1
201

Universidad Nacional Autónoma de México

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
A C A T L A N

DESARROLLO DE UN PAQUETE COMPUTACIONAL DE REGRESION LINEAL MULTIPLE.

T E S I S
Que para obtener el Título de
L I C E N C I A D O E N A C T U A R I A
p r e s e n t a n

RENE DORANTES RODRIGUEZ SAN MIGUEL
AMBROSIO DE JESUS TORRES MENDOZA

MEXICO, D. F. 1991

I N D I C E

INTRODUCCION

El objetivo de esta tesis es elaborar e implantar un paquete computacional de Regresión Lineal Múltiple que sirva para facilitar el análisis en la elaboración de tareas, trabajos y proyectos de estudiantes, profesores o egresados de la E.N.E.P. Acatlán y que además sirva de apoyo didáctico en las materias de Estadística, Regresión Lineal y en aquellas en las que se necesiten este tipo de herramienta.

La tesis se divide en dos partes: la primera contiene al primer capítulo "Bases Teóricas" y su objetivo es servir como material de consulta y sustento teórico del paquete. En la primera sección del capítulo I "Regresión Lineal Múltiple", se presenta el planteamiento del modelo así como el análisis de varianza, la carencia de ajuste, las pruebas de hipótesis y los intervalos de confianza entre otros temas. En la sección 1.2 "Violación de los Supuestos de Modelo", se tratan temas más específicos como el caso de la multicolinealidad entre las variables explicativas, la heteroscedasticidad en las varianzas de las perturbaciones, la autocorrelación entre las perturbaciones y el análisis de residuales. En los temas de multicolinealidad, heteroscedasticidad y autocorrelación se incluyen los siguientes puntos: a) Concepto, b) Consecuencias, c) Técnicas de detección y d) Posibles soluciones a la violación al supuesto. La parte teórica se termina en la seccion 1.3 donde se incluyen los métodos numéricos para resolver el sistema de ecuaciones y para obtener los valores de las distribuciones probabilísticas necesarias para el análisis, tales como la normal y su inversa, la t-Student y su inversa.

En la segunda parte de la tesis se incluyen los capítulos II y III. El primero proporciona la información referente a las opciones del paquete, y su modo de operación, la lectura, simulación y edición de datos, las transformaciones de las variables, la estructura del sistema y sus características. En el último capítulo se hacen algunas aplicaciones del paquete a manera de ejemplos, pretendiendo cubrir la mayor parte de los puntos del paquete para un mejor entendimiento y guía en su forma de uso.

CAPITULO I

BASES TEORICAS

1.1 R E G R E S I O N L I N E A L M U L T I P L E

1.1.1 MODELO DE REGRESION LINEAL MULTIPLE

El modelo de Regresión Lineal Múltiple expresa la relación de una variable dependiente o de respuesta Y con p variables independientes o explicativas $X_1, X_2, \ldots, X_p$. Tal relación está expresada en la siguiente ecuación lineal:

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \varepsilon \qquad (1.1.1.1)$$

donde β_0, $\beta_1, \ldots, \beta_p$ son llamados los coeficientes de regresión y ε es la perturbación o error aleatorio.

Para n valores observados tanto para $\underset{\sim}{Y}$ como para X_i (con i=1,n) se tiene el siguiente sistema

$$
\begin{array}{l}
Y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \beta_3 X_{13} + \ldots + \beta_p X_{1p} + \varepsilon_1 \\
Y_2 = \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \beta_3 X_{23} + \ldots + \beta_p X_{2p} + \varepsilon_2 \\
Y_3 = \beta_0 + \beta_1 X_{31} + \beta_2 X_{32} + \beta_3 X_{33} + \ldots + \beta_p X_{3p} + \varepsilon_3 \\
\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\
Y_n = \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \beta_3 X_{n3} + \ldots + \beta_p X_{np} + \varepsilon_n
\end{array}
$$

Para fines de un manejo más eficiente y simplificado del modelo, se utiliza la siguiente notación matricial:

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_n \end{bmatrix}_{n\times 1} \qquad X = \begin{bmatrix} 1 & X_{11} & X_{12} \ldots X_{1p} \\ 1 & X_{21} & X_{22} \ldots X_{2p} \\ \cdots & \cdots & \cdots \\ 1 & X_{n1} & X_{n2} \ldots X_{np} \end{bmatrix}_{n\times(p+1)}$$

$$\underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdots \\ \beta_p \end{bmatrix}_{(p+1)\times 1} \qquad \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_n \end{bmatrix}_{n\times 1}$$

Lo que reduce el modelo a la forma:

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} \qquad (1.1.1.2)$$

El método utilizado para obtener la estimación de los coeficientes de la regresión $(\hat{\underline{\beta}})$ es el de Cuadradros Mínimos, que aplicado al sistema de ecuaciones 1.1.2 genera el sistema de ecuaciones:[1]

$$X'X\hat{\underline{\beta}} = X'\underline{Y} \qquad (1.1.1.3)$$

conocidas como ecuaciones normales. Al resolver para $\hat{\underline{\beta}}$ se tiene:

$$\hat{\underline{\beta}} = (X'X)^{-1}\ X'Y \qquad (1.1.1.4)$$

que es válida si y sólo si $(X'X)^{-1}$ existe, para lo cual X'X debe ser de rango completo, es decir las columnas de X deben ser linealmente independientes.
Con (1.1.1.4) se define al valor ajustado de Y como:

$$\hat{\underline{Y}} = X\hat{\underline{\beta}} \qquad (1.1.5)$$

1) Seberg. G. A. F., 'Linear Regression Analisys', p. 44

Tomando en consideración los siguientes supuestos

i) $E[\underline{Y}] = X\underline{\beta}$ o $E[\underline{\varepsilon}] = \underline{0}$.

Cada población de Y correspondiente a un vector X dado está distribuida alrededor de su valor medio (su valor esperado es $X\underline{\beta}$).

ii) $COV(\underline{Y}) = \sigma^2 I$ o $COV(\underline{\varepsilon}) = \sigma^2 I$

La distribución de probabilidad de Y dado el vector X tiene la misma variación (igual varianza).

iii) Las X son constantes fijas o predeterminadas.

Los estimadores de cuadrados mínimos tienen las siguientes propiedades:[2]

$$E[\hat{\underline{\beta}}] = \underline{\beta} \qquad (1.1.1.6)$$

$$VAR[\hat{\underline{\beta}}] = \sigma^2(X'X)^{-1} \qquad (1.1.1.7)$$

y un estimador insesgado de σ^2 es:[3]

$$s^2 = (\underline{Y} - X\hat{\underline{\beta}})'(\underline{Y} - X\hat{\underline{\beta}}) / (n-p-1) \qquad (1.1.1.8)$$

1.1.2 ANALISIS DE VARIANZA Y COEFICIENTE DE DETERMINACION

El análisis de varianza es una herramienta que permite conocer cuánta variación del modelo se debe a cierta fuente y cuánta a otras y si tales variaciones son significativas. Con la ayuda de esta técnica se puede saber si las variables independientes del

2) Seberg, op. cit. p. 48
3) Idem., p. 51

modelo contribuyen a explicar significativamente a la variable de respuesta.

La diferencia entre el valor Y_i y el valor $\hat{Y}_i$ se llama residual y está denotado por $e_i = Yi - \hat{Y}i$. En notacion matricial: $\underline{e} = \underline{Y} - \hat{\underline{Y}}$.

El valor[4]

$$\underline{e}'\underline{e} = \underline{Y}'\underline{Y} - \hat{\underline{\beta}}'X'\underline{Y} = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \qquad (1.1.2.1)$$

se conoce como Suma de Cuadrados de los Residuales (SCRES).

La suma de Cuadrados Totales (SCT), está definida como:

$$\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2 \text{ y se puede calcular como:}^{(5)}$$

$$SCT = \underline{Y}'\underline{Y} - \left(\sum_{i=1}^{n} Y_i\right)^2 1/n \qquad (1.1.2.2)$$

A la diferencia de SCT - SCRES se le llama Suma de Cuadrados de la Regresión y es igual a:[6]

$$SCREG = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 = \hat{\underline{\beta}}'X'Y - \left(\sum Y_i\right)^2 1/n \qquad (1.1.2.3)$$

Con las anteriores definiciones se construye el siguiente cuadro de Análisis de Varianza.

4) Seberg, op. cit., p. 45
5) Jhonston, J., 'Econometric Methods'., p. 128
6) Idem., p. 120

Fuente de Variación	Grados de Libertad	Suma de Cuadrados	Media de Suma de Cuadrados	F
Regresión	p	SCREG	$MSCREG = \frac{SCREG}{p}$	$F = \frac{MSCREG}{s^2}$
Residual	n-p-1	SCRES	$s^2 = \frac{SCRES}{n-p-1}$	
Total	n-1	SCT		

Cuadro 1.2.1. Análisis de Varianza (ADV) para el modelo de regresión lineal múltiple.

Si $\varepsilon \sim N(\underline{0}, \sigma^2 I)$ entonces, $MSCREG/s^2$ se distribuye como una $F(p,n-p-1)$[7], lo que fundamenta la siguiente prueba de hipótesis:

$$Ho: \beta_i = 0$$

$$H1: \text{Algún } \beta_i \neq 0 \qquad i = 1,2,\ldots,p$$

Es decir, todas las variables independientes no contribuyen a explicar a Y contra la hipótesis de que al menos una si la explica.

El coeficiente de determinación r^2 es el cuadrado de la correlación entre $\underline{Y}$ y $\hat{\underline{Y}}$, y se puede calcular como:

$$r^2 = \frac{SCREG}{SCT} = 1 - \frac{SCRES}{SCT} \qquad (1.1.2.4)$$

y tiene un rango de [0,1].

El coeficiente de determinación ajustado r_a^2 se obtiene de dividir al SCRES y al SCT entre sus respectivos grados de libertad, esto es :

$$r_a = 1 - \frac{SCRES/(n-p-1)}{SCT/n-1} \qquad (1.1.2.5)$$

7) Jhonston, J., 'Econometric Methods'., p. 142

Estos coeficientes proporcionan una medida de lo que el modelo explica con relación a los valores que asume la variable Y, es decir, dan indicios sobre si el modelo planteado es adecuado o no.

1.1.3 CARENCIA DE AJUSTE: ERROR PURO Y PRUEBA ARCOIRIS

Un método para analizar si el modelo es inadecuado, es decir, si el modelo planteado no se ajusta al modelo real, es el análisis de la carencia de ajuste y el error puro.

Para aplicar este método es necesario que haya replicass en la medición de la variable Y, hechas para un mismo valor de las variables X's. Esto es con el objeto de obtener una estimación de σ^2 (que representa al error puro). Lo que se da a entender es que si se tienen dos o más observaciones idénticas, la variación de Y será debido a la variación aleatoria.

> "... Tales diferencias generalmente procionan una estimación de σ^2 que es mucho más confiable de la que se puede obtener de alguna otra fuente.[8]"

Sean m diferentes vectores de p elementos, X_j (j = 1..m) en los cuales:

Para X_1 hay n_1 observaciones $Y_{11} = Y_{12} = Y_{13} = ,\dots, = Y_{1n_1}$

... ...

Para X_2 hay n_2 observaciones $Y_{21} = Y_{22} = Y_{23} = ,\dots, = Y_{2n_2}$

Para X_j hay n_j observaciones $Y_{j1} = Y_{j2} = Y_{j3} = ,\dots, = Y_{jn_j}$

... ...

Para X_m hay n_m observaciones $Y_{m1} = Y_{m2} = Y_{m3} = ,\dots, = Y_{mn_m}$

8) Draper,N. R. and H. Smith, 'Applied Regression Analysis', . p 35

Donde:

$$\sum_{j=1}^{m} n_j = n$$

La Suma de Cuadrados del Error Puro (SCEP) se define como:[9]

$$SCEP = \sum_{j=1}^{m} \sum_{k=1}^{n_j} (Y_{jk} - \bar{Y}_j)^2 \qquad (1.1.3.1)$$

con grados de libertad $n_e = \sum_{j=1}^{m} (n_j - 1) = n-m$

El SCEP entre sus grados de libertad n_e da un estimador de σ^2.[10]

$$s_e^2 = \frac{SCEP}{n_e} \qquad (1.1.3.2)$$

independientemente de si el modelo planteado es adecuado o no. El SCRES dividido entre sus grados de libertad da un estimador de σ^2 si el modelo es correcto, y si no lo es, se obtiene σ^2 más un sesgo al que se conoce como carencia de ajuste. La suma de cuadrados de la carencia de ajuste (SCCA) resulta de la diferencia de el SCRES y el SCEP, esto es:

$$SCRES - SCEP = SCCA$$

donde[11]

$$SCCA = \sum_{j=1}^{m} n_j (\hat{Y}_j - \bar{Y}_j)^2 \qquad (1.1.3.3)$$

Con esto se puede construir el siguiente Cuadro de Análisis de Varianza.

9) Draper and Smith, op. cit. p. 36
10) Draper and Smith, op. cit. p. 36
11) Idem., p 36

Fuente de Variación	Grados de Libertad	Suma de Cuadrados	Media de Suma de Cuadrados	F
Carencia de Ajuste	m-p-1	SCCA	$MSCCA = \frac{SCCA}{m-p-1}$	$F = \frac{MSCCA}{s_e^2}$
Error Puro	n_e	SCEP	$s_e^2 = \frac{SCEP}{n_e}$	
Residual	n-p-1	SCRES		

Cuadro 1.3.1. ADV para la Carencia de Ajuste y el Error Puro.

$$\text{Si} \quad \underset{\sim}{\varepsilon} \sim N_n\,(\underset{\sim}{0}, \sigma^2\, I) \quad \text{entonces} \quad \frac{MSCCA}{s_e^2} \sim F(m-p-1, n_e).$$ [12]

Con este estadístico, se puede realizar la siguiente prueba de hipótesis.

H_0: No hay Carencia de Ajuste
H_1: Hay Carencia de Ajuste

Si en el modelo no hay repeticiones se puede aplicar la prueba Arcoiris[13] para detectar la existencia de carencia de ajuste. La prueba se basa en comparar el ajuste de un pequeño conjunto de observaciones contra el ajuste conteniendo todas éstas.

Supongamos que el modelo correcto incluye un conjunto de variables adicionales definidas por la matriz W. Entonces el modelo (1.1.1.1) queda como sigue:

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + W\,\underset{\sim}{\theta} + \underset{\sim}{\varepsilon} \qquad (1.1.3.4)$$

donde W es una matriz de (n x q) y θ es un vector de parámetros.

12) Idem., p. 37
13) Utts, J. M., "The Rainbow Test for Lack of Fit Regression", Commun. Statist.-Theor. Mathe. (11) 24. 2801-2815 (1982)

Ahora bien, la prueba de hipótesis a realizar es:

$H_o : \theta = \underline{0}$

$H : \theta \neq 0$

esto es, si el modelo ajustado es el adecuado o no lo es.

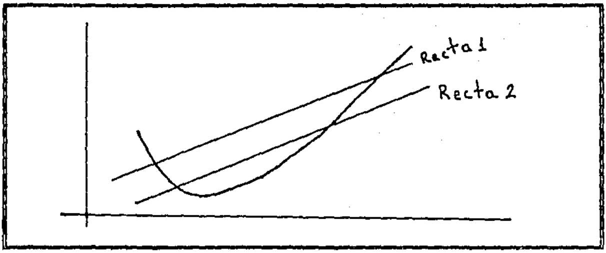


Figura 1.3.1. El modelo real es el Cuadrático, y el ajuste es Lineal.

Un ejemplo que ilustra el método se muestra en la figura 1.3.1, donde el modelo correcto es el cuadrático, pero considerando todos los puntos se ha ajustado la recta 1. Si se ajusta la recta 2, donde solo se consideran los puntos centrales, se podría obtener un ajuste no tan malo como el de la recta 1.

Sea D una matriz diagonal con n_1 unos en la diagonal principal y ceros en cualquier otro elemento, con $p+1 < n_1 < n$. Si se premultiplica $\underline{Y}$ y X por D, se obtienen n_1 renglones de estas matrices, que serán sus partes centrales, mientras que sus otros renglones serán cero. Si se asume que $E[D\underline{Y}] = DX\beta_D$ entonces se obtienen los siguientes resultados:[14]

$$\underline{\beta}_D = (X'DX)^{-1} X'D\underline{Y} \text{ , y} \qquad (1.1.3.5)$$

$$SCRES_D = Y'(D-DX(X'DX)^{-1}X'D)\underline{Y} \qquad (1.1.3.6)$$

que es solamente el ajuste de los puntos centrales.

14) Ulla, J. M., op. cit. p. 2804

Usando todos los puntos, $\hat{\beta}$ tiene el valor de (1.1.1.4) y el SCRES es igual a[15]

$$SCRES = Y'(I - X(X'X)^{-1}X')\,Y \qquad (1.1.3.7)$$

y la diferencia del SCRES y del $SCRES_D$ es igual a:[16]

$$SCRES_N = SCRES - SCRES_D \qquad (1.1.3.8)$$

definiendo la prueba estadística como:[17]

$$F = \frac{SCRES_N / n_2}{SCRES_D / (n_1 - p - 1)} \qquad (1.1.3.9)$$

donde $n_2 = n - n_1$ y F se distribuye bajo H_0 como[18] $F(n_2, n_1 - p - 1)$.

Esto se puede resumir en el Cuadro 1.3.2 de Análisis de Varianza.

Para determinar los puntos centrales que se utilizarán en la prueba, se propone escoger a los elementos más pequeños de la diagonal de la matriz $X'(X'X)^{-1}X'$. Se recomienda usar aproximadamente la mitad del total de puntos.

15) Utts, J.M., op. cit. p. 2804
16) Idem., p. 2804
17) Idem., p. 2804
18) Idem., p. 2806

Fuente de Variación	Grados de Libertad	Suma de Cuadrados	Media de Suma de Cuadrados	F
Regrsión con n datos	p	SCREG		
Residuales con n-n datos	$n_2 = n-n_1$	$SCRES_N$	$MSCRES_N = \frac{SCRES_N}{n_2}$	$F = \frac{MSCRES_N}{s_D^2}$
Residuales con n datos	n_1-p-1	$SCRES_D$	$s_D^2 = \frac{SCRES_D}{(n_1-p-1)}$	
Total	n-1	SCT		

Cuadro 1.3.2. ADV de la prueba Arcoiris.

1.1.4 PRUEBA DE HIPOTESIS

1.1.4.1 Individuales

Las pruebas de hipótesis individuales que suelen emplearse son las siguientes:

i) $H_o: \beta_j = \beta_j^o$
vs.
$H_1: \beta_j \neq \beta_j^o$

ii) $H_o: \beta_j \geq \beta_j^o$
vs.
$H_1: \beta_j < \beta_j^o$

iii) $H_o: \beta_j \leq \beta_j^o$
vs.
$H_1: \beta_j > \beta_j^o$

Para algún j, j=1,2,...,p

Donde β_j^0 es el valor hipotético de β_j. Si se supone que $\underline{Y} \sim N(X\underline{\beta},\sigma^2 I)$, entonces el estadístico de esta prueba está dado por:[19]

$$t = \frac{\hat{\beta}_j - \beta_j}{\sqrt{v_{jj}}\ s} \qquad (1.1.4.1)$$

donde v_{jj} es el j-ésimo elemento de la diagonal de $(X'X)^{-1}$, y t se distribuye como t(n-p-1).

1.1.4.2 Conjuntas

Si se desea hacer la prueba de hipótesis de que todas la betas son cero, contra la alternativa de que al menos una no lo es, esto es:

$$H_1: \beta_j = 0 \qquad j=1,2,\dots,p$$
$$\text{vs}$$
$$H_o: \beta_j \neq 0 \qquad \text{para algún } j,\ j=1,2,\dots,p$$

entonces el estadístico de prueba es:[20]

$$F = \frac{SCREG\ /p}{s^2} \qquad (1.1.4.2)$$

donde F se distribuye como una F(p,n-p-1).

1.1.4.3 Para valorar la contribución de un subconjunto de variables a la regresión

Esta prueba de hipótesis es global (abarca los casos anteriores) en la que se prueba:

$$H_o: \beta_j = 0 \qquad j = r+1,\ r+2,\dots,p$$
$$\text{vs.}$$
$$H_1: \beta_j \neq 0 \qquad \text{para algún } j = r+1,\ r+2,\dots,p$$

19) Jhonston, J. op. cit. 138
20) Idem., p. 142

Es una prueba que involucra dos modelos. La hipótesis a probar es que el modelo consta solamente de r variables explicativas (r<p), contra la alternativa de que son p.

La prueba de hipótesis se hace considerando el cociente:

$$F = \frac{\dfrac{(SCRES_r - SCRES)}{(p-r)}}{s^2} \qquad (1.1.4.3)$$

donde $SCRES_r$ es la suma de cuadrados de los residuales considerando r variables explicativas. F se distribuye como $F(r,n-p-1)$[21] si los errores están normalmente distribuidos.

A la diferencia $SCRES_r$-SCRES se le denomina $SCREG(X_{r+1},\dots,X_p / X_0,\dots,X_r)$[22] que es la suma de cuadrados de la regresión de $X_{r+1},X_{r+2},\dots X_p$ dado que en el modelo están X_0, $X_1,\dots,X_r$.

Por la razón de que la SCT es la misma para los dos modelos involucrados en la regresión, la resta $SCRES_r$ - SCRES es equivalente a la resta SCREG - $SCREG_r$ donde SCREG es la suma de cuadrados debida a la regresión considerando r variables, por lo que 1.4.3 se transforma en:

$$F = \frac{SCREG(X_{r+1},X_{r+2},\dots,X_p / X_0,X_1,\dots,X_r)}{s^2(p-r)}$$

$$= \frac{SCREG - SCREG_r}{s^2(p-r)} \qquad (1.1.4.4)$$

La utilizacion de la suma de cuadrados debido a la regresión en lugar de la suma de cuadrados de los residuales, se debe a que por lo general resulta de mayor interés en el análisis de varianza,

21) Draper and Smith, op. cit. p

22) Idem., p

observar en cuánto aumenta lo explicado al incorporar p-r variables adicionales.

Fuente de Variación	Grados de Libertad	Suma de Cuadrados	Media de Suma de Cuadrados	F
Regresión con r var.	r	$SCREG_r$		
$SCREG(X_{r+1},X_p / X_0,X_1,...,X_r)$	p-r	$SCREG-SCREG_r$	$\frac{SCREG-SCREG_r}{p-r}$	$F = \frac{SCREG-SCREG_r}{(p-r)s^2}$
Regresión con p var.	p	SCREG		
Residuales de las p var.	n-p-1	SCRES	$s^2 = \frac{SCRES}{n-p-1}$	
Total	n-1	SCT		

Cuadro 1.4.1. ADV de la Contribución de un Suconjunto de Variables a la Regresión

1.1.4.4 Combinaciones Lineales de los Estimadores (Prueba de Hipótesis General)

Esta prueba de hipótesis es el caso más general, siendo la hipótesis nula: $C\underline{\beta} = \underline{\gamma}$, donde C es una matriz de constantes de rx(p+1), con r<p+1, y $\underline{\gamma}$ es un vector columna de (rx1).

El estadístico asociado a la prueba es:[23]

$$F = \frac{(C\hat{\underline{\beta}} - \underline{\gamma})'(CVC)^{-1}(C\hat{\underline{\beta}} - \underline{\gamma})}{s^2 r} \qquad (1.1.4.5)$$

donde $V = (X'X)^{-1}$ y $F \sim F(p,n-p-1)$.

23) Seberg, op. cit. p.

1.1.5 NIVEL OBSERVADO DE SIGNIFICANCIA (Valor P):

El "valor P" es la probabilidad de obtener un valor muestral tan grande como el que se observó suponiendo que Ho es verdadero. Por ejemplo:

$$\text{Si} \quad H_o: \beta_j = 0 \quad \text{vs.} \quad H_1: \beta_j \neq 0$$

para algún j, j=1,2,...,p, entonces de 1.4.1

$$t = \frac{\hat{\beta}_j}{\sqrt{v_{jj}}\, s} \tag{1.1.5.1}$$

que es al valor indicado en las figuras 1.4.1.a y 1.4.1.b. Y para este mismo valor le corresponde cierta probabilidad (valor P) que en este caso es la zona sombreada.

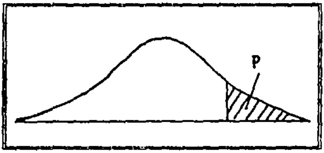


Figura 1.4.1.a.
Valor P de Una Cola.

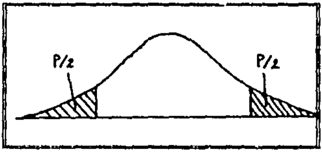


Figura 1.4.1.b.
Valor P de Dos Colas.

1.1.6 INTERVALOS DE CONFIANZA

1.1.6.1 Individuales

El intervalo de confianza (IC) para β_j al $100(1-\varepsilon)\%$ tal que $P[t_{-\varepsilon/2} < t < t_{\varepsilon/2}] = 1-\varepsilon$ está dado por:[24]

$$IC = \beta_j \pm t_{\varepsilon/2}\, s\sqrt{v_{jj}} \tag{1.1.6.1}$$

24) Wonnacott, R. and Wonnacott, T. 'Regression: a Second Course in Statistics', p. 432

1.1.6.2 Conjuntos

a. El intervalo de confianza conjunto para todas las betas al $100(1-\varepsilon)\%$ está dado por:[25]

$$(\underline{\beta} - \hat{\underline{\beta}})'X'X(\underline{\beta} - \hat{\underline{\beta}}) \leq s^2(p+1)\ F_{(p+1,\, n-p-1,\, 1-\varepsilon)} \qquad (1.1.6.2)$$

b. Sea $V = (X'X)^{-1}$ entonces el intervalo de confianza conjunto para un subconjunto de r-betas al $100(1-\varepsilon)\%$ está dado por:[26]

$$(\underline{\beta}_r - \hat{\underline{\beta}}_r)'\ V_r^{-1}(\underline{\beta}_r - \hat{\underline{\beta}}_r) < s^2(p+1)\ F_{(p+1,\, n-p-1,\, 1-\varepsilon)} \qquad (1.1.6.3)$$

con $r < p+1$ y donde $\underline{\beta}_r$ y $\hat{\underline{\beta}}_r$ son los vectores con las r-betas que han sido seleccionadas para construir el intervalo conjunto. V_r son las r-columnas de V correspondientes a las r-betas elegidas.

1.1.6.3 Para combinaciones lineales de los estimadores

Las combinaciones lineales de las betas son generadas por el producto $C\hat{\underline{\beta}}$, donde $C_{r\times p+1}$ es una matriz de constantes de rango $r \leq p+1$. Así, el intervalo de confianza al $100(1-\varepsilon)\%$ en su caso más general se obtiene de[27]:

$$(C\underline{\beta} - C\hat{\underline{\beta}})'\ (CVC)^{-1}(C\underline{\beta} - C\hat{\underline{\beta}}) \leq s^2 r\ F_{(r,\, n-p-1,\, 1-\varepsilon)} \qquad (1.1.6.4)$$

1.1.6.4 Para la varianza de σ^2

El intervalo de confianza para σ^2 al $100(1-\alpha)\%$ tal que:

$$P[\chi^2_{1-\alpha/2} < \frac{(n-p-1)\ s^2}{\sigma^2} < \chi^2_{\alpha/2}] = 1-\alpha$$ está dado por:[28]

$$[\frac{(n-p-1)\ s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-p-1)\ s^2}{\chi^2_{1-\alpha/2}}] \qquad (1.1.6.5)$$

25) Wonnacott, R. and Wonnacott, T. op. cit., pp. 426-431
26) Wonnacott, R. and Wonnacott, T. op. cit., pp. 431
27) Idem., pp. 427,428
28) Gujarati, D., 'Econometrha Basica', pp. 73,74

1.1.7 PREDICCION

1.1.7.1 Predicción e intervalo de confianza para $E[Y/X_o]$:

Dado un vector $C = (1, X_{o1}, X_{o2}, \ldots, X_{op})$ el estimador de la media de Y_o está dado por:

$$\hat{\mu} = C\hat{\beta} \qquad (1.1.7.1)$$

$$\underline{\mu}_o = C\underline{\beta} \qquad (1.1.7.2)$$

Si $\underline{Y} \sim N(X\underline{\beta}, \sigma^2 I)$ entonces, $C\hat{\underline{\beta}} \sim N(C\underline{\beta}, \sigma^2 C(X'X)^{-1}C')$ y por lo tanto:[20]

$$t = \frac{C\hat{\underline{\beta}} - C\underline{\beta}}{s\sqrt{C(X'X)^{-1}C'}} \qquad (1.1.7.3)$$

y el intervalo de confianza al $100(1-\varepsilon)\%$ está dado por:

$$C\hat{\underline{\beta}} \pm t_{\varepsilon/2}\, s\sqrt{C(X'X)^{-1}C'} \qquad (1.1.7.4)$$

1.1.7.2 Predicción e intervalo de confianza para Y/X_o:

Sea C como se definió anteriormente, el valor real de Y_o es $C\underline{\beta}+\varepsilon_o$, la diferencia entre este valor y $\hat{\underline{Y}}_o$ (con $\hat{\underline{Y}}_o = C\hat{\underline{\beta}}$) es:

$$d = \hat{\underline{Y}}_o - \underline{Y}_o \qquad (1.1.7.5)$$

así, podemos deducir que[30] $d \sim N(\underline{0}, \sigma^2(1+C(X'X)^{-1}C')$ por lo tanto

$$t = \frac{\hat{\underline{Y}}_o - Y_o}{s\sqrt{1+C(X'X)^{-1}C'}} \qquad (1.1.7.6)$$

20) Jhonston, J. op. cit., p. 153
30) Idem., p. 154

El intervalo de confianza para Y/X_o al $100(1-\varepsilon)\%$ es:[31]

$$\hat{Y}_o \pm t_{\varepsilon/2}\, s \sqrt{1+C(X'X)^{-1}C'} \qquad (1.1.7.7)$$

1.1.8 VARIABLES DICOTOMICAS Y CATEGORICAS

Dentro de las características de las variables independientes existe la posibilidad de que éstas sean de tipo cualitativo, es decir, representan características no cuantitativas. Por ejemplo, la raza, el sexo, la religión etc. Para poder expresar este tipo de variables dentro del modelo, se requiere utilizar variables dicótomas, las cuales se denominan así porque sólo asumen dos valores: 0 y 1. La categoría indica los niveles en que se subdivide una variable, por ejemplo, el sexo tiene dos categorías, y se puede definir a $D_i=0$ si es hombre y $D_i=1$ si es mujer, donde D_i es una variable dicótoma. Si una variable tiene r categorías, entonces se necesitan r-1 variables dicotómas.[32]

1.1.9 SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES

En el análisis de Regresión Múltiple la mayoría de las veces se espera que un subconjunto de las variables independientes explique adecuadamente los datos, factor que en un momento puede simplificar el análisis. Es por esta razón que se trata muchas veces de seleccionar los mejores subconjuntos de variables. El término "mejor" es en cierta forma subjetivo ya que no hay un único procedimiento para escoger el mejor subconjunto y cierto juicio personal, en la mayoría de los casos, es necesario.

31) Idem., p. 154

32) Se puede consultar la manera de construir las variables dicotomas cuando una variable tiene r categorias en Wonnacott, R. and Wonnacott, T. op. cit., pp. 110

El procedimiento sugerido por Furnival,[33] en el cual se basa el paquete, usa el criterio de seleccionar los subconjuntos con mínima suma de cuadrados de los residuales (SCRES), sin calcular todos los subconjuntos posibles de regresiones. Para esto se basa en la siguiente desigualdad:

$$SCRES(A) \leq SCRES(B) \qquad (1.1.9.1)$$

donde A es cualquier conjunto de variables independientes y B es un subconjunto de A. Esto es, que es imposible reducir la suma de cuadrados de los residuales de una regresión eliminando una o más variables.

Mediante el cálculo la SCRES de subconjuntos con un pequeño número de variables y en comparación con la SCRES de subconjuntos con un número mayor de variables, se puede evitar el calcular la SCRES de algunas regresiones que resultarían de eliminar variables del subconjunto con mayor número de variables.

1.1.9.1 Estadístico Cp-Mallow

El estadístico Cp-Mallows se define como:

$$Cp = \frac{RSSp}{s} - (n - 2p) \qquad (1.1.9.2)$$

Si el modelo no tiene carencia de ajuste entonces:

$$E(Cp) = p \qquad (1.1.9.3)$$

El estadístico de Mallows es útil en la selección de regresiones de subconjuntos de variables. Se Consideran las más adecuadas a las que tengan un valor de Cp muy aproximado o menor a p.

33) Furnival, George and Wilson, Robert. 'Regression by Leaps and Bounds', Technometrics, vol 16 No. 4 Noviembre 1974.

1.1.10 S U M A R I O

En este sección se trataron los siguientes temas: el modelo, sus estimadores y características; pruebas de hipótesis; carencia de ajuste; intervalos de confianza; predicción; variables categórica y selección de subconjuntos de variables.

Estos conceptos constituyen la báse teórica del paquete de regresión lineal múltiple. En la siguiente sección se tratarán temas referentes a la violación de los supuestos del modelo.

1.2 VIOLACION DE LOS SUPUESTOS DEL MODELO

1.2.1 INTRODUCCION

En la sección 1.1 se mostraron los conceptos básicos del análisis de regresión, las estimaciones, pruebas de hipótesis y otros temas. En esta sección se mostrarán cuatro temas.

Los tres primeros tratan problemas a la violación de supuestos del modelo definido, que son:

* Multicolinealidad entre las variables explicativas.
* Heteroscedasticidad en la varianza de las perturbaciones.
* Autocorrelación entre las perturbaciones.

En el cuarto tema se analiza con más detalle los residuales, para detectar si existen problemas en los datos, tales como observaciones influyentes o patrones sistemáticos.

1.2.2 MULTICOLINEALIDAD

1.2.2.1 Concepto

La existencia de la multicolinealidad es una violación a uno de los supuestos del modelo de Regresión Lineal: no existe ninguna relación exacta entre las variables explicativas.

La multicolinealidad puede ser perfecta o imperfecta. En la multicolinalidad perfecta se da una relación lineal exacta entre dos o más variables explicativas, es decir, son linealmente dependientes (ld).

$$X_j = \lambda_1 X_1 + \lambda_2 X_2 + \dots + \lambda_p X_p \qquad (1.2.2.1)$$

donde λ_i es una constante. Si tal relación no es exacta, o sea, que las variables explicativas no están intercorrelacionadas perfectamente se tiene una multicolinealidad imperfecta.

1.2.2.2 Consecuencias

En el caso de multicolinealidad perfecta los coeficientes de la regresión son indeterminados y sus varianzas son indefinidas. Si la multicolinealidad es imperfecta pero alta, las consecuencias son las siguientes:

a. Los errores estándar de los estimadores tienden a ser mayores conforme aumenta la colinealidad entre las variables
b. Debido a lo anterior, los intervalos de confianza de los estimadores tienden a ser grandes de lo que enrealidad son. En consecuencia, la probabilidad de aceptar una hipótesisfalsa crece.
c. Se dan cambios significativos en los coeficientes estimados cuando una variable es agregada o eliminada.
d. Se dan cambios significativos en los coeficientes estimados cuando un dato es alterado o suprimido.
e. Se puede obtener un coeficiente de determinación alto con pocos o casi ninguno de los coeficientes estimados estadísticamente significativos en forma individual.
f. Los signos algebraicos de los coeficientes estimados no están conforme a las expectativas que se hayan hecho a priori con base en los postulados teóricos.

1.2.2.3 Técnicas de detección

Algunas de las formas que ayudan a descubrir la existencia de la multicolinealidad son:

a. La multicolinealidad puede estar presente cuando se tiene un coeficiente de correlación alto entre las variables explicativas y ninguno o pocos de los coeficientes estimados son individualmente significativos.

b. La existencia de altos coeficientes de correlación entre las variables explicativas da indicios sobre la multicolinealidad. Por el contrario, si tales coeficientes no son altos no implica la no existencia de la multicolinealidad. Es adecuado por tanto revisar la matriz de correlación. Esta matriz contiene en su elemento a_{ij} la correlación entre las variables X_i y X_j.

c. Otra manera, es realizar una regresión de cada X_i con las restantes X_j $j=1,\ldots,p$, $i\neq j$. Esto da una idea sobre la existencia significativa o no de una relación entre esa X_i y las restantes. Para probar la significancia se aplica el análisis de varianza (Tabla 1.2.1) sólo que para una variable menos, es decir, para p-1 y siendo X_i la variable dependiente.

1.2.2.4 Posibles soluciones al problema

a. Si se conoce de alguna forma la combinación lineal de las variables explicativas que son colineales, se puede aplicar para superar el problema de la multicolinealidad. Por ejemplo, sea la ecuación:

$$Y = \beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3+ \quad \ldots \quad +\beta_p X_p \qquad (1.2.2.2)$$

y supóngase que el coeficiente $\beta_2=2/3\beta_1$, sustituyendo en (1.2.2.2) se obtiene

$$Y = \beta_0+\beta_1 X_1+2/3\beta_1 X_2+\beta_3 X_3+ \quad \ldots \quad +\beta_p X_p \qquad (1.2.2.3)$$

$$Y = \beta_0+\beta_1(X_1+2/3X_2)+\beta_3X_3+ \quad ... \quad +\beta_pX_p \qquad (1.2.2.4)$$

si $X_1+(2/3)X_2$ se considera como una nueva variable X_{p+1} entonces se tiene

$$Y = \beta_0+\beta_1X_{p+1}+\beta_3X_3+ \quad ... \quad +\beta_pX_p \qquad (1.2.2.5)$$

Una vez estimados los coeficientes de (1.2.2.5) se puede calcular β_2 por la relación entre éste y β_1. Quedando resuelta la ecuación original (1.2.2.2).

b. Otra posible solución es eliminar una de las variables colineales. Sin embargo, al hacer esto se puede caer en el problema del sesgo de especificación, esto es, una especificación incorrecta del modelo, ya que se pueden omitir variables de importancia.

c. Otra posibilidad que en algunas ocasiones disminuye la multicolinealidad es tomar otra muestra con las mismas variables o tan solo aumentar el tamaño de la muestra.

1.2.3 HETEROSCEDASTICIDAD

1.2.3.1 Concepto

La heteroscedasticidad es la violación al supuesto de que:

$$E\,[\varepsilon_i^2] = \sigma^2 \qquad (1.2.3.1)$$

esto es, la variación de cada perturbación ε_i es una constante igual a σ^2. Así la heteroscedasticidad se presenta cuando

$$E\,[\varepsilon_i^2] = \sigma_i^2 \qquad (1.2.3.2)$$

donde σ_i^2 no es constante para toda i.

1.2.3.2 Consecuencias

a. A pesar de que los coeficientes estimados siguen siendo insesgados, su varianza no es la misma.

b. Debido a lo anterior, los intervalos de confianza para los coeficientes estimados tienden a ser más grandes de lo que debieran ser.

c. Al ocurrir el punto anterior las pruebas t y F tienden a dar más significancia a los parámetros de lo que realmente tienen y, por lo tanto existe más riesgo de inferir equivocadamente

1.2.3.3 Técnicas de detección

a. Método gráfico. Un análisis gráfico de los residuales estimados al cuadrado (usados como una aproximación de los ε_i) puede dar indicios sobre la existencia de la heteroscedasticidad. Al graficar los e_i^2 contra la $\hat{Y}_i$ o contra las variables explicativas, se trata de identificar la presencia de algún patrón sistemático (ver Análisis de Residuales sección 4 de este capítulo.)

b. Prueba de Park. En esta prueba se propone que σ_i^2 es una función de la variable explicativa X_j que se expresa como:

$$\sigma_i^2 = \sigma^2 X_{ij}^{\beta} e^{\upsilon_i} \qquad (1.2.3.3)$$

o

$$\ln \sigma_i^2 = \ln \sigma^2 + \beta \ln(X_{ij}) + \upsilon_i \qquad (1.2.3.4)$$

siendo υ_i el término estocástico.

Otra vez se usan los e_i^2 como aproximación de los σ_i^2 por lo que (1.2.3.4) queda:

$$\ln e_i^2 = \ln \sigma^2 + \beta \ln(X_{ij}) + \upsilon_i$$

$$\ln e_i^2 = \alpha + \beta \ln(X_{ij}) + \upsilon_j \qquad (1.2.3.5)$$

si β es estadísticamente significativa, se sugiere la existencia de heteroscedasticidad.

c. Prueba de Glejser. En esta prueba al igual que en la anterior, se sugiere que σ_i^2 es una función de las variables explicativas. El propone hacer una regresión de $|e_i|$ vs. alguna función de X_j como:

$$|e_i| = \beta_0 + \beta_1 X_j + v_i \tag{12.3.6}$$

$$|e_i| = \beta_0 + \beta_1 \sqrt{X_j} + v_i \tag{1.2.3.7}$$

$$|e_i| = \beta_0 + \beta_1 (1/X_j) + v_i \tag{1.2.3.8}$$

$$|e_i| = \beta_0 + \beta_1 (1/\sqrt{X_j}) + v_i \tag{1.2.3.9}$$

Una ventaja de la aproximación de Glejser es que se puede distinguir entre heteroscedasticidad pura (si $\beta_0 = 0$ y $\beta_1 \neq 0$) y heteroscedasticidad mixta ($\beta_0 \neq 0$ y $\beta_1 \neq 0$). Lo mismo que para la prueba Parker, si β_1 es estadísticamente significativa, los datos sugieren la presencia de heteroscedasticidad

d. Prueba de Correlación de Rango de Spearman. Para realizar esta prueba, se necesita calcular el coeficiente de correlación de rango de Spearman, que se define como:

$$r_s = 1 - 6 \left[\frac{\Sigma\, d_i^2}{N(N^2-1)} \right] \tag{1.2.3.10}$$

donde d_i es la diferencia en los rangos atribuida a dos características diferentes del i-ésimo individuo o fenómeno, en este caso entre los rangos de $|e_i|$ y alguna X_j, y donde N es número de individuos o fenómenos clasificados.

Si $N > 8$, la significancia de r_s se puede probar mediante la prueba t:[1]

$$t = \frac{r_s\sqrt{N-2}}{\sqrt{1-r^2}} \qquad (1.2.3.11)$$

con N-2 grados de libertad.

Si el valor de t calculada es mayor que el valor t de tablas, se acepta la hipótesis de heteroscedasticidad. La prueba se hace para cada una de las variables explicativas que haya en el modelo.

1.2.3.4 Posibles soluciones al problema

El método relativamente más sencillo para tratar la heteroscedasticidad es el de cuadrados mínimos ponderados. Este método consiste en ponderar cada uno de los e_i^2 de tal forma que se cumpla $E(e_i^2) = \sigma^2$, para realizar esto se utiliza una matriz W = diagonal $[w_1, w_2, \ldots, w_n]$. Donde w_i es la ponderación en el caso i y es un valor constante.

Si se conoce las diferentes σ_i^2 entonces

$$w_i = \frac{1}{\sigma_i^2} \qquad (1.2.3.12)$$

Si no se conocen se pueden hacer suposiciones sobre las σ_i^2 y transformar el modelo, tales supuestos serián que w_i fuera igual a: $1/X_i, \sqrt{X_i}$.

Las ecuaciones normales a resolver serían[2]

$$X'WX\hat{\beta} = X'WY \qquad (1.2.3.13)$$

y

$$\hat{\beta} = (X'WX)^{-1}X'WY \qquad (1.2.3.14)$$

1) Gujarati, D., op. cit. p. 201
2) Seberg, G. op. cit., p. 137

1.2.4 AUTOCORRELACION

1.2.4.1 Concepto

La autocorrelación es la existencia de la correlación entre las perturbaciones y la no existencia de autocorrelación se expresa como:

$$E(\varepsilon_i,\varepsilon_j) = 0 \qquad i\neq j$$

Algunas causas de que exista autocorrelación son:

a. Si en el modelo propuesto se ha excluido una variable explicativa de gran importancia, podría provocar que las perturbaciones presentaran algún patrón sistemático, debido a que éstas expresarán la influencia ejercida por la variable excluida. A este problema se le llama sesgo de especificación por exclusión de variables.

b. El sesgo de especificación debido a una forma funcional incorrecta en el modelo provoca autocorrelación. Por ejemplo, si el modelo ajustado es una recta y los datos presentan un comportamiento cuadrático, entonces las perturbaciones captan el efecto sistemático del término cuadrático ausente.

c. Cuando una de las variables explicativas es el valor rezagado o retrasado de la variable dependiente, se produce lo que se llama Autorregresión. Esta puede estar presente en las perturbaciones y se define la autoregresión de orden 1 para la perturbación i como:[3]

$$\varepsilon_i = \rho\,\varepsilon_{i-1} + \upsilon_i \qquad (1.2.4.1)$$

3) Jhonston, J., op. cit., p. 244

donde υ_i es la variación aleatoria, que para toda i se cumple:[4]

$$E(\upsilon_i) = 0$$

$$E(\upsilon_i \upsilon_{i+h}) = \sigma^2 \qquad h = 0 \qquad (1.2.4.2)$$

$$E(\upsilon_i \upsilon_{i+h}) = 0 \qquad h = 0$$

El subíndice h es una constante en la que i+h esta entre 1 y n. ρ es el coeficiente de correlación donde $|\rho| < 1$. Si es positivo entonces se dice que las X's estan positivamente correlacionadas o que existen una autocorrelación positiva. Si es negativo entonces se dice que hay autocorrelación negativa. En ocaciones los datos son series medidas en el tiempo, por lo que en ocaciones se utilizan los términos correlacion serial positiva o negativa.

La autorregresión de orden k para la variable dependiente se puede expresar como:[5]

$$Y_i = \phi_1 Y_{i-1} + \phi_2 Y_{i-2} + \dots + \phi_k Y_{i-k} + \upsilon_i \qquad (1.2.4.3)$$

donde υ_i es igual al caso anterior y $\phi_1, \phi_2, \dots, \phi_k$ son constantes.

d. El utilizar un promedio móvil[6] en la información puede producir autocorrelación. Así, el promedio móvil de orden q es:[7]

$$\varepsilon_i = \phi_1 \varepsilon_{i-1} + \phi_2 \varepsilon_{i-2} + \dots + \phi_q \varepsilon_{i-q} + \upsilon_i \qquad (1.2.4.4)$$

donde υ_i cumple las propiedades de (1.2.4.2).

4) Wonnacott and Wonnacott, op. cit., p. 248
5) Wonnacott and Wonnacott, op. cit., p. 248
6) Definicipn de 'Promedio Mpvil', ver Wonnacott and W., o
7) Stewart and Wallis, 'Introductory Econometrics', p. 217

1.2.4.2 Consecuencias

Si se aplica el método de cuadrados mínimos ordinarios cuando existe autocorrelación, las consecuencias son que:

a. Los estimadores serán ineficientes y por lo tanto los intervalos de confianza serán más anchos de lo real y la prueba de significancia será menos potente.

b. La varianza residual estimada tiende a subestimar a la verdadera varianza, lo mismo sucede con las varianzas de los estimadores, lo cual afecta a las pruebas t y F.

c. Para una sola muestra tiende a dar una visión distorcionada de los verdaderos valores poblacionales.

1.2.4.3 Técnicas de detección

a. El método gráfico es una forma sencilla de detectar alguna tendencia sistemática en las perturbaciones. Pero como no se conocen las perturbaciones (ε's), se toma a los residuales como aproximación.

b. La prueba Durbin-Watson está definida por el estadístico d:[8]

$$d = \frac{\sum_{i=2}^{n} (e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2} \qquad (1.2.4.5)$$

Se puede demostrar que d tenderá a ser pequeña cuando exista autocorrelación serial positiva y tenderá a ser grande en caso de autocorrelación serial negativa. Existe el problema de determinar la distribución de d, ya que ésta depende de los valores de X en una muestra, pero Durbin y Watson establecieron a d_U(superior) y d_L(inferior) como los límites para el nivel de significancia de d.[9]

8) Jhonston, J., op. cit. pp. 251

9) Mas detalles sobre la Prueba de hipótesis ver: Idem.,

Para aplicar esta prueba se debe considerar que las X's son no estocásticas, que no se incluyan valores rezagados de Y como variable explicativa, y que las perturbaciones se generen mediante un esquema autorregresivo de primer orden, es decir:

$$\varepsilon_t = \rho\, \varepsilon_{t-1} + v_t \qquad (1.2.4.6)$$

Una aproximación para d está dada por:[10]

$$d = 2\left[1 - \frac{\Sigma\, e_t\, e_{t-1}}{\Sigma\, e_t^2}\right] = 2\,(1 - \rho) \qquad (1.2.4.7)$$

c. La Prueba de Von Neuman está definida por:[11]

$$\frac{\delta^2}{s^2} = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2 \,/(n-1)}{\sum_{t=1}^{n} (e - \bar{e})^2 / n} \qquad (1.2.4.8)$$

donde $\bar{e} = 0$ en cuadrados mínimos ordinarios y para n grande, se aproxima a una distribución normal con[12]

$$E\left[\frac{\delta^2}{s^2}\right] = \frac{2n}{n-1} \qquad (1.2.4.9)$$

$$Var\left[\frac{\delta^2}{s^2}\right] = \frac{4n^2(n-2)}{(n+1)(n-1)^3} \qquad (1.2.4.10)$$

Como $\frac{\delta^2}{s^2}$ $N(\frac{2n}{n-1}, \frac{4n^2(n-2)}{(n+1)(n-1)^3})$ se puede aplicar la prueba de hipótesis de que no hay correlación serial en los residuales. En esta prueba, para que (1.2.4.9) y (1.2.4.10) sean verdaderos se necesita que los valores e_t se distribuyan independientemente.

10) Gujarati, D., op. cit. p. 232
11) Jhonston, J., op. cit., p. 250
12) Idem., p. 250

d. Prueba de Corridas en la Secuencia de Signos de los Residuales. Una corrida es una secuencia ininterrumpida de signos, más "+" para residuales positivos y menos "-" para residuales negativos. La longitud de la corrida es el número de elementos que contiene. Sea n el número total de signos, n_1 el número de signos más, n_2 el número de signos menos y u el número de corridas. Si u es un número pequeño en una secuencia de residuales entonces puede indicar correlación serial positiva. En cambio, si u es grande, entonces puede indicar correlación serial negativa. Para saber si es pequeña o grande en comparación con el número de corridas esperadas en una secuencia aleatoria de n datos se aplica lo siguiente:

Si $n_1>10$ y $n_2>10$, entonces el número de corridas u se distribuye aproximadamente normal con:[13]

$$\mu = \frac{2n_1 n_2}{n_1+n_2} + 1 \qquad (1.2.4.11)$$

$$\sigma^2 = \frac{2n_1 n_2(2n_1 n_2 - n_1 n_2)}{(n_1\ n_2)^2(n_1+n_2-1)} \qquad (1.2.4.12)$$

Así

$$z = \frac{u - \mu \pm 1/2}{\sigma} \qquad (1.2.4.13)$$

donde (1/2) es la corrección por continuidad. Se usa 1/2 positivo para la cola inferior, es decir, probar que u es pequeña, y se usa -1/2 (negativo) para la cola superior, es decir, para probar que u es grande. Para $n_1<10$ o $n_2<10$, Frieda Swed y C. Eisenhart han elaborado tablas especiales que dan los valores críticos de u corridas para valores (n_1, n_2) y para estos mismos valores, la distribución acumulativa del número total de corridas.[14]

13) Draper and Smith, op.cit. p. 157-162

14) Fried S. Sved y c. Eisenhart 'Tables for testing randomnes grouping a Secuence Alternatives', Annals of Mathematical, vol.14, 1943.

1.2.4.4. Posibles soluciones al problema

Las perturbaciones ε_i son no observables, pero si se supone que siguen una tendencia autorregresiva de primer orden entonces:

$$\varepsilon_i = \rho\varepsilon_{i-1} + \upsilon_i \qquad (1.2.4.14)$$

donde $|\rho|<1$ y υ_i sigue los supuestos del modelo, al igual que cumple con (1.2.4.2). Ahora bien, si se conoce a ρ (coeficiente de correlación) y si el modelo planteado es $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \varepsilon_i$ entonces se aplica:

$$(Y_i - \rho Y_{i-1}) = \beta_0(1-\rho) + \beta_1(X_{i1} - \rho X_{i-1,1}) + \beta_2(X_{i2} - \rho X_{i-1,2}) + \ldots + \beta_p(X_{ip} - \rho X_{i-1,p}) + \upsilon_i$$

$$(1.2.4.15)$$

y como υ_i satisface los supuestos del modelo, se obtendrán estimadores con todas las propiedades óptimas. A ésta regresión se le conoce como ecuación de diferencias generalizadas. Como la primera observación no se puede transformar (Y_1, X_1) por la ecuación anterior, se puede aplicar la siguiente transformación $Y_1\sqrt{1-\rho^2}$ y $X_1\sqrt{1-\rho^2}$.

Si no se conoce ρ, se pueden hacer suposiciones sobre su valor. Si ρ es igual a cero, no existe autocorrelación. Si ρ vale uno la ecuación (1.2.4.15) se reduce a:

$$(Y_i - Y_{i-1}) = +\beta_1(X_{i1} - X_{i-1,1}) + \beta_2(X_{i2} - X_{i-1,2}) + \ldots + \beta_p(X_{ip} - X_{i-1,p}) + \upsilon_i \qquad (1.2.4.16)$$

que es la ecuación de las primeras diferencias. Como ésta ecuación no tiene intercepto, se pude suponer que el modelo original es:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \beta_{p+1} t + \varepsilon_i \qquad (1.2.4.17)$$

donde t es la variable de tendencia, entonces:

$$\Delta Y_i = \beta_1 \Delta X_{i1} + \beta_2 \Delta X_{i2} + \ldots + \beta_p \Delta X_{ip} + \beta_{p+1} + \upsilon_i \qquad (1.2.4.18)$$

Si β_p es significativa, implica que en el modelo original existe un término de tendencia lineal.

Si suponemos que $\rho = -1$ entonces (1.2.4.15) se transforma a:

$$\frac{(Y_i + Y_{i-1})}{2} = \beta_0 + \frac{\beta_1 (X_{i1} + X_{i-1,1})}{2} + \frac{\beta_2 (X_{i2} + X_{i-1,2})}{2} + \ldots + \frac{\beta_p (X_{ip} + X_{i-1,p})}{2} + \frac{\upsilon_i}{2} \qquad (1.2.4.19)$$

esta ecuación se conoce como regresión de promedios móviles en dos períodos.

Se puede obtener una aproximación del valor de ρ basándose en la estimación de Durbin-Watson. Se puede demostrar que

$$\hat{\rho} = 1 - d/2 \qquad (1.2.4.20)$$

pero para muestras pequeñas la aproximación es muy burda. Para ésto Theil y Nagar sugieren:[15]

$$\hat{\rho} = \frac{N^2(1-d/2) + k^2}{N^2 - k^2} \qquad (1.2.3.22)$$

donde k es el número de coeficientes a estimar incluyendo al intercepto, y N es el número total de observaciones.

Obtenida la aproximación de ρ, se procede a utilizarla en la ecuación de diferencias generalizadas.

1.2.5 ANALISIS DE RESIDUALES

1.2.5.1 Propiedades[16]

Como se definio en el primer capítulo I sección 1.1.2, el

15) Gujarati, D., op. cit. p. 238

16) Seberg, G., op. cit. p. 162

i-ésimo residual es:

$$e_i = Y_i - \hat{Y}_i \qquad (i=1,\ldots,n) \qquad (1.2.5.1)$$

o en notación matricial

$$\underline{e} = \underline{Y} - \hat{\underline{Y}} = (I_n - P)\underline{Y} \qquad (1.2.5.2)$$

donde $P = X(X'X)^{-1}X' = \left[(p_{ii})\right]$. Una de las propiedades de los residuales es:

$$\sum_i e_i = 0 \qquad (1.2.5.3)$$

Si $E[\underline{\varepsilon}]= 0$ y $Var[\underline{\varepsilon}] =\sigma^2 I$, entonces $E[\underline{e}]=\underline{0}$ y $Var[\underline{e}]=\sigma^2(I_n - P)^2 = \sigma^2(I_n - P)$. Si ε $N_n(0,\sigma^2 I_n)$ entonces $e_i \sim N(0,\sigma^2(1-p_{ii}))$. Además, se pueden expresar los residuales en función de las perturbaciones: $\underline{e}=(I-P)\underline{\varepsilon}$.

Dada esta relación entre las perturbaciones y residuales, el análisis de estos últimos da indicios de la violación de los supuestos.

