



ARTÍCULOS

## **Sobre la Hipótesis de Normalidad de las Perturbaciones Aleatorias en los Modelos Econométricos Lineales**

Domingo Felipe Cavallo

Revista de Economía y Estadística, Tercera Época, Vol. 15, No. 1-2-3-4 (1971): 1º, 2º, 3º y 4º Trimestre, pp. 49-87.

<http://revistas.unc.edu.ar/index.php/REyE/article/view/4629>



La Revista de Economía y Estadística, se edita desde el año 1939. Es una publicación semestral del Instituto de Economía y Finanzas (IEF), Facultad de Ciencias Económicas, Universidad Nacional de Córdoba, Av. Valparaíso s/n, Ciudad Universitaria. X5000HRV, Córdoba, Argentina.

Teléfono: 00 - 54 - 351 - 4437300 interno 253.

Contacto: [rev\\_eco\\_estad@eco.unc.edu.ar](mailto:rev_eco_estad@eco.unc.edu.ar)

Dirección web <http://revistas.unc.edu.ar/index.php/REyE/index>

### Cómo citar este documento:

Cavallo, D. (1971). Sobre la Hipótesis de Normalidad de las Perturbaciones Aleatorias en los Modelos Econométricos Lineales. *Revista de Economía y Estadística*, Tercera Época, Vol. 15, No. 1-2-3-4: 1º, 2º, 3º y 4º Trimestre, pp. 49-87.

Disponible en: <http://revistas.unc.edu.ar/index.php/REyE/article/view/4629>

El Portal de Revistas de la Universidad Nacional de Córdoba es un espacio destinado a la difusión de las investigaciones realizadas por los miembros de la Universidad y a los contenidos académicos y culturales desarrollados en las revistas electrónicas de la Universidad Nacional de Córdoba. Considerando que la Ciencia es un recurso público, es que la Universidad ofrece a toda la comunidad, el acceso libre de su producción científica, académica y cultural.

<http://revistas.unc.edu.ar/index.php/index>

## SOBRE LA HIPOTESIS DE NORMALIDAD DE LAS PERTURBACIONES ALEATORIAS EN LOS MODELOS ECONÓMICOS LINEALES \*

P O R

DOMINGO FELIPE CAVALLO

Una rápida revisión de los métodos estadísticos de la econometría, en especial de algunos progresos recientes, deja la impresión de que a menudo las técnicas de inferencia son más robustas que los supuestos sobre los que se basan. Ello ha dado lugar a un creciente número de investigaciones orientadas, por un lado, a la determinación de los efectos de la inexactitud de los distintos supuestos sobre la bondad de las inferencias y por el otro, a la construcción de criterios apropiados para el contraste de los mismos supuestos en base a la información suministrada por la muestra. La dilucidación de estas dos cuestiones con referencia a varios supuestos más o menos comunes en los modelos estadísticos lineales, ha sido el punto de partida de importantes progresos teóricos y prácticos, en particular en el campo de la econometría. De esta forma han nacido nuevas técnicas de estimación aplicables a supuestos menos restrictivos, y, por consiguiente, con mayores posibilidades de aplicación exitosa a problemas reales.

En este trabajo me ocupé de la hipótesis de normalidad de las perturbaciones aleatorias en los modelos econométricos lineales. Para este caso en particular interesan, por supuesto, las dos cuestiones señaladas. Las consecuencias de la inexactitud del supuesto de nor-

---

\* Fragmento del trabajo de Tesis presentado por el autor, para optar al grado de Doctor en Ciencias Económicas en la Universidad Nacional de Córdoba.

malidad sobre la bondad de las inferencias ya ha sido estudiada en especial por A. K. Gayen (12) y R. G. Geary (11). Respecto de esta cuestión los resultados obtenidos por estos autores, ponen en evidencia la necesidad de construir un criterio apropiado para el contraste de la hipótesis de normalidad. Esto último constituye el objetivo central de este trabajo en el que presento una forma de docimar la hipótesis de normalidad de las perturbaciones aleatorias de los modelos econométricos lineales, y la utilizo para investigar la veracidad de esta hipótesis en un conjunto de modelos seleccionados como ejemplo.

El principal inconveniente para la aplicación directa en este contexto de alguna de las varias dócimas de la hipótesis de normalidad disponibles, radica en el hecho de que las perturbaciones aleatorias a que se refiere la hipótesis son no observables, y los residuos estimados que podrían reemplazarlas a los fines de la dócima, no son independientes entre sí. De ahí que éste sea el primer problema que he de encarar.

La factibilidad de abordar este problema de la forma en que se hace en este trabajo es admitida por Malinvaud<sup>1</sup> cuando señala: "Diversos tests se han construido para juzgar la normalidad de una variable aleatoria a partir de una muestra de observaciones independientes. Estos tests no se aplican directamente a nuestro problema, ya que los errores no son exactamente conocidos. Sin embargo parecen fáciles de adaptar, pues el vector de los residuos sigue una ley normal en un subespacio lineal cuando los errores son normales. Basta por tanto definir, a partir de los residuos,  $T - m$  magnitudes no correlacionadas entre sí y de la misma varianza, y probar la normalidad de la variable aleatoria correspondiente. Esta sugerencia no parece haberse utilizado en la práctica".

También Theil<sup>2</sup> sugiere esta línea de investigación cuando refiriéndose a las hipótesis convencionales de los modelos lineales,

<sup>1</sup> MALINVAUD, E., *Métodos Estadísticos de la Econometría*. Ariel - Barcelona, 1967, Cap. 8, pág. 288.

<sup>2</sup> THEIL, H., "The Analysis of Disturbances in Regression Analysis". *Journal of the American Statistical Association*, 60, 1965. Pág. 1068. Traducción propia.

escribe: "...A veces se quiere docimarlas contra algunas otras hipótesis alternativas, tales como heteroscedasticidad y no-normalidad. En todos estos casos es igualmente verdadero que el procedimiento de dócima se simplificaría considerablemente si la matriz de covarianzas de las perturbaciones estimadas fuera de la forma  $\sigma^2 I$  en lugar de  $\sigma^2 M$  ...".

Finalmente en un artículo reciente Putter<sup>3</sup> indica más expresamente que "alguna base ortonormal del espacio de los errores puede ser utilizada para docimar la normalidad, usando el test  $\chi^2$ ; tests basados en momentos muestrales, tests del tipo Kolmogorov - Smirnov o Crámer - von Mises, o el recientemente sugerido test de Shapiro - Wilk".

A pesar de que el enfoque del problema es sugerido tan claramente por estos autores no parece haber sido desarrollado en detalle el procedimiento propuesto. En particular, no conozco que hayan sido resueltos hasta ahora los interrogantes planteados al existir una infinidad de bases ortonormales del espacio de los residuos y varios tests de normalidad alternativos.

Además no se dispone de la distribución de las estadísticas de esos tests, cuando la hipótesis nula especifica el valor de la media poblacional, pero no el de la varianza.

Cada uno de estos aspectos son investigados en el presente trabajo para llegar a aplicar una dócima de normalidad a un conjunto de modelos econométricos particulares.

#### 1 — DETERMINACION DE LAS MAGNITUDES SOBRE LAS QUE SE APLICARA LA DOCIMA DE NORMALIDAD.

La naturaleza no observable de las perturbaciones aleatorias de los modelos econométricos lineales y el hecho de que los residuos estén correlacionados entre sí, nos lleva a perseguir una transformación de estos últimos que eliminando la correlación asegure, bajo la hipó-

<sup>3</sup> PUTTER, J., "Orthonormal Bases of Error Spaces and their Use for Investigating the Normality and Variances of Residuals". *American Statistical Association Journal*, 1967. Pág. 1033. Traducción propia.

tesis de normalidad, la independencia estocástica de las variables resultantes. El objetivo de este punto, consiste en precisar cuál de las infinitas transformaciones posibles se adoptará en definitiva.

1.1.1. Una transformación adecuada de los residuos.

El denominado "modelo lineal general"<sup>4</sup>, aplicado al conjunto de observaciones muestrales y utilizando notación matricial, puede expresarse como:

$$\begin{aligned}
 Y &= X\beta + \epsilon \\
 E(\epsilon) &= 0 \\
 V(\epsilon\epsilon') &= \sigma^2 I
 \end{aligned}$$

donde

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Así planteado, el vector Y (n x 1) es una observación de un vector aleatorio, X (n x p) es una matriz de cantidades conocidas, β (p x 1) es el vector de parámetros desconocidos y ε (n x 1) es un vector de variables aleatorias no observables. Para que el modelo sea de "rango completo" es necesario incorporar la exigencia de que el rango de X sea p ≤ n.

Incorporando la hipótesis de que la distribución de la variable es normal, o, en términos del modelo aplicado al conjunto de observaciones, el vector ε tiene distribución N (0, σ² I), se demuestra que los estimadores

$$\begin{aligned}
 \hat{\beta} &= (X'X)^{-1} X'Y \\
 \hat{\sigma}^2 &= \frac{Y'(I - X(X'X)^{-1}X')Y}{n - p}
 \end{aligned}$$

<sup>4</sup> Una formulación en detalle puede encontrarse en Graybill (10).

tienen las siguientes propiedades:

- (1) Son insesgados
- (2) Son consistentes
- (3) Son eficientes
- (4) Son suficientes
- (5)  $\hat{\beta}$  se distribuye  $N(\beta, \sigma^2 (X'X)^{-1})$
- (6) Son completos
- (7) Son insesgados de mínima varianza
- (8)  $\frac{(n-p)\hat{\sigma}^2}{\sigma^2}$  se distribuye como una  $\chi^2_{(n-p)}$
- (9)  $\hat{\beta}$  y  $\hat{\sigma}^2$  son independientes.

