right] = 2^{s-1}(s-1)!\sum_{i=1}^{n-k}\nu_i^s$$

En particular se tiene que:

$$
\begin{aligned}
&(1.)\quad K_1(u) = 2^{1-1}(1-1)!\sum_{i=1}^{n-k}\nu_i = \sum_{i=1}^{n-k}\nu_i \\
&(2.)\quad K_2(u) = 2^{2-1}(2-1)!\sum_{i=1}^{n-k}\nu_i^2 = 2\sum_{i=1}^{n-k}\nu_i^2 \\
&(3.)\quad K_3(u) = 2^{3-1}(3-1)!\sum_{i=1}^{n-k}\nu_i^3 = 8\sum_{i=1}^{n-k}\nu_i^3 \\
&(4.)\quad K_4(u) = 2^{4-1}(4-1)!\sum_{i=1}^{n-k}\nu_i^4 = 48\sum_{i=1}^{n-k}\nu_i^4
\end{aligned}
\qquad (42)
$$

etc.

Luego,

$$E(r) = \frac{E(u)}{E(v)} = \mu_1' = \frac{1}{n-k}\sum_{i=1}^{n-k}\nu_i = \bar{\nu}$$

$$E(r^2) = \frac{E(u^2)}{E(v^2)} = \mu_2' = \frac{2}{(n-k)(n-k+2)}\sum_{i=1}^{n-k}\nu_i^2$$

Para obtener los momentos de r con respecto a su media E(r), se tiene que:

$$r - \mu_1' = \frac{\sum_{i=1}^{n-k}(\nu_i - \bar{\nu})\zeta_i^2}{\sum_{i=1}^{n-k}\zeta_i^2} = \frac{u'}{v} \qquad (43)$$

Donde, nuevamente se tiene que, $(r - \mu_1')$ y v son independientes, por lo que, los momentos de $r - \mu_1'$ son iguales a los momentos de u' divididos entre los momentos de v.

Obteniendo los momentos de u' usando los cumulantes se tiene que:

$$K_s(u') = 2^{s-1}(s-1)!\sum_{i=1}^{n-k}(\nu_i - \bar{\nu})^s \qquad (44)$$

En particular, tendremos que:

$$
\begin{aligned}
&(1.)\quad K_1(u') = 2^{1-1}(1-1)!\sum_{i=1}^{n-k}(\nu_i - \bar{\nu})^1 = \sum_{i=1}^{n-k}(\nu_i - \bar{\nu}) \\
&(2.)\quad K_2(u') = 2^{2-1}(2-1)!\sum_{i=1}^{n-k}(\nu_i - \bar{\nu})^2 = 2\sum_{i=1}^{n-k}(\nu_i - \bar{\nu})^2
\end{aligned}
$$

Luego, se tiene que:

$$(1.)\quad Var(r) = \sigma^2 = \frac{2\sum_{i=1}^{n-k}(\nu_i - \bar{\nu})^2}{(n-k)(n-k+2)}$$

$$(2.)\quad m_3 = \frac{8\sum_{i=1}^{n-k}(\nu_i - \bar{\nu})^3}{(n-k)(n-k+2)(n-k+4)} \qquad (44a)$$

$$(3.)\quad m_4 = \frac{48\sum_{i=1}^{n-k}(\nu_i - \bar{\nu})^4 + 12\{\sum_{i=1}^{n-k}(\nu_i - \bar{\nu})^2\}^2}{(n-k)(n-k+2)(n-k+4)(n-k+6)}$$

son los Momentos Centrados de r ; los cuales se refieren a la regresión hacia el origen, de k variables aleatorias independientes.

Si el modelo estudiado incluye un término constante o sea, que si la regresión calculada incluye una media fija y si como es usual, esperamos distinguir las restantes variables independientes, entonces, $k = k' + 1$ en la expresión anterior, siendo k' el número de variables independientes adicionadas a la constante.

Las expresiones dadas anteriormente nos permiten calcular los momentos de r cuando los $\{\nu_i\}$ son conocidos pero, como en la práctica esto no siempre sucede, entonces, para estos casos no es posible calcular los momentos de r mediante el procedimiento anterior. Bajo esta última situación debemos expresar la suma de potencias $\Sigma\, \nu_i^s$ mediante cantidades conocidas o sea mediante la matríz A y las variables independientes.

Para poder calcular los momentos de r , bajo el hecho de que no se conocen los valores de los $\{\nu_i\}$ haremos uso del concepto de traza de una matríz y de algunas propiedades importantes de la misma.

DEFINICION

Sea S una Matríz Cuadrada, se define la traza de S, que se denota por tr(S), como la suma de los elementos de su diagonal principal.

i.e. $\forall\ S = (s_{ij}) \in M_{nxn}$; $tr(S) = \sum_{i=1}^{n} s_{ii}$

PROPIEDADES :

1.) $\forall\ S \wedge T \in M_{nxn}$; $tr(S + T) = tr(S) + tr(T)$ (45)

2.) $\forall\ S \wedge T \in M_{nxn}$; $tr(T \times S) = tr(S \times T)$

TEOREMA 3:

Sean S y T dos Matrices Cuadradas y sea $g \in \mathbb{N}$, entonces,

$$tr\{(S+T)^g\} = tr(S^g) + \binom{g}{1}tr(S^{g-1}T) + \binom{g}{2}tr(S^{g-2}T^2) + \ldots + tr(T^g)$$

Haciendo uso de los conceptos y propiedades enunciadas arriba tenemos que, si $\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_m$ son las raices propias de S, entonces, $tr(S) = \sum_{i=1}^{m}\sigma_i$ y que en general se tiene que $tr(S^q) = \sum_{i=1}^{m}\sigma_i^q$

De lo anterior, para nuestro caso tenemos que:

$$tr(M \cdot A)^q = \sum_{i=1}^{n-k}\nu_i^q \tag{46}$$

puesto que $\nu_1, \nu_2, \nu_3, \ldots, \nu_{n-k}$, en conjunto con k ceros son las raíces propias de M·A.

Sabemos que, $M = I_n - X(X'X)^{-1}X'$ luego, en los casos en los cuales las variables independientes son constantes conocidas, algunas veces es posible construir la matríz M directamente y así, obtener la media y la varianza de r en forma bastante directa, pero, en los modelos en los cuales las variables independientes toman valores arbitrarios se hace necesario efectuar una reducción, así:

a.) Para poder calcular la media de r se hace necesario efectuar la siguiente reducción:

$$\begin{aligned}\sum_{i=1}^{n-k}\nu_i = tr(M \cdot A) &= tr\left[\{I - X(X'X)^{-1}X'\}A\right] \\ &= tr\{A - X(X'X)^{-1}X'A\} \\ &= tr(A) - tr\left[X(X'X)^{-1}X'A\right] \quad \text{por P. 1} \\ &= tr(A) - tr\left[X'AX(X'X)^{-1}\right] \quad \text{por P. 2}\end{aligned} \tag{47}$$

Para el cálculo de la varianza:

$$\sum_{i=1}^{n-k} \nu_i^2 = tr[(MA)^2] = tr[(\{I - X(X'X)^{-1}X'\}A)^2]$$

$$= tr[\{A - X(X'X)^{-1}X'A\}^2]$$

$$= tr(A^2) - 2tr[X'A\,X(X'X)^{-1}] + tr\{X'AX(X'X)^{-1}\}^2 \qquad (48)$$

De igual forma se encuentran expresiones equivalentes para $\Sigma\nu_i^3$, $\Sigma\nu_i^4$, etc.

Si las variables independientes son ortogonales, se tiene que X'X es una matríz diagonal; i.e. las expresiones anteriores pueden ser simplificadas así:

$$tr[X'AX(X'X)^{-1}] = \sum_{i=1}^{k} \frac{x_i'Ax_i}{x_i'x_i} \qquad (49)$$

donde x_i representa al vector de la muestra de la i-ésima variable independiente. Se observa que cada término de la sumatoria tiene la forma de r en término de una de sus variable independientes. De manera similar se encuentran expresiones para:

$$tr[X'A^2X(X'X)^{-1}] = \sum_{i=1}^{k} \frac{x_i'A^2x_i}{x_i'x_i} , \quad \text{etc.} \qquad (50)$$

Así, se tiene que:

$$tr[(X'AX(X X)^{-1})^2] = \sum_{i=1}^{k}\left[\frac{x_i'Ax_i}{x_i'x_i}\right]^2 + 2\sum_{i\neq j}^{k}\frac{(x_i'Ax_j)^2}{x_i'x_ix_j'x_j}$$

Por lo tanto, cuando los vectores de regresión son ortogonales, se tiene que:

$$\Sigma\nu_i = tr(A) - \sum_{i=1}^{k}\frac{x_i'Ax_i}{x_i'x_i}$$

$$\Sigma\nu_i^2 = tr(A^2) - 2\sum_{i=1}^{k}\frac{x_i'A^2x_i}{x_i'x_i} + \sum_{i=1}^{k}\left(\frac{x_i'Ax_i}{x_i'x_i}\right)^2 + 2\sum_{i\neq j}^{k}\frac{(x_i'Ax_j)^2}{x_i'x_ix_j'x_j} \qquad (51)$$

De estos resultados se tiene que:

1.) La Media de r está dada por:

$$E(r) = \frac{1}{n-k}\sum_{i=1}^{n-k}\nu_i = \frac{1}{n-k}\left[tr(A) - \sum_{i=1}^{k}\frac{x_i'Ax_i}{x_i'x_i}\right] \tag{52}$$

2.) La Varianza de r está dada por:

$$Var(r) = \frac{2}{(n-k)(n-k+2)}\sum_{i=1}^{n-k}(\nu_i-\bar{\nu})^2 = \frac{2}{(n-k+2)}\left[\frac{1}{n-k}\sum_{i=1}^{n-k}\nu_i^2 - \bar{\nu}^2\right]$$

$$= \frac{2}{(n-k)(n-k+2)}\left[tr(A^2) - 2\sum_{i=1}^{k}\frac{x_i'A^2x_i}{x_i'x_i} + \sum_{i=1}^{k}\left(\frac{x_i'Ax_i}{x_i'x_i}\right)^2 + \right.$$

$$\left. + 2\sum_{i\neq j}^{k}\frac{(x_i'Ax_j)^2}{x_i'x_ix_j'x_j}\right] - \frac{2\bar{\nu}^2}{(n-k+2)} \tag{53}$$

Cuando X es particionado en dos ó más conjuntos de variables ortogonales, podemos aplicar resultados similares.