1.2.5.2 Gráfica de residuales[17]

Primero, se transformará la varianza de los residuales para que sea aproximadamente la unidad. Una forma es:

$$\frac{1}{n}\sum_i^n Var(e_i) = \sigma^2\frac{(n-p-1)}{n} \qquad (1.2.5.4)$$

si se estima σ^2 por s^2, entonces :

$$c_i = \frac{e_i}{\left[\frac{s^2(n-p-1)}{n}\right]^{z}} \qquad (1.2.5.5)$$

Sin embargo, dado que $Var[e_i] =\sigma^2(1-p_{ii})$ los residuales

17) Seberg, G., op. cit. pp. 103,104

"estudentizados" son:

$$d_i = \frac{e_i}{s(1-p_{ii})^{1/2}} \qquad (1.2.5.6)$$

por lo que d_i y c_i se distribuyen aproximadamente como N(0,1).

a. Gráfica sobre Papel de Probabilidad Normal. Sean $d_1 < d_2 < \ldots < d_n$ los n residuales estudentizados. Para n grande la gráfica de los d_i vs. $\frac{i-1/2}{n}$ sobre papel de probabilidad normal, muestra características de normalidad de los d_i. Esta gráfica debe ser la recta y=x sí los d_i están normalmente distribuidos con media cero y varianza uno.

b. Residuales contra $\hat{Y}_i$. Con la gráfica de d_i vs. $\hat{Y}_i$, se pueden descubrir algunos patrones como el que algunos residuales pueden ser más grandes en valor absoluto que los demás. Si la varianza de ε_i es constante, entonces se espera que la variabilidad de los residuales sea uniforme, si no lo es, se pensará que existe algún problema en las ε_i. Un modelo inadecuado produce un comportamiento sistemático en la gráfica.

c. Los residuales contra un factor omitido. Si se cree que un factor que afecta a Y debe de ser incluído como variable explicativa, se pueden graficar los residuales contra éste factor, si se conocen sus valores. Por ejemplo, supongase que el factor omitido es X_t (con t = 1,n) y al graficar contra las d_i se observa una relación.

d. Los residuales contra cada regresor. Aquí, se puede observar alguna tendencia anormal de las variables X_j, por ejemplo, si se debe incluír una variable explicativa en forma cuadrática, logarítmica, exponencial o simplemente transformada.

e. Gráfica de X_j contra X_k. Muestra la correlación que existe entre estas variables explicativas.

f. Gráfica parcial de los residuales. Si se desea estudiar una relación más estrecha entre los residuales y el regresor X_j, una alternativa es la gráfica parcial de los residuales (e_i').

$$e_i' = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \dots - \hat{\beta}_{j-1} X_{i,j-1} - \hat{\beta}_{j+1} X_{i,j+1} - \dots - \hat{\beta}_p X_{ip} \qquad (1.2.5.7)$$

$$e_i' = e_i + \hat{\beta}_j X_{ij} \qquad (1.2.5.8)$$

contra X_{ij} para $i=1,\dots,n$.

1.2.5.3 Observaciones influyentes

Las observaciones influyentes son datos de los cuales la estimación de uno o más parametros dependen excesivamente, es decir, el incluir o no una observación influyente provoca una variación muy grande en los resultados de las estimaciones. Una observación influyente puede o no tener un residual muy grande, dependiendo del modelo ajustado y del resto de los datos.

R.D. Cook propone para detectar la influencia de una observación calcular los estadísticos Di.[18] donde:

$$D_i = \frac{(\hat{\underline{\beta}}-\hat{\underline{\beta}}_i)'\ X'X\ (\hat{\underline{\beta}}-\hat{\underline{\beta}}_i)}{(p+1)s^2} \qquad (1.2.5.9)$$

($i=1,n$) donde $\hat{\underline{\beta}}_i$ es el estimador de cuadrados mínimos después de omitir la i-ésima observación. Di es comparada con una $F(p,n-p-1,1-\alpha)$, donde una Di grande implica una observación influyente.

Para efectos de cálculo Di puede ser expresada como :

$$D_i = \left[\frac{e_i}{s(1-r_{ii})^{1/2}}\right]^2 \left[\frac{r_{ii}}{1-r_{ii}}\right] \left[\frac{1}{p+1}\right] \qquad (1.2.5.10)$$

18) Draper, N. and H. Smith, op. cit. p. 170

donde r_{ii} es el i-ésimo elemento de la diagonal principal de la matriz $X(X'X)^{-1}X'$ con todos los datos.

1.2.5.4 Observaciones Discrepantes[19]

Una observación discrepante tiene un residual más grande en valor absoluto que el resto de los residuales y quiza tres o cuatro desviaciones estándar o más lejos de la media de los residuales.

En algunos casos la observacion discrepante provee información que las otras observaciones no dan.

1.2.4 S U M A R I O

Hasta esta sección se han contemplado algunos elementos básicos y específicos del análisis de regresión, en la siguiente se mostrará algunos algoritmos que permiten resolver problemas de solución númerica.

1.3 METODOS NUMERICOS.

1.3.1 INTRODUCCION

En esta sección se contempla la solución numérica de algunas operaciones o algoritmos algebráicos que se requieren en este paquete. Por ejemplo, para encontrar los estimadores de Cuadrados Mínimos Ordinarios es necesario resolver el sistema de ecuaciones lineales simultáneas: $X'X\underline{\hat{\beta}} = X'\underline{Y}$ $(Ax = b)$. Existen varios métodos para resolver este sistema, como el de Gauss-Jordan, el método iterativo de Jacobi, Gauss_Seidel, etc., pero dado que la matriz X'X es simétrica y definida positiva, se ha aplicado el método de la descomposición de Cholesky, con algunas modificaciones. Dada la solución directa por este método, se puede encontrar una aproximación más exacta a la solución aplicando la descomposición de Cholesky en forma iterativa. Este procedimiento es necesario para resolver los errores de redondeo en la primera solución. También se incluyen en este capítulo los algoritmos que valúan las funciones de distribución más usadas en el paquete, como lo es la Normal, la t-student y la distribución F.

1.3.2 METODOS NUMERICOS PARA RESOLVER EL SISTEMA DE ECUACIONES

Para encontrar la solución del sistema

$$X'X\hat{\beta} = X'\underline{Y} \qquad (1.3.2.1)$$

se utiliza el método de la descomposición de Cholesky con algunas modificaciones,[1] que requiere que la matriz X'X sea definida positiva.[2] El método se basa en un teorema expuesto por Cholesky, que menciona que si A es una matriz simétrica definida positiva,[3] entonces existe una matriz triangular inferior L real y no singular tal que:

$$LL' = A \qquad (1.3.2.2)$$

y que si los elementos de la diagonal de L son tomados para ser positivos, la descomposición es única.

Los elementos de L pueden ser determinados renglón por renglón; para el i-ésimo renglón se tiene que:

$$\sum_{k=1}^{j} l_{ik} l_{jk} = a_{ij} \qquad (1.3.2.3)$$

donde

$$l_{ij} = (a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}) / l_{jj} \qquad (j=1,\ldots,i-1) \qquad (1.3.2.4)$$

1) Martin, R.S., Peters, G. and Wilkinson J.H. 'Simmetric Decomposition of a Positive Definite Matrix'. pp. 362-383
2) En el Apendice B se prueba que X'X es definida positiva
3) Franz E. Hohn, 'Algebra de matrices', pp. 393: Una matriz A simétrica real es definida positiva si y solo si todas sus raices características son positivas.

$$\sum_{k=1}^{i} l_{ik} l_{ik} = a_{ii} \qquad (1.3.2.5)$$

y

$$l_{ii} = \left(a_{ii} - \sum_{k=1}^{i-1} l_{ik}^{2}\right)^{1/2} \qquad (1.3.2.6)$$

En estas ecuaciones se realizan n raíces cuadradas y aproximadamente $n^3/6$ multiplicaciones.

Para evitar las raíces cuadradas se hacen las siguientes modificaciones. Sea

$$L = \tilde{L}\,\mathrm{diag}(l_{ii}) \qquad (1.3.2.7)$$

donde $\tilde{L}$ es una matriz triangular inferior con unos en su diagonal. Así, de (1.3.2.2) y (1.3.2.7) se obtiene que:

$$A = \tilde{L}\,\mathrm{diag}(l_{ii})\,\mathrm{diag}(l_{ii})\,\tilde{L}' = \tilde{L}\,\mathrm{diag}(l_{ii}^{2})\,\tilde{L}' = \tilde{L}D\tilde{L}' \qquad (1.3.2.8)$$

donde D es una matriz diagonal positiva. Al transformar las ecuaciones (1.3.2.3) a (1.3.2.6) se deduce:

$$\sum_{k=1}^{j} \tilde{l}_{ik} d_k \tilde{l}_{jk} = a_{ij} \qquad (1.3.2.9)$$

$$\tilde{l}_{ij} d_j = a_{ij} - \sum_{k=1}^{j-1} \tilde{l}_{ik} d_k \tilde{l}_{jk} \qquad (j=1,\ldots,i-1) \qquad (1.3.2.10)$$

$$\sum_{k=1}^{j} \tilde{l}_{ik} d_k \tilde{l}_{ik} = a_{ii} \qquad (1.3.2.11)$$

$$d_i = a_{ii} - \sum_{k=1}^{i-1} \tilde{l}_{ik} d_k \tilde{l}_{ik} \qquad (1.3.2.12)$$

Para la solución de un sistema de ecuaciones cualquiera :

$$A\ x = b \tag{1.3.2.13}$$

se resuelve primero la siguiente ecuación para y

$$\tilde{L}\ y = b \tag{1.3.2.14}$$

teniendo la ventaja de que $\tilde{L}$ es triangular inferior con unos en su diagonal. El siguiente paso es calcular x:

$$\tilde{L}'x = D^{-1}y \tag{1.3.2.15}$$

Esto es posible por la siguiente relación:

$$Ax = \tilde{L}D\tilde{L}'x = \tilde{L}DD^{-1}y = b \tag{1.3.2.16}$$

se resuelve (1.3.2.13) y (1.3.2.14) con n^2 multiplicaciones y n cocientes. Para calcular L se utilizan aproximadamente $n^3/6$ multiplicaciones.

1.3.3 DISTRIBUCIONES

En algunas seciones del paquete se dan los estadísticos calculados como la F-calculada en la tabla ANOVA, la t-calculada en las pruebas de hipótesis individual. A este valor calculado del estadístico, le corresponde cierta probabilidad que en la sección 1.1.5 de este capítulo, se definió como valor P. Para obtener esta probabilidad se utilizan a las siguientes formulas:

1.3.3.1 Distribición Normal[4]

Sean:

$$Z(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}} \tag{1.3.3.1}$$

4) Abramovitz and Stegun,'Handbook of Mathematical Functions',p 932

$$Q(x) = Z(x)\ (b_1 t + b_2 t^2 + b_3 t^3 + b_4 t^4 + b_5 t^5) \qquad (1.3.3.2)$$

$$t = \frac{1}{1+px} \qquad (1.3.3.3)$$

donde

$p = 0.2316419 \qquad b_1 = 0.319381530 \qquad b_2 = -0.356563782$

$b_3 = 1.781477937 \qquad b_4 = -1.821255978 \qquad b_5 = 1.33.274429$

y donde $Q(x)$ es la región sombreada en la figura 3.3.1. Se conoce x y se calcula $Q(x)$.

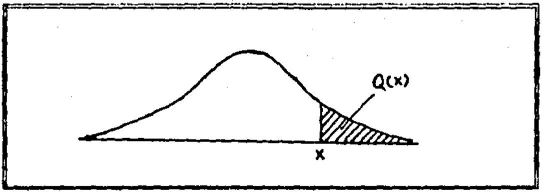


Figura 3.3.1 Area bajo una cola de la Distribición Normal

Transformación inversa:[5] Sea X_p, donde $Q(X_p)=p$, $0 < p < 0.5$. En este caso se conoce p y se desea obtener el valor de X_p.

$$X_p = t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \qquad (1.3.3.4)$$

$$t = \sqrt{\ln(1/p^2)} \qquad (1.3.3.5)$$

donde

$c_0 = 2.515517 \qquad c_1 = 0.802853 \qquad c_2 = 0.010328$

$d_1 = 1.432788 \qquad d_2 = 0.189269 \qquad d_3 = 0.001308$

5) Abramowitz and S., p 933

1.3.3.2 Distribución t de student con v grados de libertad[6]

$$\text{Sea} \qquad Q(x) = \frac{1}{4}\left(1+a_1x+ a_2x^2+ a_3x^3+ a_4x^4\right)^{-4} \qquad (1.3.3.6)$$

$$\text{donde} \qquad x = \frac{t^{2/3}\left(1-\frac{2}{9v}\right)-\frac{7}{9}}{\left[\frac{2}{9}+t^{4/3}\ \frac{2}{9v}\right]^{1/2}} \qquad (1.3.3.7)$$

y

$a_1 = 0.196854$ $\qquad$ $a_2 = 0.115194$

$a_3 = 0.000344$ $\qquad$ $a_4 = 0.019527$

Aproximación para valores grandes de v. Sea $A(t/v)$ el area sombreada en la figura 3.3.2.

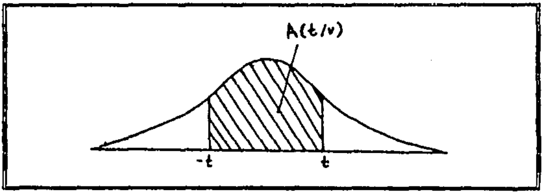


Figura 3.3.2. Area bajo una cola de la Distribución t de student.

$$A(t/v) \sim 2P(x)-1; \qquad x = \frac{t\left(1-\frac{1}{4v}\right)}{\sqrt{1+\frac{t^2}{2v}}} \qquad (1.3.3.8)$$

Aproximación para valores grandes de t y v < 5.

$$A(t/v) \sim 1-\frac{a_v}{t^v}+\frac{b_v}{t^{v+1}} \qquad (13.3.9)$$

[6] Abramovitz and S., pp. 948,949

con:

v	1	2	3	4	5
a_v	0.3183	0.4991	1.1094	3.0941	9.948
b_v	0.0000	0.0518	-0.046	-2.756	-14.05

Tabla 3.3.1. Para valores grandes de t y v < 5.

Expansion Asintótica para la Función Inversa[7]

Si $A(t_p/v) = 1-2p$ y $Q(X_p) = p$, entonces

$$t_p = X_p + \frac{g_1(X_p)}{v} + \frac{g_2(X_p)}{v^2} + \frac{g_3(X_p)}{v^3} + \frac{g_4(X_p)}{v^4} \qquad (1.3.3.10)$$

$$g_1(X) = \frac{1}{4}(X^3+X) \qquad (1.3.3.11)$$

$$g_2(X) = \frac{1}{96}(5X^5+16X^3+3X) \qquad (1.3.3.12)$$

$$g_3(X) = \frac{1}{384}(3X^7+19X^5+17X^3-15X) \qquad (1.3.3.13)$$

$$g_4(X) = \frac{1}{92160}(79X^9+776X^7+1482X^5-1920X^3-945X) \qquad (1.3.3.14)$$

1.3.3.3 Distribución Fisher (F) con v_1 y v_2 grados de libertad[8]

$$\text{Sea } Q(F) = \frac{1}{2}(1 + a_1x + a_2x^2 + a_3x^3 + a_4x^4)^{-4} \qquad (1.3.3.15)$$

la región sombreada en la figura 3.3.3. Donde

7) Abramovitz and S., p. 949
8) Abramovitz and S., p. 947

$$x = \frac{F^{1/3}\left(1-\frac{2}{9v_2}\right)-\left(1-\frac{2}{9v_1}\right)}{\left(\frac{2}{9v_1}+F^{2/3}\frac{2}{9v_2}\right)^{1/2}} \qquad (1.3.3.16)$$

y

$a_1 = 0.196854 \qquad a_2 = 0.115194$

$a_3 = 0.000344 \qquad a_4 = 0.019527$

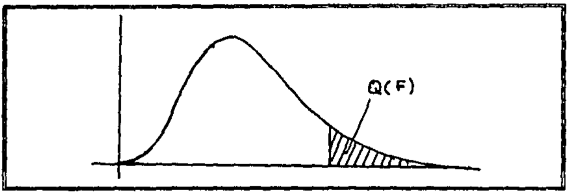


Figura 3.3.3. Area bajo una cola de la Distribución F

otro médoto es:

$$Q(F/v_1,v_2) \sim Q(x) \qquad x = \frac{F^{1/3}\left(1-\frac{2}{9v_2}\right)-\left(1-\frac{2}{9v_1}\right)}{\left(\frac{2}{9v_1}+F^{2/3}\frac{2}{9v_2}\right)^{1/2}} \qquad (1.3.3.17)$$

$Q(x)$ es igual a la ecuación (1.3.3.2).

La función inversa de F no se requiere en el paquete. El método que se utilizó para obtener las fórmulas anteriores es el de Polinomios Ortogonales de Hermite (ver anexo D).

1.3.4 ERRORES DE APROXIMACION

En la aplicación del método modificado de la descomposición de Cholesky para la solución del sistema (1.3.2.1), se observó que en ocasiones la solución era relativamente menos exacta que la derivada del método de Gauss-Jordan por lo que se decidió mejorar la solución dada por el método modificado de la descomposición de Cholesky, con la aplicación de un procedimiento iterativo en el cual la descomposición de Cholesky se usa repetidamente.[9]

Con la descomposición de la matriz definida positiva de (1.3.2.2) el procedimiento se define de la siguiente forma: se da un valor inicial de cero a la solución, esto es:

$$X^{(0)} = 0 \qquad (1.3.4.1)$$

y en cada iteración se calcula el residual correspondiente a la solución de esa iteración; para la j-ésima iteración tenemos:

$$r^{(j)} = b - AX^{(s)} \qquad (1.3.4.2)$$

También se calcula la corrección de la solución con la ayuda de la descomposición (1.3.2.2) y ésta se suma a la solución, de manera que:

$$(LL')d^{(s)} = r^{(s)} \qquad (1.3.4.3)$$

$$X^{(s+1)} = X^{(s)} + d^{(s)} \qquad (1.3.4.4)$$

1.3.5 S U M A R I O

En esta sección se explicó el método para encontrar la solución numérica, para resolver el sistema de ecuaciones y para obtener el valor de las funciones de distribución usadas en el paquete. En el próximo capítulo se tratarán temas relacionados con el uso del paquete.

9) Martin, R.S., Peters, G. and Wilkinson J.H. 'Iterative Refinement of the Solution of a Positive Definite System of Equations', pp. 203-216

C A P I T U L O I I

F O R M A D E M A N E J A R
E L P A Q U E T E D E
R E G R E S I O N L I N E A L
M U L T I P L E

2.1 INTRODUCCION

En este capítulo se da la información referente al modo de operación del sistema en general, las opciones que se contemplan en el paquete y el modo de operación de cada una de ellas, cómo introducir los datos en el paquete, la manera de crear variables por medio de la simulación, de las transformaciones y de las variables dicótomas, la forma en que esta estructurado el sistema y sus características.

2.2 OPERACION DEL SISTEMA

El Paquete de Regresión Lineal Múltiple (REGCOM) tiene una forma sencilla de manejo, por medio de menús. Se presenta en la pantalla una serie de opciones organizadas donde al seleccionar una opción oprimiendo la tecla correspondiente puede ocurrir que el sistema conduzca a otro menú o bien a que se ejecute directamente un procedimiento. Al entrar a un menú se tiene una opción de "Salida" que permite regresar al menú anterior.

Cuando en un menú algunas de las opciones aparezcan en la pantalla iluminadas con una mayor intensidad que otras, ello indica que sólo esas opciones pueden ser elegidas, mientras que si todas tienen la misma intensidad, se pueden elegir indistintamente.

Para entrar al Paquete REGCOM se teclea "REGCOM" en la pantalla de la computadora desde el sitema operativo MS-DOS (Microsoft Disk Operating Sistem)[1] y se oprime la tecla "Return" o "Enter". El paquete muestra en primera instancia las nueve opciones del Menú Obtención de Datos (figura 2.2.1).

```
                    MENU OBTENCION DE DATOS

   A   NUMERO DE DATOS DE LA REGRESION (n)
   B   REGRESION CON O SIN TERMINO INDEPENDIENTE
   C   LEER X
   D   LEER Y
   E   LEER Y y X
   F   SIMULAR X
   G   SIMULAR Y
   H   CONTINUAR

   S   SALIDA AL SISTEMA OPERATIVO (DOS)
```

Figura 2.2.1. Menú Obtención de Datos.

El objetivo de este menú es que usuario proporcione los datos con los cuales va a trabajar. Al principio de este menú sólo se pueden escoger las opciones A o la S. Con la opción A se da el número de datos de la regresión (n), es decir, el número de renglones de la matriz X y el número de elementos del vector Y. La opción S da por finalizada la sesión y regresa al sistema operativo MS-DOS.

Si se ejecuta la opción A, entonces se pueden seleccionar las opciones B, C, D, E y F, quedando aún inhibidas las opciones G y H. La opción B permite determinar que la regresión tenga o no un término independiente o no lo tenga. El manejo de los datos dentro

1) El sistema operativo MS-DOS permite el uso y la ejecucion de programas en una computadora personal.

del paquete para considerar un modelo de regresión lineal con término independiente se hace de manera que la matriz X no necesita tener la columna de unos (seccion 1.1.1, capítulo I), sólo se debe especificar desde esta opción que se desea utilizar el modelo con término independiente o no, por omisión la regresión se determina con término independiente. Las opciones C, D, E y F que se refieren a la lectura de la matriz X, a la lectura del vector Y, a la lectura de ambos y a la simulación de columnas de la matriz X respectivamente, son puntos que se explican en la seccion de Lectura y Simulación.

La opción G "Simular Y" se puede seleccionar una vez que se ha leído o simulado una columna de la matriz X con las opciones C, E, o F. Esta opción se explica en la sección de Lectura y Simulación.

La opcion H "Continuar" puede ser elegida una vez que se ha leído o simulado a la matriz X y se ha leído o simulado al vector Y, y conduce al Menú Inicial (figura 2.2.2).

```
                    M E N U   I N C I A L

   A   GRAFICAS
   B   SIMULAR  X
   C   SIMULAR  Y
   D   EDICION DE DATOS
   E   CONSTRUCCION DE VAR. DICOTOMAS
   F   TRANSFORMACIONES
   G   CAMBIO DE VARIABLE DEPENDIENTE
   H   LISTADO DE DATOS
   I   SALIDA A LA IMPRESORA          [DESACTIVADO[

   N   NUEVA REGRESION
   P   MENU PRINCIPAL
   S   SALIDA AL SISTEMA OPERATIVO (DOS)
```

Figura 2.2.2. Menú Inicial

El propósito de este menú es facilitar la aplicación de la regresión a los datos en las condiciones apropiadas, pero sin tener una relación directa con el análisis de regresión propiamente dicho.

La opción A "Gráficas" permite obtener el gráfico de cualquier columna de la matriz X contra Y, esto es, de cualquier variable explicativa contra la variable dependiente. Las opciones B, C, D, E y F son explicadas en secciones posteriores.

La opción H "Listado de Datos" lista las observaciones que contiene la matriz X y el vector Y.

La opción I "Salida a la impresora" se encuentra desactivada al iniciar el paquete, y se activa al elegirla. De esta manera, los resultados también se obtendrán por la impresora. Al seleccionarla nuevamente, se desactiva. En todos los menús del paquete, excepto el Menú de Obtención de Datos y los menús de Lectura y Simulación, aparece esta opción y en su lado derecho se despliga un mensaje que indica sí la salida a la impresora está activada o no. Con la opción N "Nueva Regresión" se regresa al Menú Obtención de datos, desde donde se puede dar comienzo a un nuevo ajuste de regresión, al tomar esta opción se eliminan de la memoria cualquier tipo de datos del ajuste previo.

La siguiente opción del Menú Inicial (P) "Menú Principal" conduce al Menú Principal, el cual contiene la especificación de los procedimientos que intervienen en forma directa en la regresión y su análisis. Las opciones que contempla este menú se explican en la sección Menú Principal.

Al igual que el Menú Obtención de Datos, este menú da la opción de finalir la sesión y regresar al sistema operativo MS-DOS.

2.3 LECTURA Y SIMULACION

El primer paso en el paquete REGCOM consiste en tener en la memoria de la computadora los valores del vector Y y de la matriz X. Para lograr esto, se tiene la posibilidad de simular o leer tanto la variable Y como la matriz X. Estas posibilidades están determinadas por las opciones C, D, E, F, y G del Menú Obtención de Datos (figura 2.2.1). Antes de que elija alguna de estas opciones, el usuario debe proporcionar el número n de datos (opción A). Al elegir cualquier opción en donde se lee o se simula a la matriz X (opciones C, E y F), se pide el número de variables explicativas, es decir, el número de columnas que se quieren leer o simular. La primera vez que se simula o lee columnas de la matriz X, éstas son guardadas en memoria en el siguiente orden: la primera columna leída o simulada se guarda en la primera columna de la matriz X; la segunda columna leída o simulada se guarda en la segunda columna de la matriz X, y así sucesivamente hasta llegar a la última columna leída o simulada en la primer opción tomada. Si se vuelve a leer o simular una o varias columnas de la matriz X, las nuevas columnas de la matriz X son guardadas en la memoria agregandose a partir de la última columna de datos que ya se tenía almacenada. Esto permite leer y simular columnas de X tantas veces como se requiera, además de que la matriz X de memoria puede tener columnas simuladas y columnas leídas.

El vector Y puede ser leído por medio de las opciones D y E, o puede ser simulado si se elige la opción G, pero para elegir ésta última se debe tener al menos ya leída o simulada una columna de la matriz X. Sí ya se obtuvo al vector Y, se tiene la posibilidad de volver a leerlo o simularlo, borrándose los datos anteriores. Pero sí Y es simulada por segunda vez , el vector obtenido será colocado después de la última columna de la matriz X.

Antes de explicar las opciones de lectura y simulación es necesario aclarar que como se explicó en la sección anterior la matriz X no necesita tener la columna de unos cuando se desea que el modelo tenga término independiente, internamente los programas la consideran cuando se desea que el modelo tenga término independiente.

2.3.1 Lectura de la matriz X

La opción C "Leer X" del Menú Obtención de Datos conduce al Menú de Lectura de X (figura 2.3.1).

```
                 MENU DE LECTURA DE X

   A    POR PANTALLA
   B    POR ARCHIVO
   S    SALIDA
```

Figura 2.3.1. Menú de Lectura de X

Después de seleccionar alguna de las opciones del Menú de Lectura de X, se debe dar el número de columnas de la matriz X que se van a leer.

Si se seleccionó la opción A "Por Pantalla" las observaciones de las variables de la matriz X se tienen que dar capturando la información en la pantalla. Los datos se dan después de que en la pantalla aparece "Renglón 1 :", capturando la primera observación de la primera columna que se va a leer seguida de una coma, luego se captura la primera observación de la segunda columna seguida de una coma, y así sucesivamente hasta llegar a la primera observación de la última columna que se va leer, la cual no debe estar seguida por una coma. Una vez que se ha escrito este valor se oprime la tecla "enter" o la tecla "return". En el siguiente renglón de la pantalla aparecerá "Renglón 2 :" y aquí se deben dar las segundas observaciones de las columnas que se están leyendo en la misma forma en que se capturó las primeras observaciones, el proceso continua hasta que se escriben las n-ésimas observaciones. Los valores que se capturan son validados y si se encuentra algún caracter no numérico (a exepción de las comas), se detecta el error y se vuelve a pedir este renglón.

Sí se seleccioná la opción B "Por Archivo" se le pedirá al usuario el nombre del archivo donde se encuentran los datos. El programa buscará el archivo, si lo encuentra y en la validación de los datos no se detecta error, en la pantalla se despliega el mensaje "Fin de Archivo" y el número de renglones leídos que debe ser igual a n, sí en la validación de los datos se encuentra error en algún registro, entonces en la pantalla aparecerá el número del registro en el que se localizó el error y el mensaje "Fin de Archivo", si el programa no encuentra el archivo donde se encuentran los datos , en la pantalla se pedirá nuevamente el nombre del archivo.

El paquete REGCOM lee archivos de tipo texto, también llamados "Standard Data File, (SDF)", que puden ser creados en un editor de textos como lo son el de TURBO Pascal, el WORDSTAR, el de SideKick, entre otros. Cada renglón del archivo es considerado como un registro y éstos deben estar en formato libre, es decir, los números separados por comas y sin espacios entre estos. Todos los archivos que se utilicen para cualquier lectura en el paquete deben considerar una observación de la varible Y como primer componente y las observaciones de cada una de las variables explicativas como los restantes componentes para cada uno de renglones que contenga el archivo.
Como el archivo de datos que se quiere leer contiene tanto a las observaciones del vector Y, como a las observaciones de la matriz X, en la opción de lectura de X por medio de archivo, el programa ignora a la primera observación de un renglón y lee las restantes observaciones hasta el número total de columnas que se especifico

al entrar a la opción. Esta característica permite que de un archivo sólo se lean una parte de las variables explicativas, aunque en realidad se tengan más.

2.3.2 Lectura del vector Y

La opción D "Leer Y" del Menú Obtención de Datos lleva al Menú de Lectura de Y (figura 2.3.2).

```
          MENU DE LECTURA  DE  Y

A    POR PANTALLA
B    POR ARCHIVO
S    SALIDA
```

Figura 2.3.2. Menú de Lectura de Y

Con la opción A "Por Pantalla", la lectura de las n observaciones del vector Y se hace capturando cada uno de los valores en la pantalla. Cuando se elige esta opción, en la pantalla se despliega una "Y" junto con un número consecutivo que indicará el número de la observación que se está leyendo, al lado derecho de el número consecutivo se capturará el valor de la observación. Sí se ingresa un caracter no numérico, se volvera a pedir la observación.

En la opción B "Por Archivo" se le pedirá al usuario el nombre del archivo donde se encuentran los datos. El programa buscará el archivo, sí lo encuentra y en la validación de los datos no se detecta error, en la pantalla se despliega el mensaje "Fin de Archivo" y el número de renglones leídos que debe ser igual a n, sí en la validación de los datos se encuentra error en algún registro, entonces en la pantalla aparecerá el número del registro en el que se localizó el error y el mensaje "Fin de Archivo", si el programa no encuentra el archivo donde se encuentran los datos, en la pantalla se pedirá nuevamente el nombre del archivo.

Para leer el vector Y de un archivo, el paquete sólo considerará el primer componente de cada renglón del archivo, ya que como se explicó en la opción de Lectura de X, los archivos de lectura deben contener tanto al vector Y como a la matriz X, siendo el primer componente de cada renglón una observación de Y.

2.3.3 Lectura de la matriz X y el vector Y

La opción E "Leer X y Y" es la más utilizada en la mayoría de las aplicaciones. Al seleccionarse esta opción aparecerá el Menú de Lectura de X y Y (figura 2.3.3).

```
                    MENU DE LECTURA DE X y Y

A   POR PANTALLA
B   POR ARCHIVO
S   SALIDA
```

Figura 2.3.3. Menú de Lectura de X y Y

Después de seleccionar alguna de las opciones de este menú, se debe dar el número de columnas de la matriz X que se van a leer. Con la opción A "Por Pantalla" la lectura del vector Y y de las columnas de la matriz X se hace capturando los datos en la pantalla. Los datos se introducen después de que en la pantalla aparece "Renglón 1 :". Se captura el valor de la primera observación del vector Y seguido de una coma, a continuación se da la primera observación de la primera columna de X que se va a leer seguida de una coma, luego se captura la primera observación de la segunda columna de X seguida de una coma, y así sucesivamente hasta llegar a la primera observación de la última columna que se va leer (que no debe ser seguida por coma), y se oprime la tecla "enter" o la tecla "return". En el siguiente renglón de la pantalla aparecerá "Renglón 2 :" y aquí se darán, la segunda observación del vector Y, seguida de las segundas observaciones de las columnas de la matriz X que se están leyendo, con el mismo orden en que se metieron las primeras y en la misma forma, el proceso continua hasta que se da la observación n-ésima del vector Y, seguida de las n-ésimas observaciones de las columnas de la matriz X.

En la opción B "Por Archivo" se le pedirá al usuario el nombre del archivo donde se encuentran los datos. El programa buscará el archivo, sí lo encuentra y en la validación de los datos no se detecta error, en la pantalla se despliega el mensaje "Fin de Archivo" y el número de renglones leídos que debe ser igual a n, sí en la validación de los datos se encuentra error en algún registro, entonces en la pantalla aparecerá el número del registro en el que se localizó el error y el mensaje "Fin de Archivo", sí el programa no encuentra el archivo donde se encuentran los datos, en la pantalla se pedirá nuevamente el nombre del archivo.

En esta opción, a diferencia de la anterior opción de lectura por archivo de X y la opción de lectura por archivo de Y, se consideran en la lectura al primer componente (observación de Y) y a las restantes componentes hasta el número de columnas que se van a leer de la matriz X para cada uno de los renglones.

Es importante por último comentar que en la lectura de un archivo, en cualquie de las opciones solamente se leerán n observaciones, aunque en el archivo existan más de éstas.

2.3.4 Simulación de la matriz X

Al seleccionar la opción F "Simular X" del Menú Obtención de Datos se entra al Menú de Simulación de X(j) (figura 2.3.4.).

```
          MENU DE SIMULACION DE X(j)

  A   SIMULACION DE UNA UNIFORME [a,b)
  B   SIMULACION DE UNA BERNOULLI(p)
  S   SALIDA
```

Figura 2.3.4. Menú de Simulación de X(j)

Antes de entrar a este menú se pide el número de columnas de la matriz X que se quieren simular. El programa se mantendrá en este menú hasta que se simule el número de columnas que se especificó o se seleccione la opción S "Salida".

En la opción A "Simulación Uniforme [a,b)" el usuario debe introducir los valores a y b cuando se pidan en la pantalla, estos valores son el límite inferior y el límite superior del intervalo de la distribución uniforme, y por las características de esta distribución, b debe ser mayor al valor de a. De esta manera la columna de la matriz X creada, tendrá una distribución uniforme en el intervalo [a,b).

En la opción B "Simulación de una Bernoulli (p)" se pide el valor p (probabilidad de ocurrencia de un evento). Como p es una probabilidad se debe dar un valor entre 0 y 1. Así, la columna de la matriz X creada, tendrá una distribución Bernoulli y las observaciones de esta columna tendrán el valor 1 o 0 (si el evento ocurrió o no).

Cada una de las columnas creadas en este menú mediante cualquiera de las dos opciones anteriores son almacenadas en la memoria una columna después de la última columna ocupada.

Esta opción también puede ser seleccionada desde el Menú Inicial (figura 2.2.2).

2.3.5 Simulación del vector Y

La opción E "Simular Y" del Menú Obtención de Datos conduce al Menú de Simulación de Y (figura 2.3.5).

SIMULACION DE Y (DISTRIBUCION NORMAL)

A	Y DISTRIBUIDA N($X\beta$, σ^2 I)
B	Y DISTRIBUIDA N($X\beta$, V)
S	SALIDA

Figura 2.3.5. Menú de Simulación de Y

Como se mencionó al principio de esta sección, para entrar a esta opción es necesario tener en memoria al menos una columna de la matriz X.

Para simular al vector Y es necesario generar aleatoriamente el vector de perturbaciones ε, mediante las opciones A o B de este menú, y una vez simulado se pedirá un vector hipotético de las estimadores con p o p+1 componentes, según sea el caso de una regresión con término independiente o sin él. Por supuesto, p es el número de columnas de datos en memoria de la matriz X. Una vez obtenido el vector de estimadores, se efectua la operación de multiplicar a la matriz X por el vector de los estimadores hipotéticos y sumarle al resultado el vector de perturbaciones generado, y este resultado será el valor simulado de Y.

La opción A "Y distribuida $N(X\beta,\sigma^2 I)$" permite crear los n valores del vector con una distribución de probabilidad Normal $N(0,\sigma^2 I)$, es decir, la esperanza del vector es igual a cero y la matriz de varianzas y covarianzas de el vector ε es $\sigma^2 I$. El usuario debe dar el valor de σ^2 al entrar a esta opción.

La opción B "Y distribuida $N(X\beta, V)$" permite crear los n valores del vector ε con una distribución de probabilidad Normal $N(0,\sigma_i^2)$, es decir, que para cada observación existe una varianza residual $(E(\varepsilon_i)=\sigma_i^2)$, no necesariamente la misma. En este caso, el usuario tiene que dar un valor de σ_i^2 para cada una de las n observaciones. Esta opción también puede ser seleccionada desde el Menú Inicial(figura 2.2.2).

2.4 TRANSFORMACIONES

En el análisis de regresión lineal simple se trabaja generalmente un modelo de combinaciones lineales de las variables explicativas. En un problema concreto, sin embargo, el comportamiento de los datos puede ser o no lineal, lo que puede conducir a un ajuste no adecuado. Existen un gran número de modelos no lineales que mediante adecuadas transformaciones de las variables pueden convertirse en lineales.

El paquete REGCOM permite llevar a cabo transformaciones de variables con lo que se evita el trabajo de calcular estas transformaciones fuera del programa.
Este procedimiento se encuentra en la opción F "Transformaciones" del Menú Inicial.

Al igual que la lectura y simulación de la matriz *X* y del vector Y, las transformaciones que se realizan son guardadas a partir de la última columna de la matriz X. Al elegir la opción de transformación de datos, en la pantalla se desplegará una ventana de ayuda (figura 2.4.1), que se mantendrá hasta que se salga de esta opción.

```
                          CARACTERES ACEPTADOS

      OPERACIONES                                         FUNCIONES

(+) SUMA              (-) RESTA              (S) SENO      (L) LOGARITMO
(*) MULTIPLICACION (/) DIVICION              (C) COSENO    (^) POTENCIA
()  PARENTESIS

    ESCRITURA DE VARIABLES : X(1),X(2),...,X(P) ;        Y=X(0) ;

    EJEMPLOS :  X(1)+155            X(0)+(X(1)*5)          L(X(1))

                                                    (T) PARA TERMINAR
```

Figura 2.4.1. Ventana de ayuda en Transformaciones

Después de que aparece la ventana de ayuda, se despliega el mensaje "Da la Transformación :" , a partir de el cual se puede dar una transformación o dar una "T" para terminar y regresar al Menú Inicial.

Una transformación se define mediante una sintaxis propia del programa, la cual aparece en forma resumida en la ventana de ayuda. Los operadores y funciones se definen de la siguiente forma:

Operador	Signo
suma	+
resta	-
division	/
multiplicacion	*

Función	Signo
exponencial	^
logaritmo natural	L
seno	S
coseno	C

Todos los caracteres alfabéticos deben darse en mayúsculas.
La forma de referirse a la i-ésima variable es de la siguiente forma: X(i), donde i es un subíndice que va de 1 hasta el número total de columnas que se tienen en memoria de la matriz X, i.e. para referirse a la segunda variable explicativa que ya tenemos en memoria, se escribe "X(2)".
Por ejemplo, Supongase quese tiene el modelo con dos variables:

$$Y = \beta_1 X_1 + \beta_2 X_2, \qquad (2.4.1)$$

pero existe una relación en la que,

$$\beta_2 = 3\beta_1, \qquad (2.4.2)$$

por lo que (2.3.2) se convierte en

$$Y = \beta_1(X_1 + 3X_2). \qquad (2.4.3)$$

Si se crea a

$$X_3 = X_1 + 3X_2, \tag{2.4.4}$$

entonces (2.4.3) se expresa como:

$$Y = \beta_1 X_3. \tag{2.4.5}$$

Supongase que en el paquete REGCOM se leyó las variables X_1 y X_2 es decir, la matriz X contiene dos columnas en memoria. Para crear la variable X_3, se da la siguiente instrucción en la opción de transformaciones:

X(1)+(3*X(2)). (2.4.6)

La asignación del resultado de la transformación es asignada automaticamente a la tercera columna de la matriz X en la memoria, ya que ésta es la columna que se encuentra inmediatamente a continuación de la última columna ocupada.

Todas las transformaciones son compiladas por el paquete, si no se encuentra error de sintaxis o de cálculo, aparece un mensaje pidiendo la siguiente transformación, de lo contrario, el programa manda un mensaje de error y pide la transformación nuevamente. Para referirse a la variable Y en las transformaciones se usa X(0), esto es Y=X(0).
Por ejemplo, supongase se quiere crear la variable X_3 de (2.4.4) como:

$$X_3 = Ln(Y).$$

Para crear esta transformación la instrucción que se da es:

L(X(0)),

y la variable así creada ocupara en memoria la tercera columna de la matriz X, suponiendo las condiciones del ejemplo anterior.

En caso de existir error en la ejecucución de las funciones u operadores, el procedimiento manda un mensaje, por ejemplo: "Logaritmo natural de cero, error en la ejecución." Es importante aclarar que la prioridad de los operadores y las funciones, la da el orden en que se den, es decir, la prioridad se da de izquierda a derecha, sin embargo, se puede asignar prioridades mediante el uso de los parentesis (ver 2.4.6).

2.5 VARIABLES CATEGORICAS

En el análisis existe la posibilidad de que una columna de X sea una variable categórica (sección 1.1.8, capítulo I). Como sus valores son cualitativos, es necesario manejarla como variables dicotomas, con un valor "0" , indicando ausencia de la cualidad y "1" indicando presencia de ésta.

Si la variable tiene s categorías entonces, se pueden construír s-1 variables dicotomas. Para esto, el programa pide: el número de columna en memoria de la matriz X donde se encuentra la variable categórica explicativa a procesar, el número (s-1) de variables dicotomas a construír y los s-1 intervalos de estas últimas. Los intervalos son "cerrados" y están definidos por un límite inferior (I) y un límite superior (S), donde I<=S y siempre son enteros. El programa no valida si se interceptan los intervalos, por lo que es necesario asegurar su correcta construcción. Una variable dicotoma tendrá el valor de "uno" si el valor de la variable categórica está dentro de su intervalo, si no lo está se le asignará el valor de "cero". La categoría que no se defina como variable dicotoma será la categoría base o de control. Las nuevas variables dicotomas serán agregadas en memoria a partir de la Oltima columna de la matriz X en el orden en que fueron definidas. Por ejemplo, sí en la matriz original hay 3 variables y una de ellas es categórica con cinco categorías, se añadirán en la memoria las variables X(4), X(5), X(6) y X(7).

Este procedimiento se encuentra en la opción E "Construcción de Var. Dicotomas" del Menú Inicial.

2.6 EDICION DE DATOS

Esta opción se encuentra en el Menú inicial en el inciso D "Edición de Datos" y conduce al Menú de Edición de Datos (figura 2.6.1).

```
                 MENU EDICION DE DATOS

A   ALTA DE UNA OBSERVACION (renglon)
B   ELIMINAR UNA OBSERVACION (renglon)
C   ALTA DE UNA VAR. EXPLICATIVA (columna)
D   ELIMINAR UNA VAR. EXPLICATIVA (columna)
E   CORREGIR UNA OBSERVACION (renglon)

S   SALIDA
```

Figura 2.6.1. Menú de Edición de Datos

A continuación se explican cada una de las opciones de este menú. La opción A "Alta de una observación (renglón)" permite añadir una observación tanto al vector Y como a cada una de las variables explicativas que se tienen en memoria. El primer dato que se pide es el renglón donde se quiere introducir la nueva observación, el renglón puede ser el uno (al principio) o hasta el renglón n+1 (al final). Con esta opción, si los datos de la variable Y y las variables de la matriz X tienen un orden de tiempo o secuencia, se puede agregar una nueva observación sin alterar la secuencia o el orden. Para introducir los datos, se da primero la observación de Y seguida de una coma, luego las observaciones de las variables explicativas que se encuentran en memoria seguidas de una coma cada una de ellas, excepto la última.

La opción B "Baja de una observación (renglón)" permite eliminar una observación de la variable respuesta y eliminar las correspondientes observaciones de cada una de las variables explicativas. Para llevar a cabo la eliminación sólo hay que dar el número del renglón que será eliminado. Si el número de observaciones es menor o igual al número de betas estimadas, no será posible eliminar más observaciónes, ya que sí el número de observaciones es menor al número de betas estimadas, las columnas de la matriz X serán linealmente dependientes y (1.1.1.4) no se cumplirá.

La opción C "Alta de una variable explicativa (columna)" permite añadir una variable explicativa a la matriz X. La nueva variable se agrega después de la última columna ocupada de la matrix X. Los valores de las observaciones de la nueva variable se capturan en la pantalla, dando cada valor y oprimiendo la tecla "enter" o "return".

La opción D "Baja de una variable explicativa (columna)" permite eliminar una variable explicativa de la memoria en la matriz X. El programa pregunta por el número de la columna de X a borrar. Una vez borrada la variable, sí está ocupaba un lugar intermedio en la matriz X, se recorren las variables posteriores una columna a la izquierda, para no dejar espacios libres. La opción no se puede usar si se tiene una sola variable explicativa.

Con la opción E "Corregir una observación", como el nombre lo indica, se corrige una observación tanto de la matriz X como del vector Y. El programa pide el número del renglón a corregir, una vez obtenido éste se mostrarán los valores actuales de la observación y se podrán introducir los nuevos valores. Estos deben darse en la pantalla, escribiendo el valor de la observación Y primero, seguido de una coma, y después los valores de las observaciones de las variables explicativas que se tengan en memoria, cada una de ellas seguida de una coma, excepto la última.

2.7 MENU PRINCIPAL

El Menu Principal contiene los procedimientos que intervienen directamente en la ejecución de una regresión y en el análisis de la misma (figura 2.7.1).

```
                M E N U   P R I N C I P A L

  A   TABLA ANOVA Y ESTIMADORES
  B   SELECCION DE VARIABLES
  C   CARENCIA DE AJUSTE
  D   PRUEBAS DE HIPOTESIS
  E   INTERVALOS DE CONFIANZA
  F   PREDICCION
  G   MULTICOLINEALIDAD
  H   HETEROSCEDASTICIDAD
  I   AUTOCORRELACION
  J   ANALISIS DE RESIDUALES
  K   SELECCION DE SUBCONJUNTOS DE VARIABLES
  L   LISTAR DATOS
  M   SALVAR  ARCHIVOS
  N   RESULTADOS POR IMPRESORA        DESACTIVADO

  O   MENU INICIAL
  S   SALIDA AL SISTEMA OPERATIVO (DOS)
```

Figura 2.7.1. Menú Principal.

Cuando se inicia el Menú Principal solamente se pueden elegir las opciones A, B, O y S. Como primer paso para realizar el análisis de la regresión es necesario saber si existe solución al sistema $X'X\hat{\beta} = X'Y$ (página 3, capítulo I). En tal caso, es posible obtener los estimadores de las betas. Este paso es realizado por la opción

A "Tabla ANOVA y Estimadores", que proporciona los estimadores, la tabla de análisis de Varianza, la r^2, la r^2ajustada y la Matriz de Covarianzas de las betas. Una vez realizada esta etapa se pueden ejecutar las demás opciones. Si no se desea que todas las columnas de la matriz X que tenemos en memoria (que pueden ser columnas leídas, simuladas, transformadas y creadas en edición), sean incluídas como variables explicativas en la regresión, se pueden seleccionar únicamente un subconjunto de ellas por medio de la opción B "Selección de Variables". Esta no tiene efecto si la matriz X sólo tiene una columna de datos.

Al entrar a la opción B se pide el número de variables (columnas), que serán incluídas en la regresion. Sí este número es menor al número de columnas de datos de la matriz X de memoria, se pide entonces que se especifique cuáles serán las columnas que se incluirán en la regresión como las variables explicativas del modelo. Sólo las columnas que se seleccionen son consideradas en el modelo de regresión durante la ejecución de las opciones del Menú Principal, en tanto no se seleccionen otras variables con esta opción o se regrese al Menú Inicial. Por supuesto, si se pide la opción B, una vez que ya se hubiera elegido la opción A, entonces, por el hecho de haber seleccionado otras variables, es necesario volver a correr la opción A "Tabla ANOVA y Estimadores". En seguida se da una explicación de las opciones del Menú Principal.

2.7.1 Carencia de Ajuste

La opción C "Carencia de Ajuste" como su nombre lo indica, realiza la prueba de Carencia de Ajuste, que verifica si las observaciones se ajustan al modelo que se planteo.

Si en los datos hay repeticiones, se aplica la prueba del error puro; de lo contrario, se utiliza la Prueba Arcoiris (sección 1.1.3, capítulo I). Sí esta última es la que se efectúa, entonces se pide al usuario el número de renglones de la matriz D, esto es, el número de datos de que constará la matriz central.

2.7.2 Pruebas de Hipótesis

La opcion D "Pruebas de Hipótesis" del Menú Principal conduce al Menú Pruebas de Hipótesis (figura 2.7.2).

```
                    MENU PRUEBAS DE HIPOTESIS

 A   PRUEBAS DE HIPOTESIS INDIVIDUALES
 B   PRUEBA DE HIPOTESIS CONJUNTA
 C   CONTRIBUCION DE UN SUBCONJUNTO DE VARIABLES
     A LA REGRESION
 D   PRUEBA DE HIPOTESIS PARA COMBINACIONES
     LINEALES DE LOS ESTIMADORES
 E   RESULTADOS POR IMPRESORA            DESACTIVADO

 S   SALIDA
```

Figura 2.7.2 Menú de Pruebas de Hipótesis

A continuación se explican las opciones A, B, C, y D de dicho menú, las opciones E y S se explican en la sección 1 de este este capítulo.

La opción A "Prueba de Hipótesis Individuales", realiza la prueba de hipótesis individual de los coeficientes de la regresión. Al entrar a esta opción se pregunta sobre el estimador al que se le aplicará la prueba. Una vez que se da el dato, el programa pide el valor contra el cual se comparará el estimador, y se calculará el estadístico t y el p-valor, los cuales se presentarán en una tabla junto con el valor de la beta estimada con su error estandar.

La opción B "Prueba de Hipótesis Conjunta" calcula el estadístico para probar la hipótesis de que todos los coeficientes de regresión asociados a las variables explicativas son cero contra la alternativa de que al menos uno no lo es (sección 1.1.4, página 13). Los resultados que se obtienen de esta prueba son el estadístico F y su respectivo valor-p. La opción C "Contribución de un Subconjunto de Variables a la Regresión", calcula el estadístico para la prueba de hipótesis de que el modelo sólo consta de r variables contra la alternativa de que las p variables deben estar en el modelo. Para realizar esta prueba se tendrán que dar los siguientes datos: el número de variables del subconjunto a probar y cuáles son esas variables. Los resultados de esta prueba son resumidos en una tabla de análisis de varianza (sección 1.1.4, página 15).

La opción D "Prueba de Hipótesis para Combinaciones Lineales de los Estimadores" realiza la prueba $C\beta = \gamma$. Donde C es una matriz de q x (p + 1) siendo q =< p y γ es un vector de q x 1. Para realizar esta prueba de hipótesis se tiene que dar el valor de q, la matriz C y el vector gamma transpuesto que es la base de la prueba. Los resultados de la prueba se dan en una tabla de análisis de varianza (sección 1.1.4).

2.7.3 Intervalos de Confianza

La opción E "Intervalos de Confianza" lleva al Menú Intervalos de Confianza (figura 2.7.3.).

```
              MENU INTERVALOS DE CONFIANZA

A    INTERVALOS DE CONFIANZA INDIVIDUALES
B    INTERVALO DE CONFIANZA PARA LA VARIANZA
C    RESULTADOS POR IMPRESORA          DESACTIVADO

O    OPCIONES
S    SALIDA
```

Figura 2.7.3. Menú de Intervalos de Confianza.

A continuación se explican las opciones A, B y O de este menú, las opciones C y S, se explican en la sección 2.2 de este capítulo.

La opción A "Intervalos de Confianza" proporciona el intervalo de confianza para un coeficiente de regresión específico. Para calcular dicho intervalo se pedirá al usuario el subíndice del coeficiente de regresión sobre el que se requiere el intervalo de Confianza.

Con la opción C "Intervalo de Confianza para la Varianza" se calcula el intervalo de confianza para la varianza. Los valores que se tiene que proporcionar son el límite inferior del intervalo de confianza y el límite superior.
La opción O "Opciones" permite cambiar el nivel de confianza para la opción A de este menú. El valor que por omisión se tiene para el nivel de confianza es del 95%.

2.7.4 Predicción

La opcion F "Predicción" del Menu Principal lleva al Menú Predicción (figura 2.7.4.).

```
                    MENU PREDICCION

A   PREDICCION E INTERVALO DE CONFIANZA
    PARA  (E[Y/Xo])
B   PREDICCION E INTERVALO DE CONFIANZA
    PARA  Y/Xo
C   RESULTADOS POR IMPRESORA         DESACTIVADO

O   OPCIONES
S   SALIDA
```

Figura 2.7.4. Menú de Predicción.

Las opciones A, B y O se explican a continuación, las restantes opciones se describen en la sección 2.2 de este capítulo. La opción A "Predicción e Intervalo de Confianza para (E[Y/Xo])" da la predicción y el intervalo de confianza de el valor medio de Y dado un vector Xo. Se deben proporcionar los valores del vector Xo.
La opción B "Predicción e Intervalo de Confianza para Y/Xo" permite obtener la predicción y el intervalo de confianza del valor Y dado un vector Xo. Los datos a proporcionar son el vector Xo.
La opción O "Opciones" permite cambiar el nivel de confianza para las opciónes A y B de este menú. El valor que por omisión se tienen para el nivel de confianza es del 95%.

2.7.5 Multicolinealidad

La opción G "Multicolinealidad" conduce al Menú Multicolinealidad (figura 2.7.5.).

```
                MENU MULTICOLINEALIDAD

 A   MATRIZ DE CORRELACION
 B   REGRESION DE UNA VARIABLE EXPLICATIVA
     CON LAS RESTANTES
 C   RESULTADOS POR IMPRESORA          DESACTIVADO

 S   SALIDA
```

Figura 2.7.5. Menú de Multicolinealidad.

Las opiones C y S se explican en la sección 2.2 de este capítulo, las restantes opciones se explican a continuación. La opción A "Matriz de Correlación" calcula los coeficientes de correlación de todas las variables y los presenta en una matriz triangular superior (sección 1.2.2.3).

La opción B "Regresión entre las X'S" toma unas variable explicativa que el usuario indique, como la variable respuesta Y calcula una regresión de ésta contra las restantes variables (sección 1.2.2.3). El único dato que se pide para que el procedimiento se ejecute es la especificación de la variable explicativa que será la variable dependiente en la regresión. Los resultados son resumidos en una tabla de análisis de varianza.

2.7.6 Heteroscedasticidad

La opción H "Heteroscedasticidad" lleva al Menú Heteroscedasticidad (figura 2.7.6.) en el que se dan los procedimientos que sirven en la detección de la heteroscedasticidad (sección 1.2.3.3).

Las opciones E y S, se explican en la sección 2.2 de este capítulo, las restantes opciones se explican a continuación.

MENU HETEROSCEDASTICIDAD		
A	PRUEBA PARK	
B	PRUEBA GLEJSER	
C	PRUEBA DE RANGO SPEARMAN	
D	GRAFICAS	
E	RESULTADOS POR IMPRESORA	DESACTIVADO
S	SALIDA	

Figura 2.7.6. Menú de Heteroscedasticidad

La opción A "Prueba de Park" realiza la Prueba de Park en la cual se propone a σ_i^2 como función de las variables explicativas y se hace una regresión del $Ln(e_i^2)$ (como aproximación de $Ln(\sigma_i^2)$) con respecto al Ln(X). Los datos que se deben proporcionar son dos: el número de variables explicativas que forman el modelo para la prueba de park y cuáles son esas variables. Los resultados del procedimiento aparecen resumidos en una tabla de análisis de varianza para observar la contribución conjunta de las variables; también se proporcionan los valores de las las betas estimadas, sus varianzas, el estadístico t individual para cada una ellas y su nivel de significancia, con objeto de que se pueda observar la contribución individual de los estimadores (sección 1.2.3.3).

La opción B "Prueba de Glejser" realiza la Prueba de Glejser, la cual, en semejanza a la prueba anterior propone a σ_i^2 como función de las variables explicativas. Se hace una regresión de $|e_i|$ con respecto a las variables explicativas transformadas. Como primer paso para que el programa lleve a cabo la prueba se tiene que dar la variable explicativa que formará parte de la prueba, a continuación se pedirá elegir una de las cuatro transformaciones para X:

i) X
ii) $\sqrt{X}$
iii) $1/X$
iv) $1/\sqrt{X}$.

Donde X puede ser cualquiera de las variables explicativas. Una vez que se seleccionó la transformación, el procedimiento calcula el valor de los estimadores, sus varianzas, el estadístico t y el nivel de significancia (sección 1.2.3.3).

La opción C "Prueba de Rango de Spearman" calcula el estadístico t para la Prueba de Correlacoín de Rango de Spearman (sección 1.2.3.3). El único dato que se proporciona a el programa es el subíndice de la variable contra la cual se quiere verificar la existencia de heteroscedasticidad; el programa generará el valor del estadístico t y su nivel de significancia.

En la opción D "Gráficas" se tiene la oportunidad de graficar a los residuales estimados al cuadrado contra Y, Y estimada o contra las variables explicativas (sección 1.2.3.3).

2.7.7 Autocorrelación

La opción I "Autocorrelación" conduce al Menú de Autocorrelación (figura 2.7.7).

```
                    MENU AUTOCORRELACION

A   GRAFICAS DE LOS RESIDUALES
B   ESTADISTICO DURBIN-WATSON
C   ESTADISTICO VON-NEUMAN
D   ESTIMACION DE ro (t) POR DURBIN-WATSON
E   ESTIMACION DE ro (t) POR DURBIN-WATSON
    APROXIMACION POR THEIL Y NAGAR
F   PRUEBA DE CORRIDAS EN LA SECUENCIA DE
    SIGNOS DE LOS RESIDUALES
G   RESULTADOS POR IMPRESORA          DESACTIVADO

S   SALIDA
```

Figura 2.7.7. Menú de Autocorrelación

Las opciones G y S, se explican en la sección 1 de este capítulo, las restantes opciones se explican a continuación.

Con la opción A "Gráficas de los Residuales" se pueden obtener las gráficas de los residuales estimados contra Y, Y estimada o contra las variables explicativas (sección 1.2.4.3).