Estas propiedades permiten realizar estimaciones por intervalos y también, docimar hipótesis nulas respecto a  $\beta$  y  $\sigma^2$ .

Se denomina vector de residuos "e" a la diferencia:

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

que puede obtenerse una vez que ha sido estimado el vector  $\hat{\beta}$ .

Es fácil demostrar que:

$$E(e) = 0$$

$$y \quad V(e) = E(ee') = \sigma^2 [I - X(X'X)^{-1}X']$$

El vector "e" por ser la imagen de una transformación lineal<sup>5</sup> singular del vector Y de distribución  $N(0, \sigma^2 I)$ , tiene una distribución normal singular o degenerada. Es decir, e se distribuye

$$N\{0, \sigma^2 [I - X(X'X)^{-1}X']\}.$$

Mediante una transformación ortogonal del vector de los residuos, es posible conseguir un nuevo vector E tal que

$$E = \begin{bmatrix} Z \\ 0 \end{bmatrix}$$

<sup>5</sup> La matriz de la transformación es singular. No obstante la distribución de e será todavía normal. Véase T. W. Anderson (2), págs. 26 y 27.

donde  $Z_{(n-p) \times 1}$  tendrá distribución normal no singular, con media 0 y matriz de varianzas y covarianzas  $\sigma^2 I$ . Consecuentemente, los elementos del vector  $Z$  pueden considerarse  $n-p$  observaciones independientes provenientes de una población normal univariante de media 0 y desviación estandar  $\sigma$ .

Vale la pena destacar que no existe una única transformación ortogonal de  $e$ . Nótese no obstante, que una transformación ortogonal de un vector  $Z$  en particular, dejará invariante su distribución:

$$\begin{aligned} \text{Si} \quad Z & \text{ ——— } N(0, \sigma^2 I) \\ \text{para } C \text{ ortogonal } C'Z & \text{ ——— } N(C'0, C' \sigma^2 I C) \\ C'Z & \text{ ——— } N(0, \sigma^2 I) \end{aligned}$$

Es decir que cualquiera sea la transformación ortogonal del vector de los residuos que nos permita obtener el nuevo vector  $E$ , la distribución de este último permanecerá invariante; a pesar de ello, el hecho de que existan infinitas transformaciones ortogonales posibles introduce una complicación. Nuestro objetivo es contrastar la hipótesis de normalidad de los residuos para lo cual procuramos obtener, a partir de ellos, un vector de variables aleatorias independientes con el cual definiremos una estadística para el test. La existencia de infinitas transformaciones ortogonales susceptibles de conducirnos a un vector que reúna tales condiciones, nos obligará a exigir que el test sea invariante ante transformaciones ortogonales del vector  $E$ .

Otra vía alternativa de solución consiste en encontrar una transformación ortogonal particular y única que por alguna razón sea la más conveniente. Esta es la que seguiremos en la presente investigación.

Los elementos del vector

$$e = [I - X(X'X)^{-1}X'] Y = [I - X(X'X)^{-1}X'] \varepsilon$$

constituyen los "mejores estimadores lineales insesgados" (MELI)<sup>6</sup>

<sup>6</sup> Utilizaremos la expresión MELI para referirnos a los "mejores estimadores lineales insesgados". Es decir, MELI sustituye a la expresión BLUE ("Best linear unbiased estimator") que se utiliza en la literatura inglesa.

de los verdaderos valores de las perturbaciones  $\varepsilon$ . La demostración puede verse en Theil (27).

Dado que nuestro objetivo es contrastar una hipótesis referida a  $\varepsilon$ , es razonable valernos de los "mejores estimadores lineales insesgados" de sus valores, pero además por imperativo de nuestro objetivo, debemos agregar ahora la condición de que tenga matriz de covarianzas escalar. En otros términos debemos ahora encontrar los "mejores estimadores lineales insesgados y de matriz de covarianza escalar" de  $\varepsilon$ .

El vector  $Z$  formado por estimadores de los elementos de  $\varepsilon$ , que reúnen esa condición, es el siguiente:

$$Z = A'Y, \quad y$$

$$A' = [-P D^{1/2} P' (X_1' X_1^{-1}) \quad P D^{1/2} P']$$

donde  $P$  es una matriz ortogonal tal que:

$$P' M_{11} P = D \quad P' P = I,$$

siendo  $D$  diagonal y sus elementos las raíces características de  $M_{11}$ . A su vez:

$$M_{11} = I - X_1 (X' X)^{-1} X_1'$$

y  $X_1$  surge de particionar a  $X$  de la siguiente forma:

$$X = \begin{pmatrix} X_0 \\ X_1 \end{pmatrix}$$

donde  $X_0$  es  $p \times p$ , y  $X_1$  es  $(n-p) \times p$ .

La demostración de que este vector  $Z$  contiene "los mejores estimadores lineales insesgados y de matriz de covarianza escalar" de los elementos de  $\varepsilon$ , puede verse también en Theil (27).

### 1.2. Un ejemplo numérico.

Un ejemplo numérico simple contribuirá a aclarar el método de cálculo.

Supongamos que para el modelo  $y = \beta x + \varepsilon$  se dispone de la siguiente información

$$\begin{array}{c|ccc} y & 0 & 1 & -1 \\ \hline x & -1 & 1 & 2 \end{array}$$

Si decidimos estimar el valor de la perturbación aleatoria para las dos últimas observaciones, podemos escribir

$$y = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \quad X = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \quad X_0 = [-1] \quad X_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Comenzamos calculando  $X'X$ :

$$X'X = [-1 \quad 1 \quad 2] \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = 6$$

y por consiguiente:

$$(X'X)^{-1} = \frac{1}{6}$$

Podemos ahora calcular  $M_{11}$

$$M_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix} \frac{1}{6} [1 \quad 2] = \begin{bmatrix} \frac{5}{6} & -\frac{2}{6} \\ -\frac{2}{6} & \frac{2}{6} \end{bmatrix}$$

Para calcular D debemos encontrar las raíces características de  $M_{11}$ . Para ello resolvemos la ecuación característica

$$\begin{bmatrix} \frac{5}{6} - \lambda & -\frac{2}{6} \\ -\frac{2}{6} & \frac{2}{6} - \lambda \end{bmatrix} = 0$$

que resulta ser

$$\lambda^2 - \frac{7}{6}\lambda + \frac{6}{36} = 0$$

cuyas raíces son  $\lambda_1 = 1$  y  $\lambda_2 = \frac{1}{6}$

Estamos en condiciones de formar la matriz D:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{6} \end{pmatrix}$$

Para obtener P debemos encontrar un par de vectores característicos de  $M_{11}$  y normalizarlos.

A partir de los sistemas

$$\left\{ \begin{pmatrix} \frac{5}{6} & -\frac{2}{6} \\ -\frac{2}{6} & \frac{2}{6} \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\} \begin{pmatrix} P_{11} \\ P_{21} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left\{ \begin{pmatrix} \frac{5}{6} & -\frac{2}{6} \\ -\frac{2}{6} & \frac{2}{6} \end{pmatrix} - \frac{1}{6} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\} \begin{pmatrix} P_{12} \\ P_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

y exigiendo que

$$[P_{11} \ P_{21}] \begin{pmatrix} P_{11} \\ P_{21} \end{pmatrix} = 1$$

$$[P_{12} \ P_{22}] \begin{pmatrix} P_{12} \\ P_{22} \end{pmatrix} = 1$$

$$[P_{11} \ P_{21}] \begin{pmatrix} P_{12} \\ P_{22} \end{pmatrix} = 0$$

obtenemos:

$$P_{11} = \frac{2}{\sqrt{5}}, P_{12} = \frac{1}{\sqrt{5}}, P_{21} = \frac{1}{\sqrt{5}}, P_{22} = \frac{-2}{\sqrt{5}}$$

Con lo cual disponemos también de P

$$P = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}$$

Estamos ahora en condiciones de determinar  $A_1$  y  $A_0$ :

$$A_1 = P D^{1/2} P' = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{4\sqrt{6}+1}{\sqrt{150}} & \frac{2\sqrt{6}-2}{\sqrt{150}} \\ \frac{2\sqrt{6}-2}{\sqrt{150}} & \frac{\sqrt{6}+4}{\sqrt{150}} \end{bmatrix} = \begin{bmatrix} 0,8816565 & 0,2367033 \\ 0,2367033 & 0,5266016 \end{bmatrix}$$

$$A_0 = -(X_1 X_0^{-1})' A_1 = - \begin{bmatrix} -1 \\ 2 \end{bmatrix} [-1]^{-1} \begin{bmatrix} \frac{4\sqrt{6}+1}{\sqrt{150}} & \frac{2\sqrt{6}-2}{\sqrt{150}} \\ \frac{2\sqrt{6}-2}{\sqrt{150}} & \frac{\sqrt{6}+4}{\sqrt{150}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{5}{\sqrt{150}} & -\frac{10}{\sqrt{150}} \end{bmatrix} = [0,4082499 \quad -0,8164998]$$

Por lo tanto:

$$Z = A'Y = \begin{bmatrix} 0,4082499 & 0,8816565 & 0,2367033 \\ -0,8164998 & 0,2367033 & 0,5266016 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -0,1715466 \\ 1,3431014 \end{bmatrix}$$

2 — SELECCION DE LA ESTADISTICA A UTILIZAR EN LA DOCIMA.