D- FUNCION CARACTERISTICA DE u y v

Sabemos que un método alterno para obtener los Momentos de una variable aleatoria es atraves del uso de su Función Característica, la cual siempre existe. Usaremos este hecho para encontrar la Función Característica conjunta de u y v.

$$\Phi_{u,v}(t_1,t_2) = E[\exp\{it_1u + it_2v\}]$$

$$= E[\exp\{it_1u(\zeta_1,\ldots,\zeta_{n-k}) + it_2v(\zeta_1,\ldots\ldots,\zeta_{n-k})\}]$$

$$= \int\exp\{it_1u(\zeta_1,\ldots,\zeta_{n-k})+it_2v(\zeta_1,\ldots,\zeta_{n-k})\}f(\zeta_1,\ldots,\zeta_{n-k})\cdot$$

$$\cdot d\zeta_1 d\zeta_2 \cdots\cdots d\zeta_{n-k}$$

Sabemos que $\forall$ j: $1 \leq j \leq n-k$; $\zeta_j \sim N(0,1)$ independientes, entonces,

$$f(\zeta_1,\zeta_2,\ldots,\zeta_{n-k}) = (2\pi)^{-(n-k)/2} \exp\{-\sum_{j=1}^{n-k}\zeta_j^2/2\}$$

$$\Longrightarrow \Phi_{u,v}(t_1,t_2) = (2\pi)^{-(n-k)/2} \int\cdots\int \exp\Big[it_1\sum_{j=1}^{n-k}\nu_j\zeta_j^2+it_2\sum_{j=1}^{n-k}\zeta_j^2 -$$

$$-\sum_{j=1}^{n-k}\zeta_j^2\ /2\Big]d\zeta_1 d\zeta_2\cdots\cdots d\zeta_{n-k}$$

$$= (2\pi)^{-(n-k)/2}\int\cdots\cdots\int\exp\{-\sum_{j=1}^{n-k}\zeta_j^2(1-2it_1\nu_j-2t_2i)/2\}d\zeta_1\cdot$$

$$\cdot d\zeta_2\cdots\cdots\cdots d\zeta_{n-k}$$

Si hacemos $\sigma_j = \zeta_j(1-2it_1\nu_j-2t_2i)^{1/2} \Longrightarrow \begin{cases}\sigma_j^2 = \zeta_j^2(1-2it_1\nu_j-2t_2i) \\ \wedge \\ d\sigma_j = (1-2it_1\nu_j-2t_2i)^{1/2}d\zeta_j\end{cases}$

$$\Longrightarrow (1 - 2it_1\nu_j - i2t_2)^{-1/2}\ d\sigma_j = d\zeta_j$$

$$\Longrightarrow \Phi_{u,v}(t_1,t_2) = \prod_{j=1}^{n-k}(1-2it_1\nu_j-2t_2i)^{-1/2}\ (2\pi)^{-(n-k)/2} \int\cdots\int\cdots\int$$

$$\int\exp\{-\sum_{j=1}^{n-k}\sigma_j^2/2\}\ d\sigma_1\cdot d\sigma_2\cdots\cdots\cdots\cdots d\sigma_{n-k}$$

$$= \prod_{j=1}^{n-k}(1 - 2it_1\nu_j - 2it_2)^{-1/2}$$

$$= (1-2it_2)^{1/2k}\prod_{j=1}^{n-k}(1-2it_1\nu_j-2it_2)^{-1/2}(1-2it_2)^{-1/2k}$$

$$= (1-2it_2)^{1/2k}\prod_{j=1}^{n-k}\Big[(1-2it_2)-2it_1\nu_j\Big](1-2it_2)^{k}\Big]^{-1/2}$$

$$= (1-2it_2)^{1/2k}\Big[\mathbb{I}_n(1-2it_2) - 2it_1MA\Big]^{-1/2}$$

ya que, ν_1, $\nu_2,\ldots\ldots,\nu_{n-k}$ en conjunto con k ceros son las raíces de la ecuación

$$|\mathbb{I}_n\nu - M\cdot A| = 0$$

Si consideramos el caso de una simple variable independiente, podemos obtener una expresión mucho más manejable que la anterior. Bajo estas condiciones, la Función Característica de u y v estará dada por:

$$\Phi_1(t_1,t_2) = \prod_{j=1}^{n-1} (1 - 2\theta_j it_1 - 2it_2)^{-1/2}$$

$$\Longrightarrow \frac{1}{\Phi_1^2} = \prod_{j=1}^{n-1}(1 - 2\theta_j it_1 - 2it_2) \tag{54}$$

donde $\theta_1, \theta_2, \ldots\ldots\ldots, \theta_{n-1}$ son las raíces de la ecuación:

$$\sum_{i=1}^{n} l_i^2 \prod_{j\neq i}^{n} (\theta - \lambda_j) = 0$$

por consiguiente,

$$\begin{aligned}
\prod_{j=1}^{n-1}(\theta - \theta_j) &= \sum_{i=1}^{n} l_i^2 \prod_{j\neq i}^{n}(\theta - \lambda_j) \\
&= l_1^2\prod_{j\neq 1}^{n}(\theta-\lambda_j)+l_2^2\prod_{j\neq 2}^{n}(\theta-\lambda_j)+\ldots\ldots+l_n^2\prod_{j=1}^{n-1}(\theta-\lambda_j) \\
&= \frac{l_1^2\prod_{j=1}^{n}(\theta-\lambda_j)}{(\theta - \lambda_1)}+\frac{l_2^2\prod_{j=1}^{n}(\theta-\lambda_j)}{(\theta - \lambda_2)}+\ldots\ldots+\frac{l_n^2\prod_{j=1}^{n}(\theta-\lambda_j)}{(\theta - \lambda_n)} \\
&= \prod_{j=1}^{n}(\theta-\lambda_j)\left[\frac{l_1^2}{\theta-\lambda_1} + \frac{l_2^2}{\theta-\lambda_2} + \cdots\cdots + \frac{l_n^2}{\theta-\lambda_n}\right] \\
&= \prod_{j=1}^{n}(\theta - \lambda_j)\sum_{j=1}^{n} \frac{l_j^2}{\theta - \lambda_j}
\end{aligned}$$

para todos los valores de θ excepto para $\lambda_1, \lambda_2, \ldots\ldots\ldots, \lambda_n$.

Luego, de (54) se tiene que:

$$\frac{1}{\Phi_1^2} = \prod_{j=1}^{n-1} (1-2it_2-2\theta_j it_1)\cdot \frac{(2it_1)}{(2it_1)} \qquad \forall i=1,\ldots n-1$$

$$\frac{1}{\Phi_1^2} = (2it_1)^{n-1} \prod_{j=1}^{n} \left[\frac{(1 - 2it_2) - 2\theta_j it_1}{2it_1} \right]$$

$$= (2it_1)^{n-1} \prod_{j=1}^{n} \left[\frac{1 - 2it_2}{2it_1} - \theta_j \right]$$

$$= (2it_1)^{n-1} \prod_{j=1}^{n} \left[\frac{1-2it_2}{2it_1} - \theta_j \right] \sum_{j=1}^{n} \frac{l_j^2}{\frac{1-2it_2}{2it_1} - \lambda_j}$$

$$= \prod_{j=1}^{n} (1-2\lambda_j it_1 - 2it_2) \sum_{j=1}^{n} \frac{l_j^2}{1-2it_2-2it_1\lambda_j} \qquad (56)$$

En esta última expresión se tiene que el factor de la mano izquierda corresponde a la Función Característica de u y v bajo el supuesto de que los $\{z_j\}$ son variables normales independientes; el factor de la mano derecha proporciona la modificación hecha para el ajuste de la regresión sobre una simple variable independiente.

De la expresión (56) se puede notar que:

i.) $\frac{1}{1-2it_2-2it_1}$ con j = 1, 2,....., n son las raíces propias de la matríz $B^{-1} = \{ (1-2it_2)I_n - 2it_1 A \}$

ii.) Los vectores de B^{-1} son los mismos que aquellos de A tal que $l_1, l_2, \ldots\ldots\ldots, l_n$ son los cosenos directores de el vector X relativo a esos vectores propios. Así, se tiene que

$$\sum_{j=1}^{n} \frac{l_j^2}{1-2it_2-2it_1\lambda_j} = \frac{X'B^{-1}X}{X'X}$$

iii.) Ademas, se tiene que: $\prod_{j=1}^{n} (1-2it_2-2it_1\lambda_j) = |B|$

De las observaciones anteriores se tiene que, la expresión (56) puede

reducirse a:
$$\frac{1}{\Phi_1^2} = |B| \frac{X'B^{-1}X}{X'X} \qquad (57)$$

donde, podemos notar que el segundo factor de (57) toma la forma general de r.

Efectuando una extensión directa del proceso anterior para una variable independiente, se puede demostrar que para la regresión sobre k variables independientes, se tiene que la Función Característica está dada por :

$$\frac{1}{\Phi^2} = |B_1| \prod_{s=1}^{k} \frac{X_s'B_s^{-1}X_s}{X_s'X_s} \qquad (58)$$

donde $B_s = (1 - 2it_2)I_n - 2it_1M_{s-1} \cdots\cdots\cdots M_2M_1A$

y donde los M_i $\forall$ i = 1, 2,....., k han sidos definidos como en (15) de la sección A de este capítulo.

Debido a la propiedad reproductiva de los productos $M_k \cdots\cdots M_2M_1A$ que se mencionó en la sección A de este capítulo, este resultado puede expresarse directamente como en (57).

Haciendo $t_2 = 0$ se tiene la Función Característica de u luego, la Cumulante y de aquí, los Momentos pueden ser obtenidos por medio de la expansión de log Φ.