La opción B "Estadístico Durbin-Watson" calcula, como su nombre lo indica, el estadístico Durbin-Watson (sección 1.2.4.3).
La opción C "Estadístico Von-Neuman" calcula el estadístico Von-Neuman (sección 1.2.4.3).
Con la opción D "Estimación de ro (ρ) por Durbin-Watson" se estima el coeficiente de correlacion (ρ) suponiendo la existencia de autocorrelación (sección 1.2.4.4).
La opción E "Estimación de ro (ρ) por Durbin-Watson aproximación de Theil y Nagar" es semejante a la anterior opción, lo que varia es que se aplica la aproximación de Theil y Nagar (sección 1.2.4.4).
La opción F "Prueba de Corridas en la Secuencia de Signos de los Residuales" porporciona al usuario el número de corridas y la secuencia de los signos de los residuales, el número de residuales positivos y el número de residuales negativos. Realiza el calculo de la aproximación de la media y de la varianza del número de secuencia de signos para ser comparados con la distribución Normal (sección 1.2.4.3).

```
                MENU ANALISIS DE RESIDUALES

A   GRAFICACION DE RESIDUALES EN PAPEL DE
    PROBABILIDAD NORMAL
B   GRAFICACION DE RESIDUALES ESTUDENTIZADOS
    vs. Y ESTIMADA
C   GRAFICACION DE RESIDUALES ESTUDENTIZADOS
    vs. TIEMPO DE ORDEN
D   GRAFICACION DE RESIDUALES ESTUDENTIZADOS
    vs. CADA REGRESOR
E   GRAFICACION ENTRE REGRESORES
F   GRAFICACION PARCIAL DE RESIDUALES
G   OBSERVACIONES INFLUYENTES (METODO DE COOK)
H   RESULTADOS POR IMPRESORA          DESACTIVADO

O   OPCIONES
S   SALIDA
```

Figura 2.7.8. Menú de Análisis de Residuales

2.7.8 Análisis de Residuales

La opción J "Analisis de Residuales" lleva al Menú de Análisis de Residuales (figura 2.7.8.). Las opciones que contiene este menú pueden dar indicios sobre un comportamiento anormal del modelo, violatorio de los supuestos del modelo de regresión (sección 1.2.5, capítulo I).

Las opciones H y S se explican en la sección 1 de este capítulo, las restantes opciones se explican a continuación.

La opción A "Gráfica de Residuales en Papel de Probabilidad Normal" da la gráfica de los residuales estudentizados sobre una escala de probabilidad normal (sección 1.2.5.2).

La opción B "Gráfica de Residuales Estudentizados versus Y Estimada" muestra la gráfica de los residuales estudentizados contra la Y estimada (sección 1.2.5.2).

La opción C "Gráfica de Residuales Estudentizados versus Tiempo de Orden" grafica los residuales estudentizados contra el orden de los datos (sección 1.2.5.2).

La opción D "Gráfica de Residuales Estudentizados versus cada Regresor" permite graficar los residuales estudentizados contra cualquier variable explicativa (sección 1.2.5.2). El usuario debe dar el subíndice de la variable contra la que se quiere graficar.

Con la opción E "Gráfica entre Regesores" se puede graficar una variable explicativa contra cualquiera de las restantes (sección 1.2.5.2). La información que pide el procedimiento son los sudíndices de las variables explicativas a graficar.

La opción F "Gráfica Parcial de Residuales" da la gráfica parcial de los residuales (sección 1.2.5.2). El dato que se debe dar es el subíndice de la variable explicativa contra la que se va a graficar.

Con la opción G "Observaciones Influyentes" se listan el vector de residuales, el vector de residuales estudentizados y el vector Di de Cook (sección 1.2.5.3), este último indica cuáles residuales pueden ser influyentes.

La opción O "Opciones" permite cambiar el nivel de significancia de la banda de confianza o suprimirla de las gráficas de residuales estudentizados. El valor que por omisión se tiene para la banda de confianza es del 95%.

2.7.9 Selección de Subconjuntos de Variables

La opción K " Selección de Subconjuntos de Variables" selecciona de las variables explicativas del modelo, los mejores subconjuntos tomando como criterio la suma de cuadrados de los residuales que tengan estos subconjuntos en su regresión con Y (sección 1.1.9, capítulo I). Los resultados que se generan son el número de variables del subconjunto, los subíndices de las variables, la Suma de Cuadrados de los Residuales (SCRES) y el estadístico Cp-Mallows (sección 1.1.9.1).

2.7.10 Listado de Datos

Con la opción L "Listar Datos" es posible obtener un listado por pantalla y por impresora de las observaciones que se estan usando en el modelo.

2.7.11 Guardar Vectores o Matrices

Con la opción M "Guardar Vectores o Matrices se tiene la oportunidad de guardar datos que pueden ser de utilidad en el análisis de la regresion. Las datos que se pueden guardar son los siguientes: La matriz X, el vector Y, el vector de Y estimada y el vector de los residuales estimados; El vector de los estimadores y la matriz de covarianzas de los estimadore y por último La matriz de correlaciones y el número de datos.

2.7.12 Menú Inicial

Con la opción O "Menú Inicial" se regresa al Menú Inicial, y la matriz X vuelve al estado original que tenía antes de entrar al Menú Principal, sin importar que columnas se hayan seleccionado con la opción B "Selección de Variables" del Menú Principal.

Las opciones N y S del Menú Principal se explican en la sección 2.2, de este capítulo.

2.8 ESTRUCTURA DEL SISTEMA

El paquete REGCOM está desarrollado en el lenguaje de programacion Pascal y compilado en la version 4.0 de Turbo Pascal.
Los programas del paquete están agrupados en 14 unidades (subrutinas) y un Programa Principal.
La función de cada una de la unidades es el siguiente:

UNIDADES	NOMBRE	FUNCION
UNIDAD 1	GLOBAL	Unidad que es usada por las demás unidades, así como por el programa principal. Define a las constantes, los tipos de datos, los tipos de arreglos y las variables globales usadas en todo el programa.
UNIDAD 2	UTILIDAD	Unidad que contiene a todos los procedimientos y funciones usadas por las demás unidades y por el programa principal.
UNIDAD 3	OBTENER	Unidad que contiene los procedimientos de lectura y obtención de la matriz X y del vector Y.
UNIDAD 4	CARENCIA	Unidad que contiene los procedimientos de las pruebas de carencia de ajuste.
UNIDAD 5	HIPOTESI	Unidad que contiene todos los procedimientos de las pruebas de hipótesis

UNIDAD 6	INTERVAL	Unidad que contiene los procedimientos para la obtención de los diversos intervalos de confianza.
UNIDAD 7	MULTICOL	Unidad que contiene los procedimientos relativos a la multicolinealidad.
UNIDAD 8	HETEROS	Unidad que contiene los procedimientos de prueba de la heteroscedasticidad.
UNIDAD 9	AUTOCORR	Unidad que contiene los procedimientos de las pruebas relativas a la autocorrelación.
UNIDAD 10	RESIDUAL	Unidad que contiene a los procedimientos usados en el análisis de residuales.
UNIDAD 11	PREDIC	Unidad que contiene los procedimientos de predicción.
UNIDAD 12	EDITAR	Unidad que contiene a los procedimientos utilizados en la edición de los datos.
UNIDAD 13	FURNIVAL	Unidad que contiene los procedimientos que seleccionan los subconjuntos de variables variables del modelo.
UNIDAD 14	TRANSFOR	Unidad que contiene el procedimiento que crea una variable a partir de la transformación de otras.

Como podemos observar, los procedimientos están agrupados en la unidades según su función; el programa principal es el encargado de conjuntar a todas las unidades y contiene a los tres principales menús del paquete: el Menú de Obtención de X y Y, el

Menú Inicial (Menú de manejo de los datos) y el Menú Principal (Menú de la regresión lineal).

El uso de la unidades como característica y como opción de Turbo Pascal 2.0 tiene varias razones:

a. Se tiene una buena Organización del programa, al agrupar las subrutinas según su función.

b. Las unidades son entidades independientes del programa principcipal y se pueden compilar separadamente, es decir, si se corrige una unidad no es necesario compilar todas las unidades, sino únicamente la unidad que fue corregida.

c. Los programas tienen la limitación de espacio de 64K de código en el compilador de Turbo Pascal, con la creación de unidades, cada una de ellas tiene un segmento de código separado y la limitación total viene a ser la suma de memoria que la máquina y el sistema operativo soporte.
En la próximas hojas se presentan los diagramas del flujo que siguen los procedimientos del programa para el manejo de los diferentes menús.
Los procedimientos que aparecen con cuadros señalados con líneas punteadas son los procedimientos de tipo auxiliar, es decir, son programas y funciones que son utilizadas por otros procedimientos, todos ellos están incluídos en la unidad 2 "UTILIDAD".
Los procedimientos que aparecen con cuadros de líneas remarcadas son los que contienen los menús principales del programa y se encuentran en el programa principal.

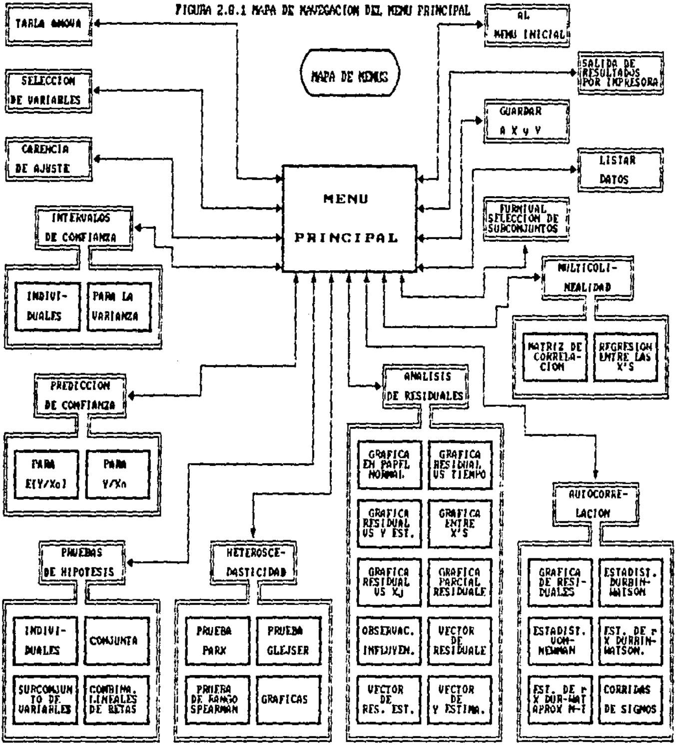


FIGURA 2.8.1 MAPA DE NAVEGACION DEL MENU PRINCIPAL

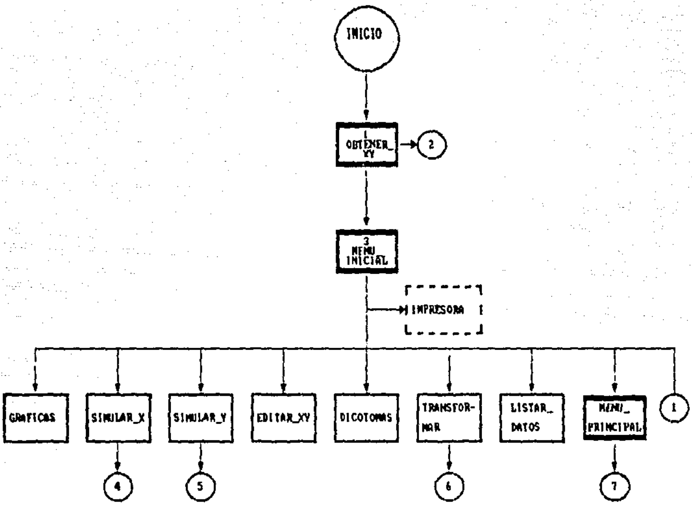


FIGURA 2.8.2. ESTRUCTURA DEL MENU INICIAL

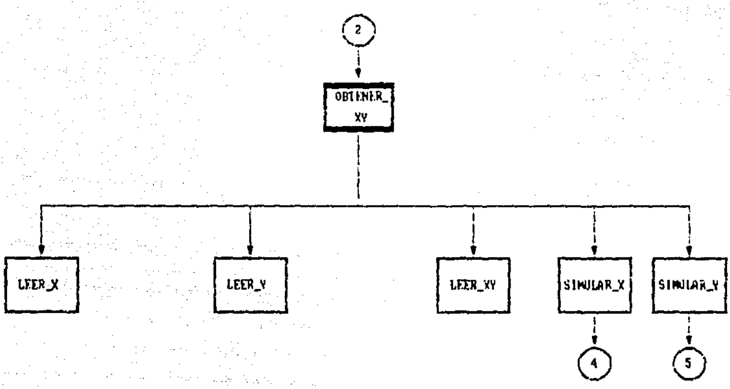


FIGURA 2.8.3. ESTRUCTURA DEL MENU OBTENCION DE DATOS

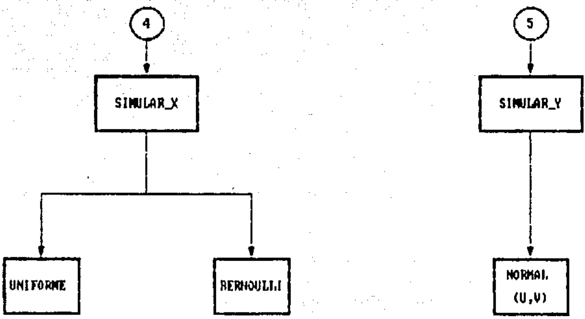


FIGURA 2.6.4. ESTRUCTURA DE LOS MENUS SIMULAR Y y SIMULAR X

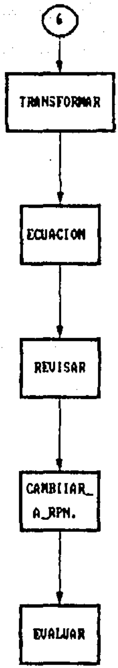


FIGURA 2.8.5. ESTRUCTURA DE TRANSFORMACION DE VARIABLES

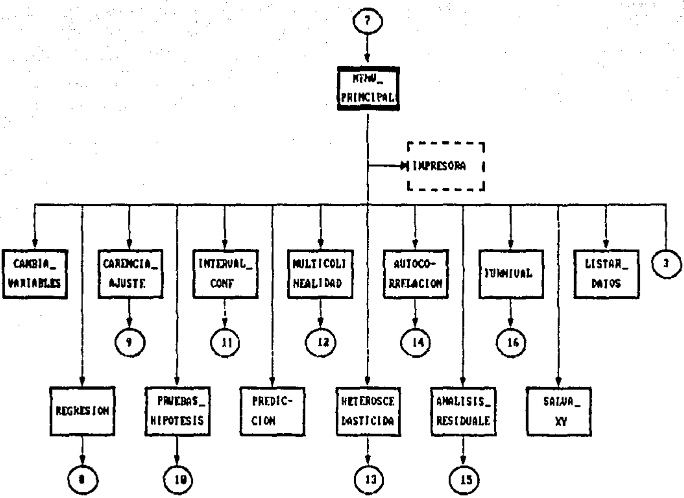


FIGURA 2.8.6. ESTRUCTURA DEL MENU PRINCIPAL

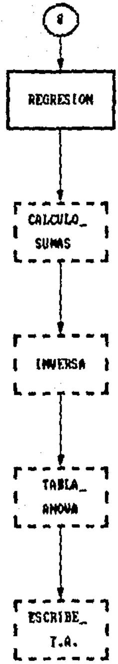


FIGURA 2.8.7. ESTRUCTURA DE LA SOLUCION DE LA REGRESION
Y DEL CALCULO DE LA TABLA ANOVA

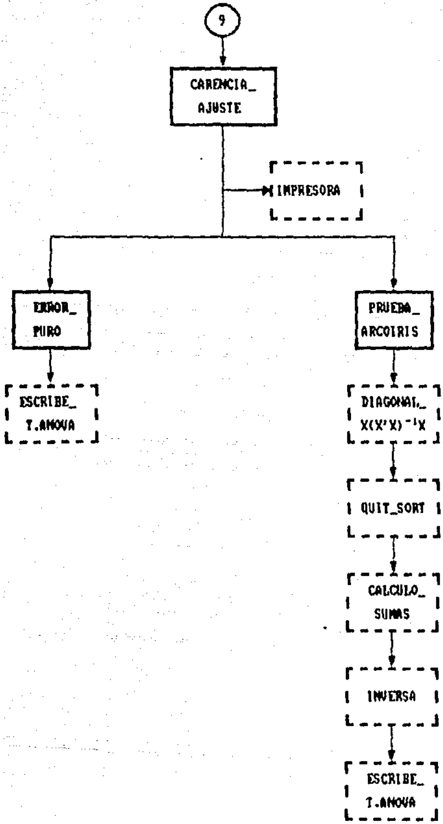


FIGURA 2.8.8. ESTRUCTURA DE LA CARENCIA DE AJUSTE

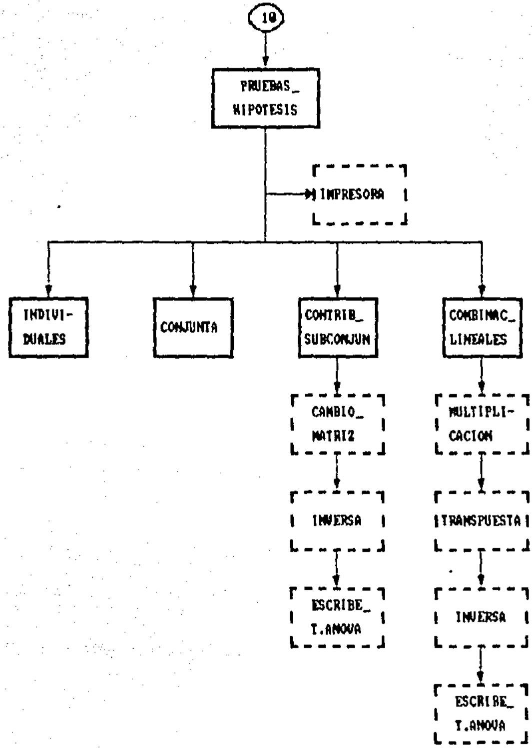


FIGURA 2.8.9. ESTRUCTURA DEL MENU PRUEBAS DE HIPOTESIS

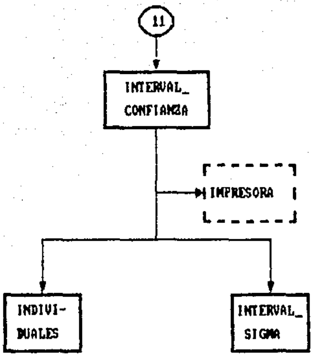


FIGURA 2.8.18. ESTRUCTURA DEL MENU DE INTERVALOS DE CONFIENZA

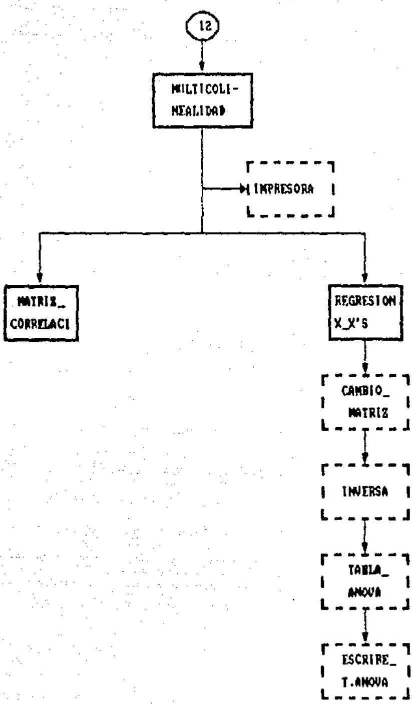


FIGURA 2.8.11. ESTRUCTURA DEL MENU DE MULTICOLINEALIDAD

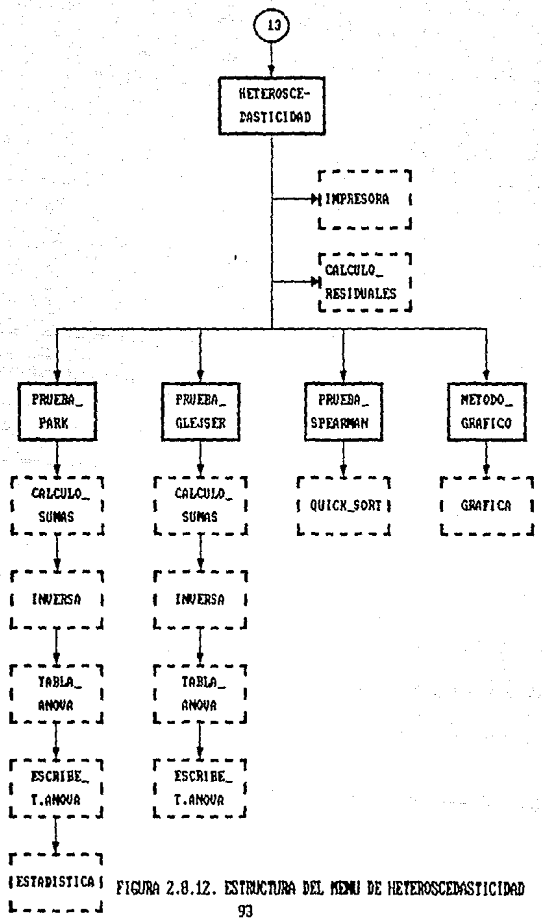


FIGURA 2.8.12. ESTRUCTURA DEL MENU DE HETEROSCEDASTICIDAD

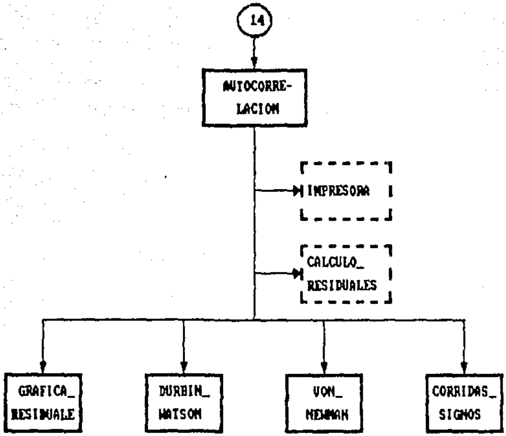


FIGURA 2.8.13. ESTRUCTURA DEL MENU DE AUTOCORRELACION

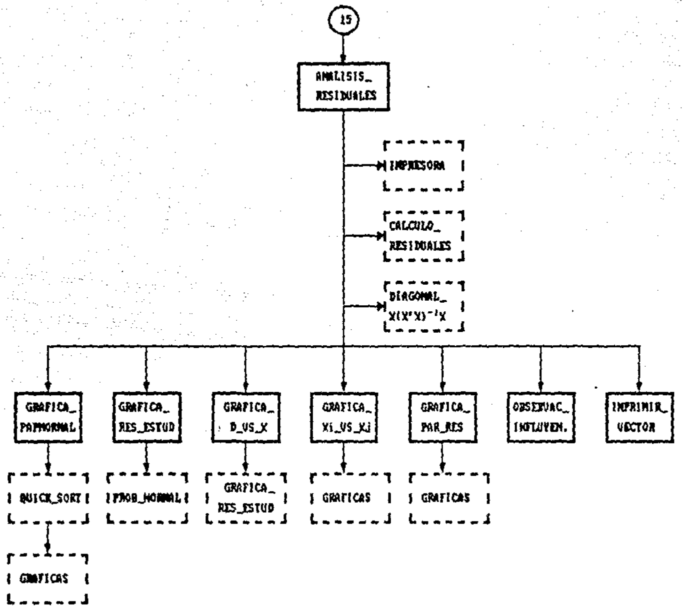


FIGURA 2.8.14. ESTRUCTURA DEL MENU DE ANALISIS DE RESIDUALES

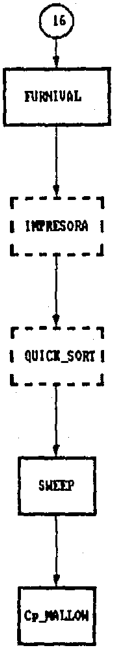


FIGURA 2.8.15. ESTRUCTURA DE FURNIVAL

2.9 LIMITACIONES

El paquete REGCOM está diseñado para microcomputadoras compatibles IBM; su limitación está determinada por la capacidad de la memoria RAM de la máquina y por las características propias del compilador Turbo Pascal versión 4.

Como se explicó en la sección anterior, el programa está dividido en unidades (subrutinas). Cada una de ellas tiene una limitación de 64K de codigo máquina, por lo que la suma de la memoria utilizada está limitada por la memoria de la máquina.

Así, la capacidad que tiene el paquete para el manejo de datos es de un maximo de 275 observaciones conjuntamente con 25 variables explicativas. La única excepción se da en la Selección de Subconjuntos de Variables donde el máximo número de variables es de diez, ya que el procedimiento ocupa demasiada memoria.

C A P I T U L O III

APLICACIONES DEL PAQUETE COMPUTACIONAL DE REGRESION LINEAL MULTIPLE. CASOS PRACTICOS.

3.1 INTRODUCCION

El propósito de este capítulo es ayudar al usuario a utilizar el paquete por medio de tres ejemplos, donde se muestra como se seleccionan las opciónes, cómo pueden introducirse los datos al paquete y la manera en que se porporcionan los resultados.

En los ejemplos se ha utilizado un "*" al principio del renglón para diferenciar los comentarios que no corresponden a los despliegues efectuados por el paquete.

3.2 EJEMPLO 1: Estudio del costo de las becas en el extranjero

Una institución que otorga becas para realizar estudios en el extranjero, está interesada en conocer las variables que influyen en el costo de las becas que autoriza. Pretende además, contar con un modelo matemático que permita estimar el presupuesto necesario,dependiendo de las becas que se otorgen. Se considera que las variables más importantes en relación al costo son: el grado y el tiempo de la beca así como el país donde se realizan los estudios. Se selecciona una muestra aleatoria de 40 de las

becas otorgadas en un año en particular y se estructuró la información para trabajar un modelo de Regresión lineal Múltiple. La información se muestra en el cuadro 3.1.

Cuadro 3.1

Observación	Costo total	Grado	Meses de Beca	Paí s
1.	8100.0 D.A.	1	15	4
2.	3240.0 "	1	6	4
3.	10170.0 "	2	12	4
4.	23650.0 "	1	24	4
5.	31053.0 "	1	24	4
6.	8345.0 "	2	12	4
7.	17515.0 "	1	16	4
8.	50941.0 "	1	36	1
9.	59736.0 "	1	36	1
10.	54350.8 "	1	28	1
11.	29238.8 "	1	17	1
12.	13465.0 "	1	10	1
13.	37025.0 "	1	24	1
14.	26731.0 "	1	16	1
15.	24921.0 "	1	24	1
16.	40246.0 "	1	24	1
17.	23527.0 "	1	21	1
18.	6174.0 "	3	6	1
19.	32021.0 "	2	24	1
20.	32067.0 "	2	24	1
21.	18298.0 "	2	12	1
22.	14904.0 "	1	18	3
23.	10736.0 "	1	12	3
24.	11564.0 "	1	13	3
25.	10736.0 "	1	12	3
26.	15195.0 "	1	12	3
27.	14874.0 "	1	12	3

Observación	Costo total		Grado	Meses de Beca	País
28.	14322.0	"	1	12	3
29.	10736.0	"	3	12	3
30.	10736.0	"	3	12	3
31.	14874.0	"	2	12	3
32.	21930.0	"	1	12	2
33.	862.0	"	1	18	2
34.	33561.0	"	1	24	2
35.	36314.0	"	1	24	2
36.	22434.0	"	1	18	2
37.	4140.0	"	1	18	2
38.	43528.0	"	1	18	2
39.	18120.0	"	2	12	2
40.	19345.0	"	2	10	2

Y = Costo total en dólares americanos (DA)

X_1 = Grado

1 = Doctorado
2 = Maestría
3 = Especialidad Academica.

X_2 = Tiempo de las becas (en meses)

X_3 = País

1 = E.U.A.
2 = Gran Bretaña
3 = Francia
4 = Otros.

* Inicialización del programa desde el sistema operativo MS-DOS.
C:>REGCOM

```
P A Q U E T E   D E

R E G R E S I O N   L I N E A L   M U L T I P L E
```

```
MENU OBTENCION DE DATOS

A > NUMERO DE DATOS DE LA REGRESION (n)
B > REGRESION CON O SIN TERMINO INDEPENDIENTE
C > LEER X
D > LEER Y
E > LEER Y y X
F > SIMULAR X
G > SIMULAR Y
H > CONTINUAR

S > SALIDA AL SISTEMA OPERATIVO (DOS)
```

» ESCOGE TU OPCION : A

* El paquete inicia con el menú Obtención de Datos, donde se podrá observar
* que las opciones A y S tienen mayor intencidad, por lo que sólo éstas se
* pueden escoger. Así, se escoge la opción A.

» CUANTOS DATOS SON: 40

* Ahora también se podran elegir las opciones B,C,D, y F. La opción B tiene
* como valor de inicio a la regresión con término independiente, por lo que
* así se dejará. Como el vector Y y la matriz X se encuentran en un solo
* archivo se escoge la opción E.

» ESCOGE TU OPCION : e

```
MENU DE LECTURA DE X y Y

A > POR PANTALLA
B > ARCHIVO DE TEXTO (SDF)

S > SALIDA
```

» OPCION : b

» CUANTAS COLUMNAS CONTIENE LA MATRIZ X : 3
» INTRODUCE EL NOMBRE DEL ARCHIVO: becas.txt

» FIN DE ARCHIVO, RENGLONES LEIDOS: 40

* PARA CONTINUAR OPRIMA ENTER (<--┘)

PARA SALIR INTRODUCIR [0]

* Las tres columnas leídas de la matriz X corresponden respectivamente al
* Grado, Meses Becas y País. La variable dependiente es el costo total de la
* beca. Una vez que se haya tecleado el (Enter), se regresa al menú Obtención
* de Datos, donde las opciones G y H ya se pueden escoger. Así, se elige la
* opción H, que conduce al Menú Inicial.

* ESCOGE TU OPCION : h

```
MENU INCIAL

A ) GRAFICAS
B ) SIMULAR X
C ) SIMULAR Y
D ) EDICION DE DATOS
E ) CONSTRUCCION DE VAR. DICOTOMAS
F ) TRANSFORMACIONES
G ) CAMBIO DE VARIABLE DEPENDIENTE
H ) LISTADO DE DATOS
I ) SALIDA A LA IMPRESORA          DESACTIVADO

N ) NUEVA REGRESION
P ) MENU PRINCIPAL
S ) SALIDA AL SISTEMA OPERATIVO (DOS)
```

* OPCION : a

* Siempre una gráfica nos puede dar una idea del comportamiento de nues-
* tros datos, para esto escogemos la opción A:

```
X(1)  -->  [1]  vs.  Y OBSERVADA
X(2)  -->  [2]  vs.  Y OBSERVADA
X(3)  -->  [3]  vs.  Y OBSERVADA
```

* CUAL GRAFICAS QUIERES ? 2

PARA SALIR INTRODUCIR [0]

* Debido a que las columnas X(1) y X(3) contiene a variables
* categóricas sólo se graficará a X(2).

```
GRAFICA DE:  X(2) vs. Y OBSERVADA
          ------------------------------------------------------------------
59736.00+                                                               *I
        I                                                                I
53194.44+                                                   *            I
        I                                                               *I
46652.89+                                                                I
        I                    *                                           I
40111.33+                                        *                       I
        I                                        *                       I
33569.78+                                        *                       I
        I                  *                     *                       I
27028.22+                *                                               I
        I                    *      *            *                       I
20486.67+      *   *                                                     I
        I          *         *                                           I
13945.11+      *   *         *                                           I
        I          * *                                                   I
 7403.56+*         *      *                                              I
        I*                   *                                           I
  862.00+                    *                                           I
         -+--------+---------+---------+---------+---------+---------+--
              8.73     13.27     17.82     22.36     26.91     31.45     36.00
         6.00       * PARA CONTINUAR OPRIMA ENTER ( <___| )
                    * QUIERES OTRA GRAFICA ? (S/N) n
```

* Al no desear otra gráfica, se regresará al Menú Inicial donde se
* seleccionará la opción E para construir de las variables X(1) y X(3) sus
* respectivas variables dicotomas.

```
    | E ) CONSTRUCCION DE VAR. DICOTOMAS                    |

    * OPCION : e


* QUIERES CONSTRUIR UNA(S) VARIABLE(S) DICOTOMICA(S), A PARTIR DE UNA
* VARIABLE CATEGORICA QUE NO ESTA CON VALORES CEROS Y UNOS (S/N) ? : s

* DAME EL No. DE LA VARIABLE CATEGORICA A CONSTRUIR :1

* CUANTAS VARIABLES DICOTOMAS NECESITAS CONSTRUIR: 2
```

* Iniciamos con X(1), y como ésta tiene 3 categorias entonces, el número de
* dicotomas sera 2.

```
* AL DARME LOS INTERVALOS DE LAS CATEGORIAS
* DEBES CONSIDERAR QUE LOS LIMITES SON INCLUYENTES

* DAME EL INTERVALO QUE DEFINA A LA 1a. VAR. DICOTOMA
* LIMITE INFERIOR: 1
* LIMITE SUPERIOR: 1
* DAME EL INTERVALO QUE DEFINA A LA 2a. VAR. DICOTOMA
* LIMITE INFERIOR: 2
```

* LIMITE SUPERIOR: 2

* LAS VARIABLES DICOTOMAS FUERON INCLUIDAS EN LAS 2 ULTIMAS
* COLUMNAS DE X, EN EL ORDEN EN QUE FUERON DEFINIDAS

El grupo base es especialidad Académica, por lo que sólo se definirá al
Doctorado y la Maestría. El valor de los y límites [1,1] para el Doctorado
es el valor como esta definida en la variable X(1), es decir, 1 si es
doctorado. De igual forma para la maestría, donde en X(1) el 2 equivale a
la maestría. La nueva variable que representa al grado de Doctor es X(4), y
X(5) para la Maestría.

* DESEAS CONSTRUIR MAS VAR. DICOTOMAS
* A PARTIR DE UNA VAR. CATEGORICA (S/N) ? : s
* DAME EL No. DE LA VARIABLE CATEGORICA A CONSTRUIR :3

* CUANTAS VARIABLES DICOTOMAS NECESITAS CONSTRUIR :3

Se construirán 3 variables dicotomas para X(3), debido a que ésta tiene
cuatro categorias.

* AL DARME LOS INTERVALOS DE LAS CATEGORIAS
* DEBES CONSIDERAR QUE LOS LIMITES SON INCLUYENTES

* DAME EL INTERVALO QUE DEFINA A LA 1a. VAR. DICOTOMA
* LIMITE INFERIOR: 1
* LIMITE SUPERIOR: 1
* DAME EL INTERVALO QUE DEFINA A LA 2a. VAR. DICOTOMA
* LIMITE INFERIOR: 2
* LIMITE SUPERIOR: 2
* DAME EL INTERVALO QUE DEFINA A LA 3a. VAR. DICOTOMA
* LIMITE INFERIOR: 3
* LIMITE SUPERIOR: 3

* LAS VARIABLES DICOTOMAS FUERON INCLUIDAS EN LAS 3 ULTIMAS
* COLUMNAS DE X, EN EL ORDEN EN QUE FUERON DEFINIDAS

En este caso el grupo base es Otros Paises.
Ahora la variable EUA es X(6), Gran Bretaña es X(7) y X(8) es Francia.

* DESEAS CONSTRUIR MAS VAR. DICOTOMAS
* A PARTIR DE UNA VAR. CATEGORICA (S/N) ? : n

Al no tener más variables categoricas, se regresará al Menú Inicial, donde
se pedirá el listado de los datos.

| H) LISTADO DE DATOS |

* OPCION : h

	Y	X(1)	X(2)	X(3)	X(4)
1	8100.0000	1.0000	15.0000	4.0000	1.0000
2	3240.0000	1.0000	6.0000	4.0000	1.0000
3	10170.0000	2.0000	12.0000	4.0000	0.0000

4	23450.0000	1.0000	24.0000	4.0000	1.0000
5	31053.0000	1.0000	24.0000	4.0000	1.0000
6	8345.0000	2.0000	12.0000	4.0000	0.0000
7	17515.0000	1.0000	16.0000	4.0000	1.0000
8	50941.0000	1.0000	36.0000	1.0000	1.0000
9	59736.0000	1.0000	36.0000	1.0000	1.0000
10	54350.8000	1.0000	28.0000	1.0000	1.0000
11	29238.6000	1.0000	17.0000	1.0000	1.0000
12	13465.0000	1.0000	10.0000	1.0000	1.0000
13	37025.0000	1.0000	24.0000	1.0000	1.0000
14	26731.0000	1.0000	16.0000	1.0000	1.0000
15	24921.0000	1.0000	24.0000	1.0000	1.0000
16	40246.0000	1.0000	24.0000	1.0000	1.0000
17	23527.0000	1.0000	21.0000	1.0000	1.0000
18	6174.0000	3.0000	6.0000	1.0000	0.0000
19	32021.0000	2.0000	24.0000	1.0000	0.0000
20	32067.0000	2.0000	24.0000	1.0000	0.0000
21	18298.0000	2.0000	12.0000	1.0000	0.0000
22	14904.0000	1.0000	18.0000	3.0000	1.0000

* PARA CONTINUAR OPRIMA ENTER (<—┘)

	Y	X(1)	X(2)	X(3)	X(4)
23	10736.0000	1.0000	12.0000	3.0000	1.0000
24	11564.0000	1.0000	13.0000	3.0000	1.0000
25	10736.0000	1.0000	12.0000	3.0000	1.0000
26	15135.0000	1.0000	12.0000	3.0000	1.0000
27	14874.0000	1.0000	12.0000	3.0000	1.0000
28	14322.0000	1.0000	12.0000	3.0000	1.0000
29	10736.0000	3.0000	12.0000	3.0000	0.0000
30	10736.0000	3.0000	12.0000	3.0000	0.0000
31	14874.0000	2.0000	12.0000	3.0000	0.0000
32	21930.0000	1.0000	12.0000	2.0000	1.0000
33	862.0000	1.0000	18.0000	2.0000	1.0000
34	33561.0000	1.0000	24.0000	2.0000	1.0000
35	36314.0000	1.0000	24.0000	2.0000	1.0000
36	22434.0000	1.0000	18.0000	2.0000	1.0000
37	4140.0000	1.0000	18.0000	2.0000	1.0000
38	43528.0000	1.0000	18.0000	2.0000	1.0000
39	18120.0000	2.0000	12.0000	2.0000	0.0000
40	19345.0000	2.0000	10.0000	2.0000	0.0000

* PARA CONTINUAR OPRIMA ENTER (<—┘)

	X(5)	X(6)	X(7)	X(8)
1	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000
3	1.0000	0.0000	0.0000	0.0000
4	0.0000	0.0000	0.0000	0.0000
5	0.0000	0.0000	0.0000	0.0000
6	1.0000	0.0000	0.0000	0.0000
7	0.0000	0.0000	0.0000	0.0000
8	0.0000	1.0000	0.0000	0.0000
9	0.0000	1.0000	0.0000	0.0000
10	0.0000	1.0000	0.0000	0.0000
11	0.0000	1.0000	0.0000	0.0000
12	0.0000	1.0000	0.0000	0.0000
13	0.0000	1.0000	0.0000	0.0000

14	0.0000	1.0000	0.0000	0.0000
15	0.0000	1.0000	0.0000	0.0000
16	0.0000	1.0000	0.0000	0.0000
17	0.0000	1.0000	0.0000	0.0000
18	0.0000	1.0000	0.0000	0.0000
19	1.0000	1.0000	0.0000	0.0000
20	1.0000	1.0000	0.0000	0.0000
21	1.0000	1.0000	0.0000	0.0000
22	0.0000	0.0000	0.0000	1.0000

* PARA CONTINUAR OPRIMA ENTER (<—┘)

	X(5)	X(6)	X(7)	X(8)
23	0.0000	0.0000	0.0000	1.0000
24	0.0000	0.0000	0.0000	1.0000
25	0.0000	0.0000	0.0000	1.0000
26	0.0000	0.0000	0.0000	1.0000
27	0.0000	0.0000	0.0000	1.0000
28	0.0000	0.0000	0.0000	1.0000
29	0.0000	0.0000	0.0000	1.0000
30	0.0000	0.0000	0.0000	1.0000
31	1.0000	0.0000	0.0000	1.0000
32	0.0000	0.0000	1.0000	0.0000
33	0.0000	0.0000	1.0000	0.0000
34	0.0000	0.0000	1.0000	0.0000
35	0.0000	0.0000	1.0000	0.0000
36	0.0000	0.0000	1.0000	0.0000
37	0.0000	0.0000	1.0000	0.0000
38	0.0000	0.0000	1.0000	0.0000
39	1.0000	0.0000	1.0000	0.0000
40	1.0000	0.0000	1.0000	0.0000

* PARA CONTINUAR OPRIMA ENTER (<—┘)

* Como se pude observar, cuando X(1) vale 1, X(4) (variable del Doctorado)
* tiene el valor de 1.
* Cuando X(1) vale 2, X(5) (variable de la Maestria) vale 1, pero si X(1)
* vale 3 entoces, X(4) y X(5) tendrán el valor de 0.
* Lo mismos ocurre con X(2), respecto a X(7), X(8) y X(9). Ahora, con las
* nuevas variables se podrá correr una regresión con la opción Menú Principal

| P) MENU PRINCIPAL |

* OPCION : p

M E N U   P R I N C I P A L

A) TABLA ANOVA Y ESTIMADORES
B) SELECCION DE VARIABLES
C) CARENCIA DE AJUSTE
D) PRUEBAS DE HIPOTESIS
E) INTERVALOS DE CONFIANZA
F) PREDICCION
G) MULTICOLINEALIDAD
H) HETEROSCEDASTICIDAD
I) AUTOCORRELACION

```
| J ) ANALISIS DE RESIDUALES                                         |
| K ) SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES             |
| L ) LISTAR DATOS                                                   |
| M ) GUARDAR INFORMACION EN UN ARCHIVO                              |
| N ) RESULTADOS POR IMPRESORA          DESACTIVADO                  |
|                                                                    |
| O ) MENU INICIAL                                                   |
| S ) SALIDA AL SISTEMA OPERATIVO (DOS)                              |
 ====================================================================

  * ESCOGE TU OPCION : b
```

* Al entrar al Menú Principal sólo las opciones A, B, O y S pueden ser
* seleccionadas. Para que las opciones de la C a la N se puedan elegir,
* primero hay que correr la opción A, es decir, obtener los estimadores de
* las betas y saber si el sistema tiene solución o no.
* Se ha escogido las opción B para seleccionar únicamente las columnas de X
* que entrarán a la regresión.

```
* CUANTAS DE LAS 8 COLUMNAS DE LA MATRIZ X SERAN
* INCLUIDAS EN LA REGRESION ? 6

* DAME LA  1a. COLUMNA: 2
```
* Variable tiempo.
```
* DAME LA  2a. COLUMNA: 4
* DAME LA  3a. COLUMNA: 5
```
* Variables dicotomas de grado.
```
* DAME LA  4a. COLUMNA: 6
* DAME LA  5a. COLUMNA: 7
* DAME LA  6a. COLUMNA: 8
```
* Variables dicotomas de pais

* Ahora sólo se consideran 6 variables donde:
* X(1) es meses de beca,
* X(2) es doctorado,
* X(3) es maestría,
* X(4) es E.U.A.,
* X(5) es Gran Bretaña y
* X(6) es Francia. Ahora se obtienen los estimadores.

```
| A ) TABLA ANOVA Y ESTIMADORES                                      |

  * OPCION : a

                         BETA ESTIMADA

                   B[ 0]=-10436.8587140
                   B[ 1]=  1465.5400259
                   B[ 2]=  1919.8781553
                   B[ 3]=  2893.8787010
                   B[ 4]=  8834.5797521
                   B[ 5]=  5471.7399950
                   B[ 6]=  3052.8978069
```

* PARA CONTINUAR OPRIMA ENTER (<┘)

TABLA DE ANALISIS DE VARIANZA

TABLA ANOVA

FUENTE DE VARIACION	g.l.	SUMA DE CUADRADOS	MEDIA DE CUADRADOS	F	VALOR-P
REGRESION	6	5899256642.000000	983209440.340000	15.960790	0.000000
RESIDUALES	33	2032851174.500000	61601550.742000		
TOTAL	39	7932107816.500000			

R Cuadrada (Coeficiente de determinacion) = 0.7437187
R Cuadrada Ajustada = 0.6971221

* PARA CONTINUAR OPRIMA ENTER (<┘)

MATRIZ DE COVARIANZAS

	X(0)	X(1)	X(2)	X(3)	X(4)
X(0)	37623004.7510	-458565.8945	-20749549.9160	-24012935.2180	-7763915.8251
X(1)		49766.8399	-352986.538	2 -145275.7995	-302977.8824
X(2)			27331851.0460	24076466.4220	3597668.4016
X(3)				31771583.2800	3141053.6023
X(4)					15242560.5670
X(5)					
X(6)					

* PARA CONTINUAR OPRIMA ENTER (<┘)

	X(5)	X(6)
X(0)	-6301343.9059	-14872985.4280
X(1)	-61262.5479	107865.1842
X(2)	352192.4395	3819513.3763
X(3)	712256.0629	5827237.6478
X(4)	9218672.2549	8569261.6897
X(5)	15762024.3410	8761617.9504
X(6)		16406869.5340

* PARA CONTINUAR OPRIMA ENTER (<┘)

* Los valores absolutos de Y son demasiados grandes respecto a los valores
* de las X, para esto se puede trabajar a Y en miles de D.A. utilizando las
* Tranformaciones del Menú Inicial.

| 0) MENU INICIAL |

* ESCOGE TU OPCION : 0

```
* Al regresar al Menú Inicial se consideran nuevamente a las 8 columnas de X
* en el orden original.
* Ahora se corre Tranformaciones del Menú Inicial.

        | F ) TRANSFORMACIONES                                  |

        * OPCION : f

                            CARACTERES ACEPTADOS

             OPERACIONES                              FUNCIONES

   (+) SUMA              (-) RESTA           (S) SENO           (L) LOGARITMO
   (*) MULTIPLICACION    (/) DIVICION        (C) COSENO         (^) POTENCIA
    () PARENTESIS

 ESCRITURA DE VARIABLES : X(1),X(2),..,X(P) ;      Y=X(0) ;

 EJEMPLOS :  X(1)+155          ^X(0)+(X(1)*5)          L(X(1))

                                                     (T) PARA TERMINAR
--------------------------------------------------------------------------
* Cada vez que se accesa a esta opción, se muestra la ventana de ayuda
* anterior.

* DA LA TRANSFORMACION :
X(0)/1000

* ESTA TRANSFORMACION SE ENCUENTRA EN LA MATRIZ X EN LA COLUMNA  9

* PARA CONTINUAR OPRIMA ENTER ( <--┘ )
T

* Después del <enter> para continuar, se introdujo una <T> para regresar al
* Menú Inicial, y de éste se elige la opción G para colocar a X(9) como
* variable dependiente.

        | G ) CAMBIO DE VARIABLE DEPENDIENTE                    |

        * OPCION : g
        * SE DESEA CAMBIAR LA VARIABLE DEPENDIENTE ORIGINAL ? (S/N)
        * EN QUE COLUMNA DE LA MATRIZ X SE ENCUENTRA
        * LA NUEVA VARIABLE DEPENDIENTE : 9

* La variable X(9) y Y han intercambiado sus valores, es decir en Y tenemos
* a X(9) y en esta columna tenemos a Y. Ahora se pasa al Menú Principal otra
* vez.

        | P ) MENU PRINCIPAL                                    |

        * OPCION : p
```

* Nuevamente se tiene que selecionar las variables que entran en
* la regresión y pedir la tabla anova y los estimadores.

| B) SELECCION DE VARIABLES |

* ESCOGE TU OPCION : b

* CUANTAS DE LAS 9 COLUMNAS DE LA MATRIZ X SERAN
* INCLUIDAS EN LA REGRESION ? 6

* DAME LA 1a. COLUMNA: 2
* DAME LA 2a. COLUMNA: 4
* DAME LA 3a. COLUMNA: 5
* DAME LA 4a. COLUMNA: 6
* DAME LA 5a. COLUMNA: 7
* DAME LA 6a. COLUMNA: 8

| A) TABLA ANOVA Y ESTIMADORES |

* ESCOGE TU OPCION : a

BETA ESTIMADA

B[0]=	-10.4366587
B[1]=	1.4655400
B[2]=	1.9198782
B[3]=	2.8938787
B[4]=	8.8845795
B[5]=	5.4717400
B[6]=	3.0528979

* PARA CONTINUAR OPRIMA ENTER (<—┘)

TABLA DE ANALISIS DE VARIANZA

TABLA ANOVA

FUENTE DE VARIACION	g.l.	SUMA DE CUADRADOS	MEDIA DE CUADRADOS	F	VALOR-P
REGRESION	6	5899.256642	983.209440	15.960790	0.000000
RESIDUALES	33	2032.851175	61.601551		
TOTAL	39	7932.107817			

R Cuadrada (Coeficiente de determinacion)= 0.7437197
R Cuadrada Ajustada = 0.6971221

* PARA CONTINUAR OPRIMA ENTER (<—┘)

* Como se puede observar los nuevos estimadores son los mismos que los
* anteriores divididos entre 1000 y los valores de la F y de las R's son
* exactamente los mismos. Comenzaremos con la prueba de carencia de ajuste.

MATRIZ DE COVARIANZAS

	X(0)	X(1)	X(2)	X(3)	X(4)
X(0)	37.6230	-0.4586	-20.7495	-24.0129	-7.7639
X(1)		0.0488	-0.3630	-0.1453	-0.3030
X(2)			27.3319	24.0765	3.8977
X(3)				31.7716	3.1411
X(4)					15.2426
X(5)					
X(6)					

* PARA CONTINUAR OPRIMA ENTER (<—┘)

	X(5)	X(6)
X(0)	-8.3013	-14.8730
X(1)	-0.0613	0.1079
X(2)	0.3522	3.8195
X(3)	0.7123	5.8272
X(4)	9.2187	8.5693
X(5)	15.7620	8.7615
X(6)		16.4069

* PARA CONTINUAR OPRIMA ENTER (<—┘)

C) CARENCIA DE AJUSTE

* ESCOGE TU OPCION : c

CARENCIA DE AJUSTE

TABLA ANOVA

FUENTE DE VARIACION	g.l.	SUMA DE CUADRADOS	MEDIA DE CUADRADOS	F	VALOR-P
CARENCIA DE AJUSTE	18	653.915314	36.328684	0.395182	0.992813
ERROR PURO	15	1378.934861	91.928991		
RESIDUAL	33	2032.851175			

PARA CONTINUAR OPRIMA ENTER (<--┘)

Como los datos presentan valores repetidos en los valores de las X's, la
Prueba de Carencia de ajuste aplicada es la de Error Puro. Si no existieran
estas repeticiones se aplicaria la Prueba Arcoiris.
En seguida regresa al Menú Principal, donde se selecionarán las Pruebas de
Hipótesis.

| D) PRUEBAS DE HIPOTESIS |

ESCOGE TU OPCION : d

MENU PRUEBAS DE HIPOTESIS

A) PRUEBAS DE HIPOTESIS INDIVIDUALES
B) PRUEBA DE HIPOTESIS CONJUNTA
C) CONTRIBUCION DE UN SUBCONJUNTO DE VARIABLES
    A LA REGRESION
D) PRUEBA DE HIPOTESIS PARA COMBINACIONES
    LINEALES DE LOS ESTIMADORES
E) RESULTADOS POR IMPRESORA   DESACTIVADO

S) SALIDA

ESCOGE TU OPCION : a

HIPOTESIS NULA	HIPOTESIS ALTERNATIVAS
Ho: ß[i] = K	H1: ß[i] <> K,
Ho: ß[i] = K	H1: ß[i] > K
Ho: ß[i] = K	o H1: ß[i] < K

SOBRE CUAL DE LAS B[i] (i= 0,...,P) QUIERES HACER LA PRUEBA [1]
DAME EL VALOR DE "K" PARA LA ß[1] =0
QUIERES HACER OTRA PRUEBA S/N...? s
SOBRE CUAL DE LAS B[i] (i= 0,...,P) QUIERES HACER LA PRUEBA [1]
DAME EL VALOR DE "K" PARA LA ß[1] =1
QUIERES HACER OTRA PRUEBA S/N...? n

	BETA ESTIMADA	ERROR ESTANDARD DE BETA ESTIMADA	t-CALCULADA	VALOR-P
ß[1]=	1.4655400	0.2208322	6.6364431	0.000
ß[1]=	1.4655400	0.2208322	2.1081170	0.021

En el primer caso se muestra la prueba de hipótesis individual de que ß[1]
es cero contra de que no lo es. Y en el segundo caso se compara ß[1] contra
1 donde, se observa que el Valor-P de es 0.021 .

‡ Al no desear otra prueba, regresará al Menú de Pruebas de Hipótesis donde
‡ se pide la prueba de hipótesis Conjunta.

```
        » ESCOGE TU OPCION : b

                 PRUEBA DE HIPOTESIS CONJUNTA

                  F-CALCULADA          VALOR-P
                  15.9607904070         0.000

» PARA CONTINUAR OPRIMA ENTER ( <——┘ )
```

‡ Para regresar al Menú Principal se elige S y de éste se obtendrán los
‡ Intervalos de confianza.

```
        » ESCOGE TU OPCION : s

        » ESCOGE TU OPCION : e
```

```
 ┌─────────────────────────────────────────────────┐
 │           MENU INTERVALOS DE CONFIANZA          │
 │                                                 │
 │ A ) INTERVALOS DE CONFIANZA INDIVIDUALES        │
 │ B ) INTERVALO DE CONFIANZA PARA LA VARIANZA     │
 │ C ) RESULTADOS POR IMPRESORA       DESACTIVADO  │
 │                                                 │
 │ O ) OPCIONES                                    │
 │ S ) SALIDA                                      │
 └─────────────────────────────────────────────────┘
```

```
        » ESCOGE TU OPCION : a
```

‡ La opción "O" permite cambiar la confiabilidad, de inicio es 0.95%.

```
» CUAL DE LAS B[i] (i= 0,...,P) QUIERES EL INTERVALO DE CONFIANZA [0 ]
» QUIERES HACER OTRA PRUEBA S/N...? S
» CUAL DE LAS B[i] (i= 0,...,P) QUIERES EL INTERVALO DE CONFIANZA [1 ]
» QUIERES HACER OTRA PRUEBA S/N...? S
» CUAL DE LAS B[i] (i= 0,...,P) QUIERES EL INTERVALO DE CONFIANZA [2 ]
» QUIERES HACER OTRA PRUEBA S/N...? S
» CUAL DE LAS B[i] (i= 0,...,P) QUIERES EL INTERVALO DE CONFIANZA [3 ]
» QUIERES HACER OTRA PRUEBA S/N...? S
» CUAL DE LAS B[i] (i= 0,...,P) QUIERES EL INTERVALO DE CONFIANZA [4 ]
» QUIERES HACER OTRA PRUEBA S/N...? S
» CUAL DE LAS B[i] (i= 0,...,P) QUIERES EL INTERVALO DE CONFIANZA [5 ]
» QUIERES HACER OTRA PRUEBA S/N...? S
» CUAL DE LAS B[i] (i= 0,...,P) QUIERES EL INTERVALO DE CONFIANZA [6 ]
» QUIERES HACER OTRA PRUEBA S/N...? N
```

INTERVALO DE CONFIANZA INDIVIDUAL

BETA ESTIMADA	ERROR ESTANDAR DE BETA ESTIMADA	t-CALCULADA	Int. de Confianza
β[0]= -10.4368587	6.1337594	-1.7015435	(-22.91899 , 2.04528)
β[1]= 1.4655400	0.2208322	6.6364431	(1.01615 , 1.91493)
β[2]= 1.9198782	5.2279902	0.3672306	(-8.71903 , 12.55878)
β[3]= 2.8938757	5.6366287	0.5134059	(-8.57660 , 14.36436)
β[4]= 8.8245798	3.9041722	2.2756629	(0.93963 , 16.82953)
β[5]= 5.4717400	3.9701416	1.3782229	(-2.60746 , 13.55094)
β[6]= 3.0528978	4.0505394	0.7537015	(-5.18991 , 11.29570)

* ESCOGE TU OPCION : s

! Una vez fuera del Menú de Intervalos, se llamará al Menú de Predicion.

F > PREDICCION

* ESCOGE TU OPCION : f

MENU PREDICCION

A > PREDICCION E INTERVALO DE CONFIANZA
PARA μ =(E[Y/Xo])
B > PREDICCION E INTERVALO DE CONFIANZA
PARA Y/Xo
C > RESULTADOS POR IMPRESORA DESACTIVADO

O > OPCIONES
S > SALIDA

* OPCION = b

! Acontinuación se calcularán algunas predicciones para un Xo dado.
! La opción "O" permite cambiar la confiabilidad, de inicio es 0.95%.

* PREDICCION E INTERVALO DE CONFIANZA PARA Y/Xo.