Existen varias estadísticas que permiten docimar la hipótesis de normalidad de una población a partir de una muestra aleatoria simple. En este punto se analizará la posibilidad del empleo de las mismas para docimar la normalidad de las perturbaciones aleatorias de un modelo lineal.

2.1. *Distribución de una variable poblacional y distribución de la muestra.*

Supongamos que  $F(\chi)$  es la función de distribución de una variable aleatoria  $\chi$ . En particular, nos interesa en esta investigación el importante caso en que  $\chi$  es una variable normal, o sea en el que

$$F(\chi) = \int_{-\infty}^{\chi} \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(\chi - \mu)^2}{2 \sigma^2}} d\chi \quad (2.1.1)$$

donde  $\mu$  es la media de  $\chi$  y  $\sigma$  su desviación standard.

Una muestra de esta población estará constituida por  $n$  observaciones  $\chi_1, \chi_2, \dots, \chi_n$ . Llamaremos aquí distribución de la muestra a la sencilla distribución discreta que se obtiene asignando probabilidad  $1/n$  a cada una de estas observaciones muestrales. Si ordenamos en forma creciente las observaciones y simbolizamos este ordenamiento mediante  $y_1, y_2, \dots, y_n$ , podemos expresar la función de distribución de la muestra de la siguiente forma:

$$F^*(x) = \begin{cases} 0 & \text{para } x \leq y_1 \\ \frac{1}{n} & \text{para } y_1 < x \leq y_2 \\ \frac{2}{n} & \text{para } y_2 < x \leq y_3 \\ \vdots & \vdots \\ 1 & \text{para } x > y_n \end{cases} \quad (2.1.2)$$

En el Gráfico 1 se ha representado, como ejemplo, la función de distribución de una muestra particular de tamaño 8.

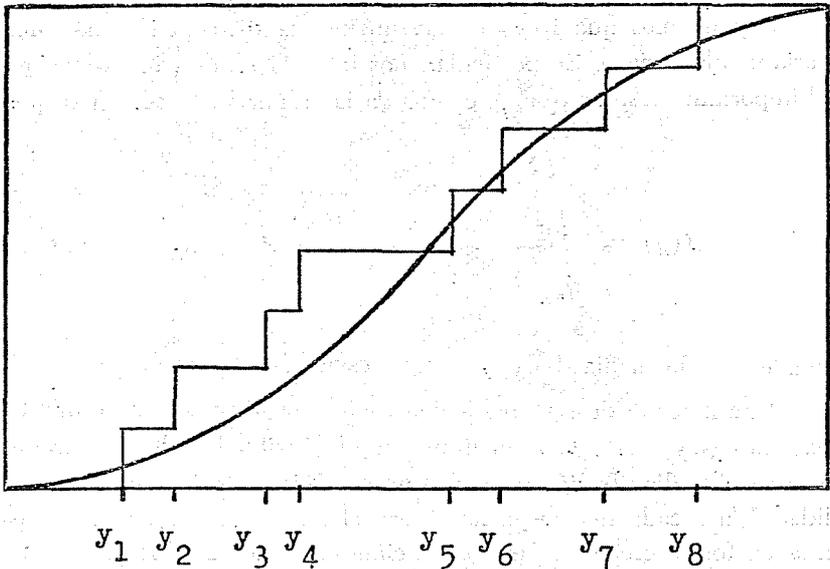


GRAFICO 1

Frecuentemente la normalidad de  $\chi$  y consecuentemente la expresión (2.1.1) para  $F(\chi)$ , constituirán sólo una hipótesis, y para el caso en que sea verdadera,  $F^*(\chi)$  puede considerarse como su "imagen estadística". Para valores grandes de  $n$  puede esperarse que la distribución de la muestra constituya una aproximación de la distribución de la variable poblacional.

Cuando se desea docimar la hipótesis que postula a  $F(\chi)$  en base a las observaciones reales, parece natural utilizar alguna medida de la "divergencia" entre  $F(\chi)$  y  $F^*(\chi)$  y basar la dócima en la distribución en el muestreo de tal medida. En esta línea de trabajo, han sido propuestas varias dócimas de bondad de ajuste que analizaremos a fin de seleccionar la más adecuada para ser aplicada al problema que investigamos en este trabajo.

## 2.2. Posibilidades de aplicación de las estadísticas disponibles en la docimacia de la hipótesis de normalidad de las perturbaciones aleatorias.

En el punto 1 logramos una transformación del vector de los residuos que conduce a un nuevo vector de variables aleatorias que mantienen la misma media y varianza pero que están incorrelacionadas.

Cuando se establece la hipótesis de normalidad de las perturbaciones aleatorias del modelo, además de las hipótesis convencionales de esperanza nula y matriz de covarianzas escalar, los residuos son también normales si la primer hipótesis es verdadera y la transformación ortogonal de los residuos (en particular la que conduce a los mejores estimadores lineales insesgados y de matriz de covarianza escalar) nos lleva a un nuevo vector de variables que siguen siendo normales. Por consiguiente el vector

$$Z = A' Y$$

donde  $A'$  cumple las condiciones establecidas en el punto 1, contiene un conjunto de  $(n - p)$  variables aleatorias normales de media 0 y desviación estándar desconocida  $\sigma$  y, por consiguiente,

puede considerarse como una muestra aleatoria de tamaño  $(n - p)$  proveniente de una población normal  $(0, \sigma)$ .

En lo sucesivo representaremos esta muestra por  $Z_1, Z_2, \dots, Z_{n-p}$  y una vez que ha sido ordenada en forma creciente por  $v_1, v_2, \dots, v_{n-p}$ .

En el análisis de la posibilidad de aplicar a esta muestra las diversas estadísticas disponibles, debemos tener especialmente en cuenta que para la distribución poblacional hipotética se conoce la media y se desconoce la desviación standard.

La estadística  $\chi^2$  de Pearson (20),  $W$  de Shapiro y Wilk (25) y el cociente  $\mu$  entre rango y desviación standard (6), pueden aplicarse aun cuando  $F(x)$  no esté completamente especificada.

Si bien originariamente la  $\chi^2$  se aplicó al caso en que la distribución hipotética estaba completamente especificada, R. A. Fisher demostró que si para la estimación de los parámetros desconocidos se emplea el método de la mínima  $\chi^2$ , entonces la única diferencia con el caso de la especificación completa estriba en que se reduce el número de grados de libertad de la  $\chi^2$  en tantas unidades como parámetros desconocidos haya.

El procedimiento para la aplicación de  $\chi^2$  en nuestro caso particular, requerirá agrupar en  $r$  intervalos las observaciones  $v_1, v_2, \dots, v_{n-p}$  y estimar  $\sigma^2$  a partir de

$$\sum_{i=1}^r \frac{V_i - n p_i}{p_i} \frac{\partial p_i}{\partial \sigma^2} = 0 \quad (2.2.1)$$

donde  $V_1, \dots, V_r$  representan las frecuencias de los  $r$  intervalos y  $p_i = P(S_i)$  donde  $S_i$  es la  $i$ -ésima parte de una partición finita del espacio de la variable aleatoria " $v$ ". De (2.2.1) resulta

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (v_i - \bar{v})^2}{n - p}$$

Con  $\hat{\sigma}^2$  se pueden calcular las probabilidades que corresponden a cada intervalo.

$$p_i = P(S_i) = \int_{V \in S_i} \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{v^2}{2\hat{\sigma}^2}} dv \quad \text{para } i = 1, 2, \dots, r$$

Con estos elementos se podrá computar

$$\chi^2_{\text{obs}} = \sum_{i=1}^r \frac{(V_i - n p_i)^2}{n p_i}$$

el que deberá compararse con el punto crítico de una  $\chi^2$  con  $r - 2$  grados de libertad para el nivel de significación elegido. Si el valor observado es mayor al punto crítico corresponderá rechazar la hipótesis de normalidad.

La estadística  $W$  de Shapiro y Wilk ha sido propuesta específicamente para el caso en que se desconoce la media y la varianza de la distribución normal hipotética. Dado que en nuestro caso se sabe que la media es 0 deberá estimarse  $\sigma$  a partir de

$$v_i = \sigma s_i \quad i = 1, 2, \dots, n - p$$

lo que equivale a asignar a  $\sigma$  el valor que minimiza la forma cuadrática

$$(V - \sigma M)' \Omega^{-1} (V - \sigma M) \quad V = [v_i]$$

de donde resulta

$$\hat{\sigma} = \frac{M' \Omega^{-1} V}{M' \Omega^{-1} M}$$

$S^2$  calculado mediante la simple suma de los valores de  $v_i$  elevados al cuadrado, será ahora un estimador insesgado de  $\sigma^2$ .