E- SELECCION DEL CRITERIO DE LA PRUEBA

Para decidir sobre el criterio apropiado para la prueba, es importante considerar el conjunto de hipótesis alternativas contra las cuales se va a discriminar. La clase de alternativa que vamos a considerar es tal que el correlograma de los errores se reduzca aproximadamente exponencialmente con separación creciente de las observaciones. Un modelo que se ajusta convenientemente para éste tipo de hipótesis es el Proceso Estacionario de Markov:

$$\varepsilon_i = \rho\varepsilon_{i-1} + u_i \quad \forall\ i:\ i = \ldots,-1,0,1,\ldots \qquad (59)$$

donde $|\rho| < 1$ y $u_i \sim N(0,\sigma^2)$ y es independiente de $\varepsilon_{i-1},\varepsilon_{i-2},\ldots$ y de $u_{i-1},u_{i-2},\ldots$

La Hipótesis Nula H_0 es entonces, en (59): $H_0: \rho = 0$.

T. W. Anderson[10] demostró que no existe una prueba de ésta hipótesis que sea Uniformemente la más Potente contra las alternativas de (59), ademas, demostró sin embargo que, para ciertos sistemas de regresión con distribuciones cerradas de los errores a los dados por (59) se pueden obtener pruebas las cuales son uniformemente más potentes contra alternativas unilaterales de (59) y que proveen regiones del tipo B_1 para alternativas bilaterales de (59). Estos sistemas de regresión incluyen casos en los que los vectores de regresión son vectores constantes coincidiendo con los vectores propios de una matríz θ (o de combinaciones lineales de k de ellos) y en los cuales las distribuciones de los errores tienen una Función de Densidad de la forma:

$$f_u(u) = K \exp \left[-1/2\sigma^2\{(1+\rho^2)\varepsilon'\varepsilon - 2\rho\varepsilon'\theta\varepsilon\}\right] \qquad (60)$$

Para tal caso la prueba uniformemente más potente de la hipótesis $H_0: \rho = 0$ contra $H_1': \rho > 0$ está dada por $r > r_0$ donde $r = \frac{Z'\Theta Z}{Z'Z}$, siendo Z el vector de los residuos de la regresión de Mínimos Cuadrados y siendo r_0 determinado para proporcionar una región crítica de tamaño apropiado.

Para el caso en que $H_0: \rho = 0$ contra $H_1: \rho \neq 0$, el típo de prueba B_1 está dada por $r < r_2$, $r > r_3$, donde r_2 y r_3 están determinados tal que proporcionan una región crítica de tamaño apropiado y que satisfacen la relación:

$$\int_{r_2}^{r_3} rf(r)dr = E[r]\int_{r_2}^{r_3} f(r)dr \qquad (61)$$

donde f(r) es la Función de Densidad en el caso nulo.

Sabemos del Corolario 1 de la sección A de este capítulo que cualesquiera que sean los vectores de regresión, se tiene que:

$$r_L \leq r \leq r_U$$

Ahora, r_L y r_U tienen distribuciones idénticas en el caso nulo con distribuciones de r obtenida de los residuos de la regresión sobre ciertos vectores propios de la matríz A. . Así, si hacemos $\theta = A$ en (60) podemos decir que, cuando el límite inferior r_L (ó superior r_U) del corolario es obtenido, el estadístico $r = (Z'AZ)/Z'Z$ proporciona una prueba uniformemente la más potente contra la alternativa unilateral de (58) y la cual es del típo B_1 contra la alternativa bilateral.

La distribución de los errores para el Proceso Estacionario de Markov (59) tiene Función de Densidad:

$$f_u(u) = K \exp\left[-1/2\sigma^2\{(1+\rho^2)\sum_{i=1}^{n}\varepsilon_i^2-\rho^2(\varepsilon_1^2+\varepsilon_n^2)-2\rho\sum_{i=2}^{n}\varepsilon_i\varepsilon_{i-1}\}\right] \quad (62)$$

Si en la ecuación (60) hacemos $\varepsilon'\Theta\varepsilon = \sum_{i=2}^{n}\varepsilon_i\varepsilon_{i-1}$ se tiene que:

$$f_u(u) = K \exp\left[-1/2\sigma^2\{(1+\rho^2)\sum_{i=1}^{n}\varepsilon^2-2\rho\sum_{i=2}^{n}\varepsilon_i\varepsilon_{i-1}\}\right] \quad (63)$$

pero, si en la ecuación (60) hacemos $\varepsilon'\Theta\varepsilon=\sum_{i=1}^{n}\varepsilon^2-\frac{1}{2}\sum_{i=2}^{n}(\varepsilon_i-\varepsilon_{i-1})^2$

obtendremos que:

$$f_u(u) = K \exp\left[-1/2\sigma^2\{(1+\rho^2)\sum_{i=1}^{n}\varepsilon_i^2-\rho(\varepsilon_1^2+\varepsilon_n^2)-2\rho\sum_{i=2}^{n}\varepsilon_i\varepsilon_{i-1}\}\right] \quad (64)$$

En vista de que las expresiones (63) y (64) están muy próximas a la expresión (62), podemos asegurar que casi cualquier valor de θ nos proporcionará un buen estadístico r para probar la hipótesis $H_o:\rho=0$ contra las alternativas de (60). Entre los dos estadísticos que proporciona el valor de $\varepsilon'\Theta\varepsilon$ dado anteriormente, no existe mucha diferencia, sin embargo se escoge una ligera modificación del segundo debido a razones de conveniencia en el cálculo (y debido a su similaridad con el estadístico de Von Neumann δ^2/S^2 (1941)). |10|

Nuestro estadístico a adoptar está definido por:

$$d = \frac{\sum_{i=2}^{n} (Z_i - Z_{i-1})^2}{\sum_{i=1}^{n} Z_i^2} \tag{65}$$

Este estadístico es un caso especial del estadístico general, $r = (Z'AZ)/Z'Z$ visto anteriormente, en el cual

$$A = 1/2Ad = \begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & & 2 & -1 \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} 1/2 \tag{66}$$

y donde tomamos a $\Theta = I - 1/2Ad$ para obtener la función de densidad (64).

Como los vectores propios de la matríz Ad y Θ en la ecuación del parrafo anterior, son los mismos, entonces, cuando el vector de regresión es un vector propio de Ad, el estadístico d de (65) provee una Prueba Uniformemente más Potente contra las alternativas unilaterales de (64). En particular, la prueba dada por d es Uniformemente la más potente cuando los límites r_L y r_U son alcanzados.

La principal alternativa al uso de d ó un estadístico relacionado como un criterio de la prueba ha sido el uso de una de las estadísticas circulares, tales como:

$$r_c = \frac{\sum_{i=1}^{n} Z_i Z_{i-1}}{\sum_{i=1}^{n} Z_i^2} \qquad d_c = \frac{\sum_{i=1}^{n} (Z_i - Z_{i-1})^2}{\sum_{i=1}^{n} Z_i^2} \tag{67}$$

en los cuales tomamos $Z_0 \equiv Z_n$.

T. W. Anderson demostró en 1948 que r_c y d_c proveen pruebas uniformemente más potente contra alternativas unilaterales en una población circular con función de densidad:

$$f\ (u) = K\ \exp\left[-1/2\sigma^2\{(1+\rho^2)\sum_{i=1}^{n}\varepsilon^2-2\rho\sum_{i=1}^{n}\varepsilon_i\varepsilon_{i-1}\}\right] \qquad (68)$$

En la práctica se presentan con mucha frecuencia casos en los cuales los vectores de regresión no son vectores propios, luego surge la interrogante siguiente. ¿ Qué tan buenos son estos estadísticos como criterio de la prueba para estos casos?.

Evidentemente podemos esperar que:

1º La Potencia de la Prueba disminuya cuando los vectores de regresión se desvían de los vectores propios.

2º Los Coeficientes de Mínimos Cuadrados no son estimaciones de Máxima Verosimilitud en el caso nulo.

3º Cualquier prueba basada sobre residuos de Mínimos Cuadrados no será siempre razón de Verosimilitud para la prueba.

En contraposición con los 3 puntos anteriores se tiene que:

1ºAun se tiene una Prueba válida, aunque posiblemente de potencia reducida.

2º Es deseable por conveniencia tener una prueba basada sobre Residuos de Mínimos Cuadrados, aun cuando no sea una prueba óptima.

3º El estadístico r se encontrará necesariamente entre los límites r_L y r_U y cuando los mismos son alcanzados, la prueba será óptima.

F- ALGUNOS RESULTADOS ESPECIALES

Sabemos que el estadístico d de (65) es un caso especial de r de la ecuación (38) de este capítulo, donde $A = 1/2A_d$ está dada en (66). en vista de que ésta matríz A_d es simétrica, entonces, ella está completamente especificada por el tríangulo superior izquierdo, por lo tanto podemos escribir:

$$A_d \doteq \begin{matrix} 1 & -1 & 0 \\ & 2 & -1 \\ & & 2 \end{matrix} \qquad (69)$$

Para obtener los momentos de d, debemos primero obtener las potencias de A_d; así se tiene que:

$$A_d^2 \doteq \begin{matrix} 2 & -3 & 1 & 0 \\ & 6 & -4 & 1 \\ & & 6 & -4 \\ & & & 6 \end{matrix} \qquad A_d^3 \doteq \begin{matrix} 5 & -9 & 5 & -1 & 0 & 0 \\ & 19 & -15 & 6 & -1 & 0 \\ & & 20 & -15 & 6 & -1 \\ & & & 20 & -15 & 6 \\ & & & & 20 & -15 \\ & & & & & 20 \end{matrix}$$

$$A_d^4 \doteq \begin{matrix} 14 & -28 & 20 & -7 & 1 & 0 & 0 \\ & 62 & -55 & 28 & -8 & 1 & 0 \\ & & 70 & -56 & 28 & -8 & 1 \\ & & & 70 & -56 & 28 & 1 \end{matrix} \qquad (70)$$