* DAME EL VECTOR Xo A PREDECIR

! La predicción e intervalo de confianza para una maestria por doce meses
! en Francia se calcula de la siguiente manera:

Xo[1]= 1
Xo[2]= 12
Xo[3]= 0
Xo[4]= 1
Xo[5]= 0
Xo[6]= 0
Xo[7]= 1

```
* EL VALOR ESTIMADO DE μ ES  =     13.0964
* EL INTERVALO DE CONFIANZA ES = (   -4.6913,    30.8841)
* PARA CONTINUAR OPRIMA ENTER ( <—┘ )

        * OPCION = b

* PREDICCION E INTERVALO DE CONFIANZA PARA  Y/Xo.

* Si se desea realizar los mismos cálculos para una maestria en Gran
* Bretaña por doce meses, los valores de Xo son:

    * DAME EL VECTOR Xo A PREDECIR

Xo[ 1]= 1
Xo[ 2]= 12
Xo[ 3]= 0
Xo[ 4]= 1
Xo[ 5]= 1
Xo[ 6]= 0
Xo[ 7]= 0

* EL VALOR ESTIMADO DE μ ES  =     18.9281
* EL INTERVALO DE CONFIANZA ES = (    1.3958,    36.4604)
* PARA CONTINUAR OPRIMA ENTER ( <—┘ )

        * OPCION = s

* Para grabar los datos como se encuentran actualmente (con la variable Y en
* miles, con X(1) tiempo de beca y  con las variables dicotomas) escogemos la
* opcion M.

        | M ) GUARDAR INFORMACION EN UN ARCHIVO                |

        * ESCOGE TU OPCION : m

        ┌──────────────────────────────────────────────────────┐
        │      MENU PARA GUARDAR INFORMACION EN ARCHIVO        │
        │                                                      │
        │ A ) GUARDAR X, Y, Y ESTIMADA Y RESIDUALES ESTIMADOS  │
        │ B ) GUARDAR ESTIMADORES Y MATRIZ DE COVARIANZAS      │
        │ C ) GUARDAR MATRIZ DE CORRELACIOES Y (n)             │
        │                                                      │
        │ S ) SALIDA                                           │
        └──────────────────────────────────────────────────────┘
        * OPCION : a

* INTRODUCE EL NOMBRE DEL ARCHIVO: becas1.txt
* OPERACION REALIZADA

* PARA CONTINUAR OPRIMA ENTER ( <—┘ )
```

| S) SALIDA |

* OPCION : s

* Con la opción "s" del Menú Guardar Información se pasa al Menú Principal.
* Por último se obtendrá la matriz de correlación entrando a la opción de
* Multicolinealidad.

| G) MULTICOLINEALIDAD |

* ESCOGE TU OPCION : g

```
MENU MULTICOLINEALIDAD

A ) MATRIZ DE CORRELACION
B ) REGRESION DE UNA VARIABLE EXPLICATIVA
    CON LAS RESTANTES
C ) RESULTADOS POR IMPRESORA        DESACTIVADO

S ) SALIDA
```

* OPCION = a

MATRIZ DE CORRELACION

	X(0)	X(1)	X(2)	X(3)	X(4)	X(5)	X(6)
X(0)	1.00000	0.83542	0.24262	-0.10072	0.52429	0.00982	-0.37432
X(1)		1.00000	0.33591	-0.18084	0.44456	-0.01444	-0.37668
X(2)			1.00000	-0.81184	-0.01761	0.06369	-0.03233
X(3)				1.00000	0.02621	0.02953	-0.14434
X(4)					1.00000	-0.39538	-0.42366
X(5)						1.00000	-0.31109
X(6)							1.00000

* PARA CONTINUAR OPRIMA ENTER (<—┘)

* Como se puede observar los coeficientes de correlación de mayor interes
* son: X(0) con X(1) y X(2) con X(3).

* OPCION = s

* Desde el Menú Principal se puede dar por concluida esta corrida.

* ESCOGE TU OPCION : s
* PRESIONA <ENTER> PARA SALIR
* O CUALQUIER OTRA TECLA PARA REGRESAR AL MENU ANTERIOR

3.3 EJEMPLO 2: Análisis del costo de operación para una compañia de aviación

El siguiente ejemplo fue obtenido de la tésis "Análisis del costo de operación para una compañia de aviación".[1] La variable dependiente que se analiza es el costo de operación para la compañia Aeronaves de Mexico. Este ejemplo fue elegido porque es un problema al que se le puede aplicar la mayoria de los procedimientos y opciones que ofrece este Paquetede de Regresión Lineal Múltiple.

Es necesario aclarar que la mayoria de las coclusiones que se mencionán con respecto a los resultados, asi como la manera de llevar cierto orden en las opciones es producto del seguimiento de la tésis. Todos los comentarios van precedidos por un asterisco.

Para entrar al Paquete de Regresión Lineal Múltiple se teclea REGCOM desde el sistema opetrativo. Una vez hecho esto, aparecera la siguiente pantalla durante algunos segundos y después el Menú de Obtencion de Datos:

1) Millan Gerardo 'Analisis del costo de operacion para una comania de aviacion.',

```
┌─────────────────────────────────────────┐
│                                         │
│            P A Q U E T E   D E          │
│                                         │
│  R E G R E S I O N   L I N E A L   M U L T I P L E │
│                                         │
└─────────────────────────────────────────┘
```

```
┌──────────────────────────────────────────────────┐
│               MENU OBTENCION DE DATOS            │
│                                                  │
│ A ) NUMERO DE DATOS DE LA REGRESION (n)          │
│ B ) REGRESION CON O SIN TERMINO INDEPENDIENTE    │
│ C ) LEER X                                       │
│ D ) LEER Y                                       │
│ E ) LEER Y y X                                   │
│ F ) SIMULAR X                                    │
│ G ) SIMULAR Y                                    │
│ H ) CONTINUAR                                    │
│                                                  │
│ S ) SALIDA AL SISTEMA OPERATIVO (DOS)            │
└──────────────────────────────────────────────────┘

      » ESCOGE TU OPCION : a


» CUANTOS DATOS SON: 32
```

* A continuación se escoge la opción E, que es la de leer el vector
* Y junto con la matriz X.

```
┌────────────────────────────────────────────────────────┐
│                 MENU DE LECTURA DE X y Y               │
│                                                        │
│ A ) POR PANTALLA                                       │
│ B ) ARCHIVO DE TEXTO (SDF)                             │
│                                                        │
│ S ) SALIDA                                             │
└────────────────────────────────────────────────────────┘

      » OPCION : b


» CUANTAS COLUMNAS CONTIENE LA MATRIZ X : 9
» INTRODUCE EL NOMBRE DEL ARCHIVO: ejemplo2.txt

» PARA CONTINUAR OPRIMA ENTER ( <── )
```

* En este menú se elige la lectura de X y Y de un archivo de texto
* llamado ejemplo2.txt.

* Cuando se dan los datos por pantalla, se tiene la opción mediante un
* procedimiento del Menú Principal de guardarlos en un archivo.
* Una vez leídos los datos, se regresará al Menú Obtención de Datos,
* en el cual se seleciona la opción H con la cual se
* pasa al Menú Inicial:

| H) CONTINUAR |

* ESCOGE TU OPCION : h

```
              M E N U   I N C I A L

  A ) GRAFICAS
  B ) SIMULAR  X
  C ) SIMULAR  Y
  D ) EDICION DE DATOS
  E ) CONSTRUCCION DE VAR. DICOTOMAS
  F ) TRANSFORMACIONES
  G ) CAMBIO DE VARIABLE DEPENDIENTE
  H ) LISTADO DE DATOS
  I ) SALIDA A LA IMPRESORA              DESACTIVADO

  N ) NUEVA REGRESION
  P ) MENU PRINCIPAL
  S ) SALIDA AL SISTEMA OPERATIVO (DOS)
```

* OPCION : h

* Mediante este Menú se pide un listado de datos, el cual se
* presenta en cuatro pantallas. Esto permite verificar la exactitud de
* los datos.

	Y	X(1)	X(2)	X(3)	X(4)
1	207.5100	77.2500	13.0000	63.0000	745.0000
2	217.3800	77.5000	13.0000	72.0000	745.0000
3	270.7100	77.8300	7.0000	80.0000	886.0000
4	272.3700	78.1700	12.0000	50.0000	822.0000
5	280.3600	78.8300	12.0000	71.0000	886.0000
6	284.8800	79.2000	12.0000	63.0000	886.0000
7	288.4800	79.6350	9.0000	48.0000	821.0000
8	229.6600	79.9500	15.0000	76.0000	530.0000
9	317.2100	80.2300	14.0000	59.0000	457.0000
10	345.3900	80.5400	13.0000	51.0000	514.0000
11	350.6300	80.5800	12.0000	64.0000	560.0000
12	394.3600	80.9200	13.0000	65.0000	850.0000
13	402.5900	81.0000	13.0000	47.0000	790.0000
14	412.1800	81.2100	15.0000	62.0000	530.0000
15	423.3200	81.3400	11.0000	67.0000	778.0000
16	443.2200	81.5000	10.0000	85.0000	1065.0000
17	452.0500	81.6000	10.0000	73.0000	1065.0000
18	457.1200	81.6700	15.0000	55.0000	822.0000

19	460.0500	81.7200	14.0000	46.0000	687.0000
20	473.6400	81.7900	19.0000	44.0000	538.0000
21	490.8800	81.9700	16.0000	59.0000	1050.0000
22	495.5800	82.0300	17.0000	52.0000	1050.0000

* PARA CONTINUAR OPRIMA ENTER (<—┘)

	Y	X(1)	X(2)	X(3)	X(4)
23	567.7900	82.3200	11.0000	70.0000	913.0000
24	608.8000	82.4000	19.0000	58.0000	821.0000
25	621.4500	82.4700	16.0000	59.0000	786.0000
26	642.2300	82.3200	11.0000	78.0000	1065.0000
27	652.3200	82.4300	11.0000	67.0000	1065.0000
28	665.9900	82.6300	22.0000	57.0000	828.0000
29	690.1900	83.0000	12.0000	71.0000	792.0000
30	697.1400	83.2500	20.0000	57.0000	1130.0000
31	712.2700	83.5300	18.0000	66.0000	845.0000
32	881.2400	83.7000	15.0000	67.0000	1090.0000

* PARA CONTINUAR OPRIMA ENTER (<—┘)

	X(5)	X(6)	X(7)	X(8)	X(9)
1	0.0000	0.0000	0.0000	0.0000	8.0200
2	1.0000	0.0000	0.0000	0.0000	8.1200
3	0.0000	0.0000	0.0000	1.0000	11.3500
4	0.0000	0.0000	0.0000	0.0000	11.4500
5	1.0000	0.0000	0.0000	1.0000	11.4500
6	0.0000	0.0000	0.0000	1.0000	11.7800
7	0.0000	0.0000	1.0000	0.0000	12.4100
8	1.0000	0.0000	1.0000	0.0000	12.4200
9	0.0000	0.0000	0.0000	0.0000	12.8100
10	0.0000	1.0000	1.0000	0.0000	12.7300
11	0.0000	0.0000	0.0000	0.0000	13.3000
12	0.0000	0.0000	0.0000	1.0000	13.1600
13	0.0000	1.0000	0.0000	0.0000	13.6000
14	0.0000	0.0000	0.0000	0.0000	14.2000
15	0.0000	0.0000	0.0000	0.0000	15.3000
16	0.0000	0.0000	0.0000	0.0000	15.3400
17	0.0000	0.0000	0.0000	0.0000	15.4200
18	1.0000	0.0000	0.0000	0.0000	17.5000
19	0.0000	1.0000	0.0000	0.0000	16.1400
20	0.0000	0.0000	0.0000	0.0000	18.1900
21	0.0000	0.0000	0.0000	0.0000	18.2800
22	0.0000	0.0000	0.0000	0.0000	19.7000

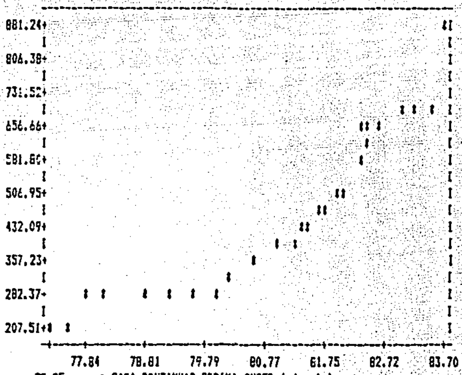
* PARA CONTINUAR OPRIMA ENTER (<—┘)

	X(5)	X(6)	X(7)	X(8)	X(9)
23	0.0000	0.0000	1.0000	1.0000	19.1500
24	1.0000	0.0000	0.0000	0.0000	20.1800
25	0.0000	0.0000	0.0000	0.0000	20.3000
26	1.0000	1.0000	0.0000	0.0000	21.1200
27	0.0000	1.0000	0.0000	0.0000	21.4200
28	1.0000	1.0000	0.0000	0.0000	21.2000
29	1.0000	1.0000	0.0000	1.0000	21.3000
30	1.0000	0.0000	0.0000	0.0000	21.4000
31	1.0000	1.0000	0.0000	0.0000	21.7000
32	1.0000	0.0000	0.0000	0.0000	21.9000

* PARA CONTINUAR OPRIMA ENTER (<—┘)

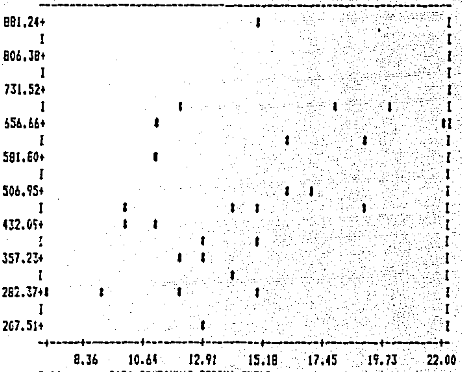
```
! Donde:
!      Y  = Costo de Operación (Por 10  pesos).
!      X1 = Fecha de evaluacion (años).
!      X2 = Seguro y renta de aviones (Por miles de dólares).
!      X3 = Sobrecargos y pilotos (sueldos en diezmiles de pesos).
!      X4 = Mantenimientos (materiales) (en dólares).
!      X5 = Apertura de nuevas rutas (dicotomica si=1, no=0).
!      X6 = Gas subsidiado (dicotomica si=1, no=0).
!      X7 = Compra de nuevas naves (dicotomica si=1, no=0).
!      X8 = Renta de al menos un avión (dicotomica si=1, no=0).
!      X9 = Salario variable (en miles de pesos).

! Regresando al Menú Inicial, se piden las gráficas de las variables
! independientes vs. Y, con el fin de dar una idea de las
! tendencias y de alguna manera observar la correlación que se da
! entre cada una de las variables independientes y la variable
! respuesta. Las variables X5, X6, X7 y X8 no se grafican por ser
! categóricas.
! Al tomar la opción A se presentan las siguientes opciones:

        |  A > GRAFICAS                                        |

        * ESCOGE TU OPCION : a

! CUAL GRAFICAS QUIERES ?

    X(1)  -->  [1]  vs.  Y OBSERVADA
    X(2)  -->  [2]  vs.  Y OBSERVADA
    X(3)  -->  [3]  vs.  Y OBSERVADA
    X(4)  -->  [4]  vs.  Y OBSERVADA
    X(5)  -->  [5]  vs.  Y OBSERVADA
    X(6)  -->  [6]  vs.  Y OBSERVADA
    X(7)  -->  [7]  vs.  Y OBSERVADA
    X(8)  -->  [8]  vs.  Y OBSERVADA
    X(9)  -->  [9]  vs.  Y OBSERVADA
PARA SALIR INTRODUCIR [0]

! Desde este menú se grafica cada una de las variables independientes
! versus la variable respuesta:
```

GRAFICA DE: X(1) vs. Y OBSERVADA



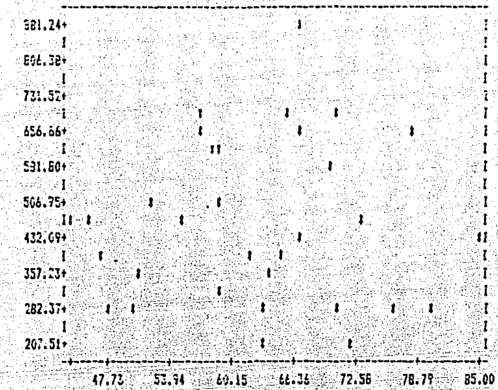
» PARA CONTINUAR OPRIMA ENTER (<—┘)

GRAFICA DE: X(2) vs. Y OBSERVADA



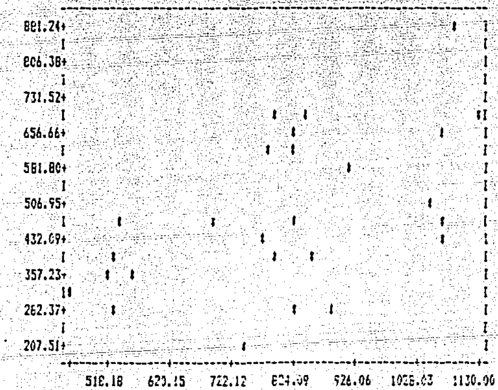
» PARA CONTINUAR OPRIMA ENTER (<—┘)

GRAFICA DE: X(3) vs. Y OBSERVADA



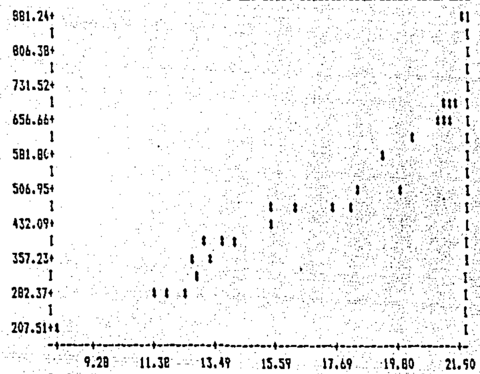
* PARA CONTINUAR OPRIMA ENTER (<—┘)

GRAFICA DE: X(4) vs. Y OBSERVADA



* PARA CONTINUAR OPRIMA ENTER (<—┘)

GRAFICA DE: X(9) vs. Y OBSERVADA



8.02 * PARA CONTINUAR OPRIMA ENTER (<—┘)

* Al salir del procedimiento de gráficas se regresa al menú
* inicial, desde el cual se elige la opcion 'P'.

```
| P ) MENU PRINCIPAL                              |

* OPCION : p
```

* El Menú Principal aparecera inmediatamente.

```
            M E N U   P R I N C I P A L

  A ) TABLA ANOVA Y ESTIMADORES
  B ) SELECCION DE VARIABLES INDEPENDIENTES
  C ) CARENCIA DE AJUSTE
  D ) PRUEBAS DE HIPOTESIS
  E ) INTERVALOS DE CONFIANZA
  F ) PREDICCION
  G ) MULTICOLINEALIDAD
  H ) HETEROSCEDASTICIDAD
  I ) AUTOCORRELACION
  J ) ANALISIS DE RESIDUALES
  K ) SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES
  L ) LISTAR DATOS
  M ) GUARDAR INFORMACION EN UN ARCHIVO
  N ) RESULTADOS POR IMPRESORA        DESACTIVADO
```

```
| 0 ) MENU INICIAL                                    |
| S ) SALIDA AL SISTEMA OPERATIVO (DOS)               |
|_____________________________________________________|
```

* ESCOGE TU OPCION : a

* Desde este menú se seleciona A, con
* lo cual se calculará la regresión con las nueve variables
* independientes y la variable Y leidas del archivo 'ejemplo2.txt'.

BETA ESTIMADA

B[0]=	-2549.0058910
B[1]=	31.6413962
B[2]=	-0.9924096
B[3]=	0.4984224
B[4]=	0.0573195
B[5]=	39.5108514
B[6]=	14.9907920
B[7]=	-24.9862828
B[8]=	0.2122514
B[9]=	22.8603610

* PARA CONTINUAR OPRIMA ENTER (<—┘)

TABLA DE ANALISIS DE VARIANZA

TABLA ANOVA

FUENTE DE VARIACION	g.l.	SUMA DE CUADRADOS	MEDIA DE CUADRADOS	F	VALOR-P
REGRESION	9	841601.750910	93511.305656	37.009179	0.000000
RESIDUALES	22	55587.526001	2526.705727		
TOTAL	31	897189.276910			

R Cuadrada (Coeficiente de determinacion)= 0.9380426
R Cuadrada Ajustada = 0.9126964

MATRIZ DE COVARIANZAS DE LAS BETAS

	X(0)	X(1)	X(2)	X(3)	X(4)
X(0)	1471514.3839	-19562.2822	-63.5467	65.3538	-20.3879
X(1)		262.3325	-5.8308	-2.6137	0.2345
X(2)			24.2942	3.5663	0.1139
X(3)				1.6068	-0.0053
X(4)					0.0044
X(5)					
X(6)					
X(7)					
X(8)					
X(9)					

* PARA CONTINUAR OPRIMA ENTER (<—┘)

	X(5)	X(6)	X(7)	X(8)	X(9)
X(0)	-5099.2501	-72.6764	3198.3100	-4421.8569	8148.1466
X(1)	95.3263	-11.8206	-60.1078	56.9960	-108.0906
X(2)	-75.5818	44.7228	43.6850	23.5606	-6.8686
X(3)	-18.4841	8.2339	6.4478	-4.5938	0.3552
X(4)	-0.1628	0.4250	0.5127	0.0311	-0.2275
X(5)	703.9946	-184.6069	-136.7287	-27.5731	-29.5563
X(6)		595.4507	64.9566	33.4288	-34.3252
X(7)			873.2731	-23.7561	10.8675
X(8)				650.5217	-23.7450
X(9)					56.0573

* PARA CONTINUAR OPRIMA ENTER (<—┘)

* De la tabla ANOVA se observa que la F calculada igual a 37.009,
* es mayor que una F(0.05,9,22) por lo que se rechaza la hipóte-
* sis de que $\beta 1=\beta 2=\beta 3=...=\beta 9=0$. También se observa un coeficiente
* de determinación de aproximadamente 94%.
* Ahora se elige la opción de análisis de residuales.

| J) ANALISIS DE RESIDUALES |

» ESCOGE TU OPCION : J

MENU ANALISIS DE RESIDUALES

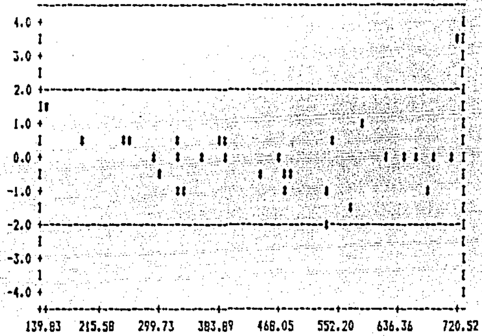
A) GRAFICACION DE RESIDUALES EN PAPEL DE PROBABILIDAD NORMAL
B) GRAFICACION DE RESIDUALES ESTUDENTIZADOS vs. Y ESTIMADA
C) GRAFICACION DE RESIDUALES ESTUDENTIZADOS vs. TIEMPO DE ORDEN
D) GRAFICACION DE RESIDUALES ESTUDENTIZADOS vs. CADA REGRESOR
E) GRAFICACION ENTRE REGRESORES
F) GRAFICACION PARCIAL DE RESIDUALES
G) OBSERVACIONES INFLUYENTES (METODO DE COOK)
H) VECTOR DE RESIDUALES
I) VECTOR DE RESIDUALES ESTUDENTIZADOS

```
| J > VECTOR DE Y ESTIMADA
| K > RESULTADOS POR IMPRESORA        DESACTIVADO
|
| O > OPCIONES
| S > SALIDA
```

» OPCION = b

! Pasamos a la opción B.

GRAFICA DE: Y ESTIMADA vs. RESIDUALES ESTUDENTIZADOS
CON BANDA DE CONFIANZA DEL 95.00%



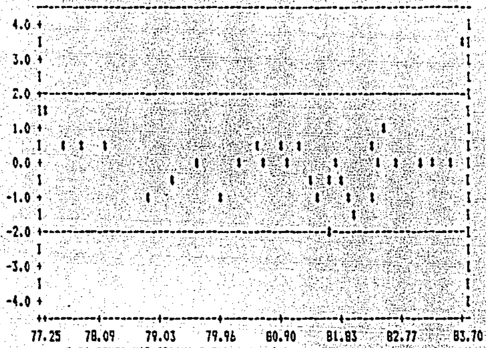
» PARA CONTINUAR OPRIMA ENTER (<—┘)

! En ella se observa cierta no linealidad, que podría llevar a una transformación
! de Y. Se selecionará del Menú de Análisis de Residuales la opción de Graficación
! de Residuales Estudentizados contra cada regresor, para X(1), X(2), X(3), X(4) y X(9):

```
| D > GRAFICACION DE RESIDUALES ESTUDENTIZADOS     |
|     vs. CADA REGRESOR                            |
```

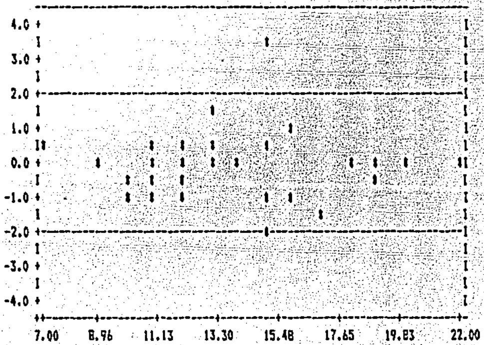
» OPCION = d

GRAFICA DE: X(1) vs. RESIDUALES ESTUDENTIZADOS
CON BANDA DE CONFIANZA DEL 95.00%



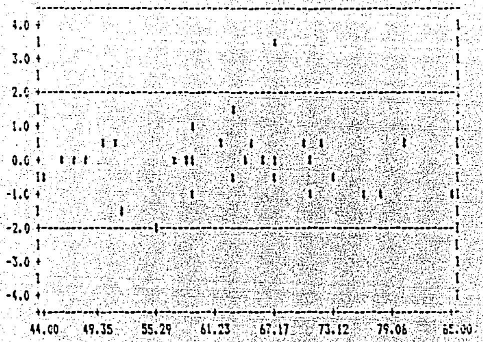
* PARA CONTINUAR OPRIMA ENTER (<—┘)

GRAFICA DE: X(2) vs. RESIDUALES ESTUDENTIZADOS
CON BANDA DE CONFIANZA DEL 95.00%



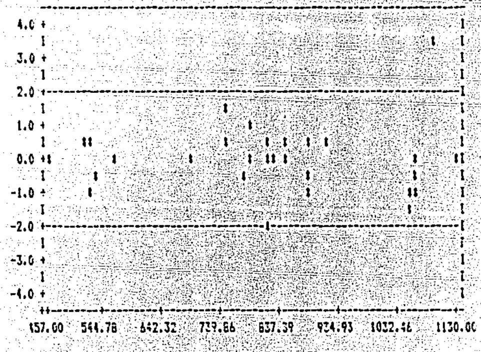
* PARA CONTINUAR OPRIMA ENTER (<—┘)

GRAFICA DE: X(3) vs. RESIDUALES ESTUDENTIZADOS
CON BANDA DE CONFIANZA DEL 95.00%



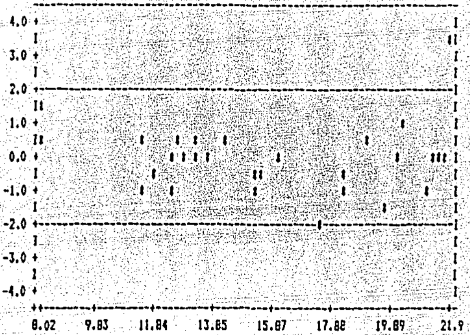
* PARA CONTINUAR OPRIMA ENTER (<—┘)

GRAFICA DE: X(4) vs. RESIDUALES ESTUDENTIZADOS
CON BANDA DE CONFIANZA DEL 95.00%



* PARA CONTINUAR OPRIMA ENTER (<—┘)

GRAFICA DE: X(9) vs. RESIDUALES ESTUDENTIZADOS
CON BANDA DE CONFIANZA DEL 95.00%



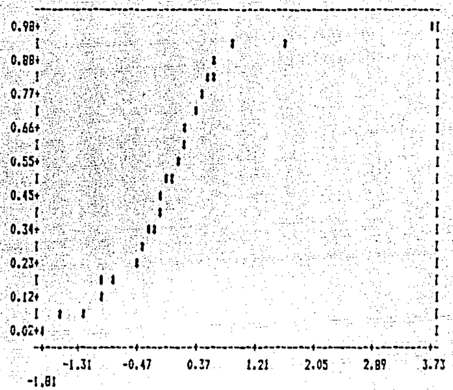
* PARA CONTINUAR OPRIMA ENTER (<—┘)

* En todas esta gráficas es importante observar el punto que
* pertenece a la observacion numero 32, ya que es un punto que
* guarda una mayor distancia en relación con los restantes, por
* lo que pudiéramos tener un punto discrepante.
* Otra caracteristica del modelo que se estudia es la verificación
* de la normalidad de los errores, con la
* gráfica de los Residuales Estudentizados en Papel de Probabilidad
* Normal, otra de las opciones del Menú de Residuales :

| A) GRAFICACION DE RESIDUALES EN PAPEL DE PROBABILIDAD NORMAL |

* OPCION : a

GRAFICA EN PAPEL NORMAL



* PARA CONTINUAR OPRIMA ENTER (<—┘)

* Gráfica que muestra la "no-normalidad" (en la normalidad
* se espera que los puntos caigan aproximadamente sobre una línea recta).
* El problema de la "no-normalidad" podría solucionarse con
* una transformación de la variable dependiente.
* Continuando con el análisis, se elige de este mismo menú la opción 6.

| 6) OBSERVACIONES INFLUYENTES (METODO DE COOK) |

* OPCION : g

RESIDUAL	R. ESTUDENTIZADO	Di-CALCULADO
67.67550	-1.81273	0.14819
23.35246	-1.59663	0.15060
13.67655	-1.24346	0.12036
26.08786	-0.99431	0.04619
-40.66388	-0.96699	0.05075
-11.89689	-0.98044	0.03511
-2.88131	-0.77564	0.67311
-42.88742	-0.48008	0.02879
-6.92289	-0.35312	0.00473
23.00035	-0.36105	0.01088
-6.15979	-0.27281	0.00202
13.67167	-0.20770	0.00281
6.94403	-0.16226	0.00164
20.57546	-0.15556	0.00071
-18.22130	-0.13669	0.00037

-30.71303	-0.08293	0.00047
-20.89493	0.06592	0.00016
-77.23795	0.09495	0.00039
-9.04835	0.12167	0.00054
-15.04671	0.17249	0.00150
-45.36072	0.18671	0.00083

* PARA CONTINUAR OPRIMA ENTER (<—┘)

RESIDUAL	R. ESTUDENTIZADO	Di-CALCULADO
-70.53955	0.23769	0.00185
22.76838	0.34254	0.01361
-7.39025	0.35481	0.00305
38.34299	0.46031	0.00595
-39.80214	0.57510	0.02035
4.94269	0.60548	0.02163
7.26828	0.61853	0.02798
2.44417	0.64758	0.03903
9.94404	0.85373	0.03593
4.23097	1.59635	0.07715
160.72174	3.72609	0.55232

* PARA CONTINUAR OPRIMA ENTER (<—┘)

\# Se nota que el residual de la observación 32 es alto en comparación
\# con los demás, por lo que se piensa que esta observación
\# es discrepante.
\# De el Menú de Análisis de Residuales se regresa al Menú
\# Principal donde, para continuar con el análisis de el modelo y de
\# los datos, se elige la opción Selección de los Mejores
\# Subconjuntos de Variables'.

```
| S ) SALIDA                                              |

* OPCION : s

| K ) SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES|

* ESCOGE TU OPCION : k
```

\# Que presenta las siguientes dos pantallas :

```
        SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES

MEJOR SUBCONJUNTO CON  1 VARIABLES
  VARIABLES  :   9
  SCRES=  78275.5567   Cp_MALLOWS=    2.9793

MEJOR SUBCONJUNTO CON  2 VARIABLES
  VARIABLES  :   9, 5
  SCRES=  70915.6566   Cp_MALLOWS=    2.0664
```

```
MEJOR SUBCONJUNTO CON  3 VARIABLES
  VARIABLES  :   9, 1, 5
  SCRES=  64524.2217   Cp_MALLOWS=    1.5369

MEJOR SUBCONJUNTO CON  4 VARIABLES
  VARIABLES  :   9, 1, 5, 4
  SCRES=  59614.0718   Cp_MALLOWS=    1.5936

MEJOR SUBCONJUNTO CON  5 VARIABLES
  VARIABLES  :   9, 1, 5, 4, 7
  SCRES=  58050.8486   Cp_MALLOWS=    2.9749

* PARA CONTINUAR OPRIMA ENTER : (<--' )


         SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES

MEJOR SUBCONJUNTO CON  6 VARIABLES
  VARIABLES  :   9, 1, 5, 4, 7, 6
  SCRES=  56815.3675   Cp_MALLOWS=    4.4857

MEJOR SUBCONJUNTO CON  7 VARIABLES
  VARIABLES  :   9, 1, 5, 4, 7, 6, 8
  SCRES=  55587.7010   Cp_MALLOWS=    6.0001

MEJOR SUBCONJUNTO CON  8 VARIABLES
  VARIABLES  :   9, 1, 5, 4, 7, 6, 3, 2
  SCRES=  55587.7007   Cp_MALLOWS=    8.0001

* PARA CONTINUAR OPRIMA ENTER : (<--' )
```

* Con los datos que ofrece esta opción se selecionarán un
* conjunto de variables que parezca más adecuado, con el que se
* continuará el análisis. Ya que el sistema selecciona al
* mejor subconjunto con una variable, al mejor subconjunto con
* dos variables, etcetera. Se tiene ocho subconjuntos entre los
* cuales elegir. Para esta selección se utiliza el estadístico
* Cp-Mallows que esta calculado en cada subconjunto. El
* criterio sugerido involucra identificar el subconjunto de
* variables explicativas con Cp mínimo y a la vez Cp próximo a p.
* Bajo este criterio elegimos al subconjunto formado por las
* variables X(9), X(1), X(5) y X(9).
* Para calcular la regresión del modelo con estas variables,
* del Menu Principal tomamos la opción Selección de Variables.

```
     |  B ) SELECCION DE VARIABLES INDEPENDIENTES          |

     * ESCOGE TU OPCION : b
```

* CUANTAS DE LAS 9 COLUMNAS DE LA MATRIZ X SERAN
* INCLUIDAS EN LA REGRESION ? 3

* DAME LA 1a. COLUMNA: 1
* DAME LA 2a. COLUMNA: 5
* DAME LA 3a. COLUMNA: 9

* Se elige la opción "A" para que el sistema calcule las betas estimadas
* para este nuevo modelo.

| A > TABLA ANOVA Y ESTIMADORES |

* ESCOGE TU OPCION : a

BETA ESTIMADA

B[0]=	-1944.7163709
B[1]=	24.1060032
B[2]=	40.5086151
B[3]=	27.3471772

* PARA CONTINUAR OPRIMA ENTER (<—┘)

TABLA DE ANALISIS DE VARIANZA

TABLA ANOVA

FUENTE DE VARIACION	g.l.	SUMA DE CUADRADOS	MEDIA DE CUADRADOS	F	VALOR-P
REGRESION	3	832665.055010	277555.018340		
				120.443770	0.000000
RESIDUALES	28	64524.221901	2304.436496		
TOTAL	31	897189.276910			

R Cuadrada (Coeficiente de determinacion)= 0.9280818
R Cuadrada Ajustada = 0.9203763

* PARA CONTINUAR OPRIMA ENTER (<—┘)

MATRIZ DE COVARIANZAS DE LAS BETAS

	X(0)	X(1)	X(2)	X(3)
X(0)	1172942.5096	-15668.6477	-4073.0430	6178.3399
X(1)		209.5158	55.8864	-83.5557
X(2)			378.5336	-36.7417
X(3)				38.0260

* PARA CONTINUAR OPRIMA ENTER (<—┘)

* En este modelo, como el anterior, se rechaza la hipótesis de
* que β1=β2=β3=,...,=0. El coeficiente de determinación es de
* 0.928, muy cercano al que se tenia en el modelo anterior.
* Con la opcion LISTAR DATOS, obtenemos las observaciones
* unicamente de las variables X(1), X(5) y X(9) que participan
* en este modelo.

| L) LISTAR DATOS |

* ESCOGE TU OPCION : 1

	Y	X(1)	X(2)	X(3)
1	207.5100	77.2500	0.0000	8.0200
2	217.3800	77.5000	1.0000	8.1200
3	270.7100	77.8300	0.0000	11.3500
4	272.3700	78.1700	0.0000	11.4500
5	280.3600	78.8300	1.0000	11.4500
6	284.8800	79.2000	0.0000	11.7800
7	288.4800	79.6300	0.0000	12.4100
8	289.6600	79.9500	1.0000	12.4200
9	317.2100	80.2300	0.0000	12.8100
10	345.3900	80.5400	0.0000	12.7300
11	350.6300	80.5800	0.0000	13.3000
12	394.3600	80.9200	0.0000	13.1600
13	402.5900	81.0000	0.0000	13.6000
14	412.1800	81.2100	0.0000	14.2000
15	423.3200	81.3400	0.0000	15.3000
16	443.2200	81.5000	0.0000	15.3400
17	452.0500	81.6000	0.0000	15.4200
18	457.1200	81.6700	1.0000	17.5000
19	460.0500	81.7200	0.0000	16.1400
20	473.6400	81.7900	0.0000	18.1900
21	490.8800	81.9700	0.0000	18.2800
22	495.5600	82.0300	0.0000	19.7000

* PARA CONTINUAR OPRIMA ENTER (<—┘)

	Y	X(1)	X(2)	X(3)
23	567.7900	82.3200	0.0000	19.1500
24	608.8000	82.4000	1.0000	20.1800
25	621.4500	82.4700	0.0000	20.3000
26	642.2300	82.3200	1.0000	21.1200
27	652.3200	82.4300	0.0000	21.4200
28	665.9900	82.6300	1.0000	21.2000
29	690.1900	83.0000	1.0000	21.3000
30	697.1400	83.2500	1.0000	21.4000

31	712.2700	83.5300	1.0000	21.7000
32	831.2400	83.7000	1.0000	21.7000

* PARA CONTINUAR OPRIMA ENTER (<—┘)

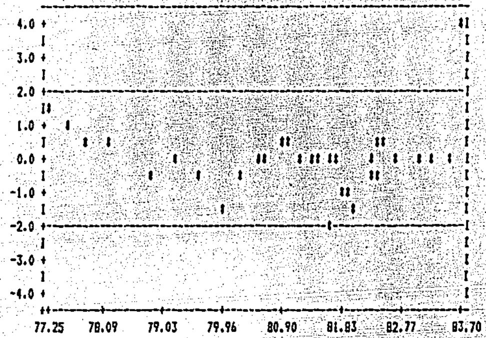
! Hay que tomar en cuenta que la nueva X(1) corresponde a
! la anterior X(1), la nueva X(2) corresponde a la anterior
! X(5) y X(3) corresponde a la anterior X(9).
! Se selecciona la opción Análisis de Residuales.

```
MENU ANALISIS DE RESIDUALES

A ) GRAFICACION DE RESIDUALES EN PAPEL DE
    PROBABILIDAD NORMAL
B ) GRAFICACION DE RESIDUALES ESTUDENTIZADOS
    vs. Y ESTIMADA
C ) GRAFICACION DE RESIDUALES ESTUDENTIZADOS
    vs. TIEMPO DE ORDEN
D ) GRAFICACION DE RESIDUALES ESTUDENTIZADOS
    vs. CADA REGRESOR
E ) GRAFICACION ENTRE REGRESORES
F ) GRAFICACION PARCIAL DE RESIDUALES
G ) OBSERVACIONES INFLUYENTES (METODO DE COOK)
H ) VECTOR DE RESIDUALES
I ) VECTOR DE RESIDUALES ESTUDENTIZADOS
J ) VECTOR DE Y ESTIMADA
K ) RESULTADOS POR IMPRESORA        DESACTIVADO

O ) OPCIONES
S ) SALIDA
```

* OPCION = d

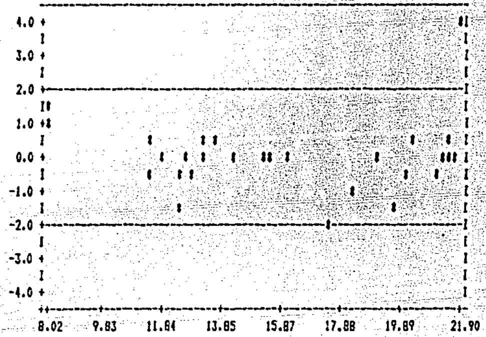
! En el Menú de Análisis de Residuales se eligió la opcion 'D)
! Graficación de Residuales Estudentizados vs. cada Regresor' y
! se graficarán los residuales estudentizados contra X(1) y X(3).

GRAFICA DE: X(1) vs. RESIDUALES ESTUDENTIZADOS
CON BANDA DE CONFIANZA DEL 95.00%



» PARA CONTINUAR OPRIMA ENTER (<—┘)

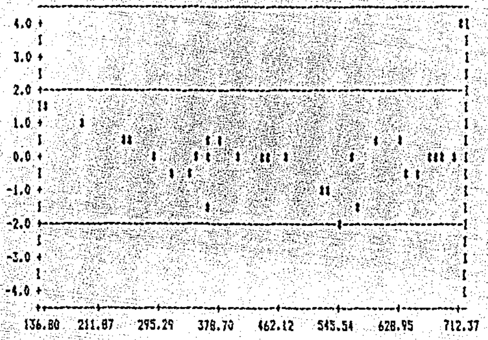
GRAFICA DE: X(3) vs. RESIDUALES ESTUDENTIZADOS
CON BANDA DE CONFIANZA DEL 95.00%



» PARA CONTINUAR OPRIMA ENTER (<—┘)

* Tambien se graficará los Residuales Estudentizados contra la
* Y estimada.

GRAFICA DE: Y ESTIMADA vs. RESIDUALES ESTUDENTIZADOS
CON BANDA DE CONFIANZA DEL 95.00%



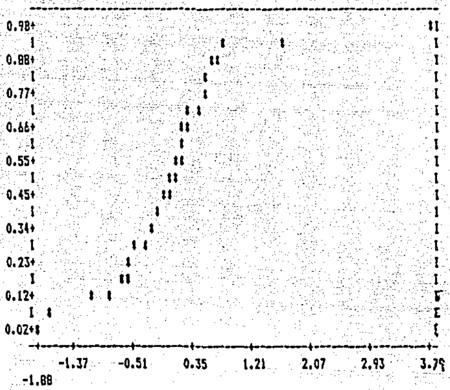
» PARA CONTINUAR OPRIMA ENTER (<┘)

* Comparando estas gráficas con las gráficas que se obtuvieron
* cuando el modelo contaba con las nueve variables se nota que
* tienen una gran semejanza, por lo que tambien se pueden inferir
* las mismas observaciones hechas anteriormente.
* Para este modelo volvemos a probar la 'normalidad' al generar la
* gráfica en papel de probabilidad normal.

| A > GRAFICACION DE RESIDUALES EN PAPEL DE PROBABILIDAD NORMAL |

» OPCION = a

GRAFICA EN PAPEL NORMAL



* PARA CONTINUAR OPRIMA ENTER (<——┘)

* Comparando esta gráfica con la anterior se observa una ligera mejora en
* la normalidad del modelo, ya que los puntos se aproximan más a una
* recta.
* Por último, haciendo una comparación entre los dos modelos presentados en
* este ejemplo, podemos decir que a pesar de que en muchos conceptos son
* parecidos, como el comportamiento de sus residuales, los coeficientes de
* determinacion, las gráficas de probabilidad normal, el punto discrepante
* que en ambos modelos se muestra, etcetera. Hay tambien diferencias
* relevantes como la simplificacion de uno respecto al otro en el manejo
* de variables.

3.4 EJEMPLO 3: Simulación de la variable dependiente

En el siguiente ejemplo se simula la variable dependiente conociendo la estimación de los valores de las betas la varianza y el valor de la variable explicativa.[1]

Y es igual al consumo familiar semanal.
X es igual al ingreso familiar por semana.

Cuadro 3.2

Observación	Y	X
1.	70	80
2.	65	100
3.	90	120
4.	95	140
5.	110	160
6.	115	180
7.	120	200
8.	140	220
9.	155	240
10.	150	260

Con estas observaciones se obtuvierón las siguientes estimaciones para las betas y la varianza:

β_o = 24.454
β_1 = 0.509
y
σ^2 = 42.15

1) Los valores fueron tomados de libro de:Damodar Gujarati, 'Econometria Basica' p. 43

```
P A Q U E T E   D E

R E G R E S I O N   L I N E A L   M U L T I P L E
```

```
MENU OBTENCION DE DATOS

A ) NUMERO DE DATOS DE LA REGRESION (n)
B ) REGRESION CON O SIN TERMINO INDEPENDIENTE
C ) LEER X
D ) LEER Y
E ) LEER Y y X
F ) SIMULAR X
G ) SIMULAR Y
H ) CONTINUAR

S ) SALIDA AL SISTEMA OPERATIVO (DOS)
```

```
* ESCOGE TU OPCION :a
```

* Como primer paso se proporciona el número de casos.

```
* CUANTOS DATOS SON: 10
```

* Sólo se leerá la variables X, y la variable dependiente se simulará.

```
* ESCOGE TU OPCION :c
```

```
MENU DE LECTURA DE X

A ) POR PANTALLA
B ) ARCHIVO DE TEXTO (SDF)

S ) SALIDA
```

```
* OPCION :a

* CUANTAS COLUMNAS SERAN LEIDAS : 1

PARA SALIR INTRODUCIR  [0]
```

* Como los datos son pocos se introducirán por pantalla.

```
* TIENES QUE DARME LA MATRIX X POR RENGLONES EN FORMATO LIBRE
```

```
* CON  1 COMPONENTES
* "NO" INCLUYAS LA COLUMNA DE UNOS PARA EL TERMINO INDEPENDIENTE

* RENGLON 1 :
80
* RENGLON 2 :
100
* RENGLON 3 :
120
* RENGLON 4 :
140
* RENGLON 5 :
160
* RENGLON 6 :
180
* RENGLON 7 :
200
* RENGLON 8 :
220
* RENGLON 9 :
240
* RENGLON 10 :
260

* PARA CONTINUAR OPRIMA ENTER ( <——┘ )
```

* Una vez que se tiene a X, se procede a simular el valor de Y.
* Para esto, se pide los valores para las betas y la σ^2 que son tomados
* del resultado del libro, página 46 (ver referencia).
* Es decir, teniendo a las β's, a las X's y las ϵ's se resuelve la siguiente
* ecuación: $Y = \beta X + \epsilon$.
* La σ^2 se pide para calcular los valores simulados como Normal de ϵ.

```
     | 6 > SIMULAR Y                          |

     * ESCOGE TU OPCION : h

     +------------------------------------------+
     |   SIMULACION DE  Y  ( DISTRIBUCION NORMAL )
     |
     | A > Y DISTRIBUIDA N( Xβ, σ²I)
     | B > Y DISTRIBUIDA N( Xβ, V)
     |
     | S > SALIDA
     +------------------------------------------+
     * OPCION : a

* DAME EL VALOR DE σ² = 42


* QUIERES LISTAR LOS ERRORES SIMULADOS  ..? (S/N) n

* DAME EL VECTOR β DE 2 COMPONENTES
```

$\beta[0]= 24$
$\beta[1]= 0.5$

* Estos son los valores a priori de σ^2 y de las betas.

» PARA CONTINUAR OPRIMA ENTER (<—┘)

* Obtenidos los valores de la variable dependiente se pasa al Menú Inicial.

| H) CONTINUAR |

» ESCOGE TU OPCION : h

```
              M E N U   I N C I A L

 A ) GRAFICAS
 B ) SIMULAR  X
 C ) SIMULAR  Y
 D ) EDICION DE DATOS
 E ) CONSTRUCCION DE VAR. DICOTOMAS
 F ) TRANSFORMACIONES
 G ) CAMBIO DE VARIABLE DEPENDIENTE
 H ) LISTADO DE DATOS
 I ) SALIDA A LA IMPRESORA          DESACTIVADO

 N ) NUEVA REGRESION
 P ) MENU PRINCIPAL
 S ) SALIDA AL SISTEMA OPERATIVO (DOS)
```

» OPCION : h

	Y	X(1)
1	56.0189	80.0000
2	78.1378	100.0000
3	77.7548	120.0000
4	95.9703	140.0000
5	104.3806	160.0000
6	100.2167	180.0000
7	127.9068	200.0000
8	144.0970	220.0000
9	138.3944	240.0000
10	153.5652	260.0000

» PARA CONTINUAR OPRIMA ENTER (<—┘)

* Como se podrá observar, el rango de los valores simulados de Y es
* ligeramente más grande.
* Ahora de mostrará la gráfica de X contra la variable dependiente.

| A) GRAFICAS |

* OPCION : a

GRAFICA DE X(1) vs Y OBSERVADA

```
        ------------------------------------------------------------------
 153.57+                                                                *I
       I                                                                 I
 142.73+                                                  *              I
       I                                                         *       I
 131.89+                                                                 I
       I                                           *                     I
 121.05+                                                                 I
       I                                                                 I
 110.21+                                                                 I
       I                            *                                    I
  99.37+                                    *                            I
       I                     *                                           I
  88.53+                                                                 I
       I                                                                 I
  77.70+      *       *                                                  I
       I                                                                 I
  66.86+                                                                 I
       I                                                                 I
  56.02+*                                                                I
        -+---------+---------+---------+---------+---------+---------+---------+-
            96.36    123.64    150.91    178.18    205.45    232.73    260.00
     80.00
```

* PARA CONTINUAR OPRIMA ENTER (<—┘)

* QUIERES OTRA GRAFICA ? (S/N)

* Se pasará al Menú Principal para calcular las nuevas betas estimadas.

| P) MENU PRINCIPAL |

* OPCION : p

M E N U   P R I N C I P A L

A) TABLA ANOVA Y ESTIMADORES
B) SELECCION DE VARIABLES
C) CARENCIA DE AJUSTE
D) PRUEBAS DE HIPOTESIS
E) INTERVALOS DE CONFIANZA
F) PREDICCION
G) MULTICOLINEALIDAD
H) HETEROSCEDASTICIDAD
I) AUTOCORRELACION
J) ANALISIS DE RESIDUALES
K) SELECCION DE SUBCONJUNTOS DE VARIABLES
L) LISTAR DATOS
M) GUARDAR INFORMACION EN UN ARCHIVO
N) RESULTADOS POR IMPRESORA   DESACTIVADO

| 0) MENU INICIAL |
| S) SALIDA AL SISTEMA OPERATIVO (DOS) |

* ESCOGE TU OPCION : a

BETA ESTIMADA

B[0]=	18.8800422
B[1]=	0.5221425

* PARA CONTINUAR OPRIMA ENTER (<—┘)

* Beta cero es menor en un 20 por ciento que el valor a priori, mientras
* que beta uno es muy cercana a la beta cero original.

TABLA DE ANALISIS DE VARIANZA

TABLA ANOVA

FUENTE DE VARIACION	g.l.	SUMA DE CUADRADOS	MEDIA DE CUADRADOS	F	VALOR-P
REGRESION	1	8996.881906	8996.881906	168.129964	0.000005
RESIDUALES	8	428.091779	53.511472		
TOTAL	9	9424.973685			

R Cuadrada (Coeficiente de determinacion)= 0.9545790
R Cuadrada Ajustada = 0.9489014

* PARA CONTINUAR OPRIMA ENTER (<—┘)

MATRIZ DE COVARIANZAS DE LAS BETAS

	X(0)	X(1)
X(0)	52.2142	-0.2757
X(1)		0.0016

* PARA CONTINUAR OPRIMA ENTER (<—┘)

* Se analizará por medio de la gráfica en Papel normal la Normalidad del
* modelo.

| 3) ANALISIS DE RESIDUALES |

```
* ESCOGE TU OPCION : j

 ----------------------------------------------------------------
|            MENU ANALISIS DE RESIDUALES                         |
|                                                                |
| A ) GRAFICACION DE RESIDUALES EN PAPEL DE                      |
|     PROBABILIDAD NORMAL                                        |
| B ) GRAFICACION DE RESIDUALES ESTUDENTIZADOS                   |
|     vs. Y ESTIMADA                                             |
| C ) GRAFICACION DE RESIDUALES ESTUDENTIZADOS                   |
|     vs. TIEMPO DE ORDEN                                        |
| D ) GRAFICACION DE RESIDUALES ESTUDENTIZADOS                   |
|     vs. CADA REGRESOR                                          |
| E ) GRAFICACION ENTRE REGRESORES                               |
| F ) GRAFICACION PARCIAL DE RESIDUALES                          |
| G ) OBSERVACIONES INFLUYENTES (METODO DE COOK)                 |
| H ) RESULTADOS POR IMPRESORA          DESACTIVADO              |
|                                                                |
| O ) OPCIONES                                                   |
| S ) SALIDA                                                     |
 ----------------------------------------------------------------

* OPCION = a
```

GRAFICA EN PAPEL NORMAL

```
      -------------------------------------------------------------------
 0.95+                                                                  *I
     I                                                                   I
 0.85+                                                            *      I
     I                                                                   I
 0.75+                                                     *             I
     I                                                                   I
 0.65+                                                   *               I
     I                                                                   I
 0.55+                                              *                    I
     I                                                                   I
 0.45+                                       *                           I
     I                                                                   I
 0.35+                                *                                  I
     I                                                                   I
 0.25+                           *                                       I
     I                                                                   I
 0.15+                         *                                         I
     I                                                                   I
 0.05+*                                                                  I
      -+------+---------+---------+---------+---------+---------+---------+-
          -1.52     -1.01     -0.49      0.02      0.53      1.05      1.56
     -1.83
```

* PARA CONTINUAR OPRIMA ENTER (<──┘)

* Como se puede observar todos los puntos (con excepción del primero) se

aproximan a una linea recta.

```
        | B ) GRAFICACION DE RESIDUALES ESTUDENTIZADOS        |
        |     vs. Y ESTIMADA                                  |

        # OPCION = b

GRAFICA DE Y ESTIMADA  vs.  RESIDUALES ESTUDENTIZADOS
CON BANDA DE CONFIANZA DEL  95.00%
    -------------------------------------------------------------------
 4.0 +                                                                 I
     I                                                                 I
 3.0 +                                                                 I
     I                                                                 I
 2.0 +-----------------------------------------------------------------I
     I                                                        *        I
 1.0 +       *                                                         I
     I                   *     *                 *                     I
 0.0 +                                                                *I
     I           *                                                     I
-1.0 +*                                                         *      I
     I                                                                 I
-2.0 +------------------------------------------*----------------------I
     I                                                                 I
-3.0 +                                                                 I
     I                                                                 I
-4.0 +                                                                 I
     ++--------+---------+---------+---------+---------+---------+---------+-
    60.65    72.91     86.53    100.15    113.77    127.39    141.02    154.64
```

Esta última gráfica no muestra indicios de algún patrón sistemático.

```
        | S ) SALIDA                                          |
        |======================================================|

        # OPCION = s

        | S ) SALIDA AL SISTEMA OPERATIVO (DOS)                |
        |======================================================|

        # ESCOGE TU OPCION : s

        # PRESIONA <ENTER> PARA SALIR
        > O CUALQUIER OTRA TECLA PARA REGRESAR AL MENU ANTERIOR
```

Se dá fin a la sesion del Menú de Residuales y del Menú Principal
con la (S), y para confirmar el último se introduce <enter>.

C O N C L U S I O N E S

Cuando fuimos estudiantes de la carrera de Actuaría y tomamos la materia de Anális de Regresión, se nos pidió aplicar la regresión lineal múltiple a un ejemplo utilizando un paquete de computación. Sin embargo, al no poderlo conseguir en la escuela, se decidió en la clase que por equipos se creara un programa sencillo que calculara la regresión. Este fue uno de los principales motivos que nos llevó a realizar la presente tesis y que quedó expresado en nuestro objetivo general de la misma: elaborar e implementar un paquete computacional de regresión lineal múltiple que permita el análisis y la elaboración de trabajos, tareas y proyectos estudiantiles con el fin de obtener un mayor aprovechamiento en las materias de Estádistica, Análisis de Regresión y Econometría. Pensamos que el paquete REGCOM cumple con este objetivo y va un poco más allá.

El paquete cuenta con procedimientos que facilitan el manejo de los datos como son la selección de variables que intervienen en el modelo; la posibilidad de realizar transformaciones y simulaciones; la existencia de gráficas de las variables explicativas contra la variable dependiente; la construcción de variables dicotomas y la posibilidad de edición de los datos.

Estos temas fueron incluidos para que el usuario pueda manejar los datos dentro del paquete en forma sencilla y no tenga que recurrir a otras herramientas fuera del paquete.

Se intenta que el paquete REGCOM no sólo genere los calculos básicos de la regresión, por lo que se le incluyeron temas de utilidad como el análisis de residuales, las pruebas de heteroscedasticidad, las pruebas de multicolinealidad, las pruebas de autocorrelación y la selección de subconjuntos de variables, que que nos ayudan en el análisis de la regresión.

El paquete se diseñó de manera que la aplicación de los ejemplos sea relativamente sencilla. Para utilizar el paquete no se requiere tener grandes conocimientos en el área de la computación, ya que el paquete se basa en la presentación de las opciones por medio de menús. Lo que si se requiere para el mejor aprovechamiento y entedimiento de las diferentes opciones y alternativas del paquete, es un conocimento del manejo estadístico y los supuestos del análisis de regresión lineal multiple, así como la posibilidad de interpretar los resultados que se pueden obtener en cada una de las opciones del paquete .

El paquete REGCOM se creó de tal forma que la captura y creación de variables no es única; se pueden obtener las variables por distintos medios: mediante la lectura de un archivo; por la captura en pantalla; por simulación; por transformaciones de otras variables; por la edición o por medio de la creación de variables dicotomicas. La captura y creación de variables es dinámica, ya que se pueden usar y combinar repetidamente las opciones mencionadas anteriormente.