En conclusión  $W$  puede determinarse mediante cualquiera de las siguientes expresiones:

$$W = \frac{R^4 \hat{\sigma}^2}{C^2 S^2} = \frac{b^2}{S^2} = \frac{(A' V)^2}{S^2} = \frac{\left( \sum_{i=1}^{n-p} a_i v_i \right)^2}{\sum_{i=1}^{n-p} v_i^2}$$

donde

$$R^2 = M' \Omega^{-1} M$$

$$C^2 = M' \Omega^{-1} \Omega^{-1} M$$

$$A' = (a_1, \dots, a_n) = \frac{M' \Omega^{-1}}{(M \Omega^{-1} \Omega^{-1})^{1/2}}$$

$$b = \frac{R^2 \hat{\sigma}^2}{C} \quad \text{y} \quad S^2 = \sum_{i=1}^{n-p} v_i^2$$

Dado que la tabla de puntos críticos confeccionada por Shapiro y Wilk (24) corresponde al caso en que se desconocen la media y la varianza, no es exactamente aplicable al problema que ahora nos ocupa, pues los valores de la verdadera función de distribución para cada nivel de significación serán en este caso algo menor. No obstante el problema no es grave porque lo que ocurrirá en la práctica es que cuando utilicemos la tabla para un nivel de significación dado en realidad estaremos trabajando con una probabilidad de error de tipo I algo mayor que la que aquel indica. Queda no obstante la posibilidad de que por un procedimiento similar al empleado por Shapiro y Wilk se determine la tabla de valores críticos que corresponde al estadístico  $W$  en este caso particular en que sólo se desconoce la varianza.

Con respecto al cociente entre el rango y la desviación standard muestral no existe ningún problema para su aplicación en estas circunstancias y las tablas más completas de Pearson y Stephens son perfectamente utilizables.

Las estadísticas del tipo Cramer-Von Mises (5 y 18) y Kolmogorov-Smirnov (15), han sido propuestas para cuando la hipótesis específica completamente la distribución de la variable poblacional y para esta circunstancia se han obtenido sus distribuciones asintóticas y las que corresponden a algunos tamaños muestrales particulares. No obstante sería posible extender su aplicación, incluso a los casos en que se desconocen algunos parámetros poblacionales, para lo cual sería necesario determinar en forma empírica la distribución del estadístico. Para  $D_n$  de Kolmogorov y Smirnov esta tarea fue emprendida por Hubert W. Lilliefors (7) en el caso en que se desconoce la media y la varianza y se estiman a partir de la muestra. No obstante la tabla de puntos críticos no es exactamente aplicable al problema que nos interesa porque también la media ha sido estimada a partir de la muestra. A su aplicación deben realizarse las mismas advertencias formuladas para la tabla de Shapiro y Wilk de la distribución de W.

No existe ningún impedimento teórico para determinar en forma empírica la distribución de  $D_n$  de Kolmogorov - Smirnov,  $V_n$  variante de Kuiper (16),  $W^2_n$  de Cramer - Von Mises, y  $U^2_n$  variante de G. S. Watson (29), cuando sólo la varianza se estima a partir de la muestra y por consiguiente también estas estadísticas podrán aplicarse en la docimacia de la hipótesis de normalidad de las perturbaciones.

Otra circunstancia a tener en cuenta en el análisis de la aplicabilidad de las diversas dócimas al problema que nos ocupa surge porque en los problemas econométricos son frecuentes las muestras pequeñas. Esto quita posibilidades a la  $\chi^2$  por cuanto los intervalos en que deben agruparse las observaciones muestrales no deben contener, en general, menos de cinco observaciones y por consiguiente para muestras pequeñas se pierde mucha información. El resto de los criterios es aplicable aun a muestras muy pequeñas.

Por último la elección definitiva entre las diversas estadísticas debe realizarse teniendo en cuenta la potencia de los tests en que se emplean.

Para este análisis nos valdremos de los resultados presentados por Shapiro y Wilk (25) y Lilliefors (7), referidos ambos al caso en que se desconocen la media y la varianza de la distribución poblacional.

Los primeros compararon la potencia de los tests de normalidad  $W$ ,  $\chi^2$ ,  $b_1$ ,  $b_2^7$ ,  $D_n$ ,  $W_n^2$ ,  $\mu$ , el de Cramer-Von Mises ponderado por

$$F(\chi)$$

por  $\frac{F(\chi)}{1 - F(\chi)}$  y la versión de Durbin del procedimiento

$$1 - F(\chi)$$

Kolmogorov-Smirnov. Para todos ellos, excepto  $\chi^2$ , determinaron la distribución bajo la hipótesis nula en forma empírica a partir de 500 muestras de tamaño 20. Las potencias para las diferentes hipótesis alternativas fue determinada también en forma empírica a partir de 200 muestras.

Lilliefors realizó una investigación semejante mediante 500 muestras de tamaño 20 y otras 500 muestras de tamaño 10, pero sólo para comparar entre los test  $\chi^2$  y Kolmogorov-Smirnov.

Del análisis de los cuadros N<sup>o</sup> 1 y 2, surge que  $W$  es altamente sensible ante una variada gama de hipótesis alternativas y su potencia es sólo levemente superada por la  $\chi^2$  para las hipótesis alternativas normal logarítmica, Poisson y Binomial, por la  $b_2$  para cuando la hipótesis alternativa es la distribución de La Place, por la de Cramer - Von Mises ponderada para la distribución de Cauchy y por el cociente rango/desviación standard para la distribución uniforme.

En general  $b_1$  no es sensible cuando la hipótesis alternativa es simétrica como lo muestra el caso de la distribución uniforme.  $b_2$  no es sensible a las asimetrías.  $D_n$  y  $W_n^2$  tienen en general potencias inferiores a  $W$  y lo mismo ocurre para el  $W_n^2$  ponderado salvo para el caso en que la hipótesis alternativa es la distribución de Cauchy. La versión de Durbin de la estadística Kolmogorov-Smirnov en general es superada en potencia por otras estadísticas salvo para el caso de alternativas discretas. Por último  $\mu$  es poco sensible a la asimetría y puntiagudés.

<sup>7</sup>  $b_1$  y  $b_2$  son respectivamente, el tercer y cuarto momentos muestrales estandarizados, que también se emplean para docimar la normalidad.

CUADRO Nº 1

*Potencia empírica de diversos tests de normalidad al 5% de significación para varias distribuciones alternativas.*

La hipótesis nula no especifica ni la media ni la varianza de la distribución.

	$\beta_1$	$\beta_2$	W	$\chi^2$	$b_1$	$b_2$	K S	CVM	WCVM	D	$\mu$
$\chi^2$ (1)	2.83	15.0	0.98	0.94	0.89	0.53	0.44	0.44	0.54	0.87	0.10
$\chi^2$ (2)	2.00	9.0	.84	.33	.74	.34	.27	.23	.27	.42	.08
$\chi^2$ (4)	1.41	6.0	.50	.13	.49	.27	.18	.13	.16	.15	.06
$\chi^2$ (10)	0.89	4.2	.29	.07	.29	.19	.11	.10	.11	.07	.06
No cent. $\chi^2$	0.73	3.7	.59	.10	.50	.20	.19	.16	.18	.20	.10
Log. normal	6.19	113.9	.93	.95	.89	.58	.44	.48	.62	.82	.06
Cauchy	—	—	.88	.41	.77	.81	.45	.55	.98	.85	.56
Uniforme	0	1.8	.23	.11	.00	.29	.13	.09	.10	.08	.38
Logística	0	4.2	.08	.06	.12	.06	.06	.03	.05	.05	.07
Beta (2, 1)	-0.57	2.4	.35	.08	.08	.13	.08	.10	.12	.12	.23
La Place	0	6.0	.25	.17	.25	.27	.07	.07	.29	.16	.19
Poisson (1)	1.00	4.0	.99	1.00	.26	.11	.55	.22	.31	1.00	.35
Binomial, (4, 0.5)	0	2.5	.71	1.00	.02	.03	.38	.15	.17	1.00	.20
T (5,2.4)	0.79	2.2	.55	.14	.24	.20	.23	.20	.22	—	—
T (10,3.1)	0.97	2.8	.89	.32	.51	.24	.32	.30	.30	—	—

CUADRO Nº 2

Probabilidad de rechazar la hipótesis de normalidad usando la estadística  $D_n$  y la  $\chi^2$  cuando el tamaño de la muestra es 20.

Distribución poblacional	Test de Kolmogorov-Smirnov		Test de Chi-cuadrado	
	=.05	=.10	=.06	=.12
Normal	.06	.10	.06	.12
Chi-cuadrado, 3 grados de libertad	.44	.55	.20	.27
t de Student, 3 grados de libertad	.50	.58	.40	.52
Exponencial	.61	.72	.29	.41
Uniforme	.12	.22	.10	.18

En conclusión,  $W$  aparece como la estadística más apropiada para el problema que nos ocupa pues además de no plantear dificultades por el desconocimiento de la varianza y ser aplicable a muestras incluso pequeñas, presenta para una variada gama de hipótesis alternativas una alta potencia, escasamente superada en algunos casos.

### 3 — LA DISTRIBUCION DE $W$

Hasta aquí hemos conseguido una transformación de los residuos que conduce a un conjunto de variables incorrelacionadas. Cuando la hipótesis de normalidad de las perturbaciones es verdadera, este conjunto puede interpretarse como una muestra aleatoria simple proveniente de una población normal de media cero y varianza desconocida  $\sigma^2$ . Además hemos seleccionado una adaptación de la estadística  $W$  de Shapiro y Wilk a fin de plantear un criterio de dócima de la hipótesis de normalidad de las perturbaciones originales.