Asi, se tiene que si $Z' = (Z_1\ Z_2\ Z_3\ Z_4 \cdots\cdots Z_{n-2}\ Z_{n-1}\ Z_n)$, entonces:

$$\Delta Z = \begin{bmatrix} Z_1-Z_2 \\ Z_2-Z_3 \\ Z_3-Z_4 \\ \vdots \\ Z_{n-2}-Z_{n-1} \\ Z_{n-1}-Z_n \\ Z_n \end{bmatrix}; \ \Delta^2 Z = \begin{bmatrix} Z_1-2Z_2+Z_3 \\ Z_2-2Z_3+Z_4 \\ Z_3-2Z_4+Z_5 \\ \vdots \\ Z_{n-2}-2Z_{n-1}+Z_n \\ Z_{n-1}-2Z_n \\ Z_n \end{bmatrix}; \Delta^3 Z = \begin{bmatrix} Z_1-3Z_2+3Z_3-Z_4 \\ Z_2-3Z_3+3Z_4-Z_5 \\ Z_3-3Z_4+3Z_5-Z_6 \\ \vdots \\ Z_{n-2}-3Z_{n-1}+3Z_n \\ Z_{n-1}-3Z_n \\ Z_n \end{bmatrix}$$

$$\Delta^4 Z = \begin{bmatrix} Z_1-4Z_2+6Z_3-4Z_4+Z_5 \\ Z_2-4Z_3+6Z_4-4Z_5+Z_6 \\ Z_3-4Z_4+6Z_5-4Z_6+Z_7 \\ \vdots \\ Z_{n-2}-4Z_{n-1}+6Z_n \\ Z_{n-1}-4Z_n \\ Z_n \end{bmatrix} \tag{71}$$

Por otro lado tenemos que:

$$Z'A_dZ = \sum_{i=1}^{n-1}(\Delta Z_i)^2$$

$$Z'A_d^2Z = \sum_{i=1}^{n-2}(\Delta^2 Z_i)^2 + (Z_1 - Z_2)^2 + (Z_{n-1} - Z_n)^2 \tag{72}$$

$$Z'A_d^3Z = \sum_{i=1}^{n-3}(\Delta^3 Z_i)^2+4Z_1^2+9Z_2^2+Z_3^2-12Z_1Z_2-6Z_2Z_3+4Z_1Z_3+ \text{ una expresión similar en } Z_n,\ Z_{n-1},\ Z_{n-2}$$

$$Z'A_d^4Z = \sum_{i=1}^{n-4}(\Delta^4 Z_i)^2+13Z_1^2+45Z_2^2+17Z_3^2+Z_4^2-48Z_1Z_2-54Z_2Z_3-8Z_3Z_4+28Z_1Z_3+12Z_2Z_4-6Z_1Z_4+ \text{ una expresión similar en } Z_n,\ Z_{n-1},\ Z_{n-2}, Z_{n-3}$$

Luego, por la definición circular de d en (67), es deicr,

$$d_c = \frac{Z'A_{dc}Z}{Z'Z} \tag{73}$$

se tiene que, los términos de corrección desaparecen, dando,

$$Z'A_{d_c}^sZ = \sum_{i=1}^{n}(\Delta^s Z_i)^2 \text{ donde } Z_{-i} \equiv Z_{n-1} \tag{74}$$

Von Neumann encontró en 1941 que las raíces propias de A_d están dadas por: [10]

$$\lambda_j = 2\left\{1 - \cos\frac{\pi(j-1)}{n}\right\} \quad \forall j:\ j=1,\ 2,\ldots,\ n \tag{75}$$

donde las primeras 4 sumas de potencias son:

$$\mathrm{tr}\ A_d = \sum_{j=1}^{n} \lambda_j = 2(n-1) \qquad \mathrm{tr}\ A_d^3 = \sum_{j=1}^{n} \lambda_j^3 = 4(5n-8)$$
$$\mathrm{tr}\ A_d^2 = \sum_{j=1}^{n} \lambda_j^2 = 2(3n-4) \qquad \mathrm{tr}\ A_d^4 = \sum_{j=1}^{n} \lambda_j^4 = 2(35n-64) \tag{76}$$

El vector propio correspondiente a la raíz cero λ_1 es $\{1, 1, \ldots, 1\}$, el cual es el vector de regresión correspondiente a un término constante en el modelo de regresión. Luego, para regresiones con una media ajustada por lo tanto, necesitamos solo considerar los restantes $\{\lambda_j\}$ los cuales reenumeramos de conformidad, tal que,

$$\lambda_j = 2\left(1 - \cos\frac{j\pi}{n}\right) \qquad (j = 1, \ldots, n-1)$$

luego, del corolario 1 tenemos que;

$$d_L \leq d \leq d_U$$

donde,

$$d_L = \frac{\sum_{i=1}^{n-k'-1} \lambda_i \zeta_i^2}{\sum_{i=1}^{n-k'-1} \zeta_i^2} \qquad d_U = \frac{\sum_{i=1}^{n-k'-1} \lambda_{i+k'} \zeta_i^2}{\sum_{i=1}^{n-k'-1} \zeta_i^2} \tag{77}$$

siendo $k' = (k-1)$ el número de variables independientes en el modelo en adición al término constante.

Con la distribución de los errores asumida en B-2 los límites de la media de d estan dados por:

$$E(d) \leq E(d_U) = 2 - \frac{2}{n-k'-1} \sum_{j=k'+1}^{n-1} \cos\frac{\pi j}{n}$$

$$\geq E(d_L) = 2 - \frac{2}{n-k'-1} \sum_{j=1}^{n-k'+1} \cos\frac{\pi j}{n}$$

y la varianza de d en sus límites son (todo esto sin probar):

$$\mathrm{var}(d) \leq \frac{16}{(n-k'-1)(n-k'+1)} \sum_{j=1}^{1/2(n-k'-1)} \cos^2\frac{\pi j}{n} \qquad (\text{ n-k' impar})$$

$$\text{var}(d) \leq \frac{16}{(n-k'-1)(n-k'+1)} \sum_{j=1}^{\frac{1}{2}(n-k')-1} \cos^2 \frac{\pi j}{n} + \frac{8(n-k'-2)}{(n-k'-1)^2(n-k'+1)} \cdot$$

$$\cos^2 \frac{(n-k')\ \pi}{2n} \qquad \text{(con n-k' par)}$$

$$\geq \frac{16}{(n-k'-1)(n-k'+1)} \sum_{j=k'+1}^{\frac{1}{2}(n-1)} \cos^2 \frac{\pi j}{n} \qquad \text{(n impar)} \qquad (78)$$

$$\geq \frac{16}{(n-k'-1)(n-k'+1)} \sum_{j=k'+1}^{\frac{1}{2}(n-1)} \cos^2 \frac{\pi j}{n} \qquad \text{(n par)}$$

II OTROS CRITERIOS PARA PROBAR LA AUTOCORRELACION DE LOS ERRORES

A- APROXIMACION A UNA BETA.

Sabemos que el estadístico d visto en (65) es aproximadamente igual a $2(1 - \rho)$, donde $-1 \leq \rho \leq 1$ luego, $0 \leq d \leq 4$; así, d se encuentra entre 0 y 2 para $\rho > 0$ y para $\rho < 0$, d se encuentra entre 2 y 4.

Si en una prueba de autocorrelación de errores se llega a que $d_L \leq d \leq d_U$, sabemos que la prueba es inconcluyente o sea que no da resultados definitivos sobre la existencia o no de la autocorrelación que se ha estado probando, luego, debemos encontrar un estadístico que nos proporcione resultados concluyentes aunque sea en forma aproximada. Un método para lograr esto consiste en efectuar un contraste aproximado. El método propuesto por Durbin y Watson [4] consiste en transformar d de forma tal que su intervalo de variación sea aproximadamente de 0 a 1 y que además se ajuste a una Distribución Beta con la misma media y la misma varianza que d.

Supongamos que hacemos $d^{\alpha} = d/4$ de tal forma que se distribuya segun una Distribución Beta, entonces, su Función de Densidad estará dada por:

$$f(d^{\alpha}) = \frac{1}{\beta(p,q)} (d/4)^{p-1} (1 - d/4)^{q-1} \qquad (79)$$

entonces se tiene que,

$$E(d) = \frac{4p}{p+q} \qquad y \qquad var(d) = \frac{16pq}{(p+q)\ (p+q+1)} \tag{80}$$

De (80) se obtiene que,

$$p = \frac{1}{4}\ (p+q)E(d) \qquad y \qquad p+q = \frac{E(d)\ 4 - E(d)}{var(d)} - 1 \tag{81}$$

De estas ecuaciones se deducen los valores de p y q, donde los valores de E(d) y var(d) se calcúlan mediante las ecuaciones (52) y (53) respectivamente con A dado por (66).

Para contrastar la correlación positiva se hace necesario calcular el valor crítico de d en la cola inferior de la distribución. Si 2p y 2q son valores enteros, el valor aproximado puede obtenerse de las tablas de razón de varianza o z de Fisher. Si 2p y 2q no son enteros, una primera aproximación se halla usando los valores enteros más próximos, luego se obtiene ,

$$F = \frac{p(4 - d)}{qd} \tag{82}$$

donde F se distribuye con razón de varianza y luego se obtiene $z = 1/2 \cdot \log F$, la cual es la z de Fisher con $n_1 = 2q$ y $n_2 = 2p$ grados de libertad.

Si lo que nos interesa es contrastar la correlación negativa de los errores, entonces debemos reemplazar d por 4 - d y luego efec= tuar el mismo procedimiento anterior.

B- LA APROXIMACION $d^* = a + bd_U$ (83)

En 1957 E. J. Hannan propuso una aproximación al estadístico d de Durbin-Watson mediante el uso del límite d dado en (32), ya que como siempre se dá que $d \leq d_U$, la curva de distribución de d es mejor aproximada mediante la distribución de d_U. |12|

Durbin y Watson[1] propusieron en 1971 una combinación lineal de d de la forma $d^* = a + bd_U$ donde a y b son escogidos de tal forma

[1] Durbin y Watson, 1971 [5]

que d^* tenga los mismos dos primeros momentos de d para la matríz particular X en cuestión.

Esta nueva aproximación es más ventajosa que la aproximación Beta puesto que, uno puede hacer uso de los puntos de significación tabulados de d para valorar la significación i.e. si d_U^n es el punto de significación tabulado en las tablas dadas por Durbin-Watson (1951) para una prueba contra correlación serial positiva para valores específicos de n y k, entonces, se rechazará la hipótesis de independencia si se obtiene que $d^* < a + bd_U^0$.