En el paquete se consideró como una opción importante y nueva la inclusión de un proceso que creara variables por medio de la simulación. El paquete puede crear variables explicativas cuyos valores simulen una distribución de probabilidades uniforme o una distribución de probabilidad Bernoulli. La variable dependiente se puede crear simulando los valores de acuerdo a una distribución de probabilidad Normal. La razon que se tuvo para hacer esta diferenciación en cuanto a la simulación de una variable explicativa y de una dependiente se debe al supuesto de que la variable dependiente se distribuye como una Normal para el modelo de regresión y no necesariamente para las variables explicativas.

El manejo de las variables que intervendrán en el modelo de regresión es sencillo y dinámico. Con las variables obtenidas o creadas se pueden elegir las que el usuario desee que se incorporen al modelo y con ellas se calcula la regresión. De esta forma se puede correr una regresión considerando unas cuantas variables, luego correr otra regresión añadiendo o quitando variables o considerando otras y comparar los resultados. Tambien permite cambiar la variable dependiente por cualquier variable explicativa y calcular la regresión respecto a ésta y con las variables que se deseen; así mismo, el paquete permite manejar en forma sencilla un modelo con término independiente o un modelo sin él.

Todas estas caracteristicas del paquete hacen que sea una herramienta que puede permitir un buen análisis en la aplicación de la regresión a problemas concretos o teóricos.

El paquete en un inicio fue diseñado para ser compilado por la versión 3.02 de TURBO Pascal. Sin embargo, al incorporarse más procedimientos y opciones, se llegó al límite de memoria de la computadora, pudiéndose manejar menos de 10 variables y menos de 100 datos por cada una de ellas. Con objeto de administrar de manera más eficiente la memoria y aumentar el número posible de variables y el número de datos a incorporar, se recurrió a una opción que ofrece esta versión de TURBO pascal: el uso de procedimientos "overlays". Estos procedimientos tienen la caracteristica de que accesan a la memoria sólo cuando son utilizados por otros programas, y unicamente tienen asignado un área de memoria del tamaño que ocupe el mayor de ellos. Gracias a esta opción, se pudo aumentar la capacidad del número de datos y del número de variables que podian ser manejados. A pesar de esto, al avanzar en el desarrollo del sistema se encontró de nuevo el problema del límite de la memoria del programa y el número de variables y observaciones que se podian manejar llegó a quedar similar al que se tenía anteriormente.

Considerando este problema se decidio utilizar la versión 4.00 de TURBO pascal mediante la cual habia posibilidades de mejorar la capacidad de manejo del paquete. Para usar esta versión fue necesario hacer varias adecuaciones a los programas: se dejó de usar los procedimientos overlays (que no son soportados por esta versión), y los procedimientos fueron agrupados en unidades. Las unidades son procedimientos que pueden tener uno o más programas y cada una de ellas tiene la misma capacidad de memoria que el programa principal que las llama. Con el uso de la versión 4.00 y de las unidades se pudo aumentar la capacidad de manejo de variables en el paquete a 25 y la capacidad de manejo de datos a 275 por cada variable.

BIBLIOGRAFIA

Abramowitz, M. and Stegun, I. "Handbook of Mathematical Functions", Edit. Dover, USA. 1970.

Cramér Harald. "Métodos Matemáticos de Estadística", Aguilar, Espana. 1970.

Draper, N.R., and H. Smith, "Applied Regression Analysis", Jhon-Wiley and Sons, Inc., U.S.A., 1981.

Gujarati, Damodar "Econometría Basica", Mc Graw Hill, México 1986.

Jhonstonn, J., "Econometric Methods", Mc Graw Hill Book Company N.Y., 1972.

Martin, R.S., Peters, G. and Wilkinson, J.H., "Iterative Refinement of the Solution of a Positive Definitive System of Equation", Numer. Math. 8.203-216 (1966)

Martin, R.S., Peters, G. and Wilkinson, J.H., "Symmetric Decomposition of a Positive Definitive Matrix", Numer. Math. 7.362-383, (1965).

Millan Campos, Gerardo "Análisis del costo de operación para una copañia de aviación", Noviembre de 1986, Tesis, UNAM, ENEP ACATLAN

Seberg, G.A.F., "Linear Regression Analysis", John Wiley and Sons,

Stewart, M. B. and Wallis, K.F. "Introducctory Econometrics", Halten Press Book, , 1981.

Utts, M. Jessica, "The Rainbow Test for Lack of Fit in Regression", Communications in Statistics Theory and Methods, 11(24), 2801-2815. 1982

Wonnacott, R. and Wonnacott, T., "Regression: A second course in statistics", Jhon Wiley and Sons, Inc.