Para que el nuevo conjunto de variables aleatorias que se utiliza para la dócima sea único, es necesario fijar qué perturbaciones quedarán sin estimar luego de la transformación. Hemos adelantado que esta selección se realizará en base a consideraciones de potencia de la dócima mencionada, por lo cual analizaremos este punto. A continuación se utiliza el método Montecarlo para obtener una tabla de puntos críticos de  $W$  para una gama apropiada de niveles de significación y tamaños muestrales. Finalmente presentamos en forma resumida los pasos que deben seguirse para la aplicación práctica del test.

#### 3.1. Selección de las perturbaciones que no resultarán estimadas.

Bajo los supuestos convencionales del modelo lineal general, el vector de residuos

$$e = [I - X(X'X)^{-1}X']Y$$

contiene los mejores estimadores lineales insesgados de las  $n$  perturbaciones aleatorias del vector  $\epsilon$ . De entre éstas deben seleccionarse

$p$  para las que no se conseguirán "los mejores estimadores lineales inescogados de matriz de covarianzas escalar".

Nos interesa que nuestra d6cimo sea lo m6s potente posible, especialmente para aquellas hip6tesis alternativas que implican marcada asimetría o puntiagudez diferente a la normal. Por consiguiente conviene resignar la estimaci6n de aquellas perturbaciones que menos contribuyen a señalar esas dos características y que son obviamente, las que han asumido valores m6s cercanos a cero. Dado que en el momento de tomar esta decisi6n s6lo conocemos los residuos como estimaci6n de las perturbaciones, deberemos seleccionar aquellas a las que correspondan los residuos m6s cercanos a cero.

Como m6todo pr6ctico convendr6 ordenar en forma creciente los residuos en valor absoluto y eliminar como objetivo de la siguiente estimaci6n los que ocupan los  $p$  primeros lugares de dicho ordenamiento.

### 3.2. Aproximaci6n empírica de la distribuci6n de $W$ .

Shapiro y Wilk han tabulado la distribuci6n empírica  $W$  para cuando se desconocen la media y la varianza de la distribuci6n poblacional y las observaciones se obtienen directamente a partir de una muestra aleatoria simple.

El contexto al que aplicaremos aquí la adaptaci6n de la estadística  $W$  difiere del que dio lugar a la distribuci6n empírica tabulada por dichos autores. Esas diferencias son las siguientes:

- a) en nuestro caso s6lo se desconoce la varianza poblacional, ya que la hip6tesis puntualiza que la media de la distribuci6n es 0.
- b) las observaciones se generan mediante transformaciones lineales sobre una muestra aleatoria simple de tamaño  $n$  a partir de la distribuci6n de una variable aleatoria "y" que depende de un conjunto de  $p$  cantidades conocidas  $\chi_1, \chi_2, \dots, \chi_p$  y de  $p$  par6metros desconocidos  $\beta_1, \beta_2, \dots, \beta_p$ .

La primera diferencia determina que la distribuci6n de  $W$  depende de  $n - p$  y n6 del tamaño de la muestra original de observaciones de  $y$ .

Para conseguir una aproximación de la distribución en la hipótesis nula de  $W$  con estas variantes, hemos aplicado el método Montecarlo. Trabajamos con el modelo

$$y_i = \beta \chi_i + \varepsilon_i \quad \text{para } i = 1, \dots, n$$

haciendo asumir a  $n$  todos los valores enteros entre 3 y 51. Para  $\chi_i$  adoptamos a priori los valores

$$\frac{i - \frac{n+1}{2}}{\frac{n-1}{2}}$$

y a  $\varepsilon_i$  se asignaron números aleatorios normales extraídos de las tablas de la "Rand Corporation". Este proceso se repitió 1.000 veces para cada tamaño muestral.

Para cada muestra aleatoria generada de esta forma se siguieron los siguientes pasos:

a) se calcularon los valores de  $y_i$  (vector  $Y$ ).

b) se estimó  $\hat{\beta} = \frac{\sum y_i \chi_i}{\sum \chi_i^2} = (X'X)^{-1} X'Y$

c) se calcularon los residuos  $e_i = y_i - \hat{\beta} \chi_i$  (vector  $e$ ).

d) se determinó el orden del mínimo  $|e_i|$ , al que llamamos  $e_n$ .

e) se formó el vector  $X_1$  con todos los  $e_i$  excepto  $e_n$ .

f) se calculó  $(X'X)^{-1} = \frac{1}{\sum \chi_i^2}$ .

g) se calculó  $M_{11} = I - X_1 (X'X)^{-1} X_1'$

h) se determinó las  $m = n - 1$  raíces características de  $M_{11}$  ( $\lambda_1, \lambda_2, \dots, \lambda_m$ ) y su matriz de vectores propios normalizados ( $P$ ).

i) se formó la matriz  $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & \\ 0 & 0 & \lambda_m \end{bmatrix}$ .

j) se calculó  $A_1 = P D^{1/2} P'$ .

k) se calculó  $A_0 = -(X_1 X_1^{-1}) A_1 = -\left( X_1 \frac{1}{\chi^2} \right)' A_1$

l) se determinó  $Z = A' Y = [A'_0 A'_1] \begin{bmatrix} Y_0 \\ Y_1 \end{bmatrix}$

m) se ordenaron  $m = n - 1$  elementos de  $Z$  obteniendo

$$v_1, v_2, \dots, v_m$$

n) se calculó  $b = \sum_{i=1}^k a_{m-i+1} (v_{m-i+1} - v_i)$

siendo  $m = n - 1 = 2k$  ó  $m = n - 1 = 2k + 1$  según que  $n - 1$  sea par o impar. Los coeficientes  $a_{m-i+1}$  fueron extraídos en las tablas de Sarhan y Greenberg completadas por Shapiro y Wilk.

ñ) se calculó  $S^2 = \sum_{i=1}^m v_i^2$

o) se determinó  $W = \frac{b^2}{S^2}$

Al cabo de todas las repeticiones del experimento fue posible disponer de los valores de  $W$  para cuando su función de distribución asume los valores 0.01, 0.02, 0.05, 0.10. Estos valores para el caso en que se dispone desde 2 hasta 50 grados de libertad aparecen en el cuadro N° 3.

Además, para evaluar el grado de aproximación obtenido con esta simulación se procedió a calcular el momento de orden uno en base a la distribución empírica de  $W$  para los diversos grados de libertad. En el cuadro N° 4 aparecen, junto al verdadero momento de orden uno calculado a partir de

$$E(W) = \frac{E(b^2)}{E(S^2)}$$

## CUADRO N° 3

*Puntos críticos para el test de normalidad de las perturbaciones  
aleatorias basado en W.*

Para muestras de 2 a 50 grados de libertad.

m	0,01	0,025	0,05	0,10
2	0,535	0,562	0,632	0,670
3	0,502	0,504	0,511	0,526
4	0,440	0,530	0,561	0,594
5	0,548	0,572	0,609	0,644
6	0,594	0,619	0,656	0,688
7	0,625	0,651	0,688	0,718
8	0,655	0,680	0,715	0,744
9	0,679	0,703	0,736	0,763
10	0,702	0,725	0,757	0,782
11	0,720	0,742	0,772	0,796
12	0,737	0,759	0,787	0,809
13	0,751	0,772	0,799	0,820
14	0,766	0,785	0,811	0,831
15	0,779	0,798	0,822	0,840
16	0,791	0,809	0,831	0,849
17	0,800	0,817	0,839	0,856
18	0,810	0,825	0,847	0,863
19	0,817	0,832	0,853	0,868
20	0,824	0,839	0,859	0,874
21	0,831	0,845	0,864	0,879
22	0,838	0,851	0,869	0,883
23	0,842	0,856	0,864	0,887
24	0,847	0,860	0,877	0,891
25	0,852	0,864	0,881	0,893
26	0,856	0,869	0,884	0,897
27	0,860	0,872	0,888	0,900
28	0,864	0,875	0,891	0,902
29	0,867	0,879	0,894	0,904
30	0,870	0,881	0,896	0,907
31	0,872	0,884	0,899	0,910
32	0,876	0,886	0,901	0,912
33	0,879	0,889	0,903	0,913
34	0,881	0,892	0,906	0,915
35	0,884	0,894	0,907	0,917

m	0,01	0,025	0,05	0,10
36	0,887	0,896	0,909	0,919
37	0,889	0,899	0,911	0,920
38	0,892	0,901	0,913	0,922
39	0,894	0,903	0,915	0,924
40	0,896	0,905	0,917	0,925
41	0,898	0,906	0,918	0,927
42	0,900	0,908	0,920	0,928
43	0,901	0,910	0,921	0,929
44	0,903	0,913	0,923	0,930
45	0,905	0,913	0,924	0,932
46	0,907	0,915	0,924	0,932
47	0,908	0,916	0,926	0,934
48	0,910	0,918	0,927	0,934
49	0,910	0,918	0,928	0,936
50	0,911	0,919	0,928	0,936

CUADRO N° 4

*Momento teórico de orden 1 y la aproximación obtenida por el método Monte Carlo, obtenida para la distribución de W bajo la hipótesis nula.*

Se especifica la media pero no la varianza

GRADOS DE LIBERTAD	$\mu_1$	$\lambda$ $\mu_1$
3	0,6090	0,5999
4	0,6759	0,6772
5	0,7220	0,7300
6	0,7560	0,7572
7	0,7819	0,7900
8	0,8027	0,8107
9	0,8126	0,8132
10	0,8337	0,8294
11	0,8457	0,8409
12	0,8558	0,8560
13	0,8648	0,8632
14	0,8726	0,8731
15	0,8795	0,8800
16	0,8856	0,8870
17	0,8912	0,8920
18	0,8963	0,8957
19	0,9007	0,9009
20	0,9046	0,9070

$$E(b^2) = \frac{R^4}{C^2} E(\hat{\sigma}^2) = \frac{R^4}{C^2} \{ \text{Var}(\hat{\sigma}) + E(\hat{\sigma})^2 \} =$$

$$= \sigma^2 [R^2 (R^2 + 1) / C^2]$$

$$E(S^2) = m \sigma^2$$

$$E(W) = \frac{R^2 (R^2 + 1)}{C^2 m}$$

Para el tamaño  $n = 21$ , o lo que es lo mismo para 20 grados de libertad, se procedió a calcular con cada muestra aleatoria el valor de las estadísticas  $W^2_n$ ,  $U^2_n$ ,  $D^2_n$ ,  $V^2_n$ ,  $\mu$ ,  $b_1$  y  $b_2$ . De esta forma se obtuvieron los puntos críticos para sus respectivos tests de normalidad. Estos valores aparecen en el cuadro N° 5.