Del hecho de que,

$$E(d) = \frac{P}{n-k} \quad y \quad Var(d) = \frac{2}{(n-k)(n-k+2)}\left[Q - PE(d)\right]$$

con

$$P = tr(A) - tr\left[X'AX(X'X)^{-1}\right] \quad y \tag{84}$$

$$Q = tr(A^2) - 2tr\left[X'A^2X(X'X)^{-1}\right] + tr\left[\{X'AX(X'X)^{-1}\}^2\right]$$

donde $tr(A) = 2(n-1)$ y $tr(A^2) = 2(3n-4)$

se tiene que,

$$E(d) = \frac{2(n-1) - tr\left[X'AX(X'X)^{-1}\right]}{n-k} \quad y$$

$$var(d) = \frac{2}{(n-k)(n-k+2)}\left[2(3n-4) - 2tr\left[X'A^2X(X'X)^{-1}\right] + \right. \tag{85}$$

$$+ tr\left[\{X'AX(X'X)^{-1}\}^2\right] - 4(n-1)^2 - 4(n-1)tr\left[X'AX(X'X)^{-1}\right]$$

$$\left. + \{tr\, X'AX(X'X)^{-1}\}^2\right]$$

Para calcular P y Q notamos que el (i,j)-ésimo elemento de X'AX es,

$$\sum_{t=1}^{n-1} \Delta x_{it} \Delta x_{jt} \tag{86a}$$

y el (i,j)-ésimo elemento de $X'A^2X$ es,

$$\sum_{t=1}^{n-2} \Delta^2 x_{it}\, \Delta^2 x_{jt} + (x_{i2} - x_{i1})(x_{j2} - x_{j1}) + (x_{in} - x_{i,n-1})(x_{jn} - x_{j,n-1}) \tag{86b}$$

donde como vimos anteriormente, Δ es el operador de las primeras diferencias.

De las ecuaciones (40), (44a) y (76) se tiene que;

$$E(d_u) = \frac{1}{n-k}\left[2(n-1) - 2\sum_{j=1}^{k-1}\{1 - \cos\frac{\pi j}{n}\}\right]$$

$$= \frac{1}{n-k}\left[2(n-k) + 2\sum_{j=1}^{k-1}\cos\frac{\pi j}{n}\right]$$

$$= 2 + \frac{2}{n-k}\sum_{j=1}^{k-1}\cos\frac{\pi j}{n} \qquad (87)$$

y

$$\mathrm{var}(d_u) = \frac{4\left[(n-2) - 2\sum_{j=1}^{k-1}\cos^2\frac{\pi j}{n} - 2\left[\sum_{j=1}^{k-1}\cos\frac{\pi j}{n}\right]^2/(n-k)\right]}{(n-k)(n-k+2)} \qquad (88)$$

Luego, los valores de a y b son determinados mediante las relaciones,

$$E(d) = a + bE(d_u)$$
$$\mathrm{var}(d) = b\ \mathrm{var}(d_u) \qquad (89)$$

POTENCIA DE LA PRUEBA

DE

DURBIN - WATSON

Dado el modelo de Regresión Lineal,

$$Y = X\beta + \varepsilon \qquad (1)$$

en el cual el vector ε de las perturbaciones es tal que, sus componentes ε_i siguen un Proceso Autoregresivo de Primer Orden; para probar la hipótesis H_o: $\rho = 0$, Durbin y Watson propusieron una prueba basada sobre el estadístico,

$$d = \frac{\sum_{i=2}^{n} (\hat{\varepsilon}_i - \hat{\varepsilon}_{i-1})^2}{\sum_{i=1}^{n} \hat{\varepsilon}_i^2} \qquad (2)$$

en el cual ,

$$\hat{\varepsilon} = Y - X\,\hat{\beta} \qquad \text{con} \qquad \hat{\beta} = (X'X)^{-1}X'Y \qquad (3)$$

Bajo la hipótesis H_o Durbin y Watson encontraron un par de variables aleatorias límites dadas por (77 del Cap II), cuyas distribuciones no dependen de X y tal que, para cualquier d^* se tiene que,

$$Pr(d_u < d^*) \leq Pr(d < d^*) \leq Pr(d_l < d^*) \qquad (4)$$

Con esta propiedad propusieron un nivel α para la prueba de H_o: $\rho = 0$ contra la alternativa H :$\rho > 0$, tal que la prueba se podía hacer en dos partes así:

i) Se acepta H_o si $d > d_u^*$
   No hay decisión si $d_l^* < d < d_u^*$
   Se rechaza H_o si $d_l^* > d$ (5)
   y donde $Pr(d_u < d_u^*) = Pr(d_l < d_l^*) = \alpha$

ii) Si no se obtiene decisión en la prueba, entonces, se utiliza un método aproximado para obtener el único punto apropiado de significación $d^*(X)$ para la matríz X que se esté usando; de aquí que,

$$Pr(d < d^*(X)) = \alpha \qquad (6)$$

Sabemos que un Proceso Estacionario de Markov (59 Cap. II) tiene Función de Densidad dada por (62 del Cap. II), por lo que la matríz de covarianza de ε está dada por,

$$V^{-1} = (1 + \rho^2)\mathbf{I} - 2\rho\Theta + \rho(1 - \rho)C \tag{7}$$

donde

$$2\Theta = \begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{7a}$$

Ademas, Anderson demostró que si V^{-1} es reemplazado por,

$$W^{-1} = (1 + \rho^2)\mathbf{I} - 2\rho\Theta \tag{8}$$

y si los vectores de regresión,se encuentran dentro del espacio cubierto por k de los vectores característicos de W, una prueba basada sobre el estadístico d es una prueba similar Uniformemente más Potente de H_0 contra H_1.

La Función de Potencia de la prueba de Durbin-Watson la estudiaremos asumiendo que la matríz de covarianza de ε es $\sigma^2 W$ en lugar de $\sigma^2 V$. Así, el estadístico d puede escribirse como;

$$d = \frac{\hat{\varepsilon}'A\hat{\varepsilon}}{\hat{\varepsilon}'\hat{\varepsilon}} = \frac{\varepsilon'MAM\varepsilon}{\varepsilon'M\varepsilon} = \frac{u'W^{1/2}MAMW^{1/2}u}{u'W^{1/2}MW^{1/2}u} \tag{9}$$

con $u \sim N(0,\sigma^2\mathbf{I})$ y M dada por (5 Cap. II) y A dada por (66 Cap. II), o sea que,

$$A = 2(\mathbf{I} - \Theta)$$

Luego, de (4) se tiene que,

$$Pr(d < d^*) \doteq Pr\left[\frac{u'W^{1/2}MAMW^{1/2}u}{u'W^{1/2}MW^{1/2}u} < d^*\right]$$

$$\Longrightarrow \Pr(\, d < d^* \,) = \Pr\left[\, u'W^{1/2}M\,(\, A - d^*\mathbf{I}\,)\, MW^{1/2}u < 0 \right]$$

$$= \Pr\,(\sum_{i=1}^{n} \pi_i \delta_i^2 < 0\,) \qquad (10)$$

donde $\pi_1 \geq \pi_2 \geq \cdots\cdots \geq \pi_n$ son las Raíces Caracteristicas de

$$\Gamma = W^{1/2}M(\, A - d^*\mathbf{I}\,)MW^{1/2} \qquad (11)$$

y donde $\{\delta_i\}_{1<i<n}$ son tales que $\delta_i \sim N(0,\sigma^2)$

Nuestro proposito consiste en obtener límites inferiores y superiores de π_i $\forall\ 1\leq i\leq n$, los cuales son independientes de la matríz X.

Supongamos que:

$$\underline{\pi_i} \leq \pi_i \leq \overline{\pi_i} \qquad (12)$$

entonces, para todo ρ, d^*, n y k dado se tiene que:

$$\Pr\left[\sum_{i=1}^{n-k}\overline{\pi}_i\delta_i^2 < 0\right] \leq \Pr\left[\sum_{i=1}^{n-k}\pi_i\delta_i^2 < 0\right] \leq \Pr\left[\sum_{i=1}^{n-k}\underline{\pi_i}\delta_i^2 < 0\right] \qquad (13)$$

Sea p_j el Vector Característico correspondiente a la j-ésima raíz máxima de W, el cual por conveniencia denotaremos por :

$$p_j = Ch_j(W) \qquad (14)$$

Si $p_j' = (p_{1j} \cdots\cdots\cdots\cdots p_{nj})$, entonces, se tiene que:

$$\left.\begin{array}{l} p_{1j} = 1/\sqrt{n} \\ p_{ij} = \dfrac{1}{\sqrt{n/2}} \cos \dfrac{(2i-1)(j-1)}{2n} \ , \ j=2,\ldots,n \end{array}\right\} \qquad (15)$$

Sabemos por (75) cap.II que las raíces propias de A están dadas por:

$$\lambda_j = 2\left[\, 1 - \cos\frac{(j-1)}{n}\,\pi\right]$$

y como $A = 2(\mathbf{I} - \Theta)$, las raíces caracteristicas de Θ quedarán expresadas por:

$$\nu_i = Ch_i(\Theta) = \cos\frac{(i-1)}{n}\,\pi \qquad (16)$$

de las cuales se obtienen las raíces características de W y de A.

A continuación desarrollaremos 2 métodos para derivar los límites de π_i, donde uno de los métodos proporciona un límite superior y el otro proporciona un límite inferior.

METODO 1:

El primer método consiste en encontrar los límites de las raíces características de Γ ignorando el hecho de que Γ tiene k raices ceros, es decir que supondremos que Γ tiene n raíces. Esta técnica consiste en obtener las raíces características como extremos de una Forma Cuadrática Restringida. (Courant y Hibert [13])

TEOREMA 1:

Para cualquier matríz simétrica A de orden n y cualquier vector columna arbitrario α_j, con $j = 1,\ldots\ldots,i-1$

$$\max_{\substack{x'\alpha_j=0\\ j=1,\ldots,i-1}} \frac{x'Ax}{x'x} \geq Ch_i(A) \qquad\qquad \min_{\substack{x'\alpha_j=0\\ j=i+1,\ldots,m}} \frac{x'Ax}{x'x} \leq Ch_i(A) \tag{17}$$

donde $Ch_i(A)$ es la i-ésima raíz característica en órden descendente.

Demostración en Courant y Hilbert [13].

Los siguientes Lemas y Teoremas que a continuación se enuncian nos serán de gran utilidad para establecer el Límite Superior para todas las n raices de Γ .