A N E X O   A

P R O G R A M A
D E
R E G R E S I O N   L I N E A L   M U L T I P L E

```
{
        UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
    ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES ACATLAN

  PROGRAMA : 'REGRESION.PAS'
  TESIS    : DESARROLLO DE UN PAQUETE DE REGRESIONLINEAL MULTIPLE.
  AUTORES  : RENE DORANTES RODRIGUEZ SAN MIGUEL
             AMBROSIO DE JESUS TORRES MENDOZA
  FECHA    : JUNIO 1990
}

{$R-} {$B+} {$S+} {$I+} {$N-} {$M 65500,16384,655360}

PROGRAM REGRESION(INPUT,OUTPUT);

{*****      UNIDADES QUE UTILIZA Y LLAMA EL PROGRAMA.           *****}

USES
  CRT,
  PRINTER,
  GLOBAL,
  UTILERIA,
  OBTENER,
  CARENCIA,
  HIPOTESI,
  INTERVAL,
  MULTICOL,
  HETEROS,
  AUTOCORR,
  RESIDUAL,
  PREDIC,
  EDITAR,
  FURNIVAL,
  TRANSFOR;

{
   EL PROCEDIMIENTO "PRG_PRINCIPAL" CONTIENE AL MENU PRINCIPAL QUE TIE-
  NE LAS SIGUIENTES OPCIONES:

        A) TABLA ANOVA Y ESTIMADORES
        B) SELECCION DE VARIABLES
        C) CARENCIA DE AJUSTE
        D) PRUEBAS DE HIPOTESIS
        E) INTERVALOS DE CONFIANZA
        F) PREDICCION
        G) MULTICOLINEALIDAD
        H) HETEROSCEDASTICIDAD
        I) AUTOCORRELACION
        J) ANALISIS DE RESIDUALES
        K) SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES
        L) LISTAR DATOS
        M) SALVAR  X  y  Y  EN UN ARCHIVO
```

```
        N) RESULTADOS POR IMPRESORA
        O) MENU INICIAL
        S) SALIDA AL SISTEMA OPERATIVO (DOS)

  VARIABLES : CONJ_VAR : VARIABLE EN DONDE SE ALMACENAN LOS INDICES DE LAS
                         VARIABLES QUE FORMARAN PARTE DEL MODELO.
              INX      : VECTOR QUE EN PRINCIPIO GUARDA LOS INDICES DE LAS
                         VARIABLES QUE FORMARAN PARTE DEL MODELO.
              ORD      : VECTOR QUE EN PRINCIPIO CONSERVA EL ORDEN ORIGI-
                         NAL DE TODAS VARIABLES DE LA MATRIZ X.
              CAM_COL  : VARIABLE LOGICA.
                           VERDADERA.- CAMBIAR COLUMNAS DE LA MATRIZ X.
                           FALSA    .- NO CAMBIAR COLUMNAS DE LA MATRIZ X.
              CAM_Y    : VARIABLE LOGICA.
                           VERDADERA.- CAMBIAR COLUMNAS DE EL VECTOR Y.
                           FALSA    .- NO CAMBIAR COLUMNAS DE EL VECTOR Y.
              CALCULO  : VARIABLE LOGICA.
                           VERDADERA.- CALCULAR LA REGRESION.
                           FALSA    .- NO CALCULAR LA REGRESION.
}

PROCEDURE PRG_PRINCIPAL(N,AP:INTEGER;VAR X:MATRIZT1;VAR Y:MATRIZT2; VAR SNP:CHAR);

VAR  I,J,K,L                 : INTEGER;
     CONJ_VAR                : SET OF BYTE;
     INX,ORD                 : MATRIZT2IC;
     CAM_Y,CAM_COL,CALCULO   : BOOLEAN;
     OP                      : CHAR;

    {
       EL PROCEDIMIENTO "CAM_COLUMNA" CAMBIA UN SUBCONJUNTO DE COLUMNAS
     DE LA MATRIZ X A LAS PRIMERAS P POSICIONES DE LA MISMA MATRIZ.

     VARIABLES : VR : VARIABLE DE PASO.

     EXPLICACION : LO QUE HACE ESTE PROCEDIMIENTO ES CAMBIAR EL ORDEN
                   DE LAS COLUMNAS DE LA MATRIZ X POR UN NUEVO ORDEN QUE
                   SE GUARDA EN INX. EL PROCESO EMPIEZA CUANDO SE COMPA-
                   RAN LAS I-ESIMAS POSICIONES DE EL VECTOR INX Y DEL
                   VECTOR ORD (QUE GUARDA EL ORDEN ORIGINAL) SI SON DI-
                   FERENTES SE PASA A LA VARIABLE DE INDICE ORD(I) A LA
                   POSICION AP+1 (UNA COLUMNA DESPUES DE LA ULTIMA CO-
                   LUMNA OCUPADA) Y SE GUARDA EN LA POSICION ORD(I) LA
                   VARIABLE QUE TIENE LA POSICION INX(I), SE BUSCA EL
                   INDICE DE LA VARIABLE QUE SE ENCUENTRA EN AP+1 DENTRO
                   DE INX Y SE CAMBIA ESTA A SU NUEVO LUGAR NO SIN
                   ANTES HABER CAMBIADO A LA VARIABLE QUE OCUPABA ESTE
                   LUGAR A LA POSICION AP+1, ESTO NOS GENERA UNA CADENA
                   QUE SERA ROTA CUANDO SE REGRESE A LA POSICION I-ESIMA
                   Y QUE SE INICIA AVANZANDO A LA SIGUIENTE POSICION
                   HASTA ALCANZAR LA POSICION AP.
    }
```

```
PROCEDURE CAM_COLUMNA(VAR X:MATRIZT1;VAR ORD,INX:MATRIZT2IC; P,AP,N:INTEGER;VAR CAM_COL:BOOLEAN);
VAR     VR          : REAL;
        RMAS1,L,I  : INTEGER;
BEGIN
  FOR I:=1 TO AP DO
    IF ORD[I]<>INX[I] THEN
     BEGIN
       CAM_COL:=TRUE;
       RMAS1:=ORD[I];
       FOR L:=1 TO N DO
         BEGIN
           X[L,AP+1]:=X[L,ORD[I]];
           X[L,ORD[I]]:=X[L,INX[I]];
         END;
       ORD[I]:=INX[I];
       K:=I;
       REPEAT
         K:=K+1;
       UNTIL   INX[K]=RMAS1;
       WHILE RMAS1 <> ORD[K] DO
         BEGIN
           FOR L:=1 TO N DO
             BEGIN
               VR:=X[L,AP+1];
               X[L,AP+1]:=X[L,K];
               X[L,K]:=VR;
             END;
           ORD[K]:=RMAS1;
           RMAS1:=K;
           K:=0;
           REPEAT
             K:=K+1;
           UNTIL  INX[K]=RMAS1;
         END
     END;
  END;

(*****                PROGRAMA PRINCIPAL DE PRG_PRINCIPAL              *****)

BEGIN
  CALCULO   := FALSE;
  CAM_COL   := FALSE;
  P:=AP;
  IF VT1=1 THEN VT2:=P+2;
  H:=P+1-VT1;
  REPEAT
    CLRSCR;
    WRITELN(' ':10,'╔══════════════════════════════════════════════════╗');
    WRITELN(' ':10,'║             M E N U   P R I N C I P A L          ║');
    WRITELN(' ':10,'║                                                  ║');
    IF CALCULO=FALSE THEN HIGHVIDEO;
    WRITELN(' ':10,'║  A '^P' TABLA ANOVA Y ESTIMADORES                  ║');
    WRITELN(' ':10,'║  B '^P' SELECCION DE VARIABLES                     ║');
```

```
  NORMVIDEO;
  WRITELN(' ':10,'║  C '^P' CARENCIA DE AJUSTE                                   ║');
  WRITELN(' ':10,'║  D '^P' PRUEBAS DE HIPOTESIS                                 ║');
  WRITELN(' ':10,'║  E '^P' INTERVALOS DE CONFIANZA                              ║');
  WRITELN(' ':10,'║  F '^P' PREDICCION                                           ║');
  WRITELN(' ':10,'║  G '^P' MULTICOLINEALIDAD                                    ║');
  WRITELN(' ':10,'║  H '^P' HETEROSCEDASTICIDAD                                  ║');
  WRITELN(' ':10,'║  I '^P' AUTOCORRELACION                                      ║');
  WRITELN(' ':10,'║  J '^P' ANALISIS DE RESIDUALES                               ║');
  WRITELN(' ':10,'║  K '^P' SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES║');
  WRITELN(' ':10,'║  L '^P' LISTAR DATOS                                         ║');
  WRITELN(' ':10,'║  M '^P' SALVAR  X  y  Y  EN UN ARCHIVO                       ║');
  WRITELN(' ':10,'║  N '^P' RESULTADOS POR IMPRESORA                             ║');
  WRITELN(' ':10,'║                                                              ║');
  IF CALCULO=FALSE THEN HIGHVIDEO;
  WRITELN(' ':10,'║  O '^P' MENU INICIAL                                         ║');
  WRITELN(' ':10,'║  S '^P' SALIDA AL SISTEMA OPERATIVO (DOS)                    ║');
  NORMVIDEO;
  WRITELN(' ':10,'╚══════════════════════════════════════════════════════════════╝');

(*****   LA SALIDA O NO DE LOS RESULTADOS POR IMPRESORA ES CONTROLADA    *****
 *****   POR MEDIO DEL PROCEDIMIENTO "IMPRESORA".                        *****)

  IMPRESORA(IMPRESION,17,1);
  IF CALCULO THEN
    REPEAT
      GOTOXY(01,23); WRITE(' ':10,'> ESCOGE TU OPCION : ');
      SNP:=READKEY;WRITE(SNP); SNP := UPCASE(SNP);
    UNTIL SNP IN ['A','B','C','D','E','F','G','H','I','J','K','L','M','N','O','S']
  ELSE
    REPEAT
      GOTOXY(01,23); WRITE(' ':10,'> ESCOGE TU OPCION : ');
      SNP:=READKEY;WRITE(SNP); SNP := UPCASE(SNP);
    UNTIL SNP IN ['A','B','O','S'];
  CASE SNP OF
    'A' : BEGIN
            IF CALCULO=FALSE THEN
              BEGIN

(*****   LA MATRIZ MXTX Y EL VECTOR MXTY ES CALCULADO POR EL PROCEDI-    *****
 *****   MIENTO "CALCULO_SUMAS"                                          *****)

                CALCULO_SUMAS(MXTX,MXTY,X,Y,P,N);
                SOLUCIONAR:=TRUE;
                INVERTIR:=TRUE;

(*****   LA INVERSA DE MXTX Y LA SOLUCION AL SISTEMA  MXTX*MBEST=MXTY    *****
 *****   ES REALIZADO POR EL PROCEDIMIENTO "INVERSA"                     *****)

                INVERSA(MXTX,INVERSION,SOLUCIONAR,INVERTIR,H,MBEST,MXTY);
                IF INVERSION THEN
                  BEGIN
                    CALCULO:=TRUE;
```

```
(*****  LOS CALCULOS DE LAS SUMAS DE CUADRADOS Y DEL ESTADISTICO F SE  *****
 *****  REALIZAN POR MEDIO DEL PROCEDIMIENTO "TABLA_ANOVA".            *****)

                    TABLA_ANOVA(S2,S,SCREG,SCRES,SCT,MSCREG,F,R,RA,H,N,P,MBEST,MXTY,Y);
                  END;
              END;
            IF INVERSION THEN
               BEGIN
                CLRSCR;
                WRITELN('BETA ESTIMADA':47);
                WRITELN(' ':30,'┌─────────────────────┐');
                FOR K:=1 TO H DO
                WRITELN(' ':30,'│ B(',K-1+VT1:2,')=',MBEST[K]:14:7,' │');
                WRITELN(' ':30,'└─────────────────────┘');WRITELN;
                CONTINUAR;
                IF IMPRESION THEN
                   BEGIN
                     WRITELN(LST);WRITELN(LST);WRITELN(LST);
                     WRITELN(LST,'BETA ESTIMADA':45,^J,^M);
                     FOR K:=1 TO H DO
                        WRITELN(LST,' ':30,'B(',K-1+VT1:2,')=',MBEST[K]:12:7);
                  END;
                CLRSCR;GOTOXY(1,2);
                IF IMPRESION THEN
                  BEGIN
                    WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
                    WRITELN(LST,' ':25,'TABLA DE ANALISIS DE VARIANZA',^J,^M);
                  END;
                WRITELN(' ':25,'TABLA DE ANALISIS DE VARIANZA');

(*****   EL ESTADISTICO Y LA INFORMACION CONCERNIENTE A ESTE ES PRE-   *****
 *****   SENTADA EN UNA TABLA DE ANALISIS DE VARIANZA MEDIANTE EL       *****
 *****   PROCEDIMIENTO "ESCRIBE".                                       *****)

                ESCRIBE(1,P,N-H,N-1,1,SCREG,MSCREG,F,SCRES,S2,SCT,S2);
               END
            ELSE
               BEGIN
                 GOTOXY(10,23);
                 WRITE('> LA MATRIZ INVERSA DE (X'X) NO EXISTE'^g);CONTINUAR;
               END;
          END;

 (*****   DE LA MATRIZ X SE PUEDEN O NO SELECCIONAR LAS VARIABLES        *****
  *****   QUE FORMARAN PARTE DEL MODELO (A ELECCION DEL USUARIO), SI     *****
  *****   SE ELIGE SOLO UN SUBCONJUNTO DE P (P<AP) VARIABLES PARA        *****
  *****   FORMAR PARTE DEL MODELO, ENTONCES LAS COLUMNAS DE LA MA-       *****
  *****   TRIZ X A LAS CUALES SE REFIEREN ESAS VARIABLES SON CAMBIA-     *****
  *****   DAS A LAS P PRIMERAS POSICIONES DE LA MATRIZ X EN EL ORDEN     *****
  *****   QUE SE HAYAN LEIDO LOS INDICES Y DURANTE TODA LA EJECUCION     *****
  *****   DEL MENU PRINCIPAL SOLO LAS P PRIMERAS POSICIONES DE LA        *****
  *****   MATRIZ X SON CONSIDERADAS.                                     *****)

     'B' :BEGIN
```

```
IF CAM_COL THEN
  BEGIN
   FOR I:=1 TO AP DO
    BEGIN
     K:=0;
     REPEAT  K:=K+1;
     UNTIL   I=INX[K];
     ORD[I]:=K;
    END;
   FOR I:=1 TO AP DO INX[I]:=I;

(*****  LAS COLUMNAS DE LA MATRIZ X SON CAMBIADAS A LAS POSICIONES      *****
 *****  ORIGINALES QUE OCUPABAN AL ENTRAR A ESTE PROCEDIMIENTO, EL      *****
 *****  CAMBIO SE REALIZA POR MEDIO DEL PROCEDIMIENTO 'CAM_COLUMNA'.    *****)

   CAM_COLUMNA(X,INX,ORD,P,AP,N,CAM_COL);
  END;
IF AP>1 THEN
  BEGIN
   CLRSCR;      GOTOXY(1,3);
   REPEAT
     WRITELN('> CUANTAS DE LAS ',AP,' COLUMNAS DE LA MATRIZ X SERAN');
     WRITE('> INCLUIDAS EN LA REGRESION ? ');
     P:=LEER_INT;
   UNTIL (P>0) AND (P<=AP);
   WRITELN;
   IF P<AP THEN
     BEGIN
       CONJ_VAR:=[];
       FOR I:=1 TO P DO
         BEGIN
          REPEAT
            WRITE('> DAME LA  ',i,'a. COLUMNA: '); J:=LEER_INT;
          UNTIL (J IN [1..AP]) AND NOT (J IN CONJ_VAR);
          CONJ_VAR:=CONJ_VAR+[J];    INX[I]:=J;
         END;
       K:=0;
       FOR J:=1 TO AP DO
       IF NOT (J IN CONJ_VAR) THEN
          BEGIN   K:=K+1;
           INX[P+K]:=J;
          END;
       FOR I:=1 TO AP DO ORD[I]:=I;

(*****  LAS COLUMNAS DE LA MATRIZ X SON CAMBIADAS PONIENDOSE AL PRIN-   *****
 *****  CIPIO LAS COLUMNAS DE LAS VARIABLES QUE SE DESEAN EN EL MODELO  *****
 *****  EL CAMBIO SE REALIZA POR MEDIO DEL PROCEDIMIENNTO 'CAM_COLUMNA' *****)

       CAM_COLUMNA(X,ORD,INX,P,AP,N,CAM_COL);
     END
  END
ELSE  P:=1;
IF VT1=1 THEN VT2:=P+2;
H:=P+1-vt1;
```

```
        END;

{*****  LA PRUEBA DE HIPOTESIS PARA CARENCIA DE AJUSTE ES REALIZADA    *****
 *****  POR EL PROCEDIMIENTO "CARENCIA".                                *****}

      'C' :CARENCIA;

{*****  LAS OPCIONES PARA REALIZAR LAS DIFERENTES PRUEBAS DE HIPOTE-   *****
 *****  SIS SON DADAS EN EL PROCEDIMIENTO 'PRUEBAS_HIPOTESIS'.          *****}

      'D' : PRUEBAS_HIPOTESIS;

{*****  LAS OPCIONES Y LOS PROCEDIMIENTOS QUE CALCULAN LOS INTERVALOS  *****
 *****  DE CONFIANZA SE ENCUENTRAN EN EL PROCEDIMIENTO 'INT_CONF'.      *****}

      'E' : INT_CONF;

{*****  LAS PREDICIONES SON REALIZADAS POR EL PROCEDIMIENTO 'PREDIC-   *****
 *****  CION'.                                                          *****}

      'F' : PREDICCION;

{*****  LOS PROCEDIMIENTOS QUE HACEN PRUEBAS SOBRE MULTICOLINEALI-     *****
 *****  DAD SE HALLAN EN EL PROCEDIMIENTO 'MULTICOLINEALIDAD'.          *****}

      'G' : MULTICOLINEALIDAD;

{*****  LOS PROCEDIMIENTOS QUE HACEN PRUEBAS SOBRE HETEROSCEDASTI-     *****
 *****  CIDAD SE HALLAN EN EL PROCEDIMIENTO 'HETEROSCEDASTICIDAD'.      *****}

      'H' : HETEROSCEDASTICIDAD;

{*****  LOS PROCEDIMIENTOS QUE HACEN PRUEBAS SOBRE AUTOCORRELACION     *****
 *****  SE HALLAN EN EL PROCEDIMIENTO 'AUTOCORRELACION'.                *****}

      'I' : AUTOCORRELACION;

{*****  LOS PROCEDIMIENTOS USADOS EN EL ANALISIS DE RESIDUALES SE      *****
 *****  ENCUENTRAN EN EL PROCEDIMIENTO 'RESIDUALES'.                    *****}

      'J' : RESIDUALES;
      'K' : IF P<11 THEN

{*****  LA SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES SE REA-  *****
 *****  LIZA A TRAVEZ DEL PROCEDIMIENTO 'FURNIVAL'.                     *****}

             FURNIVAL
            ELSE
             BEGIN
              GOTOXY(1,22);
              WRITELN('> EL PROCEDIMIENTO ESTA LIMITADO A 10 VARIABLES');
              CONTINUAR;
             END;
```

```
(*****   EL LISTADO DE LOS N DATOS DE LAS VARIABLES DEL MODELO SE REA-  *****
 *****   LIZA POR MEDIO DEL PROCEDIMIENTO "PRINT".                      *****)

      'L' : PRINT(P);

(*****   LA MATRIZ X Y EL VECTOR Y SE GUARDAN EN UN ARCHIVO POR MEDIO    *****
 *****   DEL PROCEDIMIENTO 'SALVA'.                                      *****)

      'M' : SALVA;

(*****   LA SALIDA O NO DE LOS RESULTADOS POR IMPRESORA ES CONTROLADA    *****
 *****   POR MEDIO DEL PROCEDIMIENTO 'IMPRESORA'.                        *****)

      'N' : IMPRESORA(IMPRESION,17,2);
    END;
    IF SNP IN ['S'] THEN
       BEGIN
         GOTOXY(10,23);
         WRITELN('> PRESIONA <ENTER> PARA SALIR ');
         WRITELN('          > O CUALQUIER OTRA TECLA PARA REGRESAR AL MENU ANTERIOR ');
         SNP := READKEY;        SNP := UPCASE(SNP);
         IF SNP = ^M THEN   SNP := 'S'
       END;
  UNTIL  SNP IN ['S','O'];
  IF CAM_COL AND (SNP = 'O') THEN
     BEGIN
       FOR I:=1 TO AP DO
         BEGIN
          K:=0;
          REPEAT  K:=K+1;
          UNTIL   I=INX[K];
          ORD[I]:=K;
         END;
      FOR I:=1 TO AP DO INX[I]:=I;

(*****   LAS COLUMNAS DE LA MATRIZ X SON CAMBIADAS A LAS POSICIONES      *****
 *****   ORIGINALES QUE OCUPABAN AL ENTRAR A ESTE PROCEDIMIENTO, EL      *****
 *****   CAMBIO SE REALIZA POR MEDIO DEL PROCEDIMIENTO"CAM_COLUMNA".     *****)

      CAM_COLUMNA(X,INX,ORD,P,AP,N,CAM_COL);
    END;
END;

{
    EL PROCEDIMIENTO "OBTENER_XY" ES EL PROCEDIMIENTO CON EL CUAL EMPIEZA
   EL PROGRAMA DE REGRESION LINEAL MULTIPLE. MEDIANTE ESTE PROCEDIMIENTO
   SE DAN LAS OPCIONES PARA LA OBTENCION DE LA MATRIZX Y EL VECTOR Y.
   OPCIONES :
           A) LEER X.
           B) LEER Y.
           C) LEER Y y X.
           D) SIMULAR X.
           E) SIMULAR Y.
```

```
 VARIABLES : BX  : VARIABLE LOGICA.
                   VERDADERA.- OBTENCION DE LA MATRIZ X.
                   FALSA    .- NO OBTENCION DE LA MATRIZ X.
             BY  : VARIABLE LOGICA.
                   VERDADERA.- OBTENCION DE EL VECTOR Y.
                   FALSA    .- NO OBTENCION DE EL VECTOR Y.
                                                              }

PROCEDURE OBTENER_XY(VAR N,AP:INTEGER;VAR X:MATRIZT1;VAR Y:MATRIZT2);
VAR   OP,OP1            : CHAR;
      I,J,K,R,N1        : INTEGER;
      BX,BY,CON,BN      : BOOLEAN;
      CONJ_OP           : SET OF CHAR;

(*****                PROGRAMA PRINCIPAL DE OBTENER_XY.              *****)

BEGIN
  BX:=FALSE;
  BY:=FALSE;
  CON:=FALSE;
  SEGUNDA:=FALSE;
  BN:= FALSE;
  CONJ_OP := ['A','S'];
  GOTOXY(1,2);
  WRITELN('                              DATOS INICIALES');
  WRITELN('                              ===============');
  AP:=0;
  VT1:=0;
  VT2:=1;
  REPEAT
    CLRSCR;
    GOTOXY(1,4);
    WRITELN(' ':10,'╔══════════════════════════════════════════════╗');
    WRITELN(' ':10,'║             MENU OBTENCION DE DATOS          ║');
    WRITELN(' ':10,'║                                              ║');
    HIGHVIDEO;
    WRITELN(' ':10,'║  A '^P' NUMERO DE DATOS DE LA REGRESION (n)      ║');
    NORMVIDEO;
    IF BN THEN HIGHVIDEO;
    WRITELN(' ':10,'║  B '^P' REGRESION CON O SIN TERMINO INDEPENDIENTE ║');
    WRITELN(' ':10,'║  C '^P' LEER X                                   ║');
    WRITELN(' ':10,'║  D '^P' LEER Y                                   ║');
    WRITELN(' ':10,'║  E '^P' LEER Y y X                               ║');
    WRITELN(' ':10,'║  F '^P' SIMULAR X                                ║');
    IF NOT (BX AND BN) THEN NORMVIDEO;
    WRITELN(' ':10,'║  G '^P' SIMULAR Y                                ║');
    IF BX and BY THEN  HIGHVIDEO
    ELSE  NORMVIDEO;
    IF BX AND BN THEN NORMVIDEO;
    WRITELN(' ':10,'║  H '^P' CONTINUAR                                ║');
    WRITELN(' ':10,'║                                              ║');
    HIGHVIDEO;
    WRITELN(' ':10,'║  S '^P' SALIDA AL SISTEMA OPERATIVO (DOS)        ║');
    NORMVIDEO;
```

```
    WRITELN(' ':10,'╚══════════════════════════════════════════════╝');
    REPEAT
      GOTOXY(01,19);
      WRITE(' ':10,'> ESCOGE TU OPCION : '); OP:=READKEY;WRITE(OP);OP:=UPCASE(OP);
    UNTIL  OP IN CONJ_OP;
    CASE OP OF
      'A': BEGIN
             GOTOXY(1,22);
             IF NOT BN THEN
              BEGIN
               REPEAT
                WRITE('> CUANTOS DATOS SON: ');
                N:=LEER_INT;
               UNTIL  N>0;
               BN := TRUE;
               CONJ_OP := CONJ_OP + ['B','C','D','E','F','G','H']
              END
             ELSE
              BEGIN
               REPEAT
                WRITE('> CUANTOS DATOS SON: ');
                N1:=LEER_INT;
               UNTIL  N1>0;
               IF N1 > N THEN
               IF BX OR BY THEN
                BEGIN
                  WRITELN;
                  WRITELN('> COMO EL NUEVO VALOR DE (n)ES MAYOR, NECESITA LEER NUEVAMENTE A  Y y A X'
                  BY := FALSE;
                  BX := FALSE;
                  AP := 0;
                 CONTINUAR;
                END;
               N  := N1;
              END;
           END;
      'B': BEGIN
               REPEAT
                GOTOXY(1,22);
                WRITE('> MODELO DE REGRESION CON TERMINO INDEPENDIENTE ? (S/N)) ');
                OP1:=READKEY;
                WRITE(OP1);  OP1 := UPCASE(OP1);
               UNTIL OP1 IN ['S','N'];
               IF OP1 = 'S' THEN
                  BEGIN
                    VT1:=0;
                    VT2:=1;
                  END
               ELSE
                  VT1:=1;
           END;
      'C': BEGIN

(*****   PARA LEER SOLO LA MATRIZ X SE UTILIZA AL PROCEDIMIENTO "LEER_X" *****)
```

```
            LEER_X(N,AP,X,BX);
            WRITELN; CONTINUAR;
          END;
      'D': BEGIN
            K:=0;

(*****  PARA LEER SOLO EL VECTOR Y SE UTILIZA AL PROCEDIMIENTO "LEER_Y" *****)

            LEER_Y(N,K,Y,BY);
            WRITELN; CONTINUAR;
          END;
      'E': BEGIN

(*****  PARA LEER A LA MATRIZ X Y AL VECTOR Y SE UTILIZA EL PROCEDI-    *****
 *****  MIENTO "LEER_XY".                                              *****)

           LEER_XY(N,AP,X,Y,BX,BY);
           WRITELN; CONTINUAR;
          END;
      'F': BEGIN

(*****  PARA SIMULAR COLUMNAS DE LA MATRIZ X SE UTILIZA EL PROCEDI-    *****
 *****  MIENTO "SIMULAR_X".                                            *****)

           SIMULAR_X(X,N,AP);
           IF AP > 0 THEN  BX:=true;
           WRITELN; CONTINUAR;
          END;
      'G': BEGIN
           IF BX  THEN
             BEGIN

(*****  PARA SIMULAR AL VECTOR Y SE UTILIZA EL PROCEDIMIENTO "SIMULAR_Y" *****)

               SIMULAR_Y(Y,N,AP,SEGUNDA);
               IF SEGUNDA THEN BY:=TRUE;
             END
           ELSE WRITELN(^J^M,'> SE NECESITA AL MENOS UNA COLUMNA DE X',^g);
           WRITELN; CONTINUAR;
          END;
      'H': IF BX AND BY THEN CON:=TRUE;
      'S': BEGIN
           GOTOXY(10,23);
           WRITELN('> PRESIONA <ENTER> PARA SALIR ');
           WRITELN('          > O CUALQUIER OTRA TECLA PARA REGRESAR AL MENU ANTERIOR ');
           SNP := READKEY;        SNP := UPCASE(SNP);
           IF SNP = ^M THEN
             BEGIN
               CON := TRUE;
               SNP := 'S';
             END;
          END;
    END;
```

```
  UNTIL CON;
END;
{╔══════════════════════════════════════════════════════════════╗
 ║                                                              ║
 ║        P R O G R A M A     P R I N C I P A L                 ║
 ║                                                              ║
 ╚══════════════════════════════════════════════════════════════╝}

BEGIN                    {*** PROGRAMA PRICIPAL ***}

  CLRSCR;
  GOTOXY(1,3);
  WRITELN('                 ╔════════════════════════════════════════════╗');
  WRITELN('                 ║                                            ║');
  WRITELN('                 ║              P A Q U E T E  D E            ║');
  WRITELN('                 ║                                            ║');
  WRITELN('                 ║  R E G R E S I O N  L I N E A L  M U L T I P L E ║');
  WRITELN('                 ║                                            ║');
  WRITELN('                 ╚════════════════════════════════════════════╝');
  DELAY(1500);
SNP := ' ';

{*****  LA OBTENCION DE LA MATRIZ X Y EL VECTOR Y ES REALIZADA POR EL  *****
 *****  PROCEDIMIENTO "OBTENER_XY".                                     *****}

OBTENER_XY(N,AP,X,Y);
    IMPRESION:=FALSE;
IF SNP = 'S' THEN EXIT;
REPEAT
  REPEAT  CLRSCR;
    GOTOXY(01,04);
    WRITELN(' ':10,'╔════════════════════════════════════════╗');
    WRITELN(' ':10,'║            M E N U  I N C I A L        ║');
    WRITELN(' ':10,'║                                        ║');
    WRITELN(' ':10,'║  A '^P' GRAFICAS                        ║');
    WRITELN(' ':10,'║  B '^P' SIMULAR  X                      ║');
    WRITELN(' ':10,'║  C '^P' SIMULAR  Y                      ║');
    WRITELN(' ':10,'║  D '^P' EDICION DE DATOS                ║');
    WRITELN(' ':10,'║  E '^P' CONSTRUCCION DE VAR. DICOTOMAS  ║');
    WRITELN(' ':10,'║  F '^P' TRANSFORMACIONES                ║');
    WRITELN(' ':10,'║  G '^P' CAMBIO DE VARIABLE DEPENDIENTE  ║');
    WRITELN(' ':10,'║  H '^P' LISTADO DE DATOS                ║');
    WRITELN(' ':10,'║  I '^P' SALIDA A LA IMPRESORA           ║');
    WRITELN(' ':10,'║                                        ║');
    WRITELN(' ':10,'║  N '^P' NUEVA REGRESION                 ║');
    WRITELN(' ':10,'║  P '^P' MENU PRINCIPAL                  ║');
    WRITELN(' ':10,'║  S '^P' SALIDA AL SISTEMA OPERATIVO (DOS) ║');
    WRITELN(' ':10,'╚════════════════════════════════════════╝');

{*****  LA SALIDA O NO DE LOS RESULTADOS POR IMPRESORA ES CONTROLADA  *****
 *****  POR MEDIO DEL PROCEDIMIENTO "IMPRESORA".                      *****}

    IMPRESORA(IMPRESION,15,1);
    REPEAT
```

```
      GOTOXY(10,22);
      WRITE('> OPCION : ');
      SNP := READKEY;WRITELN(SNP);
      SNP := UPCASE(SNP);
   UNTIL SNP IN ['A','B','C','D','E','F','G','H','N','P','S',^M,'I'];
   CASE SNP OF
    'A': REPEAT
          IF AP > 1 THEN
           BEGIN
             CLRSCR; GOTOXY(1,24);
             WRITE('PARA SALIR INTRODUCIR [0] '); GOTOXY(1,5);
             FOR I:= 1 TO AP DO
              WRITELN('  ':5,'X[',I:1,']  --> [',I:1,']  vs.  Y OBSERVADA');WRITELN;
              REPEAT  WRITE('> CUAL GRAFICAS QUIERES ?  ');  J:=LEER_INT;
              UNTIL (0<=J)  AND (J<=AP);
           END
          ELSE J := 1;
          IF J <> 0 THEN
           BEGIN
             FOR I:=1 TO N DO
               XT[I]:=X[I,J];

{*****   LA GRAFICA DE EL VECTOR Y VERSUS ALGUNA COLUMNA DE LA MATRIZ   *****
 *****   X ES REALIZADA POR EL PROCEDIMIENTO "GRAFICA".                  *****}

             GRAFICA(XT,Y,N);
             GOTOXY(1,1); WRITE('GRAFICA DE:  X(',j:1,') vs. Y OBSERVADA ');
             GOTOXY(20,24); CONTINUAR;
             GOTOXY(20,24); WRITE('         > QUIERES OTRA GRAFICA ? (S/N)');
             YN:=READKEY;  WRITELN(YN); YN := UPCASE(YN)
           END
          ELSE YN := 'N';
         UNTIL YN <> 'S';
     'B': BEGIN

{*****   LA SIMULACION DE LAS VARIABLES INDEPENDIENTES SE REALIZA ME-   *****
 *****   DIANTE EL PROCEDIMIENTO "SIMULAR_X".                           *****}

             SIMULAR_X(X,N,AP);
             WRITELN;WRITELN;
             WRITE('> OPERACION REALIZADA    ');

{*****   CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE   *****
 *****   LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA       *****}

             CONTINUAR;
             END;
     'C': BEGIN

{*****   LA SIMULACION DE LA VARIABLE DEPENDIENTE SE REALIZA MEDIANTE   *****
 *****   EL PROCEDIMIENTO "SIMULAR_Y".                                  *****}

             SIMULAR_Y(Y,N,AP,SEGUNDA);
            WRITELN;WRITELN;
```

```
            WRITE('> OPERACION REALIZADA    ');

(*****  CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE    *****
 *****  LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA         *****)

            CONTINUAR;
            END;

(*****  LA EDICION DE LOS DATOS ES REALIZADA POR EL PROCEDIMIENTO "E-   *****
 *****  DITAR".                                                          *****)

     'D': EDITAR_XY(X,Y,N,AP);

(*****  LA CREACION DE VARIABLES DICOTOMAS SE REALIZA MRDIANTE EL PRO- *****
 *****  CEDIMIENTO "DICOTOMAS".                                          *****)

     'E': DICOTOMAS(X,N,AP);

(*****  LAS TRANFORMACIONES DE VARIABLES SE LLEVAN A CABO MEDIANTE      *****
 *****  EL PROCEDIMIENTO 'TRANSFORMACION".                               *****)

     'F': TRANSFORMACION;
     'G':BEGIN
          GOTOXY(10,22);
          WRITE('> SE DESEA CAMBIAR LA VARIABLE DEPENDIENTE ORIGINAL ? (S/N) ');
          YN:=READKEY;  WRITELN(YN);  YN:= UPCASE(YN); WRITELN;
          IF YN='S' THEN
             BEGIN
              REPEAT
                 GOTOXY(10,22);
                 WRITELN('> EN QUE COLUMNA DE LA MATRIZ X SE ENCUENTRA');
                 WRITE(' ':10,'> LA NUEVA VARIABLE DEPENDIENTE : ');
                 READLN(I);
              UNTIL (I>0) AND (I<=AP);        WRITELN;
              FOR J:=1 TO N DO
                BEGIN
                  XT[J] := X[J,I];
                  X[J,I]:=Y[J];
                  Y[J]:=XT[J];
                END;
             END;
          END;

(*****  EL LISTADO DE LOS N DATOS DE LAS VARIABLES DEL MODELO SE REA-  *****
 *****  LIZA POR MEDIO DEL PROCEDIMIENTO "PRINT".                        *****)

     'H': PRINT(AP);

(*****  LA SALIDA O NO DE LOS RESULTADOS POR IMPRESORA ES CONTROLADA    *****
 *****  POR MEDIO DEL PROCEDIMIENTO "IMPRESORA".                         *****)

     'I' : IMPRESORA(IMPRESION,15,2);
     'N': BEGIN
          GOTOXY(10,21);
```

```
        WRITELN('> ESTA OPCION TE PERMITE COMENZAR UNA NUEVA REGRESION');
        WRITELN('          > POR LO QUE BORRA LA INFORMACION ACTUAL'); WRITELN;
        WRITELN('          > PRESIONA <ENTER> PARA CONTINUAR');
        WRITELN('          > O CUALQUIER OTRA TECLA PARA REGRESAR AL MENU ANTERIOR');
        SNP := READKEY;   SNP:= UPCASE(SNP);
        IF SNP= ^M THEN
          BEGIN
            N  := 0;
            AP := 0;

{*****   LA OBTENCION DE LA MATRIZ X Y EL VECTOR Y ES REALIZADA POR EL   *****
 *****   PROCEDIMIENTO "OBTENER_XY".                                     *****}

            OBTENER_XY(N,AP,X,Y);
            IF SNP = 'S' THEN EXIT;
          END
        ELSE SNP := 'N';
       END;
   END;
  UNTIL SNP IN ['P','S'];
  IF SNP IN ['S'] THEN
   BEGIN
     GOTOXY(10,23);
     WRITELN('> PRESIONA <ENTER> PARA SALIR ');
     WRITELN('          > O CUALQUIER OTRA TECLA PARA REGRESAR AL MENU ANTERIOR');
     SNP := READKEY;        SNP := UPCASE(SNP);
     IF SNP = ^M THEN   SNP := 'S'
   END;

{*****   EL MENU PRINCIPAL SE ENCUENTRA Y ES MANEJADO POR EL PROCEDI-    *****
 *****   MIENTO "PRG_PRINCIPAL".                                         *****}

  IF SNP IN ['P'] THEN PRG_PRINCIPAL(N,AP,X,Y,SNP);
UNTIL SNP IN ['S'];
  CLRSCR;
END.

{
     LA UNIDAD "GLOBAL" CONTIENE LA DEFINICION DE LAS CONSTANTES, VECTO-
   RES, MATRICES, ENTEROS, REALES Y DEMAS TIPOS DE VARIABLES GLOBALES USA-
   DAS POR TODAS LAS DEMAS UNIDADES DEL SISTEMA Y POR EL PROGRAMA PRINCI-
   PAL.
}

UNIT GLOBAL;
INTERFACE

CONST NMAXREG=200;
      NMAXCOL=30;

TYPE  MATRIZT1=ARRAY[1..NMAXREG,1..NMAXCOL] OF REAL;
      MATRIZT1C=ARRAY[1..NMAXCOL,1..NMAXCOL] OF REAL;
```

```
     MATRIZT2=ARRAY[1..NMAXREG] OF REAL;
     MATRIZT2C=ARRAY[1..NMAXCOL] OF REAL;
     MATRIZT2I=ARRAY[1..NMAXREG] OF INTEGER;
     MATRIZT2IC=ARRAY[1..NMAXCOL] OF INTEGER;
     R_VEC_MAT = RECORD
                   Y1:REAL;
                   X1:MATRIZT2;
                 END;
     FILNOM = STRING[12];

VAR  MXTX                                            : MATRIZT1C;
     X                                               : MATRIZT1;
     MXTY,MBEST                                      : MATRIZT2C;
     Y,XT                                            : MATRIZT2;
     VT1,VT2,I,J,AP,N,P,H                            : INTEGER;
     SEGUNDA,INVERSION,SOLUCIONAR,INVERTIR,IMPRESION : BOOLEAN;
     RA,S,S2,SCREG,SCRES,SCT,MSCREG,F,R              : REAL;
     YN,YNP,SNP                                      : CHAR;


{
  VARIABLES
  PRINCIPALES.- X        : MATRIZ DE DATOS (VARIABLES INDEPENDIENTES)
                           Y TAMBIEN MATRIZ DEL MODELO.
                Y        : VECTOR DE DATOS (VARIABLE DEPENDIENTE).
                MXTX     : MATRIZ CUADRADA QUE RESULTA DE LA MULTIPLI-
                           CACION DE LA TRANSPUESTA DE LA MATRIZ DEL
                           MODELO X POR LA MISMA MATRIZ X.
                MXTY     : VECTOR QUE RESULTA DE LA MULTIPLICACION DE
                           LA TRANSPUESTA DE LA MATRIZ DEL MODELO
                           POR EL VECTOR Y.
                MBEST    : VECTOR DONDE SE GUARDAN LAS BETAS ESTIMADAS
                XT       : VECTOR AUXILIAR PARA CUANDO SE DESEA CONSI-
                           DERAR UNA VARIABLE INDEPENDIENTE DE LA MA-
                           TRIZ DE DATOS.
                VT1      : VARIABLE QUE NOS AYUDA A MANEJAR EL MODELO
                           CON O SIN TERMINO INDEPENDIENTE.
                VT2      : VARIABLE QUE NOS AYUDA A MANEJAR EL MODELO
                           CON O SIN TERMINO INDEPENDIENTE.
                AP       : NUMERO TOTAL DE VARIABLES INDEPENDIENTES
                N        : NUMERO DE OBSERVACIONES.
                P        : NUMERO DE VARIABLES INDEPENDIENTES EN EL MO-
                           DELO.
                H        : NUMERO DE VARIABLES INDEPENDIENTES MAS UNO
                SEGUNDA  : VARIABLE QUE NOS AYUDA A SABER SI SE VA A
                           SIMULAR LA VARIABLE DEPENDIENTE UNA VEZ QUE
                           QUE YA SE TIENE ESTA.
                INVERSION : VARIABLE LOGICA.
                             VERDADERA.- EL SISTEMA TIENE INVERSA.
                             FALSA    .- EL SISTEMA NO TIENE INVERSA.
                SOLUCIONAR: VARIABLE LOGICA.
                             VERDADERA.- OBTENER SOLUCION DEL SISTEMA.
                             FALSA    .- NO OBTENER SOL. DEL SISTEMA.
                INVERTIR  : VARIABLE LOGICA.
```

```
                    VERDADERA.- OBTENER INVERSA DEL SISTEMA.
                    FALSA    .- NO OBTENER INV. DEL SISTEMA.
       IMPRESION : VARIABLE LOGICA.
                    VERDADERA.- IMPRIMIR RESULTADOS.
                    FALSA    .- NO IMPRIMIR RESULTADOS.
       SCREG     : SUMA DE CUADRADOS DE LA REGRESION
       MSCREG    : MEDIA DE LA SUMA DE CUADRADOS DE LA REGRE-
                    SION
       SCRES     : SUMA DE CUADRADOS DE LOS RESIDUALES
       SCT       : SUMA DE CUADRADOS TOTALES
       S         : DESVIACION ESTANDAR ESTIMADA
       S2        : VARIANZA ESTIMADA
       F         : ESTADISTICO F
       R         : COEFICIENTE DE DETERMINACION
       RA        : COEFICIENTE DE DETERMINACION AJUSTADO
}

IMPLEMENTATION

END.
{
   LA UNIDAD "UTILERIA" CONTIENE LAS UTILERIAS DE EL SISTEMA, QUE SON
  PROCEDIMIENTOS UTILIZADOS MAS DE UNA VEZ EN EL SISTEMA.
}

UNIT UTILERIA;
INTERFACE
  USES CRT,PRINTER,UGLOBAL;
  PROCEDURE CONTINUAR;
  FUNCTION LEER_REAL:REAL;
  FUNCTION LEER_INT:INTEGER;
  PROCEDURE LEER_VEC_REAL(VAR VEC_REAL:MATRIZT2;PP:INTEGER);
  PROCEDURE IMPRESORA(VAR IMPRESION:BOOLEAN;LIN,CON3:INTEGER);
  PROCEDURE ESTIMADA(VAR E,YEST:MATRIZT2);
  FUNCTION POT(B,E:REAL):REAL;
  FUNCTION N_P(U:REAL):REAL;
  FUNCTION TS_P(T:REAL;V:INTEGER):REAL;
  FUNCTION N_I(P:REAL):REAL;
  FUNCTION T_INV(P:REAL;V:INTEGER):REAL;
  FUNCTION F_P(F,V1,V2:REAL):REAL;
  PROCEDURE GRAFICA(VAR XX,YY:MATRIZT2; N:INTEGER);
  PROCEDURE  CALCULO_SUMAS(VAR A:MATRIZT1C; VAR B:MATRIZT2C;VAR X1:MATRIZT1;
  VAR Y1:MATRIZT2;P1,N1:INTEGER);
  PROCEDURE INVERSA(VAR A:MATRIZT1C;VAR INVERSION,SOLUCIONAR,INVERTIR:BOOLEAN;
  HL:INTEGER;VAR YSOL:MATRIZT2C;VAR B:MATRIZT2C);
  PROCEDURE QUICK(VAR A:MATRIZT2;VAR B:MATRIZT2I;VAR NELEM:INTEGER);
  PROCEDURE TABLA_ANOVA(VAR S2,S,SCREG,SCRES,SCT,MSCREG,F,R,RA:REAL;VAR H,N,P:INTE
  GER;VAR MTW,MWTY:MATRIZT2C;VAR YM:MATRIZT2);
  PROCEDURE ESCRIBE(CONTROL,P1,P2,P3,P4:INTEGER;V1,V2,V3,V4,V5,V6,V7:REAL);
  PROCEDURE ESTADISTICAS(VAR MWTW:MATRIZT1C;VAR MBESTW:MATRIZT2C;VAR SL:REAL;VAR HL:INTEGER);
  PROCEDURE CAMBIO_MATRIZ(VAR MWTW:MATRIZT1C;VAR MWTY:MATRIZT2C;VAR INW:MATRIZT2IC
  ;VAR INWY,O:INTEGER;VAR CAMBIOY:BOOLEAN);
  PROCEDURE MULTIPLICACION(VAR MR,M1,M2:MATRIZT1C;VAR RE1,CO,RE2:INTEGER);
```

```
  PROCEDURE D_XXXX(VAR VI:MATRIZT2);
  PROCEDURE PRINT(VAR P:INTEGER);

 IMPLEMENTATION

{
   EL PROCEDIMIENTO "CONTINUAR" DETIENE LA EJECUCION DEL PROGRAMA HASTA
  QUE SE OPRIMA ALGUNA TECLA.
}

PROCEDURE CONTINUAR;
BEGIN
  WRITE('> PARA CONTINUAR OPRIMA ENTER');
  TEXTATTR:=WHITE;
  HIGHVIDEO;
  WRITE(' ( <—┘ )');
  NORMVIDEO;
  READLN;
END;

{
    LA FUNCION "LEER_REAL" VALIDA LA LECTURA DE UN REAL.

  VARIABLES : M2     : VARIABLE QUE CONTIENE AL REAL A VALIDAR EN FORMA DE
                       CADENA DE CARACTERES.
              RESULT : SI EN LA VALIDACION SE ENCONTRO ERROR, ESTA VARIA-
                       BLE GUARDA LA POSICION DONDE SE ENCONTRO EL ERROR.
              R      : SI NO SE ENCUENTRA ERROR EN LA VALIDACION, ESTA VA-
                       RIABLE GUARDA EL VALOR DEL REAL.
}

FUNCTION LEER_REAL:REAL;
VAR    M2         : STRING[80];
       RESULT     : INTEGER;
       R          : REAL;
BEGIN
  REPEAT
    READLN(M2);
    VAL(M2,R,RESULT);
    IF RESULT<>0 THEN WRITELN('^':RESULT,CHAR(7),'  ERROR');
  UNTIL (RESULT=0) AND (LENGTH(M2)>0);
  LEER_REAL:=R;
END;

{
    LA FUNCION "LEER_INT" VALIDA LA LECTURA DE UN REAL.

  VARIABLES : M2     : VARIABLE QUE CONTIENE AL REAL A VALIDAR EN FORMA DE
                       CADENA DE CARACTERES.
              RESULT : SI EN LA VALIDACION SE ENCONTRO ERROR, ESTA VARIA-
                       BLE GUARDA LA POSICION DONDE SE ENCONTRO EL ERROR.
              I      : SI NO SE ENCUENTRA ERROR EN LA VALIDACION, ESTA VA-
                       RIABLE GUARDA EL VALOR DEL ENTERO.
}
```

```
FUNCTION LEER_INT:INTEGER;
VAR    M2            : STRING[10];
       RESULT,I      : INTEGER;
BEGIN
  REPEAT
    READLN(M2);
    VAL(M2,I,RESULT);
    IF RESULT<>0 THEN WRITELN('^':RESULT,CHAR(7),'  ERROR: DA UN NUMERO ENTERO');
  UNTIL (RESULT=0) AND (LENGTH(M2)>0);
  LEER_INT:=I;
END;

{
   LA FUNCION "LEER_VEC_REAL" VALIDA LA LECTURA DE UN VECTOR REAL EN FOR-
  MATO LIBRE.

  VARIABLES : M2     : VARIABLE QUE CONTIENE AL VECTOR A VALIDAR EN FORMA
                       DE CADENA DE CARACTERES.
            : M      : VARIABLE QUE CONTIENE A UN COMPONENTE DEL VECTOR
                       A VALIDAR EN FORMA DE CADENA DE CARACTERES.
              RESULT : SI EN LA VALIDACION SE ENCONTRO ERROR, ESTA VARIA-
                       BLE GUARDA LA POSICION DONDE SE ENCONTRO EL ERROR.
              R      : SI NO SE ENCUENTRA ERROR EN LA VALIDACION, ESTA VA-
                       RIABLE GUARDA EL VALOR DEL REAL.
}

PROCEDURE LEER_VEC_REAL(VAR VEC_REAL:MATRIZT2;PP:INTEGER);
VAR     K,I,RESULT,C       : INTEGER;
        M,M2               : STRING[80];
BEGIN
  I:=0;
  K:=0;
  REPEAT
    READLN(M2);
  UNTIL M2<>'';
  REPEAT
     C:=POS(',',M2);
     IF C<>0 THEN  M:=COPY(M2,1,C-1)
            ELSE M:=M2;
     K:=K+1;
     VAL(M,VEC_REAL[K],RESULT);
     IF RESULT<>0 THEN
       BEGIN
         WRITELN('^':I+RESULT,CHAR(7),'  ERROR INTRODUCIR EL REGISTRO DE NUEVO');
         READLN(M2);
         K:=0;
         I:=0;
       END
     ELSE
       BEGIN
         I:=LENGTH(M)+1+I;
         DELETE(M2,1,LENGTH(M)+1);
       END;
```

```
  UNTIL K=PP;
END;

{=====================================================================
 |   EL PROCEDIMIENTO 'PRINT' DESPLIEGA POR PANTALLA Y/O POR IMPRESORA A |
 |  LA MATRIZ X Y AL VECTOR Y.                                          |
 |                                                                      |
 |  VARIABLES : CONT : VARIABLE PARA EL CONTROL DEL NUMERO DE COLUMNAS DES-|
 |                     PLEAGADAS EN PANTALLA O EN UNA HOJA.             |
 |              NOD  : NUMERO DE PANTALLAS O DE HOJAS EN LAS QUE SE DESPLE-|
 |                     GARA TODA LA INFORMACION.                        |
 =====================================================================}

PROCEDURE PRINT(VAR P:INTEGER);
VAR     NOD,I,J,JI,JF,CONT      :INTEGER;
BEGIN
  CLRSCR;
  WRITE('Y':13,'    ');
  NOD:=P DIV 5;
  JI:=1;
  JF:=0;
  CONT:=1;
  REPEAT
  IF CONT>NOD THEN
     BEGIN
       JI:=JF+1;
       JF:=P;
     END
  ELSE
     IF CONT=1 THEN JF:=4
     ELSE
       BEGIN
         JI:=JF+1;
         JF:=JF+4;
       END;
  FOR J:=JI TO JF DO
     WRITE('X(':12,J:2,')');   WRITELN;
  FOR I:=1 TO N DO
    BEGIN
    IF I MOD 23 = 0 THEN
       BEGIN
         CONTINUAR;
         CLRSCR;
         IF CONT=1 THEN WRITE('Y':13,'    ');
         FOR J:=JI TO JF DO
         WRITE('X(':12,J:2,')');   WRITELN;
       END;
    WRITE(I:3);
    IF CONT=1 THEN WRITE('   ',Y[I]:12:4);
    FOR J:=JI TO JF DO
       WRITE('   ',X[I,J]:12:4);
    WRITELN;
    END;
    CONTINUAR;
```

```
    CLRSCR;
  CONT:=CONT+1;
  UNTIL (CONT-1)>NOD;
  IF IMPRESION THEN
   BEGIN
     WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
     WRITELN(LST,'LISTADO DE DATOS':45,^J,^M);
     WRITE(LST,'Y':13,'    ');
    JI:=1;
    JF:=0;
    CONT:=1;
    REPEAT
      IF CONT>NOD THEN
        BEGIN
          JI:=JF+1;
          JF:=P;
        END
      ELSE
        IF CONT=1 THEN JF:=4
        ELSE
         BEGIN
           JI:=JF+1;
           JF:=JF+4;
         END;
      FOR J:=JI TO JF DO
      WRITE(LST,'X(':12,J:2,')');   WRITELN(LST);
      FOR I:=1 TO N DO
        BEGIN
          WRITE(LST,I:3);
          IF CONT=1 THEN WRITE(LST,'    ',Y[I]:12:4);
          FOR J:=JI TO JF DO
            WRITE(LST,'    ',X[I,J]:12:4);
          WRITELN(LST);
        END;
      CONT:=CONT+1;
    UNTIL (CONT-1)>NOD;
 END;
END;

{
    EL PROCEDIMIENTO 'IMPRESORA' ES UN PROCEDIMIENTO QUE SE UTILIZA PARA
  ACTIVAR LA SALIDA DE RESULTADOS POR LA IMPRESORA.
}

PROCEDURE IMPRESORA(VAR IMPRESION:BOOLEAN;LIN,CON3:INTEGER);
BEGIN
  IF CON3=1 THEN
   IF IMPRESION THEN
     BEGIN
      HIGHVIDEO;
      GOTOXY(48,LIN);
      WRITE('█ ACTIVADO █');
      NORMVIDEO;
     END
```

```
  ELSE
    BEGIN
     HIGHVIDEO;
     GOTOXY(48,LIN);
     WRITE(' DESACTIVADO ');
     NORMVIDEO;
    END
  ELSE
   IF IMPRESION THEN
     IMPRESION:=FALSE
   ELSE
    BEGIN
     IMPRESION:=TRUE;
     GOTOXY(1,23);
     WRITELN('» ADVERTENCIA: LA IMPRESORA DEBE ESTAR ENCENDIDA ');
     continuar;
    END;
END;

{┌─────────────────────────────────────────────────────────────────┐
 │  EL PROCEDIMIENTO "ESTIMADA" CALCULA EL VECTOR DE RESIDUALES Y EL VEC- │
 │ TOR DE Yi ESTIMADA.                                             │
 └─────────────────────────────────────────────────────────────────┘}

PROCEDURE ESTIMADA(VAR E,YEST:MATRIZT2);
VAR J,I:INTEGER;

BEGIN
  FOR J:=1 TO N DO
    BEGIN
      IF VTI=0 THEN  YEST[J]:=MBEST[1]
               ELSE  YEST[J]:=0;
      FOR I:=1 TO P DO
        YEST[J]:=MBEST[I+1-VTI]*X[J,I]+YEST[J];
      E[J]:=Y[J]-YEST[J];
    END;
END;

{┌─────────────────────────────────────────────────────────────────┐
 │  LA FUNCION "POT" ELEVA UN NUMERO REAL B A UNA POTENCIA REAL E.  │
 │                                                                 │
 │ VARIABLES : B : NUMERO REAL A ELEVAR.                           │
 │             E : POTENCIA.                                       │
 └─────────────────────────────────────────────────────────────────┘}

FUNCTION POT(B,E:REAL):REAL;
VAR    EX : INTEGER;
BEGIN
  IF B > 0 THEN
    BEGIN
      IF E = 0 THEN POT := 1
      ELSE
        IF E > 0 THEN POT := EXP(E*LN(B))
        ELSE          POT := 1/EXP(-E*LN(B))
```

```
    END
  ELSE
    BEGIN
      IF B < 0 THEN
        BEGIN
          IF E = 0 THEN POT := 1
          ELSE
            IF E = INT(E) THEN
              BEGIN
                EX := ROUND(E);
                IF EX > 0 THEN
                  BEGIN
                    IF ODD(EX) THEN POT := -EXP(E*LN(-B))
                    ELSE            POT := EXP(E*LN(-B))
                  END
                ELSE
                  BEGIN
                    IF ODD(EX) THEN POT := -EXP(E*LN(-B))
                    ELSE            POT := EXP(E*LN(-B))
                  END
              END
            ELSE
              BEGIN
                WRITELN('BASE NEGATIVA Y EXPONENTE NO ENTERO');
                POT := 9E+10
              END
        END
      ELSE
        BEGIN
          IF E > 0 THEN POT := 0
          ELSE
            BEGIN
              WRITELN('ERROR BASE CEOR Y POTENCIA CERO O NEGATIVA');
              POT := 9E+10
            END
        END
    END
END;

{
   LA FUNCION "N_P" CALCULA LA PROBABILIDAD NORMAL PARA UN CIERTO VALOR.
  LA PROBABILIDAD SE CALCULA POR MEDIO DE UNA ECUACION POLINOMIAL.

  VARIABLES : U : VALOR AL QUE SE LE CALCULA LA PROBABILIDAD NORMAL.
              B : VECTOR QUE CONTIENE LOS COEFICIENTES DEL POLINOMIO.
}

FUNCTION N_P(U:REAL):REAL;
CONST    P = 0.2316419;
VAR      I       : INTEGER;
         T,Z,Q   : REAL;
         B       : ARRAY[1..5] OF REAL;
BEGIN
  B[1]:=  0.31938153;
```

```
  B[2]:= -0.356563782;
  B[3]:=  1.781477937;
  B[4]:= -1.821255978;
  B[5]:=  1.330274429;
  Z:=0.3989423*(EXP(-SQR(U)/2));
  T:=1/(P*U+1);
  Q:=0;
  FOR I:=1 TO 5 DO

(*****          LA FUNCION "POT" ELEVA T A LA POTENCIA I.          *****)

    Q:=Q+B[I]*POT(T,I);
  N_P:=Q*Z;
END;

{=====================================================================
   LA FUNCION "TS_P" CALCULA LA PROBABILIDAD PARA UN VALOR REAL CON BASE
  EN LA FUNCION DE DISTRIBUCION T DE STUDENT.

  VARIABLES : T  : VALOR AL QUE SE LE CALCULA LA PROBABILIDAD.
              V  : GRADOS DE LIBERTAD.
=====================================================================}

FUNCTION TS_P(T:REAL;V:INTEGER):REAL;
VAR    I          : INTEGER;
       Y,X,X1     : REAL;
       A,B        : ARRAY[1..5] OF REAL;
BEGIN
 IF (V=1)  OR  ((V=2) AND (T>4))  OR  ((V=3) AND (T>5)) THEN
    BEGIN
       A[1]:=0.3183;
       A[2]:=0.4991;
       A[3]:=1.1094;
       A[4]:=3.0941;
       A[5]:=9.948;
       B[1]:=0.0000;
       B[2]:=0.0518;
       B[3]:=-0.0460;
       B[4]:=-2.756;
       B[5]:=-14.05;

(*****       LA FUNCION "POT" ELEVA T A LA POTENCIA V Y A V+1.       *****)

       X:=A[V]/POT(T,V) + B[V]/POT(T,V+1);
       K:=4;
   END
 ELSE
   IF V<25 THEN
      BEGIN
        K:=2;
        A[1]:= 0.196854;
        A[2]:= 0.115194;
        A[3]:= 0.000344;
        A[4]:= 0.019527;
```

```
        Y:=2/(9*V);

{*****          LA FUNCION "POT" ELEVA T A LA POTENCIA 2/3.          *****}

        X:=(POT(T,2/3)*(1-Y) )- 7/9;
        X1:=SQRT((2/9)+POT(T,4/3)*Y);
        X:=X/X1;
        Y:=1;
        FOR I:=1 TO 4 DO
          Y:=Y+(A[I]*POT(X,I) );

{*****          LA FUNCION "POT" ELEVA Y A LA POTENCIA -4.           *****}

        X:=POT(Y,-4)/4
      END
   ELSE
      BEGIN
        K:=3;
        X:=T*(1-1/(4*V))/SQRT(1+SQR(T)/(2*V));

{*****  LA FUNCION "N_P" CALCULA LA PROBABILIDAD PARA X EN BASE A LA  *****
 *****  NORMAL.                                                      *****}

        X:=N_P(X);
       END;
TS_P:=X;
END;

{
   LA FUNCION "N_I" CALCULA EL VALOR INVERSO DE UNA PROBABILIDAD CON BA-
  SE EN LA FUNCION DE DISTRIBUCION NORMAL.

  VARIABLES : P : PROBABILIDAD DE LA QUE SE DESEA OBTENER SU INVERSO.
}

FUNCTION N_I(P:REAL):REAL;
CONST    C0 = 2.515517;
         C1 = 0.802853;
         C2 = 0.010328;
         D1 = 1.432788;
         D2 = 0.189269;
         D3 = 0.001308;
VAR  I          : INTEGER;
     T,NU,DE    : REAL;
BEGIN
  T:= SQRT(LN( 1/SQR(P)));
  NU:= C0 + T*(C1+C2*T);
  DE:= 1+ T*(D1+ T*(D2+ T*D3));
  N_I :=T-NU/DE;
END;

{
   LA FUNCION "T_INV" CALCULA EL VALOR INVERSO DE UNA PROBABILIDAD CON
  BASE EN LA FUNCION DE DISTRIBUCION T DE STUDENT.
```

```
 VARIABLES : P : PROBABILIDAD DE LA QUE SE DESEA OBTENER SU INVERSO.
             V : GRADOS DE LIBERTAD.
                                                                   }

FUNCTION T_INV(P:REAL;V:INTEGER):REAL;
VAR   XP   :REAL;
      SN   :CHAR;
   FUNCTION G1(X:REAL):REAL;
      BEGIN
        G1:=(POT(X,3)+X)/4.0;
      END;
   FUNCTION G2(X:REAL):REAL;
      BEGIN

(*****           LA FUNCION "POT" ELEVA X A LA POTENCIA 3            *****)

        G2:=(5*POT(X,5)+16*POT(X,3)+3*X)/96.0;
      END;
   FUNCTION G3(X:REAL):REAL;
      BEGIN

(*****   LA FUNCION "POT" ELEVA X A LA POTENCIA 5 Y A LA POTENCIA 3.   *****)

        G3:=(3*POT(X,7)+19*POT(X,5)+17*POT(X,3)-15*X)/384.0;
      END;
   FUNCTION G4(X:REAL):REAL;
      BEGIN

(*****         LA FUNCION "POT" ELEVA X A LAS POTENCIAS 7,5,3          *****)

        G4:=(79*POT(X,9)+776*POT(X,7)+1482*POT(X,5)-1920*POT(X,3)-945*X)/92160.0;
      END;
BEGIN

(*****   LA FUNCION "N_I" CALCULA EL VALOR INVERSO DE LA PROBABILIDAD    *****
 *****   P CON BASE EN LA FUNCION DE DISTRIBUCION NORMAL.                *****)

   XP:=N_I(P);

(*****   LA FUNCION "POT" ELEVA V A LA POTENCIA 3 Y A LA POTENCIA 4.     *****)

   T_INV:=XP+(G1(XP)/V)+(G2(XP)/SQR(V))+(G3(XP)/POT(V,3))+(G4(XP)/POT(V,4));
END;

{
   LA FUNCION "F_P" CALCULA LA PROBABILIDAD PARA UN VALOR REAL CON BASE
  EN LA FUNCION DE DISTRIBUCION FISHER.

  VARIABLES : F  : VALOR AL QUE SE LE CALCULA LA PROBABILIDAD.
              V1 : GRADOS DE LIBERTAD.
              V2 : GRADOS DE LIBERTAD.
                                                                   }
```

```
FUNCTION F_P(F,V1,V2:REAL):REAL;
VAR   Z,Z1,D,N   : REAL;
BEGIN
   Z1:=2/(9*V1);
   Z:=2/(9*V2);

(*****          LA FUNCION "POT" ELEVA F A LA POTENCIA 1/3.          *****)

   N:=POT(F,1/3)*(1-Z)-(1-Z1);

(*****          LA FUNCION "POT" ELEVA F A LA POTENCIA 2/3.          *****)

   D:=SQRT(Z1+POT(F,2/3)*Z);

(*****   LA FUNCION "N_P" CALCULA LA PROBABILIDAD PARA N/D EN BASE A   *****
 *****   LA NORMAL.                                                    *****)

   F_P:=N_P(N/D);
END;

{
    EL PROCEDIMIENTO "CALCULO_SUMAS" OBTIENE LA MATRIZ MXTX Y AL VECTOR
   MXTY A PARTIR DE LA MATRIZ X Y EL VECTOR Y.

   VARIABLES : SUMAX  : SUMA DE LA X(I) PARA UNA VARIABLE J.
               SUMAY  : SUMA DE LAS Y(I).
               SUMAXX : SUMA DE LOS CUADRADOS DE LOS X(I) PARA UNA VARIA-
                        BLE J.
               SUMAXY : SUMA DE LOS PRODUCTOS DE X(I) PARA UNA VARIABLE J
                        POR LAS Y(I).
}

PROCEDURE  CALCULO_SUMAS(VAR A:MATRIZT1C; VAR B:MATRIZT2C;VAR X1:MATRIZT1;VAR Y1:MATRIZT2;P1,
N1:integer);
VAR   K,I,L                        :INTEGER;
      SUMAX,SUMAXY,SUMAXX,SUMAY    :REAL;
BEGIN
  FOR K:=1 TO P1 DO
    BEGIN
      SUMAX:=0;
      SUMAXY:=0;
      FOR I:=1 TO N1 DO
        BEGIN
          SUMAX:=SUMAX+X1[I,K];
          SUMAXY:=SUMAXY+X1[I,K]*Y1[I]
        END;
      FOR L:=K TO P1 DO
        BEGIN
          SUMAXX:=0;
          FOR I:=1 TO N1 DO
            SUMAXX:=SUMAXX+X1[I,K]*X1[I,L];
          A[K+1-VT1,L+1-VT1]:=SUMAXX;
          A[L+1-VT1,K+1-VT1]:=SUMAXX
        END;
```

```
      A[VT2,K+1-VT1]:=SUMAX;
      A[K+1-VT1,VT2]:=SUMAX;
      B[K+1-VT1]:=SUMAXY
    END;
  SUMAY:=0;
  A[VT2,VT2]:=N1;
  FOR I:=1 TO N1 DO
    SUMAY:=SUMAY+Y1[I];
  B[VT2]:=SUMAY;
END;

{
    EL PROCEDIMIENTO "INVERSA" OBTIENE LA INVERSA DE UNA MATRIZ DEFINIDA
   POSITIVA Y/O OBTIENE LA SOLUCION DE UN SISTEMA DONDE INTERVIENE ESTA
   MATRIZ. EL PROCEDIMIENTO SE BASA EN EL METODO DE CHOLESKI. LA SOLUCION
   DEL SISTEMA SE OBTIENE INTERACTIVAMENTE PARA TENER UNA MEJOR APROXI-
   MACION.

   VARIABLES : A    : MATRIZ A INVERTIR.
               HL   : NUMERO DE COLUMNAS Y RENGLONES DE LA MATRIZ A.
               YSOL : VECTOR SOLUCION DEL SISTEMA.
               RES  : VECTOR DE RESIDUOS.
               EPS  : NUMERO REAL POSITIVO MAS CERCANO A CERO.
}

PROCEDURE INVERSA(VAR A:MATRIZT1C;VAR INVERSION,SOLUCIONAR,INVERTIR:BOOLEAN;HL:INTEGER;
VAR YSOL:MATRIZT2C;VAR B:MATRIZT2C);
VAR   I,J,S,T,K,I1,J1,L,D2                     :INTEGER;
      X,Y,Z,D1,D0,XMAX,BBMAX,SI,SEP,EPS        :REAL;
      RES                                      :MATRIZT2C;

     {
        EL PROCEDIMIENTO "SYMSOL" OBTIENE LA SOLUCION DEL SISTEMA :
      A * YSOL = B
     }

PROCEDURE SYMSOL(VAR A:MATRIZT1C;VAR B,YSOL:MATRIZT2C);
VAR   IS,IS1,KS     :INTEGER;
      YPASO         :REAL;
BEGIN
 FOR IS:=1 TO HL DO
   BEGIN
     YPASO:=B[IS];
     FOR KS:=IS-1 DOWNTO 1 DO
       YPASO:=YPASO-A[IS+1,KS]*YSOL[KS];
     YSOL[IS]:=YPASO;
   END;
 FOR IS:=HL DOWNTO 1 DO
   BEGIN
     IS1:=IS+1;
     YPASO:=YSOL[IS]*A[IS1,IS];
     FOR KS:=IS+1 TO HL DO
       YPASO:=YPASO-A[KS+1,IS]*YSOL[KS];
     YSOL[IS]:=YPASO;
```

```
    END;
  END;

(*****              PROGRAMA PRINCIPAL DE INVERSION.                *****)

BEGIN
 I:=0;
 INVERSION:=TRUE;
 WHILE INVERSION AND (I<HL) DO
   BEGIN
     I:=I+1; I1:=I+1;
     FOR J:=1 TO I DO
       BEGIN
         J1:=J+1;
         X:=A[J,I];
         IF I=J THEN
            BEGIN
              FOR K:=J-1 DOWNTO 1 DO
                BEGIN
                  Y:=A[I1,K];
                  Z:=Y*A[K+1,K];
                  A[I1,K]:=Y*A[K+1,K];
                  X:=X-Y*Z;
                END;
              IF X=0 THEN INVERSION:=FALSE
                     ELSE IF X<>0 THEN A[I1,I]:=1/X;
            END
          ELSE
            BEGIN
              FOR K:=J-1 DOWNTO 1 DO
                X:=X-A[I1,K]*A[J1,K];
              A[I1,J]:=X;
            END
       END
   END;

(*****                  SOLUCION DEL SISTEMA.                       *****)

 IF INVERSION THEN
   BEGIN
     IF SOLUCIONAR THEN
       BEGIN
         SEP:=0.000005;
         REPEAT
           EPS:=SEP;
           SEP:=SQR(SEP);
         UNTIL SEP+1=1;
         FOR I:=1 TO HL DO
           BEGIN
             YSOL[I]:=0;
             RES[I]:=B[I];
           END;
         L:=0;
         D0:=3.0;
```

```
    REPEAT
       SYMSOL(A,RES,RES);
       L:=L+1;
       D2:=0;
       D1:=0;
       FOR I:=1 TO HL DO
         YSOL[I]:=YSOL[I]+RES[I];
       XMAX:=0;
       BBMAX:=0;
       FOR I:=1 TO HL DO
         BEGIN
           IF ABS(YSOL[I])>XMAX THEN XMAX:=ABS(YSOL[I]);
           IF ABS(RES[I])>BBMAX THEN BBMAX:=ABS(RES[I]);
           SI:=B[I];
           FOR K:=1 TO I-1 DO
           SI:=SI-A[K,I]*YSOL[K];
           FOR K:=I TO HL DO
             SI:=SI-A[I,K]*YSOL[K];
           RES[I]:=SI;
         END;
       IF BBMAX>D1*XMAX THEN D1:=BBMAX/XMAX;
       IF BBMAX>2*EPS*XMAX THEN D2:=1;
       IF D1>D0/2 THEN D2:=2;
       D0:=D1;
     UNTIL D2<>1;
   END;

(*****                   INVERSION DE LA MATRIZ.                    *****)

  IF INVERTIR THEN
    BEGIN
      FOR I:=2 TO HL DO
        BEGIN
          I1:=I+1;
          FOR J:=2 TO I DO
             BEGIN
               J1:=J-1;
               X:=-A[I1,J1];
               FOR K:=I-1 DOWNTO J DO
                 X:=X-A[I1,K]*A[K+1,J1];
               A[I1,J1]:=X;
             END
        END;
      FOR J:=1 TO HL DO
      FOR I:=J TO HL DO
        BEGIN
          I1:=I+1;
          X:=A[I1,J];
          IF I<>J THEN
            BEGIN
               FOR K:=I1 TO HL DO
               X:=X+A[K+1,I]*A[K+1,J]
            END
          ELSE
```

```
          FOR K:=I1 TO NL DO
            BEGIN
              Y:=A[K+1,J];
              Z:=A[K+1,K]*Y;
              A[K+1,J]:=A[K+1,K]*Y;
              X:=X+Y*Z;
            END;
          A[I1,J]:=X;
        END;
      END;
   END;
END;

{
    EL PROCEDIMIENTO "QUICK" ORDENA UN VECTOR EN FORMA ASCENDENTE MEDIAN-
   TE EL METODO QUICK_SORT. CONFORME SE ORDENA EL VECTOR SE CAMBIA OTRO
   VECTOR QUE GUARDA EL ORDEN ORIGINAL DEL PRIMERO, DANDO POR RESULTADO QUE
   AL TERMINO DE LA ORDENACION SE OBTENGA TANTO EL ORDEN ORIGINAL COMO EL
   NUEVO.

   VARIABLES : A : VECTOR A ORDENAR.
               B : VECTOR QUE CONTIENE EN PRINCIPIO EL ORDEN ORIGINAL DE A.
}

PROCEDURE QUICK(VAR A:MATRIZT2;VAR B:MATRIZT2I;VAR NELEM:INTEGER);
VAR   I   :INTEGER;
PROCEDURE SORT(L,R:INTEGER);
VAR   I,J,VO   :INTEGER;
      W,X      :REAL;
BEGIN
 I:=L;
 J:=R;
 X:=A[(L+R)DIV 2];
 REPEAT
   WHILE A[I]<X DO
     I:=I+1;
   WHILE X<A[J] DO
     J:=J-1;
   IF I<=J THEN
     BEGIN
       W:=A[I];
       VO:=B[I];
       A[I]:=A[J];
       B[I]:=B[J];
       A[J]:=W;
       B[J]:=VO;
       I:=I+1;
       J:=J-1
     END
 UNTIL I>J;
 IF L<J THEN
   SORT(L,J);
 IF I<R THEN
   SORT(I,R)
```

```
END;
BEGIN
 FOR I:=1 TO NELEM DO
  B[I]:=I;
 SORT(1,NELEM);
END;

{┌──────────────────────────────────────────────────────────────────────┐
 │  EL PROCEDIMIENTO 'TABLA_ANOVA' CALCULA LAS SUMAS DE CUADRADOS, SUS  │
 │ MEDIAS, EL ESTADISTICO F Y LOS COEFICIENTES DE DETERMINACION.        │
 │                                                                      │
 │ VARIABLES :  SCREG    : SUMA DE CUADRADOS DE LA REGRESION            │
 │              MSCREG   : MEDIA DE LA SUMA DE CUADRADOS DE LA REGRE-   │
 │                         SION                                         │
 │              SCRES    : SUMA DE CUADRADOS DE LOS RESIDUALES          │
 │              SCT      : SUMA DE CUADRADOS TOTALES                    │
 │              S        : DESVIACION ESTANDAR ESTIMADA                 │
 │              S2       : VARIANZA ESTIMADA                            │
 │              F        : ESTADISTICO F                                │
 │              R        : COEFICIENTE DE DETERMINACION                 │
 │              RA       : COEFICIENTE DE DETERMINACION AJUSTADO        │
 └──────────────────────────────────────────────────────────────────────┘}

PROCEDURE TABLA_ANOVA(VAR S2,S,SCREG,SCRES,SCT,MSCREG,F,R,RA:REAL;VAR H,N,P:INTEGER;
VAR MTW,MWTY:MATRIZT2C;VAR YM:MATRIZT2);
VAR  O,T            :INTEGER;
     SUMAY2,MBXTY   :REAL;

 BEGIN
   SCRES:=0;
   SCT:=0;
   SCREG:=0;
   SUMAY2:=0;
   FOR O:=1 TO N DO
     SUMAY2:=SUMAY2+SQR(YM[O]);
   IF VT1=0 THEN SCT:=SUMAY2-(SQR(MWTY[VT2])/N)
           ELSE SCT:=SUMAY2;
   MBXTY:=0;
   FOR T:=1 TO H DO
     MBXTY:=MBXTY+MTW[T]*MWTY[T];
   IF VT1=0 THEN SCREG:=MBXTY-(SQR(MWTY[VT2]))/N
           ELSE SCREG:=MBXTY;
   SCRES:=SCT-SCREG;
   IF (N-H)=0 THEN WRITELN('DIVISION ENTRE CERO')
              ELSE S2:=SCRES/(N-H);
   S:=SQRT(S2);
   MSCREG:=SCREG/P;
   R:=SCREG/SCT;
   F:=MSCREG/S2;
   RA:=1-(SCRES*(N-1))/(SCT*(N-H));
END;

{┌──────────────────────────────────────────────────────────────────────┐
 │  EL PROCEDIMIENTO 'ESCRIBE' SE ENCARGA DE PRESENTAR POR PANTALLA Y/O │
```

```
  POR IMPRESORA AL ESTADISTICO F PARA ALGUNA PRUEBA Y LA INFORMACION CON-
  CERNIENTE A ESTE EN UNA TABLA DE ANALISIS DE VARIANZA.

  VARIABLES : P1,..,P4 : VARIABLES DE LOS GRADOS DE LIBERTAD.
              V1,V4,V7 : SUMAS DE CUADRADOS.
              V2,V5    : MEDIAS DE LAS SUMAS DE CUADRADOS.
              V7       : SUMA DE CUADRADOS TOTAL.
              V3       : ESTADISTICO F-CALCULADO.
              CONTROL  : VARIABLE QUE SIRVE PARA ESCOGER ENTRE LOS DIVER-
                         SOS TITULOS.
 }

PROCEDURE ESCRIBE(CONTROL,P1,P2,P3,P4:INTEGER;V1,V2,V3,V4,V5,V6,V7:REAL);
TYPE  STAR=STRING[13];

     {
        EL PROCEDIMIENTO "C1" ESCRIBE EN LA PANTALLA Y/O EN LA IMPRESORA
      LA TABLA ANOVA DE ACUERDO A CIERTOS TITULOS.

      VARIABLES : T1,..,T7 : TITULOS.
     }

PROCEDURE C1(T1,T2,T3,T4,T5,T6,T7:STAR);
BEGIN
  WRITELN('TABLA ANOVA':45);WRITELN;
  WRITELN('_____________________________________________________________________');
  WRITELN;
  WRITELN('FUENTE DE       g.l.      SUMA DE        MEDIA DE           F          P-VALOR');
  WRITELN('VARIACION                 CUADRADOS      CUADRADOS');
  WRITELN('_____________________________________________________________________');
  WRITELN;
  IF CONTROL>2 THEN
    BEGIN
      WRITELN(T6,' ',P4:4,' ':5,V7:12:6);
      WRITELN(T7);
      WRITELN
    END;
 WRITELN(T1,' ',P1:4,' ':5,V1:12:6,' ':3,V2:12:6);
 WRITELN(T2);
 IF V3>0 THEN
   WRITELN(' ':54,V3:12:6,'   ',F_P(V3,P1,P2):8:6)
 ELSE
   WRITELN(' ':54,V3:12:6);
 WRITELN(T3,' ',P2:4,' ':5,V4:12:6,' ':3,V5:12:6);
 WRITELN(T4);
 WRITELN('.........................................................................');
 WRITELN(T5,' ',P3:4,' ':5,V6:12:6);  WRITELN;WRITELN;
 IF CONTROL=1 THEN
   BEGIN
     WRITELN('   R CUADRADA (Coeficiente de determinacion)= ',R:12:7);
     WRITELN('   R AJUSTADA = ',RA:12:7)
   END;
 IF IMPRESION THEN
 BEGIN
```

```
    WRITELN(LST,'TABLA ANOVA':45);
    WRITELN(LST,'_____________________________________________________________________________');
    WRITELN(LST);
    WRITELN(LST,'FUENTE DE      g.l.      SUMA DE        MEDIA DE             F          P-VALOR');
    WRITELN(LST,'VARIACION                CUADRADOS      CUADRADOS');
    WRITELN(LST,'_____________________________________________________________________________');
    WRITELN(LST);
    IF CONTROL>2 THEN
      BEGIN
        WRITELN(LST,T6,' ',P4:4,' ':5,V7:12:6);
        WRITELN(LST,T7);
        WRITELN(LST)
      END;
    WRITELN(LST,T1,' ',P1:4,' ':5,V1:12:6,' ':3,V2:12:6);
    WRITELN(LST,T2);
    IF V3>0 THEN
      WRITELN(LST,' ':54,V3:12:6,'    ',F_P(V3,P1,P2):8:6)
    ELSE
      WRITELN(LST,' ':54,V3:12:6);
    WRITELN(LST,T3,' ',P2:4,' ':5,V4:12:6,' ':3,V5:12:6);
    WRITELN(LST,T4);
    WRITELN(LST,'.............................................................................');
    WRITELN(LST,T5,' ',P3:4,' ':5,V6:12:6);  WRITELN(LST,^J,^M);
    IF CONTROL=1 THEN
      BEGIN
        WRITELN(LST,'   R CUADRADA (Coeficiente de determinacion)= ',R:12:7);
        WRITELN(LST,'   R AJUSTADA = ',RA:12:7,^J,^M)
      END;
  END;
END;

{*****                    PROGAMA PRINCIPAL DE ESCRIBE.                    *****}

BEGIN
 CASE CONTROL OF
 1:C1('REGRESION    ','              ','RESIDUALES   ','              ','TOTAL
   ','           ','            ');
 2:C1('CARENCIA     ','DE AJUSTE    ','ERROR        ','PURO         ','RESIDUAL     '
 ,'           ','            ');
 3:C1('RESIDUALES   ','N-N1 DATOS   ','RESIDUALES   ','N1 DATOS     ','TOTAL        '
 ,'REGRESION    ','N1 DATOS     ');
 4:C1('REGRESION    ','DE P-R VARS. ','RESIDUALES   ','DEL TOTAL    ','TOTAL        '
 ,'REGRESION    ','DE R VARS.   ');
 END;
 CONTINUAR;
END;

{
 |  EL PROCEDIMIENTO 'ESTADISTICA' CALCULA EL ESTADISTICO T PARA PRUEBAS  |
 |  DE HIPOTESIS INDIVIDUALES DE LOS ESTIMADORES.                         |
 |                                                                        |
 |  VARIABLES : T      : ESTADISTICO T-CALCULADO.                         |
 |              SBEST  : ESTIMADOR DE LA DESVIACION ESTANDAR DEL ESTIMADOR.|
 |              MBESTN : VECTOR SOLUCION.                                 |
```

```
|           SL     : ESTIMADOR DE LA DESVIACION ESTANDAR DEL MODELO.   |
|______________________________________________________________________|}

PROCEDURE ESTADISTICAS(VAR MWTW:MATRIZT1C;VAR MBESTW:MATRIZT2C;VAR SL:REAL;VAR HL:INTEGER);
VAR   C           :INTEGER;
      E           :REAL;
      T,SBEST     :MATRIZT2;

BEGIN
  WRITELN;
  FOR C:=1 TO HL DO
    BEGIN
      SBEST[C]:=SL*SQRT(MWTW[C+1,C]);
      E:=0.0;
      IF SBEST[C]<>0 THEN
      T[C]:=(MBESTW[C]-E)/SBEST[C]  ELSE WRITELN('DIVISION ENTRE CERO')
    END;  WRITELN;WRITELN;WRITELN;
  WRITELN('-----------------------------------------------------------------');
  WRITELN('BETA ESTIMADA      ERROR ESTANDARD        t-CALCULADA    P-VALOR');
  WRITELN('                   DE BETA ESTIMADA                          ');
  WRITELN('.................................................................');clreol;
  WRITELN;
  FOR C:=1 TO HL DO
    WRITELN('B[',C-1:2,']=',MBESTW[C]:12:7,' ':2,SBEST[C]:12:7,' ':11,T[C]:12:7,
    ts_p(t[C],n-HL):6:3);
  WRITELN;WRITELN;WRITELN;WRITELN;
  IF IMPRESION THEN
    BEGIN
      WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
      WRITELN(LST,'-----------------------------------------------------------------');
      WRITELN(LST,'BETA ESTIMADA      ERROR ESTANDARD        t-CALCULADA    P-VALOR');
      WRITELN(LST,'           .       DE BETA ESTIMADA                          ');
      WRITELN(LST,'.................................................................');
      WRITELN(LST);
      FOR C:=1 TO HL DO
       WRITELN(LST,'B[',C-1:2,']=',MBESTW[C]:12:7,' ':2,SBEST[C]:12:7,' ':11,T[C]:12:7,
       ts_p(t[C],n-HL):6:3);
    END;
END;

{______________________________________________________________________
|  EL PROCEDIMIENTO "CAMBIO_MATRIZ" SELECCIONA DE LA MATRIZ MXTX CIERTOS |
| COLUMNAS Y CIERTOS RENGLONES. SELECCIONA TAMBIEN CIERTO ELEMENTOS DEL  |
| VECTOR MXTY.                                                           |
|                                                                        |
| VARIABLES : MWTW    : MATRIZ DONDE SE GUARDAN LAS COLUMNAS Y RENGLONES |
|                       SELECCIONAS DE MXTX.                             |
|             MWTY    : VECTOR DONDE SE GUARDAN LOS ELEMENTOS SELECCIONA-|
|                       DOS DE MXTY O UNA COLUMNA SELECCIONADA DE MXTX.  |
|             INW     : VECTOR QUE CONTIENE LAS COLUMNAS Y RENGLONES A   |
|                       SER SELECCIONADOS.                               |
|             Q       : NUMERO DE RENGLONES Y COLUMNAS A SER SELECCIONADAS|
|             CAMBIOY : VARIABLE QUE CONTROLA EL SELECCIONAR EN MWTY A E-|
|                       LEMENTOS DE MXTY O UNA COLUMNA DE MXTX.          |
```

```
 |              INWY    : COLUMNA DE MXTX SELECCIONADA PARA SER GUARDADA EN |
 |                        MWTY.                                             |
 |__________________________________________________________________________|}

PROCEDURE CAMBIO_MATRIZ(VAR MWTW:MATRIZT1C;VAR MWTY:MATRIZT2C;VAR INW:MATRIZT21C;VAR INWY,Q:INTEGER
VAR KH,KI,KJ:INTEGER;

BEGIN
   FOR KI:=1 TO Q DO
     BEGIN
       FOR KJ:=KI TO Q DO
         BEGIN
           MWTW[KI+1-VT1,KJ+1-VT1]:=MXTX[INW[KI]+1-VT1,INW[KJ]+1-VT1];
           MWTW[KJ+1-VT1,KI+1-VT1]:=MXTX[INW[KI]+1-VT1,INW[KJ]+1-VT1];
         END;
       MWTW[VT2,KI+1-VT1]:=MXTX[VT2,INW[KI]+1-VT1];
       MWTW[KI+1-VT1,VT2]:=MXTX[VT2,INW[KI]+1-VT1];
       IF CAMBIOY THEN
         BEGIN
           IF INW[KI]<INWY+1-VT1 THEN
              MWTY[KI+1-VT1]:=MXTX[INW[KI]+1-VT1,INWY+1-VT1]
           ELSE
              MWTY[KI+1-VT1]:=MXTX[INWY+1-VT1,INW[KI]+1-VT1];
         END
       ELSE
         MWTY[KI+1-VT1]:=MXTY[INW[KI]+1-VT1];
       END;
       MWTW[VT2,VT2]:=N;
   IF CAMBIOY THEN
     MWTY[VT2]:=MXTX[VT2,INWY+1-VT1]
   ELSE
     MWTY[VT2]:=MXTY[VT2];
END;

{ __________________________________________________________________________
 |                                                                          |
 |   EL PROCEDIMIENTO "MULTIPLICACION" OBTIENE EL RESULTADO DE LA MULTIPLI- |
 |  CACION DE DOS MATRICES.                                                 |
 |                                                                          |
 |  VARIABLES : MR  : MATRIZ RESULTANTE DE LA MULTIPLICACION.              |
 |              M1  : MATRIZ QUE PRE-MULTIPLICA.                            |
 |              M2  : MATRIZ POS-MULTIPLICADA.                              |
 |              RE1 : REGLONES DE LA MATRIZ M1.                             |
 |              CO  : COLUMNAS DE LA MATRIZ M1 Y RENGLONES DE LA MATRIZ M2. |
 |              RE2 : COLUMNAS DE LA MATRIZ M2.                             |
 |__________________________________________________________________________|}

PROCEDURE MULTIPLICACION(VAR MR,M1,M2:MATRIZT1C;VAR RE1,CO,RE2:INTEGER);
VAR   CS,CN,CL     :INTEGER;
      M            :REAL;
BEGIN
  FOR CS:=1 TO RE1 DO
    FOR CN:=1 TO RE2 DO
      BEGIN
        M:=0.0;
```

```
        FOR CL:=1 TO CO DO
          M:=M+M1[CS,CL]*M2[CL,CN];
        MR[CS,CN]:=M;
      END;
  END;

{=======================================================================
|    EL PROCEDIMIENTO "GRAFICA" GRAFICA DOS VECTORES.                  |
|                                                                      |
|   VARIABLES : XX   : VECTOR A GRAFICAR.                              |
|               YY   : VECTOR CONTRA EL CUAL SE GRAFICA.               |
|               MAXX : VALOR MAXIMO DEL VECTOR XX.                     |
|               MINX : VALOR MINIMO DEL VECTOR XX.                     |
|               MAXY : VALOR MAXIMO DEL VECTOR YY.                     |
|               MINY : VALOR MINIMO DEL VECTOR YY.                     |
|               DX   : ORDENADA.                                       |
|               DY   : ABSCISA.                                        |
|               ESCX : VALOR QUE SE CALCULA PARA ESTANDARIZAR LOS VALORES
|                      DE XX EN UNA ESCALA EN LA QUE SEA POSIBLE SU GRAFI-
|                      CACION EN LA PANTALLA.                          |
|               ESCY : VALOR QUE SE CALCULA PARA ESTANDARIZAR LOS VALORES
|                      DE YY EN UNA ESCALA EN LA QUE SEA POSIBLE SU GRAFI-
|                      CACION EN LA PANTALLA.                          |
=======================================================================}

PROCEDURE GRAFICA(VAR XX,YY:MATRIZ2; N:INTEGER);
VAR      I,J,K,DX,DY                          : INTEGER;
         MAXX,MINX,ESCX,NR,MAXY,MINY,ESCY     : REAL;
BEGIN
  MAXX:=XX[1];
  MINX:=XX[1];
  MAXY:=YY[1];
  MINY:=YY[1];
  FOR J:=2 TO N DO
    BEGIN
      IF XX[J] > MAXX THEN MAXX:=XX[J]
                      ELSE IF XX[J] < MINX THEN MINX:=XX[J];
      IF YY[J] > MAXY THEN MAXY:=YY[J]
                      ELSE IF YY[J] < MINY THEN MINY:=YY[J];
    END;
  CLRSCR;
  FOR I:= 10 TO 78 DO
    BEGIN
         GOTOXY(I,2); WRITE('-');
         GOTOXY(I,22); IF I MOD 10 = 7 THEN    WRITE('+')
                                       ELSE    WRITE('-');
    END;
  FOR J:= 3 TO 21 DO
    BEGIN
      GOTOXY(10,J);
      WRITE('I');
      GOTOXY(78,J);
      WRITE('I');
    END;
```

```
  ESCX:=(MAXX-MINX)/66.0;
  ESCY:=(MAXY-MINY)/18.0;
  FOR J:=1 TO N DO
    BEGIN
      DY:=ROUND((MAXY-YY[J])/ESCY)+3;
      DX:=ROUND((XX[J]-MINX)/ESCX)+11;
      GOTOXY(DX,DY);
      WRITE('*');
    END;
  GOTOXY(1,3);
  WRITE(MAXY:9:2,'+');
  FOR I:=4 TO 20 DO
    IF I MOD 2 = 1 THEN
      BEGIN
        NR:=MAXY-((I-3)*ESCY);
        GOTOXY(1,I);
        WRITE(NR:9:2,'+');
      END;
  GOTOXY(1,21);
  WRITE(MINY:9:2,'+');
  GOTOXY(5,24);
  WRITE(MINX:9:2);
  GOTOXY(11,22);
  WRITE('+');
  K:=7;
  REPEAT
    K:=K+10;
    GOTOXY(K-6,23);
    NR:=((K-11)*ESCX)+MINX;
    WRITE(NR:9:2);
  UNTIL K=67;
  GOTOXY(71,23);
  WRITE(MAXX:9:2);
END;

{
    EL PROCEDIMIENTO "D_XXXX" CALCULA LA DIAGONAL PRINCIPAL DE LA MATRIZ
   X(X'X)^-'X'.

   VARIABLES : V1 : DIAGONAL DE X(X'X)^-'X'.
                                                                          }

PROCEDURE D_XXXX(VAR V1:MATRIZT2);
VAR  I,J,K      :INTEGER;
BEGIN
 FOR K:=1 TO N DO
   BEGIN
     IF VT1=0 THEN
        BEGIN
          V1[K]:=MXTX[2,1];
          FOR J:=2 TO H DO
            V1[K]:=V1[K]+MXTX[J+1,1]*X[K,J-1]*2
        END
     ELSE
```

```
      V1[K]:=0;
    FOR I:= 2-VT1 TO H DO
      BEGIN
        V1[K]:=V1[K]+MXTX[I+1,I]*SQR(X[K,I-1+VT1]);
        FOR J:=I+1 TO H DO
           V1[K]:=V1[K]+MXTX[J+1,I]*X[K,I-1+VT1]*X[K,J-1+VT1]*2;
      END;
   END;
END;
{
    LA UNIDAD "OBTENER" CONTIENE LOS PROCEDIMIENTOS MEDIANTE LOS CUALES
  SE OBTIENEN LA MATRIZ X Y EL VECTOR Y.
                                                                        }

 UNIT OBTENER;
 INTERFACE
   USES CRT,UGLOBAL,UUTIL;
   PROCEDURE LEER_X(VAR N,AP:INTEGER;VAR X MATRIZT1;VAR BX:BOOLEAN);
   PROCEDURE LEER_Y(VAR N,P:INTEGER;VAR Y:MATRIZT2;VAR BY:BOOLEAN);
   PROCEDURE LEER_XY(VAR N,AP:INTEGER;VAR X:MATRIZT1;VAR Y:MATRIZT2;VAR BY,BX:BOOLEAN);
   PROCEDURE SIMULAR_X(VAR X:MATRIZT1;VAR N,AP:INTEGER);
   PROCEDURE SIMULAR_Y(VAR Y:MATRIZT2;VAR N,P:INTEGER;VAR SEGUNDA:BOOLEAN);
   PROCEDURE DICOTOMAS(VAR X:MATRIZT1;VAR N,P:INTEGER);
   PROCEDURE SALVA;

 IMPLEMENTATION

{
    EL PROCEDIMIENTO "EXISTE_SN" VERIFICA LA EXISTENCIA DE UN ARCHIVO EN
  EL DISCO.

  VARIABLES : ARCH    : ARCHIVO A LEER.
              EXISTE  : VARIABLE LOGICA, VERDADERA SI EXISTE EL ARCHIVO,
                        FALSA SI NO EXISTE.
                                                                        }

PROCEDURE EXISTE_SN(VAR NOM_ARCH:FILNOM; EXISTE:BOOLEAN);
VAR    ARCH       : FILE;
       EXISTE1    : BOOLEAN;
BEGIN
    EXISTE1:=FALSE;
    REPEAT
      REPEAT
        GOTOXY(01,16);
        WRITE('> INTRODUCE EL NOMBRE DEL ARCHIVO: ');
        READLN(NOM_ARCH);
      UNTIL LENGTH(NOM_ARCH)>0;
      ASSIGN(ARCH,NOM_ARCH);
      {$I-} RESET(ARCH); {$I+}
      IF IORESULT = 0 THEN
          IF EXISTE THEN EXISTE1:=TRUE
                    ELSE
                       WRITELN(^G,'YA EXISTE ESTE ARCHIVO ')
```

```
      ELSE
         IF EXISTE THEN WRITELN(^G,'NO EXISTE ESTE ARCHIVO ')
                   ELSE
                     BEGIN
                        {$I-} REWRITE(ARCH); {$I+}
                        IF IORESULT <> 0 THEN WRITELN(^G,'NOMBRE ILEGAL')
                                         ELSE EXISTE1:=TRUE;
                     END;
   UNTIL EXISTE1;
END;

{┌──────────────────────────────────────────────────────────────────────┐
 │  EL PROCEDIMIENTO 'LEER_XY_ARCH' LEE A LA MATRIZ X Y AL VECTOR Y DE  │
 │ UN ARCHIVO CON FORMATO DEFINIDO (REAL,ARREGLO), UN COMPONENTE DE Y SE-│
 │ GUIDO DE VARIOS COMPONENTES DE X.                                    │
 │                                                                      │
 │ VARIABLES : ARCHIVO_1 : ARCHIVO DE DATOS.                            │
 │             REGIS_1   : VARIABLE QUE CONTIENE UN REGISTRO DEL ARCHIVO.│
 │             NOMBRE    : NOMBRE DEL ARCHIVO DONDE ESTAN LOS DATOS.    │
 └──────────────────────────────────────────────────────────────────────┘}

PROCEDURE LEER_XY_ARCH(VAR X:MATRIZT1;VAR Y:MATRIZT2;VAR N,P,AP:INTEGER);
VAR ARCHIVO_1  : FILE OF R_VEC_MAT;
    REGIS_1    : R_VEC_MAT;
    NOMBRE     : FILNOM;
    I          : INTEGER;
BEGIN
      EXISTE_SN(NOMBRE,TRUE);
      ASSIGN(ARCHIVO_1,NOMBRE);
      RESET(ARCHIVO_1);
      I:=1;
      REPEAT
        {$I-}  READ(ARCHIVO_1,REGIS_1); {$I+}
        IF IORESULT<>0 THEN
           WRITELN(^G,'ERROR EN EL REGISTRO: ',I)
        ELSE
           BEGIN
               Y[I]:=REGIS_1.Y1;
               FOR J:=1 TO P DO
                  X[I,AP+J]:=REGIS_1.X1[J];
               I:=I+1
           END;
        IF EOF(ARCHIVO_1) THEN
           BEGIN
               WRITELN(^G,^J,^M,'FIN DE ARCHIVO, RENGLONES LEIDOS: ',I-1);
               N:=I-1;
               CLOSE(ARCHIVO_1)
           END;
      UNTIL (I=N+1) ;
END;

{┌──────────────────────────────────────────────────────────────────────┐
 │  EL PROCEDIMIENTO "LEER_XY_TXT" LEE A LA MATRIZ X Y AL VECTOR Y DE UN │
 │ ARCHIVO DE TIPO TEXTO. SE LEE UN COMPONENTE DE Y SEGUIDO DE VARIOS COM-│
```

```
  PONENTES DE X.

  VARIABLES : ARCHIVO_4 : ARCHIVO DE DATOS.
              REGIS_1   : VARIABLE QUE CONTIENE UN REGISTRO DEL ARCHIVO.
              NOMBRE    : NOMBRE DEL ARCHIVO DONDE ESTAN LOS DATOS.
              RESULT    : SI EN LA VALIDACION SE ENCONTRO ERROR, ESTA VA
                          RIABLE GUARDA LA POSICION DONDE SE ENCONTRO EL
                          ERROR.
                                                                        }

PROCEDURE LEER_XY_TXT(VAR X:MATRIZT1;VAR Y:MATRIZT2;VAR N,P,AP:INTEGER);
VAR ARCHIVO_4          : TEXT;
    REGIS_4,M          : STRING[255];