CUADRO N° 5

*Puntos críticos para los tests de normalidad de las perturbaciones aleatorias basados en  $W^2_n$ ,  $V^2_n$ ,  $D^2_n$ ,  $U^2_n$ ,  $u$ ,  $b_1$  y  $b_2$ . Para muestras de veinte grados de libertad.*

	Límite inferior		Límite superior	
	0,01	0,05	0,05	0,01
$W^2_n$	0,026	0,037	0,460	0,731
$U^2_n$	0,020	0,028	0,184	0,293
$V^2_n$	0,158	0,166	0,361	0,438
$u$	3,02	3,23	4,52	4,87
$b_1$	-0,54	-0,53	0,53	0,54
$b_2$	1,25	1,80	5,21	5,73
$D_n$			0,194	0,236

Con estos elementos es posible determinar a partir de otra simulación, la potencia de la dócima a que conduce  $W$  en comparación con la de  $W^2_n$ ,  $U^2_n$ ,  $D^2_n$ ,  $V^2_n$ ,  $\mu$ ,  $b_1$  y  $b_2$ . Un análisis semejante, pero realizado en base a los resultados obtenidos por Shapiro y Wilk, nos llevó a elegir  $W$  como la estadística más conveniente para

el problema que nos ocupa. Con los resultados de esta nueva simulación estaremos en condiciones de verificar esa conclusión para los datos de potencia de las distintas d6cimas aplicadas directamente al contraste de la hip6tesis de normalidad de las perturbaciones de un modelo lineal.

Se realizaron simulaciones compuestas por 500 muestras cada una para las siguientes distribuciones de  $\epsilon$ :

- a)  $\chi^2(1) - 1$
- b)  $\chi^2(2) - 2$
- c)  $\chi^2(10) - 10$
- d)  $\chi^2(1)$
- e)  $\chi^2(2)$
- f)  $\chi^2(10)$
- g) Log. normal
- h) uniforme  $(-1, 1)$
- i)  $\chi - 1$ , donde  $\chi$  es Poisson de media 1
- j)  $m - 2$ , donde  $m$  es binomial  $\left(4, \frac{1}{2}\right)$
- k) binomial  $\left(4, \frac{1}{2}\right)$

En cada uno de estos casos se computaron los valores de  $W$ ,  $\chi^2$ ,  $W^2_n$ ,  $U^2_n$ ,  $D^2_n$ ,  $V^2_n$ ,  $b_1$  y  $b_2$  y se calcul6 el porcentaje de veces que result6 rechazada la hip6tesis nula a los niveles del 5% y del 1% de significaci6n. En el cuadro N9 6 aparecen los resultados, los que confirman plenamente las conclusiones que hab6bamos elaborado en base a los resultados de Shapiro y Wilk para el caso en que la hip6tesis nula no especifica ni la media ni la varianza de la distribuci6n poblacional.

### 3.3. Resumen del procedimiento para la aplicaci6n de la d6cima.

Para docimar la hip6tesis de normalidad de las perturbaciones a partir del modelo

CUADRO N° 6

*Potencia empírica de los tests de normalidad basados en  $W_n^2$ ,  $U_n^2$ ,  $D_n^2$ ,  $V_n^2$ ,  $u$ ,  $b_1$ ,  $b_2$  y  $W$  al 5 y al 1% de significación para varias distribuciones alternativas.*

La hipótesis nula especifica la media pero no la varianza de la distribución.

	$W_n^2$		$U_n^2$		$D_n^2$		$V_n^2$		$u$		$b_1$		$b_2$		$W$	
	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01
$\chi^2$ (1) — 1	0,41	0,29	0,42	0,29	0,38	0,27	0,41	0,30	0,09	0,03	0,85	0,62	0,72	0,53	0,89	0,65
$\chi^2$ (2) — 2	0,40	0,23	0,40	0,20	0,35	0,20	0,41	0,27	0,08	0,03	0,77	0,45	0,61	0,42	0,77	0,44
$\chi^2$ (10) — 10	0,31	0,15	0,33	0,12	0,22	0,12	0,32	0,10	0,05	0,01	0,51	0,31	0,30	0,23	0,63	0,31
$\chi^2$ (1)	0,45	0,27	0,38	0,23	0,45	0,34	0,51	0,27	0,12	0,07	0,87	0,69	0,81	0,69	0,92	0,73
$\chi^2$ (2)	0,39	0,20	0,41	0,17	0,42	0,33	0,43	0,30	0,12	0,03	0,79	0,53	0,40	0,31	0,81	0,53
$\chi^2$ (10)	0,35	0,18	0,39	0,19	0,25	0,09	0,38	0,17	0,07	0,01	0,53	0,31	0,19	0,10	0,67	0,35
log. normal	0,12	0,05	0,10	0,03	0,10	0,03	0,11	0,05	0,03	0,00	0,70	0,61	0,10	0,03	0,90	0,69
uniforme(-1, +1)	0,40	0,21	0,32	0,17	0,39	0,21	0,43	0,27	0,35	0,23	0,00	0,00	0,17	0,12	0,16	0,09
x, Poisson (1)	0,52	0,20	0,47	0,23	0,50	0,22	0,57	0,31	0,33	0,22	0,23	0,10	0,09	0,03	0,95	0,72
X — 1	0,30	0,17	0,23	0,12	0,32	0,21	0,38	0,20	0,27	0,22	0,10	0,02	0,08	0,04	0,93	0,67
m, Binomial (4, 1/2)	0,26	0,15	0,17	0,09	0,30	0,12	0,19	0,07	0,15	0,09	0,03	0,00	0,03	0,00	0,34	0,18
m — 1	0,22	0,12	0,16	0,05	0,17	0,09	0,17	0,03	0,12	0,05	0,00	0,00	0,03	0,00	0,20	0,09

$$Y = X\beta + \epsilon$$

$$E(\epsilon) = 0$$

$$E(\epsilon\epsilon') = \sigma^2 I$$

deben seguirse los siguientes pasos:

- a) computar  $e = [I - X(X'X)^{-1}X']Y$
- b) encontrar los  $p$  desvíos más cercanos a cero

$$e_1, e_2, \dots, e_p.$$

- c) formar los vectores  $X_0 = \begin{bmatrix} X_1 \\ \vdots \\ X_p \end{bmatrix}$  y  $X_1 = \begin{bmatrix} X_{p+1} \\ \vdots \\ X_n \end{bmatrix}$

- d) computar  $M_{11} = I - X_1(X'X)^{-1}X_1'$
- e) encontrar las raíces características de  $M_{11}$  y sus vectores característicos. Es decir determinar  $\lambda_i$  y  $P_i$  tal que

$$M_{11}P_i = \lambda_i P_i \quad \lambda_i = 1, \dots, n-p$$

$$P_i' P_i = I$$

- f) formar las matrices  $D = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_{n-p} \end{bmatrix}$  y

$$P = [P_1 \ P_2 \ \dots \ P_{n-p}]$$

y calcular

$$A_1 = P D^{1/2} P'$$

$$A_0 = (X_1 X_0^{-1})' A_1$$

- g) calcular  $Z = [A_0' A_0'] \begin{bmatrix} Y_0 \\ Y_1 \end{bmatrix}$

- h) ordenar los elementos de  $Z$  de menor a mayor

$$v_1 \quad v_2 \quad \dots \quad v_{n-p}$$

- i) calcular

$$b = \sum_{i=1}^k a_{m-i+1} (v_{m-i+1} - v_i)$$

para K tal que  $m = n - p = 2k$

$$o \quad m = n - p = 2k + 1$$

Los valores de  $a_{m-i+1}$  se obtienen del cuadro N° 7.

j) calcular

$$S^2 = \sum_{i=1}^m v_i^2$$

k) determinar finalmente

$$W_{ob} = \frac{b^2}{S^2}$$

Si  $F(W_{ob})$  es menor al nivel de significación preestablecido para la décima, corresponde rechazar la hipótesis de normalidad. El valor de  $F(W_{ob})$  se encuentra en el cuadro N° 3.