LEMA 1:

Sean B y C matrices de pxm $(p \geq m)$, tal que $rang(B - C) = r$. Entonces, para cualquier matríz simétrica A, $i = 1,\ldots\ldots,m-r$

i) $Ch_{i+r}(B'AB) \leq Ch_i(C'AC)$ (18)

ii) $Ch_i(B'AB) \geq Ch_{i+r}(C'AC)$

DEMOSTRACION:

i) Sean $\alpha_{i+r+1},\ldots,\alpha_m$ los vectores característicos de B'AB correspondientes a las m-i+r menores raíces características. La ecuación $(B - C)x = 0$ impone r restricciones sobre X, ya que $\text{rang}(B - C) = r$, entonces, para $i+r \leq m$ se tiene que:

$$Ch_i(C'AC) \geq \text{mín}\left[\frac{x'C'ACx}{x'x} : x'\alpha_j=0 \text{ con } j=i+r+1,..,m;(B-C)x=0\right]$$

$$= \text{mín}\left[\frac{x'B'ABx}{x'x} : x'\alpha_j=0 \text{ con } j=i+r+1,..,m;(B-C)x=0\right]$$

$$\Longrightarrow Ch_i(C'AC) \geq \text{mín}\left[\frac{x'B'ABx}{x'x} : x'\alpha_j=0;\ j=i+r+1,\ldots, m\right]$$

$$= Ch_{i+r}(B'AB)$$

$$Ch_i(C'AC) \geq Ch_{i+r}(B'AB)$$

ii) Esta parte se demuestra intercambiando B y C en la demostración anterior, puesto que $\text{rang}(B - C) = \text{rang}(C - B)$

LEMA 2:

Sea A una matríz simétrica tal que $Ch_i(A) > 0$, entonces, para cualquier matríz real D:

$$Ch_i(D'AD) \leq Ch_i(A)\cdot Ch_1(D'D) \tag{19}$$

DEMOSTRACION:

Si $Ch_i(D'AD) \leq 0$, entonces, el resultado es trivial. Asumamos que $Ch_i(D'AD) > 0$ y sean $\alpha_1, \alpha_2,\ldots,\alpha_{i-1}$ los vectores característicos de A correspondientes a sus i-1 raíces características mayores, entonces,

$$Ch_i(D'AD) \leq \max\left[\frac{x'D'ADx}{x'x} : x'D'\alpha_j=0; j=1,\ldots\ldots,i-1;\ x\neq 0\right]$$

$$= \max\left[\frac{x'D'ADx}{x'x} : x'D'\alpha_j=0; j=1,\ldots\ldots,i-1; Dx\neq 0\right]$$

$$= \max\left[\frac{x'D'ADx}{x'x}\cdot\frac{x'D'Dx}{x'x} : x'D'\alpha_j=0; j=1,\ldots,i-1; Dx\neq 0\right]$$

$$= \max\left[\frac{Y'AY}{Y'Y}\cdot\frac{x'D'Dx}{x'x} : Y'\alpha_j=0;\ j=1,\ldots,i-1; Y=Dx\neq 0\right]$$

$$= \max\left[\frac{Y'AY}{Y'Y} : Y'\alpha_j=0;\ j=1,\ldots,i-1; Y\neq 0\right]\max\left[\frac{x'D'Dx}{x'x};\ x\neq 0\right]$$

$$= Ch_i(A)\cdot Ch_1(D'D)$$

$$\therefore\ Ch_i(D'AD) \leq Ch_i(A)Ch_1(D'D)$$

TEOREMA 2:

Sea $\Gamma = W^{1/2}M(A - d^*\mathbb{I})MW^{1/2}$ y $G = W^{1/2}(A - d^*\mathbb{I})W^{1/2}$; entonces,

i) $Ch_{i+k}(\Gamma) \leq Ch_i(G)$ (20)

ii) $Ch_i(\Gamma) \geq Ch_{i+k}(G)$

DEMOSTRACION :

i) Si hacemos $A = A - d^*\mathbb{I}$; $B = MW^{1/2}$ y $C = W^{1/2}$ y sustituimos en el Lema 1 .i se tiene que:

$$Ch_{i+k}(B'AB) = Ch_{i+k}\left[(MW^{1/2})'(A - d^*I)(MW^{1/2})\right]$$

$$\leq Ch_i\left[(W^{1/2})'(A - d^*\mathbb{I})(W^{1/2})\right]$$

$$= Ch_i(G)$$

$$\therefore\ Ch_{i+k}(\Gamma) \leq Ch_i(G)$$

ii) Utilizando el Lema 1 (ii) se prueba el resultado.

El teorema anterior nos permite obtener los límites de las raíces características de Γ. Los límites superiores de las k raíces características mayores de Γ que no estan limitadas por el teorema en su parte i se obtienen usando el Lema 2. Así se tiene que haciendo $A = A - d^*I$ y $D = MW^{1/2}$ en el Lema 2 resulta que:

$$Ch_k(\Gamma) \le Ch_k(A - d^*I)\cdot Ch_1(W^{1/2}MW^{1/2}) \tag{21}$$

Si queremos obtener Límites Superiores que no dependan de X, ademas del Lema 2 necesitaremos hacer uso del siguiente Lema.

LEMA 3:

Sea $H = W^{1/2}MW^{-1}MW^{1/2}$, entonces,

$$Ch_1(H) \le \frac{\{ Ch_1(W) + Ch_n(W)\}^2}{4Ch_1(W)Ch_n(W)} \equiv \Delta \tag{22}$$

donde $Ch_1(H) \equiv \Delta$ si las columnas de X se encuentran en el espacio generado por

$$\left\{ \frac{p_1 + p_n}{\sqrt{2}}, \ldots\ldots\ldots\ldots, \frac{p_k + p_{n-k+1}}{\sqrt{2}} \right\} \tag{23}$$

donde los $\{p_i\}$ están definidos como en (15).

DEMOSTRACION:

Sabemos que M es una matríz simétrica, idempotente y de rango n-k, entonces existe una matríz L_1 de nx(n-k) tal que:

a.) $L_1'L_1 = I$ b.) $L_1L_1' = M$ (24)

Así:

$$\begin{aligned} Ch_1(H) &= Ch_1(W^{1/2}MW^{-1}MW^{1/2}) \\ &= Ch_1(W^{1/2}L_1L_1'W^{-1}L_1L_1'W^{1/2}) \\ &= Ch_1(L_1L_1'W^{-1}L_1L_1'W) \\ &= Ch_1(L_1'L_1L_1'W^{-1}L_1L_1'WL_1) \\ &= Ch_1(L_1'W^{-1}L_1L_1'WL_1) \end{aligned}$$

$$Ch_1(H) = \text{máx}\left[\frac{x'(L_1'W^{-1}L_1L_1'WL_1)x}{x'x} \; : x \neq 0 \right]$$

$$= \text{máx}\left[\frac{x'(L_1'WL_1)x}{x'(L_1'W^{-1}L_1)x} \; : x \neq 0 \right] \tag{25}$$

Luego, por la Desigualdad de Cauchy - Schwarz se tiene que:

$$\text{máx}\left[\frac{x'(L_1'WL_1)x}{x'(L_1'W^{-1}L_1)\;x} : x \neq 0 \right] \leq \text{máx}\left[\frac{x'(L_1'WL_1)x}{x'x} \cdot \frac{x'(L_1'W^{-1}L_1)\,x}{x'x} \right]$$

$$\Longrightarrow Ch_1(H) \leq \text{máx}\left[\frac{x'(L_1'WL_1)x}{x'x} \cdot \frac{x'(L_1'W^{-1}L_1)x}{x'x} \right]$$

$$= \text{máx}\left[\sum_{i=1}^{n} Ch_i(W) z_i^2 \cdot \sum_{i=1}^{n} \frac{1}{Ch_i(W)} z_i^2 \; : \; \Sigma z_i^2 = 1 \right] \tag{26}$$

Si hacemos $Ch_i(W) = a_{n-i+1}$ y $z_i^2 = b_{n-i+1}$ y usamos la Desigualdad de Kantorovich, se tiene que:

$$\sum Ch_i(W) z_i^2 \cdot \sum \frac{1}{Ch_i(W)} z_i^2 = \sum a_{n-i+1} b_{n-i+1} \cdot \sum \frac{b_{n-i+1}}{a_{n-i+1}}$$

$$\leq \frac{(a_1 + a_n)^2}{4a_1 a_n}$$

$$= \frac{\left[Ch_1(W) + Ch_n(W)\right]^2}{4Ch_1(W) \cdot Ch_n(W)}$$

$$\equiv \Delta$$

$$\therefore \quad 1 \leq Ch_1(H) \leq \frac{\left[Ch_1(W) + Ch_n(W)\right]^2}{4Ch_1(W) \cdot Ch_n(W)} \equiv \Delta \tag{27}$$

TEOREMA 3:

Si $Ch_i(G)$ es positiva, entonces para todo $i = 1,\ldots,n$ se tiene que:

$$Ch_i(\Gamma) \leq Ch_i(G) \cdot \Delta \tag{28}$$

Las Raíces Características de G están dadas por:

$$Ch_i(G) = Ch_i\{W^{1/2}(A - d^*\mathbb{I})W^{1/2}\}$$
$$= \frac{2(1 - \nu_{n-i+1}) - d^*}{1 + \rho^2 - 2\rho\nu_{n-i+1}} \qquad (29)$$

DEMOSTRACION:

$$Ch_i(\Gamma) = Ch_i\left[W^{1/2}M(A - d^*\mathbb{I})MW^{1/2}\right]$$
$$= Ch_i\left[W^{1/2}MW^{-1}W^{1/2}W^{1/2}(A - d^*\mathbb{I})W^{1/2}W^{1/2}W^{-1}MW^{1/2}\right] \qquad (*)$$

Si hacemos $D = W^{1/2}W^{-1}MW^{1/2}$; $D' = W^{1/2}MW^{-1}W^{1/2}$ y

$G = W^{1/2}(A - d^*\mathbb{I})W^{1/2}$ en (*) se tiene que:

$Ch_i(\Gamma) = Ch_i[D'GD]$ con $D'D = W^{1/2}MW^{-1}MW^{1/2} = H$

$\Longrightarrow Ch_i(\Gamma) \leq Ch_i(G)\cdot Ch_1(D'D)$ por Lema 2

Como $D'D = H \Longrightarrow Ch_1(D'D) = Ch_1(H) \leq \Delta$

$\Longrightarrow Ch_i(\Gamma) \leq Ch_i(G)\cdot\Delta$

Por otro lado, sabemos que $A = 2(\mathbb{I} - \theta)$

$W^{-1} = (1 + \rho^2)\mathbb{I} - 2\rho\theta$ con θ dada por (7a), ademas,

$Ch_i(G) = Ch_i\left[W^{1/2}(A - d^*\mathbb{I})W^{1/2}\right] \Longrightarrow Ch_i(G) = Ch_i\left[W(A - d^*\mathbb{I})\right]$

$\Longrightarrow |W(A - d^*\mathbb{I}) - Ch_i(G)\mathbb{I}| = 0$

$\Longrightarrow |(A - d^*\mathbb{I}) - Ch_i(G)W^{-1}| = 0$

$\Longrightarrow |[2(\mathbb{I} - \theta) - d^*\mathbb{I}] - Ch_i(G)[(1 + \rho^2)\mathbb{I} - 2\rho\theta]| = 0$

$\Longrightarrow |(2 - d^*)\mathbb{I} - Ch_i(G)(1 + \rho^2)\mathbb{I} - 2\theta + 2\theta\rho Ch_i(G)| = 0$

$\Longrightarrow |[(2 - d^*) - Ch_i(G)(1 + \rho^2)]\mathbb{I} - 2\theta[1 - \rho Ch_i(G)]| = 0$

$\Longrightarrow |\theta[2 - 2\rho Ch_i(G)] - [2 - d^* - (1 + \rho^2)Ch_i(G)]\mathbb{I}| = 0$

$$\Longrightarrow \left|\theta - \frac{2 - d^* - (1 + \rho^2)Ch_i(G)}{2 - 2\rho Ch_i(G)}\mathbb{I}\right| = 0$$

$$\Longrightarrow \nu_j = \frac{2 - d^* - (1 + \rho^2)Ch_i(G)}{2 - 2\rho Ch_i(G)} \quad \text{para algun } j$$

$$\Longrightarrow 2\nu_j - 2\rho\nu_j Ch_i(G) = 2 - d^* - (1 + \rho^2)Ch_i(G)$$

$$\Longrightarrow [(1 + \rho^2) - 2\rho\nu_j]Ch_i(G) = 2 - d^* - 2\nu_j$$

$$\Longrightarrow Ch_i(G) = \frac{2(1 - \nu_j) - d^*}{(1 + \rho^2) - 2\nu_j\rho} \qquad (**)$$

Sabemos que : $\lambda_1 \geq \lambda_2 \geq \ldots\ldots \geq \lambda_{n-1} \geq \lambda_n$

con $\lambda_i = Ch_i(G) \quad \forall\ 1 \leq i \leq n$

ademas que si $\lambda_k \geq \lambda_p$ con $p > k$, se tiene que:

$$f(\lambda_k) = \frac{a - b\lambda_k}{2 - 2c\lambda_k} \quad \text{con } a = 2 - d^*;\ b = (1 + \rho^2) \ \text{ y } \ c = \rho$$

$$f(\lambda_p) = \frac{a - b\lambda_p}{2 - 2c\lambda_p}$$

$$\Longrightarrow f(\lambda_p) \geq f(\lambda_k)$$

Esto significa que a medida que λ crece, $f(\lambda)$decrece. Luego, si hacemos $f(\lambda_i) = \nu_j$ con $1 \leq i,j \leq n$, se tiene que:

$\nu_n \geq \nu_{n-1} \geq \nu_{n-2} \geq \ldots\ldots\ldots \geq \nu_2 \geq \nu_1$

por lo tanto tendremos que:

$$\nu_{n-i+1} = \frac{a - b\lambda_i}{2 - 2c\lambda_i}$$

$$\Longrightarrow Ch_i(G) = \frac{2(1 - \nu_{n-i+1}) - d^*}{1 + \rho^2 - 2\rho\nu_{n-i+1}} \qquad \text{por } (**)$$

METODO 2

Este método , para obtener las raíces características de Γ , toma en cuenta el hecho que las k primeras raíces características son nulas, por lo que son eliminadas y solo se obtendrán las restantes n - k raíces.

Sabemos que, $A = 2(\mathbb{I} - \theta)$ luego, si $\rho \neq 0$ entonces,

$$A - d^*\mathbb{I} = \frac{1}{\rho}(W^{-1} - a\mathbb{I}) \quad \text{con } a = (1-\rho)^2 + \rho d^* > 0 \qquad (1')$$

En efecto,

$$\frac{1}{\rho}(W^{-1} - a\mathbb{I}) = \frac{1}{\rho}\left[(1+\rho^2)\mathbb{I} - 2\rho\theta - \left[(1-\rho)^2 + \rho d^*\right]\mathbb{I}\right]$$

ya que $W^{-1} = (1+\rho^2)\mathbb{I} - 2\rho\theta$ y por (1')

$$\Longrightarrow \frac{1}{\rho}(W^{-1} - a\mathbb{I}) = \frac{1}{\rho}\left[(1+\rho)^2\mathbb{I} - 2\rho\theta - (1+\rho^2)\mathbb{I} + 2\rho\mathbb{I} - \rho d^*\mathbb{I}\right]$$

$$= \frac{1}{\rho}\left[2\rho\mathbb{I} - 2\rho\theta - \rho d^*\mathbb{I}\right]$$

$$= 2(\mathbb{I} - \theta) - d^*\mathbb{I}$$

$$= A - d^*\mathbb{I}$$

De aquí se tiene que:

$$\Gamma = \frac{1}{\rho}\left[W^{1/2}M(W^{-1} - a\mathbb{I})MW^{1/2}\right] \qquad (30)$$

Como queremos eliminar las k primeras raíces características de Γ, las cuales son nulas, debemos particionar la matríz Γ. Para ello, usaremos como en el Capítulo II, acapite A el hecho de que como M es una matríz simétrica, semi definida positiva e idempotente, existe una matríz ortogonal L tal que:

$$L'ML = \left[\begin{array}{c:c} I_{n-k} & 0 \\ \hdashline 0 & 0 \end{array}\right] \qquad (31)$$

Sea L' la matríz de los Vectores Propios de M tal que $L = (L_1, L_2)$, entonces,

$$M = L\left[\begin{array}{c:c} I_{n-k} & 0 \\ \hdashline 0 & 0 \end{array}\right]L' = L_1L_1' \qquad (32)$$

Así, las Raíces Características de Γ son iguales a las Raíces Características de:

$$\begin{bmatrix} \tilde{\Gamma} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\rho}\left[L_1'W^{-1}L_1 - a\mathbb{I}\right]L_1'WL_1 & 0 \\ 0 & 0 \end{bmatrix} \qquad (33)$$

En efecto, sabemos que,

$$\Gamma = \frac{1}{\rho}W^{1/2}M(W^{-1} - a\mathbb{I})MW^{1/2}$$

$$\Longrightarrow \quad \Gamma = \frac{1}{\rho}\, W^{1/2}L_1L_1'(W^{-1} - a\mathbb{I})L_1L_1'W^{1/2} \qquad \text{por (32)}$$

$$= \frac{1}{\rho}\, W^{1/2}L_1(\, L_1'W^{-1}L_1 - aL_1'\mathbb{I}L_1\,)L_1'W^{1/2}$$

$$= \frac{1}{\rho}\, W^{1/2}L_1(\, L_1'W^{-1}L_1 - a\mathbb{I}\,)L_1'W^{1/2}$$

Luego,

$$|\Gamma - \lambda I| = 0 \Longrightarrow \left| \frac{1}{\rho}\, W^{1/2}L_1(L_1'W^{-1}L_1 - a\mathbb{I})L_1'W^{1/2} - \lambda\mathbb{I} \right| = 0$$

$$\Longrightarrow \left| \frac{1}{\rho}\, W^{1/2}L_1(L_1'W^{-1}L_1 - a\mathbb{I})L_1'W^{1/2} - \lambda\mathbb{I} \right| \left| W^{1/2}L_1 \right| = 0$$

$$\Longrightarrow \left| \frac{1}{\rho}\, W^{1/2}L_1(L_1'W^{-1}L_1 - a\mathbb{I})L_1'WL_1 - \lambda\mathbb{I}W^{1/2}L_1 \right| = 0$$

$$\Longrightarrow \left| W^{1/2}L_1 \right| \cdot \left| \frac{1}{\rho}\, (L_1'W^{-1}L_1 - a\,\mathbb{I})L_1'WL_1 - \lambda\mathbb{I} \right| = 0$$

$$\Longrightarrow \left| \frac{1}{\rho}\, (L_1'W^{-1}L_1 - a\mathbb{I})L_1'WL_1 - \lambda\mathbb{I} \right| = 0$$

$$\therefore \quad \tilde{\Gamma} = \frac{1}{\rho}\, (L_1'W^{-1}L_1 - a\mathbb{I})L_1'WL_1$$

Ahora, nos corresponde obtener los límites de las raíces características de $\tilde{\Gamma}$. Para tal efecto haremos uso de dos teoremas que limitan las raíces del producto de 2 matrices y de la suma de 2 matrices en términos del producto de las raíces de las matrices individuales y de la suma de las raíces de las matrices individuales respectivas. Estos 2 teoremas se enuncian sin demostración y son

TEOREMA 4: (ANDERSON Y DAS GUPTA [13])

Sean A y B dos matrices simétricas de mxm definida positiva y semidefinida positiva respectivamente, entonces,

$$\begin{aligned} &\text{i) } Ch_{i+j-m}(A\cdot B) \geq Ch_i(A)\cdot Ch_j(B) \text{ donde } i+j \geq m+1 \\ &\qquad\qquad \text{con } i,j = 1,\dots\dots\dots,m \\ &\text{ii) } Ch_{i+j-1}(A\cdot B) \leq Ch_i(A)\cdot Ch_j(B) \text{ donde } i+j \leq m+1 \\ &\qquad\qquad \text{con } i,j = 1,\dots\dots\dots,m \end{aligned} \tag{34}$$

TEOREMA 5: (AMIR - MOEZ y FASS [14])

Sean A y B dos matrices simétricas de mxm, entonces,

$$\begin{aligned} &\text{i) } Ch_{i+j-m}(A+B) \geq Ch_i(A) + Ch_j(B) \text{ donde } i+j \geq m+1 \\ &\qquad\qquad \text{con } i,j = 1,\dots\dots\dots,m \\ &\text{ii) } Ch_{i+j-1}(A+B) \leq Ch_i(A) + Ch_j(B) \text{ donde } i+j \leq m+1 \\ &\qquad\qquad \text{con } i,j = 1,\dots\dots\dots,m \end{aligned} \tag{35}$$

TEOREMA 6:

Dadas las n - k raíces de Γ distintas de cero que son raíces de $\tilde{\Gamma}$, entonces,

$$Ch_i(\tilde{\Gamma}) \geq Ch_{i+k}(G) \tag{36}$$

Si los vectores de regresión se encuentran dentro del espacio generado por los vectores $p_{n-k+1},\dots\dots\dots\dots, p_n$, entonces, se da la igualdad.