    NOMBRE             : FILNOM;
    ERROR              : BOOLEAN;
    K,I,L,RESULT       : INTEGER;
 BEGIN
   EXISTE_SN(NOMBRE,TRUE);
   I:=1;
   L:=1;
   ASSIGN(ARCHIVO_4,NOMBRE);
   RESET(ARCHIVO_4);
   REPEAT
       ERROR:=FALSE;
       {$I-}  READLN(ARCHIVO_4,REGIS_4); {$I+}
       IF (IORESULT<>0) OR (LENGTH(REGIS_4)=0) THEN
          WRITELN(^G,'ERROR EN EL REGISTRO: ',I)
       ELSE
          BEGIN
              FOR K:=1 TO P+1 DO
                BEGIN
                  IF POS(',',REGIS_4)<>0 THEN
                     M:=COPY(REGIS_4,1,POS(',',REGIS_4)-1)
                  ELSE
                     M:=REGIS_4;
                  IF K=1 THEN
                     VAL(M,Y[I],RESULT)
                  ELSE
                     VAL(M,X[I,AP+K-1],RESULT);
                  IF RESULT=0 THEN
                     DELETE(REGIS_4,1,LENGTH(M)+1)
                  ELSE  ERROR:=TRUE;
                END;
              IF NOT ERROR THEN
                 I:=I+1
              ELSE
                 WRITELN('ERROR EN EL REGISTRO: ',L);
          END;
          IF EOF(ARCHIVO_4) THEN
             BEGIN
                WRITELN(^G,^J,^M,'FIN DE ARCHIVO, RENGLONES LEIDOS: ',I-1);
                N:=I-1;
                CLOSE(ARCHIVO_4)
             END;
```

```
        L:=L+1;
    UNTIL (I=N+1) ;
END;

{╔══════════════════════════════════════════════════════════════════════╗
 ║                                                                      ║
 ║    EL PROCEDIMIENTO 'LEER_X' REALIZA LA LECTURA DE LA MATRIZ X POR PAN-║
 ║ TALLA O POR ARCHIVO                                                  ║
 ║                                                                      ║
 ║ VARIABLES.- P  : SIRVE PARA QUE SEAN LEIDAS E INCORPORADAS A LA MATRIZ ║
 ║                  X UN CIERTO NUMERO DE VARIABLES.                    ║
 ║             B  : ES UNA VARIABLE AUXILIAR QUE GUARDA TEMPORALMENTE UN ║
 ║                  RENGLON DE X.                                       ║
 ╚══════════════════════════════════════════════════════════════════════╝}

PROCEDURE LEER_X(VAR N,AP:INTEGER;VAR x:MATRIZT1;VAR DX:BOOLEAN);
TYPE  FILNOM               : STRING[12];
VAR   P,J,I,K RESULT       : INTEGER;
      B                    : MATRIZT2;
      NOMBRE               : FILNOM;
      OP1                  : CHAR;
      REGIS,M              : STRING[255];
      ERROR                : BOOLEAN;
BEGIN
  CLRSCR; GOTOXY(01,04);
  WRITELN(' ':10,'╔══════════════════════════════════════════════╗');
  WRITELN(' ':10,'║              MENU DE LECTURA DE X            ║');
  WRITELN(' ':10,'║                                              ║');
  WRITELN(' ':10,'║ A '^P' POR PANTALLA                           ║');
  WRITELN(' ':10,'║ B '^P' ARCHIVO CON FORMATO [Real,Arreglo]     ║');
  WRITELN(' ':10,'║ C '^P' ARCHIVO DE TEXTO (SDF)                 ║');
  WRITELN(' ':10,'║ S '^P' SALIDA                                 ║');
  WRITELN(' ':10,'╚══════════════════════════════════════════════╝');
  REPEAT
     GOTOXY(01,12);
     WRITE(' ':10,'> OPCION : ');
     OP1:=READKEY;   WRITELN(OP1);
     OP1:=UPCASE(OP1);
  UNTIL OP1 IN ['A','B','C','S'];
  IF OP1='S' THEN EXIT;
    GOTOXY(10,24); WRITE('PARA SALIR INTRODUCIR  [0]');   GOTOXY(10,14);
  REPEAT
     WRITE('> CUANTAS COLUMNAS SERAN LEIDAS : '); P:=LEER_INT;
  UNTIL P>=0;
  CASE OP1 OF
   'A': BEGIN
          CLRSCR;
          WRITELN(^J,^M,^M,^M,^M);
          WRITELN('TIENES QUE DARME LA MATRIX X POR RENGLONES EN FORMATO LIBRE ');
          WRITELN('CON ',P:2,' COMPONENTES ');
          WRITELN('"NO" INCLUYAS LA COLUMNA DE UNOS PARA EL TERMINO INDEPENDIENTE ');
          FOR I:= 1 TO N DO
            BEGIN
              WRITELN('RENGLON ',I:1,' :');
```

```
(*****  EN LA LECTURA POR PANTALLA, CADA RENGLON SE LEE EN FORMATO     *****
 *****  LIBRE MEDIANTE EL PROCEDIMIENTO "LEER_VEC_REAL".              *****)

            LEER_VEC_REAL(B,P);
            FOR J:=1 TO P DO
               X[I,AP+J]:=B[J];
          END;
       END;

(*****  LA LECTURA POR ARCHIVO CON FORMATO [REAL,ARREGLO] DE LA MATRIZ *****
 *****  X  SE LLEVA A CABO MEDIANTE EL PROCEDIMIENTO "LEER_XY_ARCH".   *****)

   'B': LEER_XY_ARCH(X,B,N,P,AP);

(*****  LA LECTURA POR ARCHIVO DE TIPO TEXTO DE LA MATRIZ X SE LLEVA   *****
 *****  A CABO MEDIANTE EL PROCEDIMIENTO "LEER_XY_TXT".                *****)

   'C': LEER_XY_ARCH(X,B,N,P,AP);
   END;
   AP:=AP+P;
   BX:=TRUE;
END;

{
    EL PROCEDIMIENTO "LEER_Y" REALIZA LA LECTURA DE EL VECTOR Y POR PAN-
  TALLA O POR ARCHIVO.
}

PROCEDURE LEER_Y(VAR N,P:INTEGER;VAR Y:MATRIZT2;VAR BY:BOOLEAN);
VAR   I,K,J,RESULT  : INTEGER;
      A             : MATRIZT1;
      ARCHIVO       : TEXT;
      NOMBRE        : FILNOM;
      OP1           : CHAR;
      REGIS         : STRINF[255];
BEGIN
  CLRSCR;
  GOTOXY(01,04);
  WRITELN(' ':10,'╔══════════════════════════════════════════════╗');
  WRITELN(' ':10,'║                 LECTURA  DE  Y               ║');
  WRITELN(' ':10,'║                                              ║');
  WRITELN(' ':10,'║ A '^P' POR PANTALLA                          ║');
  WRITELN(' ':10,'║ B '^P' ARCHIVO CON FORMATO [Real,Arreglo]    ║');
  WRITELN(' ':10,'║ C '^P' ARCHIVO DE TEXTO (SDF)                ║');
  WRITELN(' ':10,'║ S '^P' SALIDA                                ║');
  WRITELN(' ':10,'╚══════════════════════════════════════════════╝');
  REPEAT
    GOTOXY(01,12);
    WRITE(' ':10,'> OPCION : ');
    OP1:=READKEY;  WRITELN(OP1);   OP1:=UPCASE(OP1);
  UNTIL OP1 IN ['A','B','C','S'];
  IF OP1='S' THEN EXIT;
  CASE OP1 OF
 'A': BEGIN
```

```
          CLRSCR;
          WRITELN(^J,^M,^M,^M' ME TIENES QUE DAR UN VALOR POR RENGLON : ');
          FOR I:= 1 TO N DO
            BEGIN
              WRITE('VALOR DE Y[',I,']= ');
              Y[I]:=LEER_REAL;
            END;
        END;

{*****   LA LECTURA POR ARCHIVO CON FORMATO [REAL,ARREGLO] DE EL VEC-   *****
 *****   TOR Y SE LLEVA A CABO MEDIANTE EL PROCEDIMIENTO "LEER_XY_ARCH".*****}

 'B': LEER_XY_ARCH(A,Y,N,P,P);

{*****   LA LECTURA POR ARCHIVO DE TIPO TEXTO DE EL VECTOR Y SE LLEVA   *****
 *****   A CABO MEDIANTE EL PROCEDIMIENTO "LEER_XY_TXT".                *****}

 'C': LEER_XY_TXT(A,Y,N,P,P);
 END;
 BY:=TRUE.
END;

{ ==================================================================
  |   EL PROCEDIMIENTO "LEER_XY" REALIZA LA LECTURA DE LA MATRIZ X Y DEL
  | VECTOR Y POR PANTALLA O POR ARCHIVO
  |
  | VARIABLES.- P  : SIRVE PARA QUE SEAN LEIDAS E INCORPORADAS A LA MATRIZ
  |                  X UN CIERTO NUMERO DE VARIABLES.
  |           · XT : ES UNA VARIABLE AUXILIAR QUE GUARDA TEMPORALMENTE UN
  |                  RENGLON DE LA FORMA (y,x1,..,xP) PARA POSTERIORMENTE
  |                  ASIGNAR A X Y A Y SUS RESPECTIVOS VALORES
  ================================================================== }

PROCEDURE LEER_XY(VAR N,AP:INTEGER;VAR X:MATRIZT1;VAR Y:MATRIZT2VAR BY,BX:BOOLEAN);
VAR   K,I,P,J     : INTEGER;
      OP,YN       : CHAR;
      XT,B        : MATRIZT2;
      NOMBRE      : FILNOM;
      A           : MATRIZT1;
BEGIN
  CLRSCR; GOTOXY(01,04);
  WRITELN(' ':10,'╔══════════════════════════════════════════════╗');
  WRITELN(' ':10,'║                MENU DE LECTURA DE X y Y      ║');
  WRITELN(' ':10,'║                                              ║');
  WRITELN(' ':10,'║ A '^P' POR PANTALLA                          ║');
  WRITELN(' ':10,'║ B '^P' ARCHIVO CON FORMATO [Real,Arreglo]    ║');
  WRITELN(' ':10,'║ C '^P' ARCHIVO DE TEXTO (SDF)                ║');
  WRITELN(' ':10,'║ S '^P' SALIDA                                ║');
  WRITELN(' ':10,'╚══════════════════════════════════════════════╝');
  REPEAT
     GOTOXY(01,12);
     WRITE(' ':10,'> OPCION : ');
     OP:=READKEY;   WRITELN(OP);  OP:=UPCASE(OP);
  UNTIL OP IN ['A','B','C','S'];
```

```
 IF OP='5' THEN EXIT;
 GOTOXY(10,24);
 WRITE('PARA SALIR INTRODUCIR  (0)');
 GOTOXY(10,14);
 REPEAT
    WRITELN;
    WRITE('> CUANTAS COLUMNAS CONTIENE LA MATRIZ X : ');
    P:=LEER_INT;
 UNTIL P>=0;
 IF P = 0  THEN EXIT;
 CASE OP OF
    'A' : BEGIN
            CLRSCR;
            WRITELN(^J,^M,^M,^M,^M);
            WRITELN('PARA DAR UNA OBSERVACION, DA EL VALOR DE Y SEGUIDO DE LAS X's EN FORMATO LIBRE
            WRITELN('ES DECIR, DA UN VECTOR RENGLON DE ',P+1:1,' COMPONENTES ');
            WRITELN('"NO" INCLUYAS LA COLUMNA DE UNOS PARA EL TERMINO INDEPENDIENTE ');
            FOR I:= 1 TO N DO
              BEGIN
               WRITELN('RENGLON ',I:1,' :');

{*****   EN LA LECTURA POR PANTALLA, CADA RENGLON SE LEE EN FORMATO     *****
 *****   LIBRE MEDIANTE EL PROCEDIMIENTO "LEER_VEC_REAL".               *****}

               LEER_VEC_REAL(XT,P+1);
               Y[I]:=XT[1];
               FOR K:=2 TO P+1 DO
                 X[I,AP+K-1]:=XT[K];
              END;
           END;

{*****   LA LECTURA POR ARCHIVO CON FORMATO (REAL,ARREGLO) DE LA MATRIZ *****
 *****   X Y EL VECTOR Y SE LLEVA A CABO MEDIANTE EL PROCEDIMIENTO      *****
 *****   "LEER_XY_ARCH".                                                *****}

    'B' : LEER_XY_ARCH(X,Y,N,P,AP);

{*****   LA LECTURA POR ARCHIVO DE TIPO TEXTO DE LA MATRIZ X Y EL VEC-  *****
 *****   TOR Y SE LLEVA A CABO MEDIANTE EL PROCEDIMIENTO "LEER_XY_TXT". *****}

    'C' : LEER_XY_TXT(X,Y,N,P,AP);
 END;
 AP:=AP+P;
 BX:=TRUE;
 BY:=TRUE;
END;

{=====================================================================
 |   EL PROCEDIMIENTO "SIMULAR_X" CREA UN NUMERO P DE COLUMNAS EN LA MA-|
 |  TRIZ X MEDIANTE LA SIMULACION.                                      |
 |                                                                      |
 |  VARIABLES : P  : NUMERO DE COLUMNAS A SIMULAR.                      |
 |              XS : VECTOR CREADO POR SIMULACION.                      |
 =====================================================================}
```

```
PROCEDURE SIMULAR_X(VAR X:MATRIZT1;VAR N,AP:INTEGER);
VAR    I,J,K,P  :  INTEGER;
       XS       :  MATRIZT2;
       OP       :  CHAR;

     {┌──────────────────────────────────────────────────────────────┐
      │  EL PROCEDIMIENTO 'SIM_UNIF CREA UN VECTOR MEDIANTE LA SIMULACION │
      │ UNIFORME.                                                    │
      │                                                              │
      │ VARIABLES : XU : VECTOR CREADO POR SIMULACION UNIFORME.      │
      │             A  : LIMITE INFERIOR DEL INTERVALO.              │
      │             B  : LIMITE SUPERIOR DEL INTERVALO.              │
      └──────────────────────────────────────────────────────────────┘}

PROCEDURE SIM_UNIF(VAR XU:MATRIZT2; N:INTEGER);
VAR   A,B    : REAL;
      I      : INTEGER;
BEGIN
  CLRSCR;
  GOTOXY(01,04);
  WRITE('DAME EL VALOR DE a = ');
  A:=LEER_REAL;
  REPEAT
     WRITE('DAME EL VALOR DE b = ');
     B:=LEER_REAL;
  UNTIL A < B;
  FOR I:=1 TO N DO
    XU[I]:=(B-A)*RANDOM+A;
END;

     {┌──────────────────────────────────────────────────────────────┐
      │  EL PROCEDIMIENTO 'SIM_BIN' CREA UN VECTOR MEDIANTE LA SIMULACION │
      │ BINARIA.                                                     │
      │                                                              │
      │ VARIABLES : XB : VECTOR CREADO POR SIMULACION BINARIA.       │
      │             P  : PROBABILIDAD DE EXITO.                      │
      └──────────────────────────────────────────────────────────────┘}

PROCEDURE SIM_BIN(VAR XB:MATRIZT2; N:INTEGER);
 VAR   I     :INTEGER;
       P     : REAL;
BEGIN
  CLRSCR;
  REPEAT
    WRITE('DAME EL VALOR p (PROBABILIDAD DE EXITO) = ');
    P:=LEER_REAL;
  UNTIL (0<P) AND (P<1);
  FOR I:=1 TO N DO
    IF (RANDOM-P)>=0 THEN  XB[I]:=0
           ELSE  XB[I]:=1;
END;

(*****                PROGRAMA PRINCIPAL DE SIMULAR_X.              *****)
```

```
BEGIN
   CLRSCR;
   WRITELN(CHAR(10));
   WRITELN('                                    SIMULACION DE   X');
   WRITELN('                                    ══════════════');
   GOTOXY(10,24);
   WRITELN('PARA SALIR INTRODUCIR [0]');
   GOTOXY(10,06);
   REPEAT
     WRITE('> CUANTAS COLUMNAS SERAN SIMULADAS : '); P:=LEER_INT;
   UNTIL P>=0;
   J:=1;
   IF P <> 0 THEN
      REPEAT
         CLRSCR; GOTOXY(01,04);
         WRITELN(' ':10,'╔══════════════════════════════════════════════╗');
         WRITELN(' ':10,'║              SIMULAR  X(',j,')                ║');
         WRITELN(' ':10,'║                                              ║');
         WRITELN(' ':10,'║ A '^P' SIMULACION DE UNA UNIFORME [a,b)      ║');
         WRITELN(' ':10,'║ B '^P' SIMULACION DE UNA BERNOULLI(p)        ║');
         WRITELN(' ':10,'║ S '^P' SALIDA                                ║');
         WRITELN(' ':10,'╚══════════════════════════════════════════════╝');
         REPEAT
            GOTOXY(01,12);
            WRITE(' ':10,'> OPCION : ');  OP:=READKEY;
            WRITELN(OP);  OP:=UPCASE(OP);
         UNTIL OP IN ['A','B','S'];
         IF OP='S' THEN P:=J-1
          ELSE
           BEGIN
             IF (OP='A') THEN  SIM_UNIF(XS,N)
             ELSE  SIM_BIN(XS,N);
             FOR I:=1 TO N DO
                X[I,AP+J]:=XS[I];
           END;
        J:=J+1;
      UNTIL J>P;
   AP:=AP+P;
END;

{╔══════════════════════════════════════════════════════════════╗
 ║  EL PROCEDIMIENTO "SIMULAR_Y" CREA UN VECTOR Y MEDIANTE SIMULACION
 ║ NORMAL.
 ║
 ║ VARIABLES : SIGMA   : VARIANZA.
 ║             SEGUNDA : VARIABLE QUE NOS AYUDA A SABER SI SE VA A SIMULAR
 ║                       LA VARIABLE DEPENDIENTE UNA VEZ QUE YA SE TIENE
 ║                       ESTA, CON EL OBJETO DE COLOCAR A LA VARIABLE OB-
 ║                       TENIDA POR ESTE PROCEDIMIENTO UNA COLUMNA DESPUES
 ║                       DE LA ULTIMA COLUMNA OCUPADA EN LA MATRIZ X.
 ╚══════════════════════════════════════════════════════════════╝}

PROCEDURE SIMULAR_Y(VAR Y:MATRIZT2;VAR N,P:INTEGER;VAR SEGUNDA:BOOLEAN);
```

```
VAR I,J,K     : INTEGER;
    GE,ER     : MATRIZT2;
    OP        : CHAR;
    SIGMA     : REAL;

FUNCTION SIM_NOR(SIGMA:REAL):REAL;
CONST   NN = 50;
VAR     J  : INTEGER;
        X  : REAL;
BEGIN
  X:=0;
  FOR J:=1 TO NN DO
      X:=X+RANDOM;
  SIM_NOR:= (X-(NN/2)) / SQRT(NN/(12*SIGMA));
END;

 BEGIN
   CLRSCR;
   GOTOXY(01,04);
   WRITELN(' ':10,'╔════════════════════════════════════════════╗');
   WRITELN(' ':10,'║    SIMULACION DE  Y  ( DISTRIBUCION NORMAL )  ║');
   WRITELN(' ':10,'║                                              ║');
   WRITELN(' ':10,'║ A '^P' Y DISTRIBUIDA N( XB, σ²I)               ║');
   WRITELN(' ':10,'║ B '^P' Y DISTRIBUIDA N( XB, V)                 ║');
   WRITELN(' ':10,'║ S '^P' SALIDA                                  ║');
   WRITELN(' ':10,'╚════════════════════════════════════════════╝');
   REPEAT
      GOTOXY(01,11);
      WRITE(' ':10,'> OPCION : ');
      OP:=READKEY;WRITELN(OP); OP:=UPCASE(OP);
      UNTIL OP IN ['A','B','S'];
   CASE OP OF
       'A' : BEGIN
               CLRSCR;
               REPEAT
                  WRITE('DAME EL VALOR DE σ² = ');
                  READLN(SIGMA);
               UNTIL SIGMA>0;
               FOR I:=1 TO N DO
                 ER[I]:=SIM_NOR(SIGMA);
             END;
       'B' :  BEGIN
                CLRSCR;
                WRITELN('TIENES QUE DARME ',N,' σ², UNA PARA CADA OBSERVACION');
                FOR I:=1 TO N DO
                  BEGIN
                    REPEAT
                       WRITE('DAME EL VALOR DE σ²(',I,') = ');
                       SIGMA:=LEER_REAL;
                    UNTIL SIGMA>0;
                    ER[I]:=SIM_NOR(SIGMA);
                  END;
              END;
     END;
```

```
      IF OP <> 'S' THEN
       BEGIN
          WRITELN;
          WRITE('QUIERES LISTAR LOS ERRORES SIMULADOS  ..? ');
          OP:=READKEY;  WRITELN(OP);
          OP:=UPCASE(OP);
          IF OP ='S' THEN
            BEGIN
            IF I MOD 19 = 0 THEN
               BEGIN
                  CONTINUAR; CLRSCR;
                  WRITELN(ER[I]:14:8,'  ');
               END;
            WRITELN(ER[I]:14:8,'  ');
          END;
        CONTINUAR;CLRSCR;
       END;
      WRITELN(CHAR(10));
      WRITELN('DAME EL VECTOR B DE ',P+1-VT1,'COMPONENTES');WRITELN;
      FOR I:=1 TO P+1-VT1 DO
        BEGIN
          WRITE('[',I-1,']= ');
          READLN(BE[I]);
        END;
      IF SEGUNDA THEN
        BEGIN
          P:=P+1;
          FOR I:=1 TO N DO
           BEGIN
             IF VT1 =0 THEN  X[I,P]:=BE[1]
                       ELSE  X[I,P]:=0;
             FOR J:=1 TO P-1 DO
               X[I,P]:=X[I,P]+BE[J+1-VT1]*X[I,J]+ER[I];
           END;
           WRITELN('ESTA SIMULACION DE Y SE ENCUENTRA EN LA COLUMNA ',P:2,' DE LA MATRIZ X');
        END
      ELSE
        FOR I:=1 TO N DO
          BEGIN
            Y[I]:=BE[1];
            FOR J:=1 TO P DO
              Y[I]:=Y[I]+BE[J+1]*X[I,J]+ER[I];
          END;
      SEGUNDA:= TRUE;
   END;
END;

{
   EL PROCEDIMIENTO "DICOTOMAS" CONSTRUYE Y CREA VARIABLES DICOTOMAS A
  PARTIR DE UNA VARIABLE CATEGORICA QUE NO ESTA CON VALORES 0,1. LAS NUE-
  VAS VARIABLES CREADAS SON ALMACENADAS AL FINAL DE LA MATRIZ X.

  VARIABLES : ND   : NUMERO DE VARIABLES CATEGORICAS A CONSTRUIR.
              LINF : LIMITE INFERIOR DEL INTERVALO DE UNA CATEGORIA.
```

```
   |            LSUP : LIMITE SUPERIOR DEL INTERVALO DE UNA CATEGORIA.           |
   |_____________________________________________________________________________|}

PROCEDURE DICOTOMAS(VAR X:MATRIZT1;VAR N,P:INTEGER);
VAR   I,J,K,ND,LINF,LSUP   : INTEGER
      OP                   : CHAR;
BEGIN
  CLRSCR;
  WRITELN;WRITELN;
  WRITELN('QUIERES CONSTRUIR UNA(S) VARIABLE(S) DICOTOMICA(S), A PARTIR DE UNA ');
  WRITE('VARIABLE CATEGORICA QUE NO ESTA CON VALORES CEROS Y UNOS (S/N) ? :' );
  OP:=READKEY; WRITELN(OP); OP:=UPCASE(OP);
  IF OP ='S' THEN
    WHILE OP ='S' DO
      BEGIN
        REPEAT
          WRITE('DAME EL No. DE LA VARIABLE CATEGORICA A CONSTRUIR ');
          K:=LEER_INT;
        UNTIL (K>0) AND (K<=P);
        REPEAT
          WRITE('CUANTAS VARIABLES DICOTOMAS NECESITAS CONSTRUIR: ');
          ND:=LEER_INT;
        UNTIL (ND)=1) AND (ND<=NMAXCOL-AP);
        WRITELN;
        WRITELN('AL DARME LOS INTERVALOS DE LAS CATEGORIAS',^J,^M,'DEBES CONSIDERAR QUE LOS LIMITES
        FOR J:=1 TO ND DO
          BEGIN
            WRITELN('DAME EL INTERVALO QUE DEFINA A LA ',J,'a. VAR. DICOTOMA');
            WRITE('LIMITE INFERIOR: ');
            LINF:=LEER_INT;
            REPEAT
              WRITE('LIMITE SUPERIOR: ');
              LSUP:=LEER_INT;
            UNTIL LSUP>=LINF;
            FOR I:=1 TO N DO
              IF (LINF<=X[I,K]) AND (X[I,K]<=LSUP)  THEN X[I,P+J]:=1
                                                    ELSE X[I,P+J]:=0;
          END;
        P:=P+ND;
        WRITELN('LAS VARIABLES DICOTOMAS FUERON INCLUIDAS EN LAS ',ND,' ULTIMAS');
        WRITELN('COLUMNAS DE X. EN EL ORDEN EN QUE FUERON DEFINIDAS');WRITELN;
        OP:='N';
        WRITE('DESEAS CONSTRUIR MAS VAR. DICOTOMAS)',^J,^M,' A PARTIR DE UNA VAR. CATEGORICA (S/N) ?
        OP:=READKEY; WRITELN(OP); OP:=UPCASE(OP);
      END;
END;

 {_____________________________________________________________________________
  |                                                                           |
  |  EL PROCEDIMIENTO "SALVA" GUARDA EN UN ARCHIVO, AL VECTOR Y Y A LA        |
  |  MATRIZ X.                                                                |
  |                                                                           |
  |___________________________________________________________________________|}

PROCEDURE SALVA;
```

```
VAR   I,J,K                    : INTEGER;
      YN,OP                    : CHAR;
      NOMBRE                   : FILNOM;
      ARCHIVO_1                : TEXT;
      VALOR,REGIS_1            : STRING(255);

BEGIN
  CLRSCR;

{*****  EL NOMBRE DEL ARCHIVO A CREAR SE DA EN EL PROCEDIMIENTO 'EXIS- *****
 *****  TE_SN'. EL CUAL ADEMAS SE ENCARGA DE CHECAR DE QUE EL NOMBRE   *****
 *****  EXISTA Y SEA VALIDO.                                           *****}

  EXISTE_SN(NOMBRE,FALSE);
  ASSIGN(ARCHIVO_1,NOMBRE);
  REWRITE(ARCHIVO_1);
  FOR I:=1 TO N DO
    BEGIN
      STR(Y[I]:1:6,REGIS_1);
      FOR J:=1 TO P DO
        BEGIN
          STR(X[I,J]:1:6,VALOR);
          REGIS_1:=REGIS_1+','+VALOR;
        END;
      WRITELN(ARCHIVO_1,REGIS_1);
    END;
  CLOSE(ARCHIVO_1);
  GOTOXY(01,19);
  WRITELN('> OPERACION REALIZADA ');
  GOTOXY(01,24);
  CONTINUAR;
END;
END.


{===================================================================
|  LA UNIDAD "CARENCIA" HACE UNA ESTIMACION DE LA VARIANZA Y DA UN ESTA-
|  DISTICO QUE NOS DA INFORMACION SOBRE LA CARENCIA DE AJUSTE DEL MODELO.
|  LA ESTIMACION Y EL ESTADISTICO SE DAN POR EL METODO DEL ERROR PURO SI
|  EN LA MATRIZ X HAY RENGLONES REPETIDOS O POR EL METODO DE LA PRUEBA
|  ARCOIRIS SI NO LOS HAY.
|
|  VARIABLES.-   NE : GRADOS DE LIBERTAD DE SCEP, Y ES LA VARIABLE CON LA
|                     QUE SABEMOS SI HAY O NO RENGLONES REPETIDOS
===================================================================}

UNIT CAREN;
INTERFACE
  USES CRT,PRINTER,UGLOBAL,UUTIL;
  PROCEDURE CARENCIA;

IMPLEMENTATION
```

```
PROCEDURE CARENCIA;
VAR NE : INTEGEP;

   {=================================================================
   |  EL PROCEDIMIENTO "ERRORPURO" HACE LA ESTIMACION DE LA VARIANZA |
   |Y DEL ESTADISTICO F= (SCCA/(M-H))/(SCEP/NE) PARA LA CARENCIA DE A-|
   |JUSTE CUANDO EXISTEN RENGLONES REPETIDOS EN LA MATRIZ X.        |
   |                                                                |
   |VARIABLES.- M      : CONTADOR DEL NUMERO DE GRUPOS CON REPETICIO-|
   |                     NES                                         |
   |            NJ     : CONTADOR DEL NUMERO DE REPETICIONES PARA UN |
   |                     GRUPO                                       |
   |            IND    : VECTOR DONDE SE GUARDAN LOS NUMEROS DE REN- |
   |                     GLONES DONDE SE ENCONTRO REPETICION         |
   |            NR     : CONTADOR DEL NUMERO TOTAL DE REPETICIONES EN|
   |                     TODOS LOS GRUPOS                            |
   |            INJ    : VECTOR DONDE SE GUARDAN LOS NJ PARA CADA GRUPO|
   |            YCONST : SUMA DE LOS VALORES Yi PARA TODAS LAS REPETI-|
   |                     CIONES DE UN GRUPO                          |
   |            YREP   : MATRIZ DONDE SE GUARDAN LOS VALORES Yi DE TO-|
   |                     DAS LAS REPETICIONES PARA CADA GRUPO        |
   |            YMED   : VECTOR DONDE SE GUARDA LA MEDIA DE CADA GRUPO|
   |            SCEP   : SUMA DE CUADRADOS DEL ERROR PURO            |
   |            SCCA   : SUMA DE CUADRADOS DE LA CARENCIA DE AJUSTE  |
   |            SE2    : MEDIA DE LA SUMA DE CUADRADOS DEL ERROR PURO|
   |            MSCCA  : MEDIA DE LA SUMA DE CUADRADOS DE LA CARENCIA|
   |                     DE AJUSTE                                   |
   |            FE     : ESTADISTICO F DE LA CARENCIA DE AJUSTE      |
   |                                                                |
   |EXPLICACION : SE COMPARA EL RENGLON I-ESIMO CON LOS N-I-1 RESTAN-|
   |    TES, SI UNO O MAS DE ESTOS RENGLONES ES IGUAL AL I-ESIMO ENTON-|
   |    CES SE SUMAN SUS Y[K] CORRESPONDIENTES Y SE GUARDAN EN YREP POR|
   |    GRUPO PARA UN POSTERIOR USO, SE INCREMENTA EL CONTADOR DE GRU-|
   |    POS M EN 1 Y SE GUARDA EN INJ[M] EL NUMERO NJ DE REPETICIONES|
   |    DE ESE GRUPO. SE GUARDA EN IND[NR] EL RENGLON DONDE SE ENCONTRO|
   |    LA REPETICION CON EL FIN DE QUE AL PASAR POR ESE RENGLON POS-|
   |    TERIORMENTE NO SE REPITA EL PROCESO Y SE SALTE AL SIGUIENTE |
   |    RENGLON. POR CADA GRUPO SE OBTIENE UNA MEDIA RESPECTO A LAS |
   |    Y[K] REPETIDAS GUARDANDOSE EN YMED. UNA VEZ QUE SE HA PASADO|
   |    POR TODOS LOS RENGLONES SE CALCULAN SCEP, SCCA, SE2, MSCCA, FE|
   |    CON OPERACIONES ALGEBRAICAS SENCILLAS.                      |
   =================================================================}

PROCEDURE ERRORPURO;
VAR IND,YMED                    : MATRIZT2;
    INJ                         : MATRIZT2I;
    YREP                        : MATRIZT1;
    M,I,NJ,L,K,J,NR             : INTEGER;
    YCONST,SCEP,SCCA,FE,SE2,MSCCA : REAL;
    BANDERA:BOOLEAN;
BEGIN
  IND[1]:=0;
  NR:=1;
  M:=0;
```

```
(*****    BUSQUEDA DE LOS RENGLONES REPETIDOS DE LA MATRIZ X          *****)

  FOR I:=1 TO N-1 DO
    BEGIN
      NJ:=1;
      YCONST:=Y[I];
      BANDERA:=TRUE
      FOR L:=1 TO NR DO
       IF I=IND[L] THEN
         BANDERA:=FALSE;
      IF BANDERA THEN
        FOR K:=I+1 TO N DO
          BEGIN
            BANDERA:=TRUE;
            FOR J:=1 TO P DO
              IF X[I,J]<>X[K,J] THEN
              BANDERA:=FALSE;

(*****   SUMA DE LAS Yi DE LOS RENGLONES REPETIDOS DE LA MATRIZ X      *****)

            IF BANDERA THEN
            BEGIN
              YCONST:=YCONST + Y[K];
              IF NJ=1 THEN
              BEGIN
               M:=M+1;
               YREP[M,NJ]:=Y[I];
              END;
              NJ:=NJ + 1;
              NR:=NR + 1;
              IND[NR]:=K;
              YREP[M,NJ]:=Y[K]
            END
          END;

(*****              OBTENCION DE LA MEDIA PARA UN GRUPO                *****)

        IF NJ<>1 THEN
        BEGIN
          YMED[M]:=YCONST/NJ;
          INJ[M]:=NJ
        END
    END;
NE:=0;
SCEP:=0;

(*****              OBTENCION DE LAS SUMAS DE CUADRADOS                *****)

FOR L:=1 TO M DO
BEGIN
  FOR K:=1 TO INJ[L] DO
  SCEP:=SCEP + SQR(YREP[L,K]-YMED[L]);
  NE:=NE + INJ[L]-1
```

```
END;
IF NE>0 THEN
 BEGIN
 CLRSCR;
 M:=N-NE;
 SCCA:=SCRES-SCEP;
 MSCCA:=SCCA/(M-H);
 SE2:=SCEP/NE;
 FE:=MSCCA/SE2;
 WRITELN;
 IF IMPRESION THEN
  BEGIN
   WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
   WRITELN(LST,' ':20,' C A R E N C I A   D E   A J U S T E ',^J,^M);
  END;
 WRITELN(' ':20,' C A R E N C I A   D E   A J U S T E ');

(*****    TODOS LOS CALCULOS ANTERIORES SON RESUMIDOS EN UNA TABLA DE    *****
 *****    ANALISIS DE VARIANZA MEDIANTE EL PROCEDIMIENTO "ESCRIBE"       *****)

 ESCRIBE(2,M-H,NE,N-H,1,SCCA,MSCCA,FE,SCEP,SE2,SCRES,SE2);
 END;
END;

     {
        EL PROCEDIMIENTO "ARCOIRIS" CALCULA LA PRUEBA ESTADISTICA F PARA
       LA CARENCIA DE AJUSTE MEDIANTE EL METODO DE LA PRUEBA "ARCOIRIS"
       CUANDO NO EXISTEN RENGLONES REPETIDOS EN LA MATRIZ X.

       VARIABLES.- X_1    : MATRIZ DE LOS DATOS CENTRALES (VARIABLES IN-
                            DEPENDIENTES).
                   Y_1    : VECTOR DE LOS DATOS CENTRALES (VARIABLE DEPEN-
                            DIENTE).
                   XDX    : MATRIZ QUE RESULTA DE MULTIPLICAR LA TRANS-
                            PUESTA DE X_1 POR LA MISMA X_1.
                   XDY    : MATRIZ QUE RESULTA DE MULTIPLICAR LA TRANS-
                            PUESTA DE X_1 POR Y_1.
                   Z      : VECTOR QUE  GUARDA LA DIAGONAL DE LA MATRIZ
                            X(X'X)^X'
                   D_SOL  : VECTOR SOLUCION PARA EL MODELO TOMANDO EN
                            CUENTA SOLO LOS DATOS CENTRALES
                   O      : VECTOR QUE NOS AYUDA A MANEJAR EL ORDEN CRE-
                            CIENTE DE OTRO VECTOR.
                   FD     : ESTADISTICO F DE LA PRUEBA "ARCOIRIS"
                   D_SCRES: SUMA DE CUADRADOS DE LOS RESIDUALES TOMANDO EN
                            CUENTA SOLO LOS DATOS CENTRALES
                   N1     : NUMERO DE DATOS CENTRALES A CONSIDERAR

       EXPLICACION : SE OBTIENE LA DIAGONAL DE LA MATRIZ X(X'X)^X' Y SE
                     GUARDA EN Z, SE ENCUENTRA EL ORDEN CRECIENTE DE LOS N
                     DATOS DE Z, TOMANDO COMO RELACION ESTE ORDEN QUE ES
                     GUARDADO EN O, SE SELECCIONAN N1 RENGLONES DE LA MA-
                     TRIZ X Y EL VECTOR Y DE DATOS QUE SE GUARDAN EN X_1 Y
                     EN Y_1 RESPECTIVAMENTE,CON ESTAS ULTIMAS MATRICES SE
```

```
    !                    CALCULAN XDX Y XDY. SE OBTIENE LA SOLUCION DEL SISTE-|
    |                    MA   XDX * B' = XDY   QUE SE GUARDA EN D_SOL Y QUE   |
    |                    NOS SIRVE PARA CALCULAR LA SUMA DE CUADRADOS DE LOS  |
    |                    RESIDUALES, VALOR QUE NOS PERMITE CALCULAR EL ESTA   |
    |                    DISTICO DE LA PRUEBA ARCOIRIS.                       |
    |_____________________________________________________________________|}

PROCEDURE ARCOIRIS;
VAR  X_1                    : MATRIZT1;
     XDX                    : MATRIZT1C;
     Z,Y_1                  : MATRIZT2;
     XDY,D_SOL              : MATRIZT2C;
     O                      : MATRIZT2I;
     FD,D_SCRES             : REAL;
     I,J,N1                 : INTEGER;

BEGIN
      K:=0;
      REPEAT
        CLRSCR;
        WRITELN;
        WRITE('DAME EL No. DE RENGLONES (N1) DE LA MATRIZ  "D": '); READLN(N1);
      UNTIL (M < N1) AND (N1 < N) ;

      {*****    MEDIANTE EL PROCEDIMIENTO "D_XXXX" SE OBTIENE LA DIAGONAL DE  *****
       *****    LA MATRIZ X(X'X)^X' Y SE GUARDA EN Z, DE LA CUAL POSTERIOR-   *****
       *****    MENTE SE LE ENCUENTRA EL ORDEN CRECIENTE MEDIANTE EL PROCEDI- *****
       *****    MIENTO "QUICK".                                               *****}

          D_XXXX(Z);
          QUICK(Z,O,N);

      {*****                   OBTENCION DE X_1 Y DE Y_1                      *****}

          FOR I:= 1 TO N1 DO
                BEGIN
                  Y_1[I]:=Y[O[I]];
                  FOR J:=1 TO P DO
                    X_1[I,J]:=X[O[I],J];
                END;

      {*****    LA MATRIZ XDX Y EL VECTOR XDY SON CALCULADOS A PARTIR DE LA    *****
       *****    MATRIZ X_1 Y EL VECTOR Y_1 QUE POSEEN N1 DATOS MEDIANTE EL     *****
       *****    PROCEDIMIENTO "CALCULO_SUMAS". POSTERIORMENTE SE ENCUENTRA     *****
       *****    SOLUCION DEL SISTEMA (XDX)(D_SOL)=XDY MEDIANTE EL PROCEDI-     *****
       *****    MIENTO "INVERSA".                                              *****}

          CALCULO_SUMAS(XDX,XDY,X_1,Y_1,P,N1);
          SOLUCIONAR:=TRUE; INVERTIR:=FALSE;
          INVERSA(XDX,INVERSION,SOLUCIONAR,INVERTIR,M,D_SOL,XDY);

      {*****              OBTENCION DE LAS SUMAS DE CUADRADOS                 *****}

          IF INVERSION THEN
```

```
            BEGIN
             D_SCRES:=0;
             FOR I:=1 TO N1 DO
                 D_SCRES:=D_SCRES+SQR(Y_1[I]);
             FOR I:=1 TO H DO
                 D_SCRES:=D_SCRES-(D_SOL[I]*DY[I]);
             FD:=((SCRES-D_SCRES)/(N-N1))/(D_SCRES/(N1-H));
              IF IMPRESION THEN BEGIN WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
                WRITELN(LST,' ':25,' P R U E B A    A R C O I R I S',^J,^M); END;
             CLRSCR;
             WRITELN(' ':25,' P R U E B A    A R C O I R I S ');

      (*****   EL ESTADISTICO Y LA INFORMACION CONCERNIENTE A ESTE ES PRE-   *****
       *****   SENTADA EN UNA TABLA DE ANALISIS DE VARIANZA MEDIANTE EL PRO- *****
       *****   CEDIMIENTO "ESCRIBE".                                         *****)

             ESCRIBE(3,N-N1,N1-H,N-1,N-H,SCRES-D_SCRES,(SCRES-D_SCRES)/(N-N1),FD,D_SCRES,D_SCRES/(
           END
          ELSE
           WRITELN('NO SE ENCONTRO LA INVERSA');
      END;

      (*****              PROGRAMA PRINCIPAL DE CARENCIA                     *****)

      BEGIN

      (*****   LA CARENCIA DE AJUSTE CUANDO HAY REPETICIONES SE CALCULA POR  *****
       *****   MEDIO DEL PROCEDIMIENTO "ERRORPURO"                           *****)

      ERRORPURO;
      IF NE=0 THEN

      (*****   LA CARENCIA DE AJUSTE CUANDO NO HAY REPETICIONES SE CALCULA   *****
       *****   POR MEDIO DEL PROCEDIMIENTO "ARCOIRIS"                        *****)

       ARCOIRIS;
      END;
      END.
{
    LA UNIDAD "HIPOTESI" DA A ESCOGER ENTRE LAS SIGUIENTES PRUEBAS DE
  HIPOTESIS:
          A) PRUEBAS DE HIPOTESIS INDIVIDUALES
          B) PRUEBA DE HIPOTESIS CONJUNTA
          C) CONTRIBUCION DE UN SUBCONJUNTO DE VARIABLES A LA REGRESION
          D) PRUEBA DE HIPOTESIS PARA COMBINACIONES LINEALES DE LOS ES-
             TIMADORES
}

UNIT HIPOTESI;
INTERFACE
  USES CRT,PRINTER,UGLOBAL,UUTIL;
  PROCEDURE PRUEBAS_HIPOTESIS;
IMPLEMENTATION
```

```
{
   EL  PROCEDIMIENTO "HIP_IND" CALCULA EL ESTADISTICO t PARA LAS
  PRUEBAS DE HIPOTESIS INDIVIDUALES DE LOS ESTIMADORES DEL MODELO.

  VARIABLES : T     : ESTADISTICO t CALCULADO PARA LAS PRUEBAS DE
                      HIPOTESIS INDIVIDUALES.
              SBEST : DESVIACION ESTANDAR DE EL ESTIMADOR.
              K     : VALOR CONTRA EL CUAL SE COMPARA EL ESTIMADOR
                      EN LA PUEBA DE HIPOTESIS.
}

PROCEDURE HIP_IND;
VAR  I,J,L        :INTEGER;
     T,SBEST,K    :REAL;
     SN           :CHAR;
BEGIN
   CLRSCR;
   WRITELN;
   WRITELN(' ':5,'HIPOTESIS NULA              HIPOTESIS ALTERNATIVAS');
   WRITELN;
   WRITELN(' ':5,' Ho: β[i] = K                  H1: β[i] <> K,');
   WRITELN(' ':5,' Ho: β[i] = K                  H1: β[i]  > K');
   WRITELN(' ':5,' Ho: β[i] = K                o H1: β[i]  < K');
   L:=7;
   IF IMPRESION THEN
     BEGIN
       WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
       WRITELN(LST,'                 PRUEBAS DE HIPOTESIS INDIVIDUALES');
       WRITELN(LST);
       WRITELN(LST,'--------------------------------------------------------------');
       WRITELN(LST,' ':5,'BETA ESTIMADA     ERROR ESTANDARD       t-CALCULADAVALOR-P');
       WRITELN(LST,' ':5,'                  DE BETA ESTIMADA                        ');
       WRITELN(LST,'..................................................................');
       WRITELN(LST);
     END;
   REPEAT
     REPEAT
        GOTOXY(01,22);
        WRITE(' > SOBRE CUAL DE LAS B[i] (i=',vt1:2,',...,P) QUIERES HACER LA
        PRUEBA  [  ]');
        GOTOXY(65,22);
        I:=LEER_INT;
     UNTIL I IN [VT1..P];
     GOTOXY(01,23);
     WRITE('> DAME EL VALOR DE "K" PARA LA β[',I:1,'] =');
     K:=LEER_REAL;  CLREOL;
     J:=I+1;
     SBEST:=S*SQRT(MXTX[J+1,J]);
     IF SBEST = 0 THEN T:=0
     ELSE T:=(MBEST[J]-K)/SBEST;
     IF L=7 THEN
       BEGIN
         GOTOXY(01,02);
         WRITELN('————————————————————————————————————————————————————————');
```

```
        WRITELN(' ':5,'BETA ESTIMADA      ERROR ESTANDARD        t-CALCULADA VALOR-P');
        WRITELN(' ':5,'                   DE BETA ESTIMADA                        ');
        WRITELN('.............................................................');
        CLREOL;
     END;
     GOTOXY(5,L);
     WRITELN('B[',I:2,']=',MBEST[j]:12:7,' ':2,SBEST:12:7,' ':11,T:12:7,'    ',
     TS_P(ABS(T),N-H):6:3);
     IF IMPRESION THEN
     WRITELN(LST,'B[',i:2,']=',Mbest[j]:12:7,' ':4,SBEST:12:7,' ':12,T:12:7,'    ',

     TS_P(ABS(T),N-H):6:3);
     GOTOXY(01,25);
     WRITE('> QUIERES HACER OTRA PRUEBA S/N...? ');
     SN:=READKEY;
     WRITE(SN);
     L:=L+1;
     IF L=21 THEN L:=7;
   UNTIL SN IN ['n','N'];
END;

     {
     EL  PROCEDIMIENTO 'HIP_CONJ' CALCULA EL ESTADISTICO F PARA LA
     PRUEBA DE HIPOTESIS CONJUNTA DE LOS ESTIMADORES DEL MODELO.

     VARIABLES : F      : ESTADISTICO F CALCULADO PARA LA PRUEBA DE HI-
                          POTESIS CONJUNTA.
     }

PROCEDURE HIP_CONJ;
VAR   F    :REAL;
BEGIN
  F:=SCREG/(P*SQR(S));
  GOTOXY(05,19);
  WRITELN('                  PRUEBA DE HIPOTESIS CONJUNTA');
  WRITELN;
  WRITELN(' ':20,'F -CALCULADA           VALOR-P');
  WRITE(' ':20,F:14:10,' ':8,F_P(F,P,N-H):6:3);
  IF IMPRESION THEN
    BEGIN
      WRITELN(LST,'        PRUEBA DE HIPOTESIS CONJUNTA');WRITELN(LST);
      WRITELN(LST,^J,^M,^M,^M,' ':20,'F -CALCULADA           VALOR-P');
      WRITE(LST,' ':20,F:14:10,' ':8,F_P(F,P,N-H):6:3);
    END;
    GOTOXY(5,25);
    CONTINUAR;
END;

     {
     EL PROCEDIMIENTO 'HIPOTESIS_GENERAL' REALIZA LA PRUEBA DE HIPO-
     TESIS QUE MUESTRA LA CONTRIBUCION DE UN SUBCONJUNTO DE VARIABLES
     A LA REGRESION.

     VARIABLES : FW       : ESTADISTICO F PARA LA PRUEBA DE HIPOTESIS
```

```
                         DE LA CONTRIBUCION DE UN SUBCONJUNTO.
           QI       : NUMERO DE VARIABLES DEL SUBCONJUNTO A PRO
                      BAR.
           SCTW     : SUMA DE CUADRADOS TOTAL DE LA REGRESION
                      PARA EL SUBCONJUNTO DE QI VARIABLES.
           SCREGW   : SUMA DE CADRADOS DE LA REGRESION PARA EL
                      SUBCONJUNTO DE QI VARIABLES.
           SCRESW   : SUMA DE CUADRADOS DE LOS RESIDUALES PARA
                      EL SUBCONJUNTO DE QI VARIABLES.
           SCRESWMT : SUMA DE CUADRADOS DE LOS RESIDUALES PARA
                      P - QI VARIABLES.
           MWTW     : MATRIZ X'X CONSIDERANDO AQUELLAS VARIABLES
                      DEL SUBCONJUNTO.
           MWTY     : VECTOR X'Y QUE CONSIDERA  A LAS VARIABLES
                      SELECCIONADAS PARA EL D
           MBESTW   : BETAS ESTIMADAS PARA EL SUBCONJUNTO DE
                      VARIABLES.
           INW      : VECTOR QUE CONTIENE LOS INDICES DE LAS
                      VARIABLES DEL SUBCONJUNTO
                                                              )

PROCEDURE HIPOTESIS_GENERAL;
VAR  INWY,QI,KH,KL,KM                                            :INTEGER;
     SUMAYW,SCTW,MBWTY,SCREGW,SCRESW,SCRESWMT,MSCRESWMT,FW       :REAL;
     CAMBIOY                                                     :BOOLEAN;
     MWTW                                                        :MATRIZT1C;
     MWTY,MBESTW                                                 :MATRIZT2C;
     INW                                                         :MATRIZT2IC;
BEGIN
  GOTOXY(1,4);
  REPEAT
    WRITE('> CUAL ES EL NUMERO DE VARIABLES DEL SUBCONJUNTO QUE',^J,^M,'> SE QUIERE PROBAR? : ');
    QI:=LEER_INT;
  UNTIL QI IN [1..P];
  WRITELN;
  WRITELN('> CUALES SON LAS VARIABLES : ');
  FOR KH:=1 TO QI DO
    REPEAT
      WRITE('> VAR ',KH,': ');
      INW[KH]:=LEER_INT;
    UNTIL INW[KH] IN [1..P];
    CAMBIOY:=FALSE;
    INWY:=1;
    KL:=QI+1-VTI;
    SOLUCIONAR:=TRUE;
    INVERTIR:=FALSE;
    CAMBIO_MATRIZ(MWTW,MWTY,INW,INWY,QI,CAMBIOY);
    INVERSA(MWTW,INVERSION,SOLUCIONAR,INVERTIR,KL,MBESTW,MWTY);
    IF INVERSION THEN
      BEGIN
        SUMAYW:=0;
        FOR KH:=1 TO N DO
        SUMAYW:=SUMAYW+SQR(Y[KH]);
        IF VTI=0 THEN SCTW:=SUMAYW-SQR(MWTY[1])/N
```

```
            ELSE SCTW:=SUMAYW;
        MBWTY:=0;
        FOR KH:=1 TO OI+1-vtl DO
        MBWTY:=MBWTY+MBESTW[KH]*MWTY[KH];
        IF VTI=0 THEN SCREGW:=MBWTY-SQR(MWTY[1])/N
                 ELSE SCREGW:=MBWTY;
        SCRESW:=SCTW-SCREGW;
        SCRESWMT:=SCRESW-SCRES;
        MSCRESWMT:=SCRESWMT/(H-OI);
        FW:=MSCRESWMT/S2;
        CLRSCR;GOTOXY(01,04);
        IF IMPRESION THEN
           BEGIN
             WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
             WRITELN(LST,' ':10,'C O N T R I B U C I O N   D E   U N  S U B C O N J U N T O');
           END;
        WRITELN(' ':10,'C O N T R I B U C I O N   D E   U N   S U B C O N J U N T O');
        ESCRIBE(4,P-OI,N-H,N-1,OI,SCRESWMT,MSCRESWMT,FW,SCRES,S2,SCT,SCREGW);
      END
    ELSE
      BEGIN
        WRITELN;
        WRITELN('> NO SE ENCONTRO LA INVERSA');WRITELN;
        CONTINUAR;
      END;
END;

     {
        EL PROCEDIMIENTO "PRUEBA_HIPOTESIS_COMBINACIONES: REALIZA LA
        PRUEBA DE HIPOTESIS PARA COMBINACIONES LINEALES DE LOS ESTIMADO-
        RES.

        VARIABLES : FR       : ESTADISTICO F PARA LA PRUEBA DE HIPOTESIS
                               DE COMBINACIONES LINEALES.
                    SCTW     : SUMA DE CUADRADOS TOTAL DE LA REGRESION
                               PARA LAS COMBINACIONES LINEALES.
                    SCREGW   : SUMA DE CADRADOS DE LA REGRESION PARA LAS
                               COMBINACIONES LINEALES.
                    SCRESW   : SUMA DE CUADRADOS DE LOS RESIDUALES PARA
                               LAS COMBINACIONES LINEALES.
                                                                        }

PROCEDURE PRUEBA_HIPOTESIS_COMBINACIONES;
VAR RO,CY,CZ                    :INTEGER;
    XT                          :MATRIZT2;
    MBESTW                      :MATRIZT2C;
    C,GA,MIITX,XM,MWTW          :MATRIZTIC;
    CV,SCREGQ,SCREGR,MSCREGR,FR :REAL;

     {
        EL  PROCEDIMIENTO "TRANSPUESTA" TRANSPONE UNA MATRIZ.

        VAIABLES MA : MATRIZ A TRANSPONER.
                 RE : RENSLONES DE LA MATRIZ MA.
```

```
|              CO : COLUMNAS DE LA MATRIZ MA.                    |
|                                                                |
 ----------------------------------------------------------------}

PROCEDURE TRANSPUESTA(VAR MA:MATRIZTIC;VAR RE,CO:INTEGER);
VAR CX,CW,XEW  :INTEGER;
BEGIN
 FOR CX:=1 TO RE DO
 BEGIN
   IF RE>CO THEN XEW:=1 ELSE XEW:=CX+1;
   FOR CW:=XEW TO CO DO
     BEGIN
       CV:=MA[CX,CW];
       MA[CX,CW]:=MA[CW,CX];
       MA[CW,CX]:=CV;
     END;
 END;
END;
BEGIN
  GOTOXY(01,04);
  WRITELN('> PRUEBA : C(BETA=GAMMA');
  WRITELN('> LA MATRIZ C ES DE O x ',H:2);WRITELN;
  REPEAT
    WRITE('> DA EL VALOR DE O (DEBE SER MENOR O IGUAL QUE ',H:2,' ) :');
    RO:=LEER_INT;
  UNTIL RO IN [1..H];
  WRITELN('> DA LA MATRIZ C  EN FORMATO LIBRE');WRITELN;
  FOR CZ:=1 TO RO DO
    BEGIN
      WRITELN('> DA EL RENGLON ',CZ:3);
      LEER_VEC_REAL(XT,H);
      FOR CY:=1 TO H DO
      C[CZ,CY]:=XT[CY];
      WRITELN;
    END;
  WRITELN('> DA LA MATRIZ GAMMA-TRANSPUESTA DE 1 x ',RO:2,' EN FORMATO LIBRE');
  LEER_VEC_REAL(XT,RO);
  FOR CZ:=1 TO RO DO
    GA[CZ,1]:=XT[CZ]:
  FOR CZ:=1 TO H DO
    FOR CY:=CZ TO H DO
      BEGIN
        MIXTX[CZ,CY]:=MXTX[CY+1,CZ];
        MIXTX[CY,CZ]:=MXTX[CY+1,CZ];
      END;
  MULTIPLICACION(MWTW,C,MIXTX,RO,H,H);
  TRANSPUESTA(C,RO,H);
  MULTIPLICACION(XM,MWTW,C,RO,H,50);
  SOLUCIONAR:=FALSE; INVERTIR:=TRUE;
  INVERSA(XM,INVERSION,SOLUCIONAR,INVERTIR,RO,MBESTW,MXTY);
  IF INVERSION THEN
    BEGIN
     FOR CZ:=1 TO RO DO
       FOR CY:=CZ TO RO DO
```

```
            BEGIN
              XM[CZ,CY]:=XM[CY+1,CZ];
              XM[CY,CZ]:=XM[CY+1,CZ];
            END;
      FOR CZ:=1 TO H DO
        MWTW[CZ,1]:=MBEST[CZ];
      CY:=1;
      TRANSPUESTA(C,H,RO);
      MULTIPLICACION(MIXTX,C,MWTW,RO,H,CY);
      FOR CZ:=1 TO RO DO
        MWTW[CZ,1]:=MIXTX[CZ,1]-GA[CZ,1];
      MULTIPLICACION(MIXTX,XM,MWTW,RO,RO,CY);
      TRANSPUESTA(MWTW,RO,CY);
      MULTIPLICACION(XM,MWTW,MIXTX,CY,RO,CY);
      SCREGR:=XM[1,1];
      MSCREGR:=SCREGR/(H-RO);
      FR:=MSCREGR/S2;
      SCREGO:=SCREG-SCREGR;
      IF IMPRESION THEN
        BEGIN
         WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
         WRITELN(LST,' ':18,'C O M B I N A C I O N E S   D E   B E T A S',^J,^M);
        END;
      CLRSCR;GOTOXY(01,03);
      WRITELN(' ':18,'C O M B I N A C I O N E S   D E   B E T A S');
      ESCRIBE(4,F-RO,N-H,N-1,RO,SCREGR,MSCREGR,FR,SCRES,S2,SCT,SCREGO);
     END
   ELSE
     BEGIN
       WRITELN;
       WRITELN('NO SE ENCONTRO LA INVERSA');WRITELN;
       CONTINUAR;
     END;
END;

      {╔══════════════════════════════════════════════════════════╗
       ║  EL PROCEDIMIENTO "PRUEBAS_HIPOTESIS" CONTIENE AL MENU DE LAS ║
       ║ PRUEBAS DE HIPOTESIS.                                        ║
       ║                                                              ║
       ╚══════════════════════════════════════════════════════════╝}

PROCEDURE PRUEBAS_HIPOTESIS;
VAR  OP  :CHAR;

BEGIN
 REPEAT
   CLRSCR;
   WRITELN;
   WRITELN(' ':10,'╔══════════════════════════════════════════╗');
   WRITELN(' ':10,'║             MENU PRUEBAS DE HIPOTESIS    ║');
   WRITELN(' ':10,'║                                          ║');
   WRITELN(' ':10,'║ A '^P' PRUEBAS DE HIPOTESIS INDIVIDUALES   ║');
   WRITELN(' ':10,'║ B '^P' PRUEBA DE HIPOTESIS CONJUNTA        ║');
   WRITELN(' ':10,'║ C '^P' CONTRIBUCION DE UN SUBCONJUNTO DE VARIABLES ║');
```

```
   WRITELN(' ':10,'|      A LA REGRESION                              |');
   WRITELN(' ':10,'| D '^P' PRUEBA DE HIPOTESIS PARA COMBINACIONES    |');
   WRITELN(' ':10,'|      LINEALES DE LOS ESTIMADORES                 |');
   WRITELN(' ':10,'| E '^P' RESULTADOS POR IMPRESORA                  |');
   WRITELN(' ':10,'|                                                  |');
   WRITELN(' ':10,'| S '^P' SALIDA                                    |');
   WRITELN(' ':10,'└──────────────────────────────────────────────────┘');

(******   LA SALIDA O NO DE LOS PESULTADOS POR IMPRESORA ES CONTROLADA   ******
 ******   POR MEDIO DEL PROCEDIMIENTO 'IMPRESORA'.                       ******)

   IMPRESORA(IMPRESION,11,1);
   REPEAT
     GOTOXY(01,16);
     WRITE(' ':10,'> ESCOGE TU OPCION : ');
     OP:=READKEY;
     WRITE(OP);
     OP:=UPCASE(OP);
   UNTIL OP IN ['A','B','C','D','E','S'];
   CASE OP OF

(******   LAS PRUEBAS DE HIPOTESIS INDIVIDUALES SON REALIZADAS POR EL    ******
 ******   PROCEDIMIENTO 'HIP_IND'.                                       ******)

     'A' : HIP_IND;

(******   LA PRUEBA DE HIPOTESIS CONJUNTA ES REALIZADA POR EL PROCEDI-   ******
 ******   MIENTO 'HIP_CONJ'.                                             ******)

     'B' : HIP_CONJ;
     'C' :BEGIN
            CLRSCR;

(******   LA PRUEBA DE HIPOTESIS GENERAL ES REALIZADA POR EL PROCEDI-    ******
 ******   MIENTO 'HIPOTESIS_GENERAL'                                     ******)

            HIPOTESIS_GENERAL;
          END;
     'D' :BEGIN
            CLRSCR;

(******   LA PRUEBA DE HIPOTESIS PARA COMBINACIONES DE BETAS ES REALI-   ******
 ******   ZADA POE EL PROCEDIMIENTO 'PRUEBA_HIPOTESIS_COMBINACIONES'     ******)

            PRUEBA_HIPOTESIS_COMBINACIONES;
          END;

(******   LA SALIDA O NO DE LOS RESULTADOS POR IMPRESORA ES CONTROLADA   ******
 ******   POR MEDIO DEL PROCEDIMIENTO 'IMPRESORA'.                       ******)

     'E' :IMPRESORA(IMPRESION,11,2);
   END;
 UNTIL OP IN ['S']
END;
```

```
END.
{
   LA UNIDAD 'INTERVAL' CONTIENE LOS PROCESOS PARA EL CALCULO DE LOS IN-
   TERVALOS DE CONFIANZA.
   TIENE COMO OPCIONES:
         A) INTERVALOS DE CONFIANZA INDIVIDUALES.
         B) INTERVALO DE CONFIANZA PARA LA VARIANZA
}

UNIT INTERVAL;
INTERFACE
   USES CRT,PRINTER,UGLOBAL,UUTIL;
   PROCEDURE INT_CONF;

IMPLEMENTATION

PROCEDURE INT_CONF;
VAR   YN,SN    : CHAR;

     {
        EL PROCEDIMIENTO 'IC_IND' CALCULA LOS INTERVALOS DE CONFIANZA
       INDIVIDUALES, ES DECIR PARA CADA UNO DE LOS ESTIMADORES.

       VARIABLES : T     : CONFIABILIDAD.
                   L_INF : LIMITE INFERIOR DEL INTERVALO.
                   L_SUP : LIMITE SUPERIOR DEL INTERVALO
                   SBEST : DESVIACION ESTANDAR DE UN ESTIMADOR
     }

PROCEDURE IC_IND;
VAR I,J,K,L                : INTEGER;
    T,L_INF,L_SUP,SBEST    : REAL;
BEGIN
  CLRSCR;
  WRITELN;
  WRITELN(' ':10,'INTERVALO DE CONFIANZA INDIVIDUAL');
  GOTOXY(01,03);
  WRITELN('-----------------------------------------------------------------------');
  WRITELN(' ','BETA ESTIMADA      ERROR ESTANDARD           t          INT. DE CONFIANZA');
  WRITELN(' ','                   DE BETA ESTIMADA                   ');
  WRITELN('.......................................................................');CLREOL;
  IF IMPRESION THEN
   BEGIN
    WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
    WRITELN(lst,' ':10,'INTERVALO DE CONFIANZA INDIVIDUAL');
    WRITELN(LST,'-----------------------------------------------------------------------');
    WRITELN(LST,' ','BETA ESTIMADA      ERROR ESTANDARD           t          INT. DE CONFIANZA');
    WRITELN(LST,' ','                   DE BETA ESTIMADA                   ');
    WRITELN(LST,'.......................................................................');
   END;
  L:=7;
  REPEAT
    REPEAT
      GOTOXY(05,22);
```

```
        WRITE('CUAL DE LAS B(i) (I=',vt1:2,',...,P) QUIERES EL INTERVALO DE CONFIANZA (  )');
        GOTOXY(70,22); I:=LEER_INT;
      UNTIL (i)=vt1) and (i<=p);
      REPEAT
        GOTOXY(05,23);
        WRITE('DAME LA CONFIABILIDAD= ');
        t:=LEER_REAL;
      UNTIL (T>0) AND (T<1);
      CLREOL;

  {*****   EL INVERSO PARA EL ESTADISTICO T DE LA CONFIABILIDAD DADA PARA *****
   *****   EL INTERVALO REQUERIDO SE OBTIENE MEDIANTE EL PROCEDIMIENTO    *****
   *****   "T_INV"                                                        *****}

      T:=T_INV((1-T)/2,N-M);
      J:=I+1;

  {*****              OBTENCION DEL INTERVALO DE CONFIANZA               *****}

      SBEST:=S*SQRT(MXTX[J+1,J]);
      L_INF:=MBEST[J]-T*SBEST;
      L_SUP:=MBEST[J]+T*SBEST;
      GOTOXY(1,L);
      WRITELN('B[',I:1,']=',MBEST[J]:12:7,' ':2,SBEST:12:7,' ':11,T:12:7,'  (',l_inf:8:5,' , ',l_sup
      IF IMPRESION THEN
       BEGIN
         WRITELN(LST,^J,^M,^M);
         WRITELN(LST,'B[',I:1,']=',MBEST[J]:12:7,' ':2,SBEST:12:7,' ':11,T:12:7,'  (',l_inf:8:5,' ,
       END;
      GOTOXY(05,25);
      WRITE('QUIERES HACER OTRA PRUEBA S/N...? ');
      READ(KBD,SN);
      WRITE(SN);
      L:=L+1;
      IF L=21 THEN L:=7;
    UNTIL SN IN ['n','N'];
  END;

      {
       EL PROCEDIMIENTO "IC_IND" CALCULA EL INTERVALO DE CONFIANZA
      PARA LA VARIANZA.

      VARIABLES : X1    : VALOR DE CHI-CUADRADA PARA CALCULAR EL LIMITE
                          INFERIOR DEL INTERVALO.
                  X2    : VALOR DE CHI-CUADRADA PARA CALCULAR EL LIMITE
                          SUPERIOR DEL INTERVALO
                                                                          }

PROCEDURE INT_SIGMA;
VAR    X1,X2     : REAL;
BEGIN
  REPEAT
    CLRSCR;
    GOTOXY(5,5);WRITELN(' ':5,'INTERVALO DE CONFIANZA DE    ' ',^J^J^M);
```

```
    WRITELN(' Pr [ X² 1-α·½  ≤  σ²  ≤ X² α·½ ] = α ');
    WRITELN;
    IF IMPRESION THEN
      BEGIN        WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
       WRITELN(LST,' ':5,'INTERVALO ED CONFIANZA DE    σ² ',^J^J^M);
       WRITELN(LST,' Pr [ X² 1-α·½  ≤  σ²  ≤ X² α·½ ] = α ');
       WRITELN(LST);
      END;
    REPEAT
      WRITE('DAME EL VALOR DE "X²" PARA (1 - α·½ ), = ');
      X1:=LEER_REAL;
    UNTIL X1>0;
    REPEAT
      WRITE('DAME EL VALOR DE "X²" PARA (α·½),       = ');
      X2:=LEER_REAL;
    UNTIL X2 > X1;
    WRITELN;WRITELN;
    WRITELN('σ² = ',SQR(S):14:10,^J^M);
    WRITELN('  IC ≡ [',SCRES/X2:14:8,',',SCRES/X1:14:8,']');
    IF IMPRESION THEN
      BEGIN
        WRITELN(LST,'σ² = ',SQR(S):14:10,^J^M);
        WRITELN(LST,'  IC ≡ [',SCRES/X2:14:8,',',SCRES/X1:14:8,']');
      END;
    GOTOXY(05,25);
    WRITE('QUIERES HACER OTRA PRUEBA S/N...? ');
    READ(KBD,SN);
    WRITE(SN);
  UNTIL SN IN ['n','N'];
END;

(*****         PROGRAMA PRINCIPAL DE INTERVALOS DE CONFIANZA          *****)

BEGIN
REPEAT
  CLRSCR;
  GOTOXY(01,04);
  WRITELN(' ':10,'╔══════════════════════════════════════════════════╗');
  WRITELN(' ':10,'║             MENU INTERVALOS DE CONFIANZA         ║');
  WRITELN(' ':10,'║                                                  ║');
  WRITELN(' ':10,'║ A '^P' INTERVALOS DE CONFIANZA INDIVIDUALES      ║');
  WRITELN(' ':10,'║ B '^P' INTERVALO DE CONFIANZA PARA LA VARIANZA   ║');
  WRITELN(' ':10,'║ C '^P' RESULTADOS POR IMPRESORA                  ║');
  WRITELN(' ':10,'║                                                  ║');
  WRITELN(' ':10,'║ S '^P' SALIDA                                    ║');
  WRITELN(' ':10,'╚══════════════════════════════════════════════════╝');
  IMPRESORA(IMPRESION,9,1);
  REPEAT
     GOTOXY(01,14);
     WRITE(' ':10,'> ESCOGE TU OPCION : ');
     YN:=READKEY;
     WRITE(YN); YN:=UPCASE(YN);
  UNTIL  YN IN ['A','B','C','S'];
  CASE YN OF
```

```
{*****  LOS INTERVALOS DE CONFIANZA INDIVIDUALES SON CALCULADOS POR     *****
 *****  EL PROCEDIMIENTO "IC_IND"                                       *****}

       'a','A' : IC_IND;

{*****  EL INTERVALO DE CONFIANZA PARA LA VARIANZA ES CALCULADO POR     *****
 *****  EL PROCEDIMIENTO "INT_SIGMA".                                   *****}

       'b','B' : INT_SIGMA;

{*****  LA SALIDA O NO DE LOS RESULTADOS POR IMPRESORA ES CONTROLADA    *****
 *****  POR MEDIO DEL PROCEDIMIENTO "IMPRESORA".                        *****}

       'c','C' :IMPRESORA(IMPRESION,9,2);
  END;
UNTIL YN IN ['s','S']
END;
END.

{==================================================================
 |  LA UNIDAD "MULTICOL" CONTIENE PROCESOS QUE REALIZAN CALCULOS O DAN |
 | INFORMACION QUE AYUDA A DETECTAR LA MULTICOLINEALIDAD.             |
 | TIENE COMO OPCIONES:                                               |
 |       A) MATRIZ DE CORRELACION                                     |
 |       B) REGRESION DE UNA VARIABLE EXPLICATIVA CON LAS RESTANTES   |
 ==================================================================}

 UNIT MULTICOL;
 INTERFACE
   USES CRT,PRINTER,UGLOBAL,UUTIL;
   PROCEDURE MULTICOLINEALIDAD;

 IMPLEMENTATION

 PROCEDURE MULTICOLINEALIDAD;
 VAR OP:CHAR;

    {=============================================================
     |  EL PROCEDIMIENTO "MATRIZ_CORRELACION" CALCULA LA TRIANGULAR SU- |
     | PERIOR DE LA MATRIZ DE CORRELACION.                         |
     |                                                             |
     | VARIABLES : MCORR : MATRIZ DE CORRELACION.                  |
     =============================================================}

PROCEDURE MATRIZ_CORRELACION;
VAR   I,J    :INTEGER;
      MCORR  :MATRIZTIC;
BEGIN
MCORR[H,H]:=1;
MCORR[1,1]:=1;
FOR I:=2-VT1 TO P DO
  BEGIN
    MCORR[1,I]:=(N*MXTY[I]-MXTX[VT2,I]*MXTY[VT2])/SQRT((N*MXTX[I,I]-SQR(MXTX[VT2,I]))*(N*SCT));
```

```
    FOR J:=I+1-VT1 TO H DO
     MCORR[I,J]:=(N*MXTX[I,J]-MXTX[VT2,I]*MXTX[VT2,J])/SQRT((N*MXTX[I,I]-SQR(MXTX[VT2,I]))*(N*MXTX[
     MCORR[I,I]:=1;
  END;
MCORR[1,H]:=(N*MXTY[H]-MXTX[VT2,H]*MXTY[VT2])/SQRT((N*MXTX[H,H]-SQR(MXTY[VT2,H]))*(N*SCT));
WRITELN(' MATRIZ DE CORRELACION':55);
WRITELN;
NOD:=H DIV 8;
JI:=1;
JF:=0;
CONT:=1;
REPEAT
  IF CONT>NOD THEN
    BEGIN
      JI:=JF+1;
      JF:=H;
  END
  ELSE
    IF CONT=1 THEN JF:=8
    ELSE
      BEGIN
        JI:=JF+1;
        JF:=JF+8;
      END;
  WRITE('          ');
  FOR J:=JI TO JF DO
       WRITE('X(',J-1+VT1:2,')    '); WRITELN;
  FOR I:=1 TO H DO
     BEGIN
       WRITE('X(',I-1+VT1:2,')  ');
       FOR J:=JI TO JF DO
        IF J<I THEN  WRITE(' ':9)
        ELSE  WRITE(' ',MCORR[I,J]:8:5);
        WRITELN;
     END;
  CONTINUAR;
  CLRSCR;
  CONT:=CONT+1;
UNTIL (CONT-1)>NOD;
IF IMPRESION THEN
 BEGIN
    WRITELN(LST);WRITELN(LST);WRITELN(LST);
    WRITELN(LST,'MATRIZ DE CORRELACION':55,^J,^M);
    JI:=1;
    JF:=0;
    CONT:=1;
    REPEAT
       IF CONT>NOD THEN
         BEGIN
           JI:=JF+1:
           JF:=H;
         END
       ELSE
         IF CONT=1 THEN JF:=8
```

```
          ELSE
           BEGIN
            JI:=JF+1;
            JF:=JF+8;
           END;
        WRITE(LST,'          ');
        FOR J:=JI TO JF DO
           WRITE(LST,'X(',J-1+VTI:2,')     ');
        WRITELN(LST);
        FOR I:=1 TO N DO
           BEGIN
             WRITE(LST,'X(',I-1+VTI:2,')  ');
             FOR J:=JI TO JF DO
              IF J<I THEN      WRITE(LST,' ':9)
              ELSE   WRITE(LST,' ',MCORR[I,J]:8:5);
               WRITELN(LST);
           END;
        CONT:=CONT+1;
     UNTIL (CONT-1)>NOD;
  END;
 END;
```

```
(*--------------------------------------------------------------------
  EL PROCEDIMIENTO 'REGRESION_X' REALIZA LA REGRESION DE UNA VA-
 RIABLE DEPENDIENTE -DEPENDIENTE EN EL MODELO ORIGINAL-, CONTRA LAS
 RESTANTES VARIABLES DEPENDIENTES.

 VARIABLES : VARIN   : INDICE DE LA VARIABLE EXPLICATIVA QUE SE
                       TOMA COMO VARIABLE DEPENDIENTE.
             INW     : VECTOR DONDE SE GUARDAN LOS INDICES DE LAS
                       VARIABLES A PROBAR.
             SCTR    : SUMA DE CUADRADOS TOTALES PARA ESTE MODELO.
             SCREGR  : SUMA DE CUADRADOS DE LA REGRESION PARA ESTE
                       MODELO.
             SCRESR  : SUMA DE CUADRADOS DE LOS RESIDUALES PARA ESTE
                       MODELO.
             MSCREGR : MEDIA DE EL SCREGR.
             FR      : ESTADISTICO F PARA ESTE MODELO.
             SR      : MEDIA DEL SCRESR.
             SR2     : SR AL CUADRADO.
             RR      : COEFICIENTE DE DETERMINACION PARA ESTE MODE-
                       LO.
             RAR     : COEFICIENTE DE DETERMINACION AJUSTADO PARA
                       ESTE MODELO.
             CAMBIOY : VARIABLE DE CONTROL EN EL PROCEDIMIENTO
                       'CAMBIO_MATRIZ'.
             MWTW    : MATRIZ CUADRADA QUE RESULTA DE LA MULTIPLICA-
                       CACION DE LA TRANSPUESTA DE LA MATRIZ DE VA-
                       RIABLES INDEPENDIENTES PARA ESTE MODELO POR
                       LA MISMA MATRIZ.
             MWTY    : VECTOR QUE RESULTA DE LA MULTIPLICACION DE
                       LA TRANSPUESTA DE LA MATRIZ DE VARIABLES
                       INDEPENDIENTES PARA ESTE MODELO POR EL VEC-
```

```
                    TOR DE LA VARIABLE DEPENDIENTE.
            MBESTW  : VECTOR DE BETAS ESTIMADAS PARA ESTE MODELO

 EXPLICACION : LA MATRIZ MWTW Y EL VECTOR MWTY SE OBTIENEN A PARTIR
               DE LA MATRIZ MXTX SELECCIONANDO DE ESTA ULTIMA LOS
               RENGLONES Y/O COLUMNAS QUE SE RELACIONAN CON LOS IN-
               DICES DE LAS P-1 VARIABLES GUARDADAS EN INW (LAS CA-
               RACTERISTICAS DE MXTX Y MXTY Y DE MWTW Y MWTY PERMI-
               TEN CALCULARLOS ASI). UNA VEZ OBTENIDAS MWTW Y MWTY
               SE OBTIENE LA SOLUCION DEL SISTEMA  MWTW * MBESTW =
               MWTY   QUE NOS SIRVE PARA CALCULAR SCREGR, SCTR,
               SCRESR, MSCREGR, SR, SR2 Y FW.
                                                                    }

PROCEDURE REGRESION_X;
VAR   KD,VARIN,KK                                      :INTEGER;
      INW                                              :MATRIZT2IC;
      CAMBIOY                                          :BOOLEAN;
      YM                                               :MATRIZT2;
      MWTY,MBESTW                                      :MATRIZT2C;
      MWTW                                             :MATRIZT1C;
      SCRESR,SCREGR,SCTR,SR,SR2,MSCREGR,FR,RR,RAR      :REAL;
BEGIN
   REPEAT
     GOTOXY(01,03);
     WRITE('CUAL ES LA VARIABLE EXPLICATIVA QUE SERA LA',^J,^M,'VARIABLE DEPENDIENTE? : ');
     VARIN:=LEER_INT;
   UNTIL VARIN IN [1..P];
   KK:=0;
    FOR KD:=1 TO P DO
    IF KD<>VARIN THEN
     BEGIN
      KK:=KK+1;
      INW[KK]:=KD;
     END;
   FOR KD:=1 TO N DO
     YM[KD]:=X[KD,VARIN];
   CAMBIOY:=TRUE;
   KK:=P-1;
   KD:=P-VT1;
   SOLUCIONAR:=TRUE;
   INVERTIR:=FALSE;

{*****   LA MATRIZ MWTW Y EL VECTOR MWTY SE OBTIENEN A MEDIANTE EL       *****
 *****   PROCEDIMIENTO "CAMBIO_MATRIZ". LA SOLUCION DEL SISTEMA PARA     *****
 *****   EL MODELO CON EL SUBCONJUNTO DE VARIABLES SE OBTIENE MEDIANTE   *****
 *****   EL PROCEDIMIENTO "INVERSA".                                     *****}

   CAMBIO_MATRIZ(MWTW,MWTY,INW,VARIN,KK,CAMBIOY);
   INVERSA(MWTW,INVERSION,SOLUCIONAR,INVERTIR,KD,MBESTW,MWTY);
   IF INVERSION THEN
    BEGIN

{*****   LOS CALCULOS DE LAS SUMAS DE CUADRADOS Y DEL ESTADISTICO F SE   *****
```

```
*****   REALIZAN POR MEDIO DEL PROCEDIMIENTO "TABLA_ANOVA".           *****)

     TABLA_ANOVA(SR2,SR,SCREGR,SCRESR,SCTR,MSCREGR,FR,RR,RAR,P,N,KK,MBESTN,MMTY,YM);
     IF IMPRESION THEN BEGIN WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
       WRITELN(LST,' ':18,'R E G R E S I O N   E N T R E   L A S   X'S',^J,^M);END;
     CLRSCR; GOTOXY(01,02);
     WRITELN(' ':18,'R E G R E S I O N   E N T R E   L A S   X'S');

(*****   EL ESTADISTICO Y LA INFORMACION CONCERNIENTE A ESTE ES PRE-    *****
 *****   SENTADA EN UNA TABLA DE ANALISIS DE VARIANZA MEDIANTE EL PRO-  *****
 *****   CEDIMIENTO "ESCRIBE".                                          *****)

    ESCRIBE(1,P-1,N-P,N-1,1,SCREGR,MSCREGR,FR,SCRESR,SR2,SCTR,S2);
   END
  ELSE
     WRITELN('NO SE ENCONTRO LA INVERSA');
END;

BEGIN              ( PROGRAMA PRINCIPAL DE MULTICOLINEALIDAD )
REPEAT
 CLRSCR;
 GOTOXY(01,04);
 WRITELN(' ':10,'╔══════════════════════════════════════════════╗');
 WRITELN(' ':10,'║              MENU MULTICOLINEALIDAD          ║');
 WRITELN(' ':10,'║                                              ║');
 WRITELN(' ':10,'║ A '^P' MATRIZ DE CORRELACION                  ║');
 WRITELN(' ':10,'║ B '^P' REGRESION DE UNA VARIABLE EXPLICATIVA  ║');
 WRITELN(' ':10,'║    CON LAS RESTANTES                         ║');
 WRITELN(' ':10,'║ C '^P' RESULTADOS POR IMPRESORA               ║');
 WRITELN(' ':10,'║                                              ║');
 WRITELN(' ':10,'║ S '^P' SALIDA                                 ║');
 WRITELN(' ':10,'╚══════════════════════════════════════════════╝');
 IMPRESORA(IMPRESION,10,1);
 REPEAT
   GOTOXY(01,15);    WRITE(' ':10,'> OPCION = ');
   OP:=READKEY;  WRITE(OP);    OP:=UPCASE(OP);
 UNTIL OP IN ['A','B','C','S'];
 CASE OP OF

(*****  LA MATRIZ DE CORRELACION ES CALCULADA MEDIANTE EL PROCEDIMIEN- *****
 *****  TO "MATRIZ_CORRELACION".                                      *****
 *****  CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE   *****
 *****  LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA       *****)

      'A':BEGIN
           CLRSCR;
           MATRIZ_CORRELACION;
           CONTINUAR
          END;

(*****  LA REGRESION DE UNA VARIABLE EXPLICATIVA CON LAS RESTANTES ES  *****
 *****  REALIZADA MEDIANTE EL PROCEDIMIENTO "REGRESION_X".             *****)

      'B':BEGIN
```

```
          CLRSCR;
          REGRESION_I;
         END;

{*****  LA SALIDA O NO DE LOS RESULTADOS POR IMPRESORA ES CONTROLADA  *****
 *****  POR EL PROCEDIMIENTO "IMPRESION"                             *****}

      'C':IMPRESORA(IMPRESION,10,2);
 END;
UNTIL OP IN ['S'];
END;
END.

{
    LA UNIDAD "RESIDUAL" CONTIENE PROCESOS QUE DAN INFORMACION EN BASE A
  LOS RESIDUALES Y QUE AYUDA A DETECTAR VIOLACIONES EN LOS SUPUESTOS DEL
  MODELO DE REGRESION LINEAL MULTIPLE.
  TIENE COMO OPCIONES:
       A) GRAFICACION DE RESIDUALES EN PAPEL DE PROBABILIDAD NORMAL
       B) GRAFICACION DE RESIDUALES ESTUDENTIZADOS vs. Y ESTIMADA
       C) GRAFICACION DE RESIDUALES ESTUDENTIZADOS vs. TIEMPO DE ORDEN
       D) GRAFICACION DE RESIDUALES ESTUDENTIZADOS vs. CADA REGRESOR
       E) GRAFICACION ENTRE REGRESORES
       F) GRAFICACION PARCIAL DE RESIDUALES
       G) OBSERVACIONES INFLUYENTES (METODO DE COOK)
       H) VECTOR DE RESIDUALES
       I) VECTOR DE RESIDUALES ESTUDENTIZADOS
       J) VECTOR DE Y ESTIMADA

  VARIABLES : E    : VECTOR DE RESIDUALES.
              YEST : VECTOR DONDE SE GUARDAN LAS Yi ESTIMADAS.
              D    : VECTOR DE RESIDUALES ESTUDENTIZADOS.
              VP   : VECTOR QUE  GUARDA LA DIAGONAL DE LA MATRIZ X(X'X)"X'
              II   : VECTOR QUE GUARDA EL TIEMPO DE ORDEN.
              CONF : VALOR DE LA BANDA DE CONFIANZA.
              BANDA: VARIABLE LOGICA PARA QUE LAS GRAFICAS TENGAN O NO
                     BANDA DE CONFIANZA.
}

UNIT RESIDUAL;
INTERFACE
  USES CRT,PRINTER,UGLOBAL,UUTIL;
  PROCEDURE RESIDUALES;
IMPLEMENTATION

PROCEDURE RESIDUALES;
VAR    I,J,K              : INTEGER;
       YN,OP1             : CHAR;
       VP,E,YEST,D,II     : MATRIZIZ2;
       NT,CONF,CONF1      : REAL;
       BANDA              : BOLEAN;

    {
      EL PROCEDIMIENTO "GRAF_RES_EST" GRAFICA LOS RESIDUALES ESTUDENTI-
```

```
  ZADOS VERSUS ALGUN VECTOR XX. SE TIENE OPCION EN ESTE PROCEDIMIEN-
  TO DE OBTENER LA GRAFICA CON BANDA DE CONFIANZA.

  VARIABLES : XX   : VECTOR CONTRA EL CUAL SE GRAFICAN LOS RESIDUALES
                     ESTUDENTIZADOS.
              MAXX : VALOR MAXIMO DEL VECTOR XX.
              MINX : VALOR MINIMO DEL VECTOR XX.
              CONF : CONFIABILIDAD.
              DX   : ORDENADA.
              DY   : ABSCISA.
              ESCX : VALOR QUE SE CALCULA PARA ESTANDARIZAR LOS VALO-
                     RES DE XX EN UNA ESCALA EN LA QUE SEA POSIBLE SU
                     GRAFICACION EN LA PANTALLA.
}

PROCEDURE GRAF_RES_EST(VAR XX:MATRIZT2; N:INTEGER);
VAR      I,J,K,DX,DY                    : INTEGER;
         OP                             : CHAR;
         MAXX,MINX,CONF,ESCX,NR         : REAL;
BEGIN
  MAXX:=XX[1];  MINX:=XX[1];
  FOR J:=2 TO N DO

    IF XX[J] > MAXX THEN MAXX:=XX[J]
                ELSE IF XX[J] < MINX THEN MINX:=XX[J];
  CLRSCR;
  FOR I:= 7 TO 78 DO
    BEGIN GOTOXY(I,3); WRITE('-');
          GOTOXY(I,21); IF I MOD 10 = 7 THEN   WRITE('+')
                                     ELSE   WRITE('-');
    END;
  FOR J:= 4 TO 20 DO
    BEGIN
      GOTOXY(7,J);
      WRITE('I');
      GOTOXY(78,J);
      WRITE('I');
    END;
  IF BANDA THEN
    FOR I:=8 TO 77 DO
      BEGIN
        GOTOXY(I,12-ROUND(CONF*2));  WRITE('-');
        GOTOXY(I,12+ROUND(CONF*2));  WRITE('-');
      END;
  ESCX:=(MAXX-MINX)/69.0;
  FOR J:=1 TO N DO
    BEGIN
      DY:=12-ROUND(D[J]*2);
      DX:=ROUND((XX[J]-MINX)/ESCX)+8;
      GOTOXY(DX,DY);
      WRITE('*');
    END;
  NR:=4.0;
  FOR I:=2 TO 10 DO
```

```
    BEGIN
      GOTOXY(1,I*2);
      WRITE(NR:5:1,' +');
      NR:=NR-1;
    END;
  GOTOXY(2,22); WRITE(MINX:9:2);
  GOTOXY(8,21); WRITE('+');
  K:=7;
  REPEAT
    K:=K+10;
    GOTOXY(K-6,22);
    NR:=((K-8)*ESCX)+MINX;
    WRITE(NR:9:2);
  UNTIL K=67;
  GOTOXY(72,22); WRITE(MAXX:9:2);
END;

     {================================================================
     |   EL PROCEDIMIENTO 'OBS_INFLUYENTE' CALCULA EL VECTOR D DE COOK  |
     |  QUE NOS AYUDA A ENCONTRAR OBSERVACIONES INFLUYENTES.           |
     |                                                                 |
     |  VARIABLES : PD : VALOR DE UN COMPONENTE DEL VECTOR D DE COOK.  |
     ================================================================}

PROCEDURE OBS_INFLUYENTE;
VAR     PD    :REAL;
        YN    :CHAR;
        I     :INTEGER;
BEGIN
  CLRSCR;
  WRITELN('       RESIDUAL       R. ESTUDENTIZADO        Di-CALCULADO');
    IF IMPRESION THEN
    BEGIN
      WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
      WRITELN(LST,'                    OBSERVACIONES INFLUYENTES'^J,^M);
      WRITELN(LST,'       RESIDUAL       R. ESTUDENTIZADO        Di-CALCULADO');
    END;
  FOR I:=1 TO N DO
    BEGIN
      IF I MOD 21 = 0 THEN
      BEGIN

{*****   CON EL PROCEDIMIENTO 'CONTINUAR' LA INFORMACION QUE CONTIENE    *****
 *****   LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA         *****}

        CONTINUAR;
        CLRSCR;
        WRITELN('       RESIDUAL       R. ESTUDENTIZADO        Di-CALCULADO');
        WRITELN;
      END;

{*****                   OBTENCION DEL VECTOR DE COOK                    *****}

      PD:=SQR(d[I])*VP[I]/( (1-VP[I])*P);
```

```
    WRITELN(E[I]:15:5,' ':6,D[I]:17:5,' ':5,PD:15:5);
    IF IMPRESION THEN WRITELN(LST,E[I]:15:5,' ':6,D[I]:17:5,' ':5,PD:15:5);
   END;

(*****   CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE    *****
 *****   LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA        *****)

  CONTINUAR;
END;

    {
     |   EL PROCEDIMIENTO "PAP_NOR" GRAFICA LOS RESIDUALES ESTUDENTIZADOS |
     | EN UNA ESCALA QUE SIMULA LA PROBABILIDAD NORMAL.                   |
     |                                                                    |
     | VARIABLES : RS : RESIDUALES ESTUDENTIZADOS.                        |
     |             VI : VECTOR CONTRA EL CUAL SE GRAFICAN LOS RESIDUALES  |
     |                  ESTUDENTIZADOS Y QUE GUARDA LA ESCALA DE LA PROBA-|
     |                  BILIDAD NORMAL.                                   |
     |             O  : MATRIZ QUE CONSERVA TANTO EL ORIGINAL ORDEN DE VI |
     |                  COMO SU ORDEN ASCENDENTE.                         |
                                                                          }

PROCEDURE PAP_NOR(VAR RS:MATRIZT2);
VAR VI  : MATRIZT2;
    O   : MATRIZT2I;
    I   : INTEGER;
 BEGIN
   FOR I:=1 TO N DO
      VI[I]:=(I-0.5)/N;

(*****   EL ORDENAMIENTO DE EL VECTOR VI EN FORMA ASCENDENTE, ASI COMO  *****
 *****   LA ASIGNACION DE EL ORDEN ORIGINAL Y LA DEL NUEVO ORDEN A EL    *****
 *****   VECTOR O ES REALIZADO POR EL PROCEDIMIENTO "QUICK".             *****)

   QUICK(RS,O,N);

(*****   LA GRAFICA DE ESTOS DOS VECTORES ES REALIZADA POR EL PROCEDI-   *****
 *****   MIENTO "GRAFICA".                                               *****)

   GRAFICA(RS,VI,N);
   GOTOXY(1,1);
   WRITE('GRAFICA EN PAPEL NORMAL ');
   GOTOXY(10,25);

(*****   CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE    *****
 *****   LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA        *****)

   CONTINUAR;
END;

    {
     |   EL PROCEDIMIENTO "D_VS_X" GRAFICA LOS RESIDUALES ESTUDENTIZADOS  |
     | CONTRA ALGUNA DE LAS VARIABLES EXPLICATIVAS.                       |
     |                                                                    |
```

```
  VARIABLES : NUM : INDICE DE LA VARIABLE EXPLICATIVA CONTRA LA CUAL
                    SE VA A GRAFICAR.
              XT  : VECTOR AL CUAL SE LE ASIGNA LA COLUMNA DE LA MA-
                    TRIZ X QUE CONTIENE LA VARIABLE EXPLICATIVA A
                    GRAFICAR.
                                                                    }

PROCEDURE D_VS_X;
    VAR   NUM,I,J : INTEGER;
          YN      : CHAR;
          XT      :MATRIZT2;
BEGIN
  REPEAT
    CLRSCR;
    GOTOXY(1,5);
    WRITELN('> CONTRA CUAL VARIABLE EXPLICATIVA DESEAS GRAFICAR?  ');WRITELN;
    FOR I:= 1 TO P DO
        WRITELN(' X(',I:1,') --> ',I:1); WRITELN;
    WRITE('> OPCION : ');
    REPEAT
      NUM:=LEER_INT;
    UNTIL (0<NUM) AND (NUM<=P);
    FOR J:=1 TO N DO
         XT[J]:=X[J,NUM];

{*****   LA GRAFICA DE LOS RESIDUALES ESTUDENTIZADOS CONTRA ALGUNA VA-  *****
 *****   RIABLE EXPLICATIVA ES HECHA POR EL PROCEDIMIENTO "GRAF_RES_EST" ****}

    GRAF_RES_EST(XT,N);
    GOTOXY(1,1);
    WRITELN('GRAFICA DE:  X(',num:1,') vs. RESIDUALES ESTUDENTIZADOS ');
    IF BANDA THEN WRITELN('CON BANDA DE CONFIANZA DEL  ',CONF1*100:5:2,'%');
    GOTOXY(10,25);

{*****   CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE    *****
 *****   LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA        *****}

    CONTINUAR;
    CLRSCR;
    GOTOXY(5,5);
    WRITE('QUIERES OTRA GRAFICA (S/N) ');
    YN:=READKEY;
    WRITELN(YN);
  UNTIL YN IN ['n','N'];
END;

      {
        EL PROCEDIMIENTO "G_ENTRE XS" EJECUTA LA GRAFICA ENTRA DOS VARIA-
      BLES EXPLICATIVAS.

      VARIABLES : NUM  : INDICE DE LA PRIMER VARIABLE EXPLICATIVA A GRA-
                         FICAR.
                  NUM1 : INDICE DE LA SEGUNDA VARIABLE EXPLICATIVA A GRA-
                         FICAR.
```

```
        XT   : VECTOR AL CUAL SE LE ASIGNA LA COLUMNA DE LA MA-
               TRIZ X QUE CONTIENE LA PRIMER VARIABLE EXPLICA-
               TIVA A GRAFICAR.
        XT1  : VECTOR AL CUAL SE LE ASIGNA LA COLUMNA DE LA MA-
               TRIZ X QUE CONTIENE LA SEGUNDA VARIABLE EXPLICA-
               TIVA A GRAFICAR.
 }

PROCEDURE G_ENTRE_XS;
VAR  NUM,NUM1,J          :  INTEGER;
     YN                  :  CHAR;
     XT1,XT              :  MATRIZT2;
BEGIN
  REPEAT
    CLRSCR;
    GOTOXY(1,5);
    WRITELN('DAME LAS DOS VAR. EXPLICATIVAS A GRAFICAR?  ');WRITELN;
    FOR I:= 1 TO P DO
        WRITELN('(RES.EST.) vs. X(',I:1,') . . . .(',I:1,') '); WRITELN;
    REPEAT
      GOTOXY(1,17);
      WRITE('DAME LA PRIMERA = ');
      NUM:=LEER_INT;
    UNTIL (0<NUM) and (NUM<=P);
    REPEAT
      GOTOXY(1,19);
      WRITE('DAME LA SEGUNDA = ');
      NUM1:=LEER_INT;
    UNTIL (0<NUM1) and (NUM1<=P) AND (NUM<>NUM1);
    FOR J:=1 TO N DO
      BEGIN
         XT[J]:=X[J,NUM];
         XT1[J]:=X[J,NUM1];
      END;

(*****   LA GRAFICA DE ESTOS DOS VECTORES ES REALIZADA POR EL PROCEDI-  *****
 *****   MIENTO "GRAFICA".                                              *****)

    GRAFICA(XT,XT1,N);
    GOTOXY(1,1);
    WRITE('GRAFICA DE X(',NUM:1,') vs  X(',NUM1:1,')');
    GOTOXY(10,25);

(*****   CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE   *****
 *****   LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA       *****)

    CONTINUAR;
    CLRSCR;
    GOTOXY(5,5);
    WRITE('QUIERES OTRA GRAFICA (S/N) ');
    YN:=READKEY;  WRITELN(YN);
  UNTIL YN IN ['n','N']; YN:='s';
END;
```

```
{ EL PROCEDIMIENTO "PAR_RES" GRAFICA LOS RESIDUALES PARCIALES CON-
  TRA ALGUNA DE LAS VARIABLES EXPLICATIVAS.

  VARIABLES : NUM    : INDICE DE LA VARIABLE EXPLICATIVA CONTRA LA
                       CUAL SE VA A GRAFICAR.
              XT     : VECTOR AL CUAL SE LE ASIGNA LA COLUMNA DE LA
                       MATRIX X QUE CONTIENE LA VARIABLE EXPLICATIVA
                       A GRAFICAR.
              EPRIMA : VECTOR DE RESIDUALES PARCIALES.                }

PROCEDURE PAR_RES;
VAR     NUM,I        : INTEGER;
        EPRIMA,XT    : MATRIZT2;
        YN           : CHAR;
BEGIN
  REPEAT
    CLRSCR;
    GOTOXY(1,5);
    WRITELN('CON CUAL DE LAS SIGUIENTES VAR. EXPLICATICAS QUIERES GRAFICAR?  ');WRITELN;
    FOR I:= 1 TO P DO
        WRITELN('(RES.EST.) vs. X(',I:1,') . . . .(',I:1,') '); WRITELN;
    REPEAT  J:=LEER_INT;
    UNTIL (0<J) AND (J<=P);
    FOR I:=1 TO N DO
       BEGIN
         EPRIMA[I]:=E[I]+MBEST[J+1]*X[I,J];
         XT[I]:=X[I,J];
       END;

{*****  LA GRAFICA DE ESTOS DOS VECTORES ES REALIZADA POR EL PROCEDI-  *****
 *****  MIENTO "GRAFICA".                                              *****}

    GRAFICA(XT,EPRIMA,N);
    GOTOXY(1,1);
    WRITE('GRAFICA PARCIAL DE RESIDUALES:  X(',j:1,')  vs.  e''');
    GOTOXY(10,25);

{*****  CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE   *****
 *****  LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA       *****}

    CONTINUAR;
    CLRSCR;
    GOTOXY(5,5);
    WRITE('QUIERES OTRA GRAFICA (S/N) ');
    YN:= READKEY;   WRITELN(YN);
  UNTIL YN IN ['n','N']; YN:='s';
END;

{ EL PROCEDIMIENTO "IMPRIM" MANDA A LA PANTALLA O LA IMPRESORA
  CUALQUIERA DE LOS SIGUIENTES VECTORES: VECTOR DE RESIDUALES, VEC-
  TOR DE RESIDUALES ESTUDENTIZADOS, VECTOR DE Y ESTIMADA.
```