### 3.4. Un ejemplo numérico.

En el capítulo I trabajamos con el modelo  $y_i = \beta x_i + \varepsilon_i$  para el que se dispone de la siguiente información:

$$\begin{array}{c|ccc} y & 0 & 1 & -1 \\ \hline x & -1 & 1 & 2 \end{array}$$

Siguiendo los lineamientos esbozados en el punto 3.3 ilustramos el procedimiento allí referido. En primer lugar estimamos  $\hat{\beta}$ :

$$\hat{\beta} = \frac{\sum y x}{\sum x^2} = \frac{1}{6}$$

Obtenemos luego el vector de los residuos

$$e = Y - \hat{\beta} X = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + \frac{1}{6} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{6} \\ \frac{7}{6} \\ -\frac{4}{6} \end{bmatrix}$$

El mínimo  $|e_i|$  es  $\left| -\frac{1}{6} \right|$  el cual ocupa el primer lugar en el vector  $e$ . Podemos entonces establecer las matrices  $X_0$  y  $X_1$ .

$$X_0 = [-1] \quad X_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

En el punto 1, hemos calculado ya

$$(X' X)^{-1} = \frac{1}{6}$$

$$M_{11} = \begin{bmatrix} \frac{5}{6} & -\frac{2}{6} \\ -\frac{2}{6} & \frac{2}{6} \end{bmatrix}$$

$$D = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{6} \end{bmatrix}$$

$$P = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 0,8816565 & 0,2367033 \\ 0,2367033 & 0,5266016 \end{bmatrix}$$

$$A_0 = [0,4082499 \quad -0,8164998]$$

$$yZ = A' Y = \begin{bmatrix} -0,1715466 \\ 1,3431014 \end{bmatrix}$$

Ordenamos ahora los elementos de  $Z$  de menor a mayor, es decir

$$-0,1715466 \quad ; \quad 1,3431014$$

Calculamos  $b = 0,7071 (1,3431014 + 0,1715466) = 1,071$  para lo cual hemos obtenido el coeficiente  $a_1 = 0,7071$  para el tamaño  $m = 2$  de la tabla de Sarhan y Greenberg que reproducimos en el cuadro N° 7.

$$\begin{aligned} \text{También calculamos } S^2 &= (1,3431014)^2 + (-0,1715466)^2 = \\ &= 1,8333 \text{ para obtener finalmente } W_{ob} = \frac{(1,071)^2}{1,8333} = 0,6257. \end{aligned}$$

Comparando  $W_{ob}$  con el valor de  $W = 0,6320$  que encontramos en la tabla N° 5 para un nivel de significación del 0,05, observamos que el primero es menor que el segundo por lo que corresponde rechazar la hipótesis nula que puntualiza que las perturbaciones  $\varepsilon_i$  tienen distribución normal de media 0.

#### 4 — ALGUNAS APLICACIONES EXTENDIDAS DE LA DÓCIMA Y LOS RESULTADOS OBTENIDOS PARA ALGUNOS MODELOS ESPECIFICOS

No está probado que las perturbaciones aleatorias de los modelos econométricos lineales se distribuyen siempre normalmente y por consiguiente no necesitamos demostrar lo contrario para poner en evidencia que cuando constituye una hipótesis crucial del modelo, quizás sea conveniente docimarla. La aplicación de la dócima a un conjunto de modelos específicos que realizamos en este capítulo, sólo persigue poner en evidencia que la verdadera distribución de las perturbaciones se aleja significativamente de la distribución normal con más frecuencia que lo que a veces se cree<sup>8</sup>.

Antes de comentar los resultados de la aplicación mencionada, mostramos cómo la dócima puede aplicarse a algunos modelos diferentes al que hemos utilizado aquí.

<sup>8</sup> Véase por ej. BARBANCHO, *Complementos de Econometría* - Ed. Ariel - Barcelona, 1966, pág. 34.

CUADRO Nº 7

*Coefficientes  $a_{m-i+1}$  para el test de normalidad basado en W. Para muestras de 2 a 50 grados de libertad.*

$i$	$m$	2	3	4	5	6	7	8	9	10	
1		0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739	
2		—	.0000	.1677	.2413	.2806	.3031	.3164	.3244	.3291	
3		—	—	—	.0000	.0875	.1401	.1743	.1976	.2141	
4		—	—	—	—	—	.0000	.0561	.0947	.1224	
5		—	—	—	—	—	—	—	.0000	.0399	
$i$	$m$	11	12	13	14	15	16	17	18	19	20
1		0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2		.3315	.3325	.3325	.3318	.3306	.3290	.3273	.3253	.3232	.3211
3		.2260	.2347	.2412	.2460	.2495	.2521	.2540	.2553	.2561	.2565
4		.1429	.1586	.1707	.1802	.1878	.1939	.1988	.2027	.2059	.2085
5		.0695	.0922	.1099	.1240	.1353	.1447	.1524	.1587	.1641	.1686
6		0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7		—	—	.0000	.0240	.0433	.0593	.0725	.0837	.0932	.1013
8		—	—	—	—	.0000	.0196	.0359	.0496	.0612	.0711
9		—	—	—	—	—	—	.0000	.0163	.0303	.0422
10		—	—	—	—	—	—	—	—	.0000	.0140

#### 4.1. Aplicación de la d6cima a modelos lineales con perturbaciones serialmente correlacionadas.

Sea el modelo

$$\begin{aligned} Y &= X\beta + \varepsilon \\ E(\varepsilon) &= 0 \\ E(\varepsilon\varepsilon') &= \sigma^2\Omega \end{aligned}$$

donde  $\Omega$  es una matriz simétrica no diagonal. Es obvio que no podr6 aplicarse directamente el procedimiento descrito en el capitulo III.

No obstante para cuando  $\Omega$  es conocida o se dispone de alguna estimaci6n, el problema es de f6cil soluci6n. Basta determinar una matriz ortogonal  $P$  tal que

$$P' \Omega P = I$$

Obtenida  $P$  puede aplicarse el procedimiento convencional para la d6cima sobre el modelo con variables transformadas:

$$(P'Y) = (P'X)\beta + P'\varepsilon$$

dado que ahora

$$\begin{aligned} E(P'\varepsilon) &= 0 \\ E(P'\varepsilon\varepsilon'P) &= \sigma^2 P' \Omega P = \sigma^2 I \end{aligned}$$

y para la hip6tesis nula "ε normal",  $P'\varepsilon$  ser6 tambi6n normal.

#### 4.2. Aplicaci6n de la d6cima a modelos lineales con perturbaciones heterosed6sticas.

En el modelo

$$\begin{aligned} Y &= X\beta + \varepsilon \\ E(\varepsilon) &= 0 \\ E(\varepsilon\varepsilon') &= \sigma^2\Lambda \end{aligned}$$

donde  $\Lambda$  es una matriz diagonal no escalar, las variables aleatorias obtenidas mediante la transformaci6n indicada en el punto 1 no tendr6n todas la misma varianza. Nuevamente no puede aplicarse en forma directa el procedimiento para la d6cima que hemos descrito en el punto anterior.

No obstante luego de transformar las variables mediante la matriz  $\Lambda^{-1/2}$  dicho procedimiento puede aplicarse sin variantes. En efecto, en el modelo transformado

$$(\Lambda^{-1/2} Y) = (\Lambda^{-1/2} X) \beta + (\Lambda^{-1/2} \epsilon)$$

se dan las condiciones

$$\begin{aligned} E(\Lambda^{-1/2} \epsilon) &= 0 \\ E(\Lambda^{-1/2} \epsilon \epsilon' \Lambda^{-1/2'}) &= \sigma^2 \Lambda^{-1/2} \Lambda \Lambda^{-1/2} = \sigma^2 I \\ \text{y } \Lambda^{-1/2} \epsilon &\text{ es normal si lo es } \epsilon. \end{aligned}$$

#### 4.3. Aplicación de la d{oc}ima a los modelos lineales multiecuacionales.

Puede desearse docimar la normalidad de las perturbaciones de cada ecuaci{on} estructural del modelo

$$B Y_t + \Gamma Z_t = \epsilon_t \quad |B| \neq 0$$

donde  $B$  y  $\Gamma$  son las matrices de coeficientes estructurales de las variables end{og}enas y predeterminadas respectivamente,  $Y_t$  y  $Z_t$  los vectores que contienen a dichas variables, y  $\epsilon_t$  un vector cuyos elementos son las perturbaciones aleatorias contempor{a}neas de cada ecuaci{on}. Si se supone adem{as} que  $E(\epsilon_t) = 0$ ,  $E(\epsilon_t \epsilon_t') = M$  donde  $M$  es constante y  $E(\epsilon_{it} \epsilon_{it+s}) = 0$  cuando  $s \neq 0$ , el procedimiento de la d{oc}ima de normalidad no podr{a} aplicarse directamente a cada ecuaci{on} de la forma estructural si contienen m{as} de una variable end{og}ena, pero s{ı} sobre cada ecuaci{on} de la forma reducida

$$Y = -B^{-1} \Gamma Z + B^{-1} \epsilon$$

En este {u}ltimo caso:

- a) existir{a} s{ol}o una variable end{og}ena en cada ecuaci{on}.
- b) las perturbaciones aleatorias de cada ecuaci{on} ser{a}n serialmente incorrelacionadas.
- c) bajo la hip{ot}esis que  $\epsilon_t$  sea un conjunto de variables normales con  $E(\epsilon_t) = 0$ , cada elemento de  $B^{-1} \epsilon_t$  tambi{e}n ser{a} una variable normal de esperanza nula.