DEMOSTRACION:

Sabemos que la matríz L'WL dada en (33) es definida positiva, luego, por el teorema 5 se tiene que:

$$Ch_{i+j-m}(\tilde{\Gamma}) = Ch_{i+j-m}\left[\frac{1}{\rho} (L_1'W^{-1}L_1 - aI)L_1'WL_1 \right]$$

$$Ch_{i+j-m}(\tilde{\Gamma}) = \frac{1}{\rho}\left\{Ch_{i+j-m}\left[L_1'W^{-1}L_1 \cdot L_1'WL_1 - aL_1'WL_1\right]\right\}$$

$$\geq \frac{1}{\rho}\left[Ch_i(L_1'W^{-1}L_1 \cdot L_1'WL_1) - aCh_j(L_1'WL_1)\right]$$

donde $m = n - k$, $i,j = 1,\ldots\ldots,m$ y como $i + j \geq m+1$

$m + 1 = n - k + 1 \Longrightarrow j \geq n - k + 1 - i = n - k - i + 1$

$$\Longrightarrow Ch_{i+j-n+k}(\tilde{\Gamma}) \geq \frac{1}{\rho}\left[Ch_i(L_1'W^{-1}L_1L_1'WL_1) - aCh_{n-k-i+1}(L_1'WL_1)\right]$$

Pero, por la Desigualdad de Kantorovich sabemos que,

$Ch_i(L_1'W^{-1}L_1L_1'WL_1) \geq 1$

ademas, por la Desigualdad de Cauchy se tiene que:

$Ch_{n-k-i+1}(L'WL) \leq Ch_{n-k-i+1}(W)$

Así, tenemos que:

$$Ch_i(\tilde{\Gamma}) \geq \frac{1}{\rho}\left[1 - aCh_{n-k-i+1}(W)\right] \quad \text{con } i = 1,\ldots,n-k$$

$$= Ch_{i+k}\left[W^{1/2}(A - d^*\mathbb{I})W^{1/2}\right]$$

$$= Ch_{i+k}(G)$$

Los dos métodos vistos nos permiten encontrar las raíces características de la matríz $\tilde{\Gamma}$ bajo el supuesto de que el vector de regresíon no incluye ningun término constante. Se demuestra que, si el vector de regresión incluye un término constante, tambien se pueden usar los métodos vistos solo que se deben hacer algunos cambios en algunos de los teoremas usados.

CONCLUSIONES

En el desarrollo de este trabajo hemos tenido que superar grandes dificultades, entre las cuales la principal de ellas ha sido la obtención de la bibliografía necesaria puesto que, en nuestro medio existe muy poco material al respecto, además, de lo relativamente reciente que resulta el estudio del problema que nos ocupa.

En el análisis de nuestro modelo supusimos que el mismo cumplía con cuatro hipótesis básicas, pero, en la realidad no siempre se cumple alguna o algunas de las hipótesis impuestas en el modelo. En el caso de que no se cumpla la primera hipótesis, entonces, se presenta el problema conocido como "Multicolinealidad"; en el caso de que sea la segunda hipótesis la que no se cumpla, entonces, se presenta el problema conocido como "Heteroscedasticidad"; si no se cumple la tercera, entonces,se presenta el problema conocido como "Variables Retardadas" y por último, si no se cumple la cuarta hipótesis, entonces, se presenta el problema que nos ocupa llamado "Autocorrelación de las Perturbaciones".

Cuando se tiene una sucesión de datos muestrales, antes de tratar de ajustar los datos a un modelo de regresión lineal debemos probar si los terminos de errores de la muestra son independientes o no. Como los errores en cualquier caso particular son desconocidos, netonces, basamos nuestra prueba sobre los residuos de la regresión calculada.

Se ha demostrado que para probar la hipótesis de independencia entre los errores, basandonos sobre la independencia de los residuos, no es posible el uso de las pruebas ordinarias de independencia, por lo tanto, se ha tenido que recurrir al uso de otras pruebas poco comunes de las cuales, presentamos tres de ellas. El primero de los criterios que presentamos llamado Criterio de Durbin-Watson se basa en el criterio del cociente de Máxima Verosimilitud de las Perturbaciones; este criterio puede conducirnos a que la prueba sea inconclusa esto es, no se obtienen resultados definitivos sobre la existencia o no de la autocorrelación de los errores. Cuando esta situación se presenta, debemos buscar otro criterio que nos conduzca a alguna conclusión; para ello proponemos dos criterios, donde no existe posibilidad de que la prueba sea inconclusa.

Sabemos que existen otros criterios para probar la Autocorrelación de las Perturbaciones que, aunque no proporcionan una Función de Potencia que sea Uniformemente la más Potente de las pruebas frente a la Prueba Bilateral de Durbin-Watson, su uso proporciona pruebas concluyentes.

Después de finalizado nuestro análisis sobre el tema que ha sido el motor de nuestro trabajo podemos conlcuir lo siguiente:

1.) Aunque el Criterio de Durbin-Watson nos proporciona una Función Potencia que es la Función de Potencia Uniformemente la más Potente, en algunos casos proporciona pruebas inconcluyentes.

2.) Todos los criterios que se han estudiado hasta la actualidad se basan en el Criterio del Cociente de los Residuos.

3.) Para poder usar el Método de los Mínimos Cuadrados en el proceso de regresión es necesario que los errores sean independientes serialmente pués, de lo contrario nuestras estimaciones en los problemas abordados serán sesgadas, lo cual nos conducirá a conlcuciones erradas.

Sería interesante efectuar un estudio detellado de los otros criterios que hasta hoy existen para probar la independencia de las perturbaciones y comparar sus Funciones de Potencia.

Nos preguntamos si es posible encontrar un criterio para probar la independencia entre las perturbaciones que siempre nos conduzca a resultados concluyentes y que además, proporcione una Función de Potencia que sea comparable con la de Durbin-Watson.

Sería prudente investigar que sucedería si además de no cumplirse la hipótesis de independencia entre las perturbaciones, no se cumplen algunas de las otras hipótesis. ¿ Serán válidas las fórmulas utilizadas para calcular las estimaciones de los parámetros? ¿ Se podría obtener algun tipo de prueba o estadístico que nos permita estudiar esta situación?. De no ser válidas las fórmulas usadas para calcular las estimaciones de los parámetros; ¿ Que método alternativo se podría presentar para efectuar dichas estimaciones de modo que las mismas sean las mejores?.

Estas y otras interrogantes planteadas quedan pendientes y rogamos al Todopoderoso que nos proporcione la capacidad y facilidad para poder deslindar algunas de ellas.

B I B L I O G R A F I A

[1] ROBB. J., MUIRHEAD "Aspects of Multivariate Statistical Theory" John Willy & Sons Inc. New York 1982.

[2] JOHN A., TILLMAN " The Power of the Durbin-Watson Test" Econometrica Vol. 43 N°5 - 6 1975.

[3] J. DURBIN and G. S. WATSON "Testing for Serial Correlation in Least Squares Regression I" Biometrica 1950 vol. 37.

[4] J. DURBIN; G. S. WATSON "Testing for Serial Correlation in Least Squares Regression II Biometrica 1951 vol. 38.

[5] J. DURBIN; G. S. WATSON "Testing for Serial Correlation in Least Squares Regression III" Biometrica 1971 vol. 58.

[6] KENNETH J. WHITE and N. E. SAVIN "The Durbin-Watson Test for Serial Correlation with extreme Sample Sizes or Many Regressors" Econométrica, 1977, vol. 45, N°8.

[7] G. S., MADDALA "Econometría" M^c Graw - Hill, 1985.

[8] ANGEL ALCAIDE "Lecturas de Econometría" Biblioteca de Ciencias Económicas Editorial Gredos S. A., Madrid, 1972.

[9] R. L. ANDERSON Ann. Math. Statist. 1942, vol. 13, N°1.

[10]	T. W., ANDERSON	Skand. Aktiar Tidskr, 1948, vol. 31 N° 88.
[11]	E. J., HANNAN	"Testing for Serial Correlation in Least Squares Regression" Biométrika, 1957, vol. 44,N°57.
[12]	T. W., ANDERSON and S. DAS GUPTA	Biométrica, 1963, vol. 50, N° 522.
[13]	G. S., WATSON and E. J., HANNAN	"Serial Correlation in Regression Analysis. II" Biométrica, 1956, N° 436.
[14]	RICHARD C., HENSHAW JR.	"Testing Single-Equation Least Squares Regression Models for Autocorrelated Disturbances" Econométrika, 1966, N° 646.
[15]	FRANKLIN A. GRAYBILL	" An Introduction to Linear Statistical Models". M^{C} Graw - Hill Book Company, Inc. 1961
[16]	RONALD J. WONNACOTT THOMAS H. WONNACOTT	" Econometrics" John Wiley & Sons. 1978.
[17]	BRUCE L. BOWERMAN RICHARD T. O'CONNELL DAVID A. DICKEY	"Linear Statistical Models" Duxbury Press. 1986.
[18]	R. COURANT & D. HILBERT	" Methods of Mathematical Physics" Interscience Publishers, Inc., 1953.