#### 4.4. *Resultados obtenidos al docimar la normalidad de las perturbaciones en algunos modelos econométricos.*

La elección de los modelos de Klein-Goldberger para los Estados Unidos de Norteamérica y de Klein, Ball, Hazlewood y Vandome para el Reino Unido, con el objeto de aplicar como ejemplo el test de normalidad, responde a las siguientes razones:

- a) se trata de dos trabajos ya clásicos sobre el tema elaborado por econométricos de nombradía.
- b) se los considera correctamente especificados y han servido de base para la elaboración de muchos otros modelos del mismo tipo.
- c) la información estadística utilizada aparece transcrita en los trabajos publicados.
- d) la hipótesis de normalidad de las perturbaciones juega un papel crucial en los dos modelos dado que los autores han estimado sus parámetros mediante el método de la máxima verosimilitud con información limitada.

Sólo hemos docimado la hipótesis de normalidad de las perturbaciones de aquellas ecuaciones estructurales de los dos modelos que contienen una sola variable endógena. Como hemos señalado en el punto anterior el test podría haberse aplicado también a todas las ecuaciones de la forma reducida del modelo. En el cuadro N° 8 se ha transcripto como ejemplo una ecuación estimada, el vector de residuos transformados, y el valor obtenido de  $W$ .

También se agrega la representación gráfica de la función de distribución hipotética y la correspondiente distribución de la muestra.

De las catorce aplicaciones realizadas cinco veces resultó rechazada la hipótesis de normalidad.

#### 4.5. *Posibles causas de discrepancias con la normalidad de las perturbaciones en los modelos econométricos.*

Las condiciones que por el teorema central del límite podrían asegurar la normalidad de las perturbaciones, son poco frecuentes

CUADRO Nº 8

Relación entre las tasas de interés de corto y largo plazo. Modelo de Klein-Goldberger.

$$i_L = 2,58 + 0,44 (i_S)_{t-3} + 0,26 (i_S)_{t-5}$$

(0,15) (0,10) (0,09)

$$R = 0,95$$

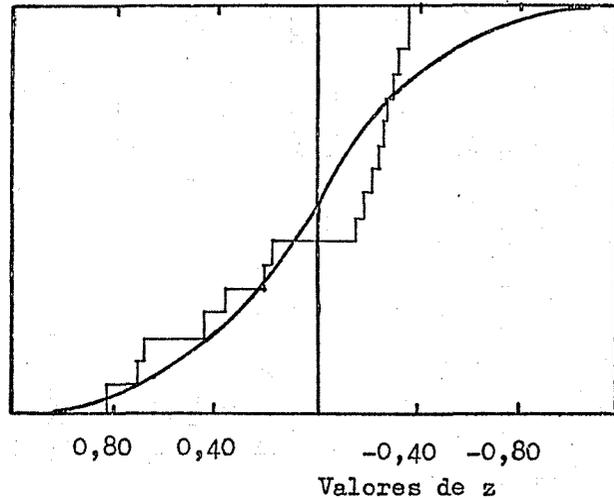
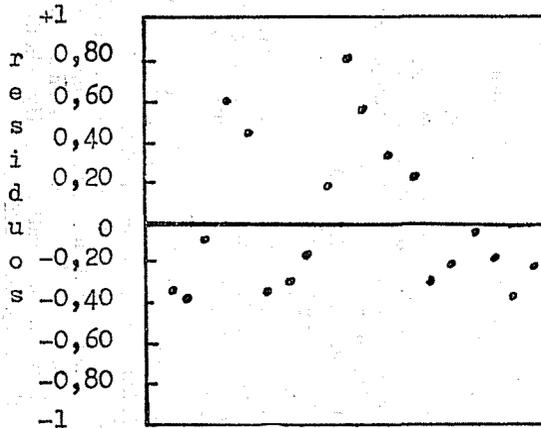
$$n - p = 17$$

$$W_{ob} = 0,792$$

$$W_{0,05} = 0,839$$

$i_L$  = rendimiento promedio sobre bonos de las empresas.

$i_S$  = rendimiento promedio sobre papeles comerciales de corto plazo.



en la práctica. Si bien puede pensarse que en general la perturbación aleatoria es el resultado de la suma de un conjunto numeroso de causas, es difícil asegurar que las mismas serán independientes entre sí, especialmente cuando constituyen fenómenos económicos. A modo de ejemplo puede pensarse en el conjunto de los precios de los sustitutos y complementos lejanos que corrientemente quedan resumidos en la perturbación aleatoria de las funciones de demanda. Es razonable pensar que en la mayor parte de los casos estarán fuertemente correlacionados entre sí.

Por otro lado, los errores de especificación pueden conducir a alejamientos importantes de la normalidad. Estos errores son frecuentes cuando por desconocimiento del fenómeno (o por carencia de información estadística) no se incluyen en el modelo en forma explícita algunas variables relevantes o no se adopta la forma funcional más cercana a la realidad.

Finalmente, errores de observación asimétricos en la variable endógena e inexistencia de la varianza de los errores, pueden constituir otras causas de no-normalidad.

#### REFERENCIAS BIBLIOGRAFICAS

- (1) AINE, "A Simple Test for Uniformity of a Circular Distribution". *Biometrika* 55, (1968).
- (2) ANDERSON, T. W., *Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, Inc. New York (1958).
- (3) BIRNBAUM, Z. W., "Numerical Tabulation of the Distribution of Kolmogorov's Statistic for Finite Sample Size". *American Statistical Association Journal*, 17, (1952).
- (4) CRAMER, H., *Métodos Matemáticos de Estadística*. Aguilar. Segunda edición, Madrid (1960).
- (5) CRAMER, H., "On the Composition of Elementary Error". *American Statistical Association Journal*. (1928).
- (6) DAVID, H. A., HARTLEY, H. O. & PEARSON, E. S., "The distribution of the ratio, in a single normal sample, of range to standard deviation". *Biometrika*, 41, (1954).
- (7) LILLIEFORS, H. W., "On the Kolmogorov - Smirnov Test for Normality with Mean and Variance Unknown". *American Statistical Association Journal*. (June 1967).
- (8) FISHER, R. A., "On the Interpretation of  $\chi^2$  from Contingency Tables, and the Calculation". *Journal of the Royal Statistical Society*, 85, (1922).
- (9) FISHER, R. A., "The Conditions under which  $\chi^2$  Measures the Discrepancy between Observation and Hypothesis". *Journal of the Royal Statistical Society*. 87, (1924).

- (10) GRAYBILL, F. A., *An Introduction to Linear Statistical Models*. Mac Graw Company, Inc. New York (1961).
- (11) GEARY, R. C., "Testing for Normality". *Biometrika*, 34. (1947).
- (12) GAYEN, A. K. "The distribution of the variance ratio in random samples of any size drawn from non-normal universes". *Biometrika*, 37, (1950).
- (13) KLEIN, L. R., BALL, R. J., HAZLEWOOD AND VANDOME, P., *An Econometric Model of the United Kingdom*. Basil Blackwell. Oxford. (1961).
- (14) KLEIN, L. R., and GOLDBERGER, A. S., *An Econometric Model of the United States 1929-1952*. North - Holland Publishing Company. Amsterdam (1955).
- (15) KOLMOGOROV, A. N., "Sulla Determinazione Empirica di una Legge Distribuzione". *Giornale dell'Istituto Italiana Attuari*, 4, (1933).
- (16) KUIPER, N. H., "Tests concerning random points on a circle". *Proc. Koninkl. Nederl. Akad. Van Wetenschappen. Series A*, 63 (1960).
- (17) MALINVAUD, E., *Métodos Estadísticos de la Econometría*. Ariel, Barcelona (1967).
- (18) MISES, R. V., *Wahrscheinlichkeits Rechnung und ihre Anwendung in der Statistik und Theoretis cheu Physik*. Leipzig - Viena (1931).
- (19) PEARSON, E. S. & STEPHENS, M. A., "The Goodness of Fit Test Based on  $W^2_n$  y  $U^2_n$ ". *Biometrika*, 49, (1962).
- (20) PEARSON, K., "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling". *Phil Mag.* V. 50. (1900).
- (21) PEARSON, E. S. & STEPHENS, M. A., "The Ratio of Range to Standard Deviation in the same Normal Sample". *Biometrika*, 51 (1965).
- (22) PUTTER, J., "Orthonormal bases of error spaces and their use for investigating the normality and variances of residuals". *American Statistical Association Journal*, (1967).
- (23) SARHAN, A. E. & GREENBERG, B. G., "Estimation of location and scale parameters by order statistics from singly and double censored samples". Part. I. *Ann. Math. Statist.* 27, (1956).
- (24) SHAPIRO, S. S. & WILK, M. B., "An Analysis of Variance Test for Normality (Complete Samples)". *Biometrika*, 52 (1965).
- (25) STEPHENS, M. A., "The Distribution of the Goodness of Fit Statist  $U^2_n$ ". *Biometrika*, 50, (1963).
- (26) SMIRNOV, "Table for Estimating the Goodness of Fit of Empirical Distributions". *Ann. Math. Stat.* Vol. 19 (1948).
- (27) THEIL, H., "The Analysis of Disturbances in Regression Analysis". *Journal of the American Statistical Association*, 60, (1965).
- (28) THE RAND CORPORATION, *A Million Random Digits with 100.000 Normal Deviates*. Collier - Macmillan, Canadá (1966).
- (29) WATSON, G. S., "Goodness of Fit Tests on a Circle". *Biometrika*, 48, (1961).