VARIABLES : A : VECTOR QUE SE MANDA A LA PANTALLA O A LA IMPRESORA Y QUE CONTIENE LOS RESIDUALES, LOS RESIDUALES ESTUDENTIZADOS O LA Y ESTIMADA.

```
PROCEDURE IMPRIM(VAR A:MATRIZT2);
VAR I,J:INTEGER;
BEGIN
     CLRSCR;
     CASE YN OF
        'H'  :  WRITELN(' ':20,' VECTOR DE RESIDUALES ');
        'I'  :  writeln(' ':10,' VECTOR DE RESIDUALES ESTUDENTIZADOS');
        'J'  :  writeln(' ':20,' VECTOR DE Y ESTIMADA');
     END;
     WRITELN('                    NUMERO','       VALOR ');WRITELN;
     FOR I:=1 TO N DO
       BEGIN
         IF I MOD 21 = 0 THEN
          BEGIN
            WRITELN;CONTINUAR; CLRSCR;
            CASE YN OF
              'H'  :  WRITELN(' ':20,' VECTOR DE RESIDUALES ');
              'I'  :  writeln(' ':10,' VECTOR DE RESIDUALES ESTUDENTIZADOS');
              'J'  :  writeln(' ':20,' VECTOR DE Y ESTIMADA');
            END;
            WRITELN('                    NUMERO','       VALOR ');WRITELN;
          END;
         WRITE('                    ',I:3);
         WRITE('    ',A[I]:12:4);
         WRITELN;
       END;
     CONTINUAR;CLRSCR;
     IF IMPRESION THEN
       BEGIN
         WRITELN(LST);WRITELN(LST);WRITELN(LST);
         CASE YN OF
              'h','H'  :writeln(LST,' ':20,' VECTOR DE RESIDUALES ');
             'i','I'  :writeln(LST,' ':10,' VECTOR DE RESIDUALES ESTUDENTIZADOS');
             'j','J'  :writeln(LST,' ':20,' VECTOR DE Y ESTIMADA');
         END;
         WRITELN(LST);
         WRITELN(LST,'                    NUMERO','      VALOR');WRITELN(LST);
         FOR I:=1 TO N DO
           BEGIN
             WRITE(LST,'                    ',I:3);
             WRITE(LST,'    ',A[I]:12:4);
             WRITELN(LST);
           END;
       END;
END;
```

```
{*****              PROGRAMA PRINCIPAL DE RESIDUALES                  *****}

BEGIN

{*****  LOS RESIDUALES Y EL VECTOR DE LAS Yi ESTIMADAS SE CALCULAN POR *****
 *****  MEDIO DEL PROCEDIMIENTO 'ESTIMADA'.                           *****}

  ESTIMADA(E,YEST);

{*****   MEDIANTE EL PROCEDIMIENTO 'D_XXXX' SE OBTIENE LA DIAGONAL DE *****
 *****   LA MATRIZ X(X'X)^X' Y SE GUARDA EN VP.                       *****}

  D_XXXX(VP);
  FOR I:=1 TO N DO
        D[I]:=E[I]/ ( S*SQRT(1-VP[I]) );
  CONF  := N_I(0.025);
  CONF1 := 0.95;
  BANDA := TRUE;
  REPEAT
     CLRSCR;
     WRITELN(' ':10,'┌──────────────────────────────────────────────────┐');
     WRITELN(' ':10,'│              MENU ANALISIS DE RESIDUALES          │');
     WRITELN(' ':10,'│                                                  │');
     WRITELN(' ':10,'│ A '^P' GRAFICACION DE RESIDUALES EN PAPEL DE       │');
     WRITELN(' ':10,'│     PROBABILIDAD NORMAL                           │');
     WRITELN(' ':10,'│ B '^P' GRAFICACION DE RESIDUALES ESTUDENTIZADOS    │');
     WRITELN(' ':10,'│     vs. Y ESTIMADA                                │');
     WRITELN(' ':10,'│ C '^P' GRAFICACION DE RESIDUALES ESTUDENTIZADOS    │');
     WRITELN(' ':10,'│     vs. TIEMPO DE ORDEN                           │');
     WRITELN(' ':10,'│ D '^P' GRAFICACION DE RESIDUALES ESTUDENTIZADOS    │');
     WRITELN(' ':10,'│     vs. CADA REGRESOR                             │');
     WRITELN(' ':10,'│ E '^P' GRAFICACION ENTRE REGRESORES                │');
     WRITELN(' ':10,'│ F '^P' GRAFICACION PARCIAL DE RESIDUALES           │');
     WRITELN(' ':10,'│ G '^P' OBSERVACIONES INFLUYENTES (METODO DE COOK)  │');
     WRITELN(' ':10,'│ H '^P' VECTOR DE RESIDUALES                        │');
     WRITELN(' ':10,'│ I '^P' VECTOR DE RESIDUALES ESTUDENTIZADOS         │');
     WRITELN(' ':10,'│ J '^P' VECTOR DE Y ESTIMADA                        │');
     WRITELN(' ':10,'│ K '^P' RESULTADOS POR IMPRESORA                    │');
     WRITELN(' ':10,'│                                                  │');
     WRITELN(' ':10,'│ O '^P' OPCIONES                                    │');
     WRITELN(' ':10,'│ S '^P' SALIDA                                      │');
     WRITELN(' ':10,'└──────────────────────────────────────────────────┘');
    IMPRESORA(IMPRESION,18,1);
    REPEAT
      GOTOXY(01,23);
      WRITE(' ':10,'> OPCION = ');
      YN:=READKEY;WRITE(YN);
      YN:=UPCASE(YN);
    UNTIL YN IN ['A','B','C','D','E','F','G','H','I','J','K','O','S'];
    CASE YN OF

{*****   LA GRAFICA DE LOS RESIDUALES ESTUDENTIZADOS EN PAPEL DE PROBA- *****
 *****   BILIDAD NORMAL ES REALIZADO POR EL PROCEDIMIENTO 'PAP_NOR'.    *****}
```

```
        'A' :  PAP_NGR(D);
        'B' :  BEGIN

{*****  LA GRAFICA DE LOS RESIDUALES ESTUDENTIZADOS VERSUS EL VECTOR    *****
 *****  DE LAS Yi ESTIMADAS ES REALIZADO POR EL PROCEDIMIENTO            *****
 *****  "GRAF_RES_EST".                                                  *****}

                GRAF_RES_EST(YEST,N);
                GOTOXY(1,1);
                WRITELN('GRAFICA DE: Y ESTIMADA  vs.  RESIDUALES ESTUDENTIZADOS');
                IF BANDA THEN WRITELN('CON BANDA DE CONFIANZA DEL  ',CONF1*100:5:2,'%');
                GOTOXY(10,25);

{*****  CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE    *****
 *****  LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA         *****}

                CONTINUAR;
             END;
        'C' :  BEGIN FOR I:=1 TO N DO
                 II[I]:=I;

{*****  LA GRAFICA DE LOS RESIDUALES ESTUDENTIZADOS VERSUS EL TIEMPO    *****
 *****  DE ORDEN ES REALIZADO POR EL PROCEDIMIENTO "GRAF_RES_EST".       ****}

               GRAF_RES_EST(II,N);
               GOTOXY(1,1);
               WRITELN('GRAFICA DE : TIEMPO DE ORDEN  vs.  RESIDUALES ESTUDENTIZADOS');
               IF BANDA THEN WRITELN('CON BANDA DE CONFIANZA DEL  ',CONF1*100:5:2,'%');
               GOTOXY(10,25);

{*****  CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE    *****
 *****  LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA         *****}

               CONTINUAR;
            END;

{*****  LA GRAFICA DE LOS RESIDUALES ESTUDENTIZADOS VERSUS ALGUNA VAR- *****
 *****  RIABLE EXPLICATIVA ES REALIZADO POR EL PROCEDIMIENTO "D_VS_X". *****}

        'D'  :   D_VS_X(D);
        'E'  :  IF P>1 THEN

{*****  LA GRAFICA ENTRE REGRESORES ES REALIZADA POR EL PROCEDIMIENTO  *****
 *****  "G_ENTRE_XS".                                                   *****}

                     G_ENTRE_XS
                  ELSE
                     WRITELN(^G,'SOLO TIENES UNA VAR. EXPLICATIVA');

{*****  LA GRAFICA PARCIAL DE RESIDUALES SE REALIZA MEDIANTE EL PROCE- *****
 *****  DIMIENTO "PAR_RES".                                             *****}

        'F'  :   PAR_RES(E);
```

```
{***** EL CALCULO DEL VECTOR D DE COOK ES HECHO POR EL PROCEDIMIENTO *****
 ***** 'OBS_INFLUYENTE'.                                             *****}

          'G'  :   OBS_INFLUYENTE;

{***** EL DESPLIEGUE DE EL VECTOR DE RESIDUALES POR PANTALLA Y/O POR *****
 ***** IMPRESORA SE REALIZA CON DEL PROCEDIMIENTO 'IMPRIM'.          *****}

          'H'  :   IMPRIM(E);

{***** EL DESPLIEGUE DE EL VECTOR DE RESIDUALES ESTUDENTIZADOS POR   *****
 ***** PANTALLA Y/O POR IMPRESORA SE REALIZA POR MEDIO DEL PROCE-    *****)
 ***** DIMIENTO 'IMPRIM'.                                            *****}

          'I'  :   IMPRIM(D);

{***** EL DESPLIEGUE DE EL VECTOR DE Y ESTIMADA POR PANTALLA Y/O POR *****
 ***** IMPRESORA SE REALIZA A TRAVEZ DEL PROCEDIMIENTO 'IMPRIM'.     *****}

          'J'  :   IMPRIM(YEST);

{***** LA SALIDA O NO DE LOS RESULTADOS POR IMPRESORA ES CONTROLADA  *****
 ***** POR MEDIO DEL PROCEDIMIENTO 'IMPRESORA'.                      *****}

          'K'  :   IMPRESORA(IMPRESION,18,2);

 ***** EN O ''P' OPCIONES SE DA LA OPORTUNIDAD DE USAR BANDAS DE CON- *****}
 ***** FIANZA EN LAS DIFERENTES GRAFICAS DE ESTE MENU, CON LA OPOR-   *****}
 ***** TUNIDAD DE SELECCIONAR TAMBIEN LA CONFIABILIDAD.              *****}

          'O'  :  BEGIN
                    CLRSCR;GOTOXY(01,05);
                    WRITE('> DESEAS LAS BANDAS DE CONFIANZA ? (S/N) :');
                    OP1 :=READKEY   ;WRITELN(OP1);    WRITELN;
                    IF OP1 IN ['S'] THEN
                      BEGIN
                        REPEAT
                          WRITE('> DAME LA CONFIABILIDAD (Formato 0.###) = ');
                          CONF:=LEER_REAL;
                        UNTIL (0<CONF) AND (CONF<1);
                        CONF1 := CONF;
                        CONF:=N_I((1-CONF)/2);
                        BANDA := TRUE;
                      END
                    ELSE  BANDA := FALSE;
                  END;
    END;
   UNTIL YN in ['S'];
   YN:='N';
END;
END.

{╔═══════════════════════════════════════════════════════════════════╗
 ║   LA UNIDAD 'PREDIC' CONTIENE LOS PROCESOS PARA LA ESTIMACION DE LOS   ║
```

```
 VALORES A PREDECIR Y PARA EL CALCULO DE LOS INTERVALOS DE CONFIANZA DE
 ESTOS VALORES.
 TIENE COMO OPCIONES:
       A) PREDICCION E INTERVALO DE CONFIANZA PARA E[Y/Xo].
       B) PREDICCION E INTERVALO DE CONFIANZA PARA Y/Xo.
 VARIABLES : XO : VECTOR A PREDECIR.
             SB : VARIANZA ESTIMADA DE E[Y/Xo].
             T  : CONFIABILIDAD.
             M  : VALOR ESTIMADO DE E[Y/Xo] .
             TC : T CALCULADA.
}

UNIT PREDIC;
INTERFACE
  USES CRT,PRINTER,UGLOBAL,UUTIL;
  PROCEDURE PREDICCION;

IMPLEMENTATION


PROCEDURE PREDICCION;
VAR   I,J,K          : INTEGER;
      SB,M,T,TC      : REAL;
      XO             : MATRIZT2C;
      YN             : CHAR;
BEGIN
   REPEAT
     CLRSCR;
     GOTOXY(01,04);
     WRITELN(' ':10,'                                                ');
     WRITELN(' ':10,'                 MENU PREDICCION                ');
     WRITELN(' ':10,'                                                ');
     WRITELN(' ':10,' A '^P' PREDICCION E INTERVALO DE CONFIANZA       ');
     WRITELN(' ':10,'     PARA  #  (E[Y/Xo])                         ');
     WRITELN(' ':10,' B '^P' PREDICCION E INTERVALO DE CONFIANZA       ');
     WRITELN(' ':10,'     PARA  Y/Xo                                 ');
     WRITELN(' ':10,' C '^P' RESULTADOS POR IMPRESORA                  ');
     WRITELN(' ':10,'                                                ');
     WRITELN(' ':10,' S '^P' SALIDA                                    ');
     WRITELN(' ':10,'                                                ');
     IMPRESORA(IMPRESION,11,1);
     REPEAT
        GOTOXY(01,16);    WRITE(' ':10,'> OPCION = ');
        YN:=READKEY;    WRITELN(YN);  YN:=UPCASE(YN);
     UNTIL YN IN ['A','B','C','S'];

(*****   LA SALIDA O NO DE LOS RESULTADOS POR IMPRESORA ES CONTROLADA   *****
 *****   POR MEDIO DEL PROCEDIMIENTO "IMPRESORA".                       *****)

     IF YN IN ['C'] THEN IMPRESORA(IMPRESION,11,2);
     IF YN IN ['A','B'] THEN
       BEGIN
         CLRSCR;
         GOTOXY(1,5);
```

```
 IF YN IN ['a','A'] THEN WRITELN('PREDICCION E INTERVALO DE CONFIANZA FARA "µ" (E [Y/Xo]
 ELSE WRITELN('PREDICCION E INTERVALO DE CONFIANZA PARA  Y/Xo. ');
 GOTOXY(5,8);
 IF IMPRESION THEN
 BEGIN
    WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
    IF YN IN ['a','A'] THEN WRITELN(LST,'PREDICCION E INTERVALO DE CONFIANZA PARA "µ"  (E
    ELSE WRITELN(LST,'PREDICCION E INTERVALO DE CONFIANZA PAFA  Y/Xo. ');
 END;
  WRITELN('DAME EL VECTOR Xo A PFEDECIR');WRITELN;
  IF VT1=0 THEN
    BEGIN
      M:=MBEST[1];
      XO[1]:=1;
    END
  ELSE M:=0;
  FOR I:=2-VT1 TO H DO
    BEGIN
      READLN(XO[I]);
      M:= XO[I]*MBEST[I] +M;
    END;
  SB:=0;
  FOR I:=1 TO H DO
    BEGIN
      SB:= SB+MXTX[I+1,I]*SQR(XO[I]);
      FOR J:=I+1 TO H DO
         SB:=SB+MXTX[J+1,I]*XO[I]*XO[J]*2;
    END;
  IF YN IN ['A'] THEN   SB:=S*SQRT(SB)
                 ELSE   SB:=S*SQRT(SB+1);
  TC:= M/SB;
  WRITELN;
  REPEAT
     WRITE('      DAME LA CONFIABILIDAD= ');
     T:=LEER_REAL;
  UNTIL (T>0) AND (T<1);
  WRITELN;WRITELN;
  T:=T_INV((1-T)/2,N-H);
  WRITELN('      EL VALOR ESTIMADO DE µ ES  = ',M:10:4);
  WRITELN('      EL INTERVALO DE CONFIANZA ES = (',M-(T*SB):10:4,',',M+(T*SB):10:4,')');
  IF IMPRESION THEN
   BEGIN
     WRITELN(LST,^J,^M);
     WRITELN(LST,'      EL VALOR ESTIMADO DE µ ES  = ',M:10:4);
     WRITELN(LST,'      El INTERVALO DE CONFIANZA ES = (',M-(T*SB):10:4,',',M+(T*SB):10:4,')
   END;
   WRITELN;

{*****  CON EL PROCEDIMIENTO "CONTINUAR" LA INFORMACION QUE CONTIENE   *****
 *****  LA PANTALLA ES RETENIDA HASTA QUE SE OPRIMA ALGUNA TECLA       *****}

     CONTINUAR;
 END;
 UNTIL YN IN ['S','s'];
```

```
END;
END.

{╔══════════════════════════════════════════════════════════════╗
 ║   LA UNIDAD "EDITAR" PERMITE EDITAR LOS DATOS DE LA MATRIZ X Y DEL  ║
 ║ VECTOR Y.                                                          ║
 ╚══════════════════════════════════════════════════════════════╝}

UNIT EDITAR;
INTERFACE
  USES CRT,UGLOBAL,UUTIL;
  PROCEDURE EDITAR_XY(VAR X:MATRIZT1;VAR Y:MATRIZT2;VAR N,P:INTEGER);

IMPLEMENTATION

PROCEDURE EDITAR_XY(VAR X:MATRIZT1;VAR Y:MATRIZT2;VAR N,P:INTEGER);
VAR I,J,K   : INTEGER;
    XAR     : MATRIZT2;
    OP      : CHAR;
BEGIN
  REPEAT
    CLRSCR;
    GOTOXY(01,03);
    WRITELN(' ':10,'╔══════════════════════════════════════════════════╗');
    WRITELN(' ':10,'║                MENU EDICION DE DATOS             ║');
    WRITELN(' ':10,'║                                                  ║');
    WRITELN(' ':10,'║ A '^P' ALTA DE UNA OBSERVACION (renglon)           ║');
    WRITELN(' ':10,'║ B '^P' ELIMINAR UNA OBSERVACION (renglon)          ║');
    WRITELN(' ':10,'║ C '^P' ALTA DE UNA VAR. EXPLICATIVA (columna)      ║');
    WRITELN(' ':10,'║ D '^P' ELIMINAR UNA VAR. EXPLICATIVA (columna)     ║');
    WRITELN(' ':10,'║ E '^P' CORREGIR UNA OBSERVACION (renglon)          ║');
    WRITELN(' ':10,'║                                                  ║');
    WRITELN(' ':10,'║ S '^P' SALIDA                                      ║');
    WRITELN(' ':10,'╚══════════════════════════════════════════════════╝');
    REPEAT
      GOTOXY(11,19);  write('> OPCION : ');
      OP:=readKEY; WRITELN(OP);
      OP := UPCASE(OP);
    UNTIL OP IN ['A','B','C','D','E','S'];
    CLRSCR;
    IF op in ['A','B','D','E'] THEN
       BEGIN
         GOTOXY(01,24);
         WRITE('PARA SALIR INTRODUCIR (0)');
       END;
    GOTOXY(01,03);
    CASE OP OF

{*****                      ALTA DE UNA OBSERVACION.                    *****}

     'A': BEGIN
            REPEAT
              WRITE('EN QUE RENGLON QUIERES INTRODUCIR LA NUEVA OBSERVACION : ');
              K:=LEER_INT;
```

```
          UNTIL (K>0) AND (K<=N+1);
          WRITELN('DAME EL VALOR DE Y SEGUIDO DE LAS VARIABLES',^J,^M,'EXPLICATIVAS EN FORMATO LIB
          WRITELN('ME TIENES QUE DAR UN VECTOR DE ',p+1:1,' COMPONENTES ');
          WRITELN(''NO' INCLUYAS EL "1" DEL TERMINO INDEPENDIENTE ');
          LEER_VEC_REAL(XAR,P+1);
          FOR I:= N DOWNTO K DO
            BEGIN
              Y[I+1]:=Y[I]:
              FOR J:=1 TO P+1 DO
                 X[I+1,J]:=X[I,J];
            END;
          Y[K]:=XAR[1];
          FOR I:=2 TO P+1 DO
              X[K,I-1]:=XAR[I];
          N:=N+1;
        END;

(*****                    BAJA DE UNA OBSERVACION.                    *****)

     'B': BEGIN
           IF N>P+1 THEN
             BEGIN
               REPEAT
                 WRITE('QUE RENGLON QUIERES ELIMINAR : ');
                 K:=LEER_INT;
               UNTIL (K>0) AND (K<=N);
               FOR I:=K TO N-1 DO
                 BEGIN
                   Y[I]:=Y[I+1];
                   FOR J:=1 TO P DO
                     X[I,J]:=X[I+1,J];
                 END;
               N:=N-1;
             END
           ELSE
           WRITELN('NO PUEDES TENER MENOS OBSERVACIONES QUE ESTIMADORES',^G);
        END;

(*****                 ALTA DE UNA VARIABLE EXPLICATIVA.               *****)

     'C': BEGIN
          CLRSCR;
          WRITELN('DAME UN VALOR POR RENGLON : ');
          FOR I:=1 TO N DO
            BEGIN
              WRITE('X[',I,',',P+1,']= ');
              X[I,P+1]:=LEER_REAL;
            END;
          P:=P+1;
         END;

(*****                 BAJA DE UNA VARIABLE EXPLICATIVA.               *****)

     'D': BEGIN
```

```
          IF P>1 THEN
            BEGIN
              REPEAT
                WRITE('QUE COLUMNA QUIERES ELIMINAR : ');
                K:=LEER_INT;
              UNTIL (K>0) AND (K<=P);
              FOR I:=1 TO N DO
                FOR J:=K TO P-1 DO
                  X[I,J]:=X[I,J+1];
              P:=P-1;
            END
          ELSE
            WRITELN('NO PUEDES BORRAR LA UNICA VAR. EXPLICATIVA QUE TIENES',^g);
        END;

(*****              CORRECION DE UNA OBSERVACION.                  *****)

      'E': BEGIN
             REPEAT
               WRITE('DAME EL NUMERO DE RENGLON A CORREGIR : ');
               K:=LEER_INT;
             UNTIL (K>0) AND (K<=N);
             WRITELN('LOS VALORES ACTUALES SON :');
             WRITE(Y[K]:6:3);
             FOR J:=1 TO P DO WRITE(',',X[K,J]:6:3);WRITELN;
             WRITELN('DAME LOS NUEVOS VALORES : ');
             LEER_VEC_REAL(XAR,P+1);
             Y[K]:=XAR[1];
             FOR I:=2 TO P+1 DO
                X[K,I-1]:=XAR[I];
           END;
    END;
     IF OP IN ['A','B','C','D','E'] THEN
       BEGIN
         WRITELN;WRITELN;
         WRITE('> OPERACION REALIZADA    ');
         CONTINUAR;
       END;
  UNTIL OP IN ['S'];
END;
END.
 {
   LA UNIDAD "FURNIVAL" SELECCIONA LA MEJOR SUMA DE CUADRADOS DE LOS RE-
   SIDUALES PARA CADA SUBCONJUNTO DE VARIABLES. EL PROCEDIMIENTO SE BASA
   EN EL METODO DE FURNIVAL.

   VARIABLES : TMI  : MATRIZ CON LA CUAL SE OBTIENEN LAS SCRES DE LOS SUB-
                      CONJUNTOS DE VARIABLES A PARTIR DE LA MATRIZ MATX
                      QUE SE PIVOTEA CONSTANTEMENTE.
               TXM  : MATRIZ CON LA CUAL SE OBTIENEN LAS SCRES DE LOS SUB-
                      CONJUNTOS DE VARIABLES A PARTIR DE LA INVERSA DE
                      LA MATRIZ MXTX LA CUAL SE PIVOTEA CONSTANTEMENTE.
               IND  : VECTOR QUE GUARDA LOS INDICES DE LAS VARIABLES PIVO-
                      TEADAS.
```

```
ORE  : VECTOR QUE CONTIENE EL ORDEN ORIGINAL DE LAS VARIA-
       BLES Y EL NUEVO QUE SE OBTUVO AL ORDENAR LAS VARIA-
       BLES.
VREG : MATRIZ QUE GUARDA EL NUMERO DE ELEMENTOS DE CADA SUB-
       CONJUNTO Y LOS INDICES DE LAS VARIABLES QUE TUVIERON
       LA MEJOR SUMA DE CUADRADOS DE LOS RESIDUALES ENTRE
       LOS SUBCONJUNTO CON IGUAL NUMERO DE VARIABLES.
RSS  : VECTOR QUE GUARDA LA MEJOR SUMA DE CUADRADOS DE LOS
       RESIDUALES PARA CADA SUBCONJUNTO
}

UNIT FURNIVAL;
INTERFACE
  USES CRT,PRINTER,UGLOBAL,UUTIL;
  PROCEDURE FURNIVAL;

IMPLEMENTATION

PROCEDURE FURNIVAL;
TYPE  MATRIZT11=ARRAY[1..NMAXCOL,1..NMAXCOL]OF INTEGER;
      TRI=ARRAY[1..122,1..11]OF REAL;
VAR   TMI,TIM                    :TRI;
      IND                        :MATRIZT2IC;
      ORE                        :MATRIZT2I;
      VREG                       :MATRIZT1I;
      MN,DIF,NVF,LL,L,M,TER,COL  :INTEGER;
      YM                         :MATRIZT2;
      RSS,YL,YS                  :MATRIZT2C;
      AVANCE                     :BOOLEAN;
      CFS,JVL,CPS1               :REAL;
      XM                         :MATRIZT1C;

    {
      EL PROCEDIMIENTO "SWEEP" PIVOTEA LA MATRIZ A.

      VARIABLES : IP : RENGLON PIVOTE.
    }

PROCEDURE SWEEP(IB,IS,IP:INTEGER;VAR A:TRI;KP:INTEGER);
VAR    LB,LI,MI   :INTEGER;
       E          :REAL;
BEGIN
   IF IB=1 THEN IB:=0;
   IF IS=1 THEN IS:=0;
   LB:=IP+1;
   FOR LI:=LB TO KP DO
    BEGIN
        A[IS+IP,LI]:=A[IB+IP,LI]/A[IB+IP,IP];
        FOR MI:=LI TO KP DO
        A[IS+LI,MI]:=A[IB+LI,MI]-A[IB+IP,MI]*A[IS+IP,LI];
    END;
END;

    {
```

```
  | EL PROCEDIMIENTO "CP_MALLOWS" CALCULA EL ESTADISTICO CP_MALLOWS. |
  |                                                                  |
  | VARIABLES : NV     : NUMERO DE VARIABLES.                        |
  |             CP     : ESTADISTICO CP-MALLOWS.                     |
  |             SCRESR : SUMA DE CUADRADOS DE LOS RESIDUALES.        |
  |__________________________________________________________________|)

PROCEDURE CP_MALLOWS(VAR SCRESR:REAL;VAR NV:INTEGER;VAR CP:REAL);
BEGIN
  CP:=SCRESR/S2-(N-2*(NV+1));
END;

{*****                 PROGRAMA PRINCIPAL DE FURNIVAL.              *****}

BEGIN
 IF P>1 THEN
   BEGIN
     CPS:=0;
     CLRSCR;

{*****                 ORDENAMIENTO DE LAS VARIABLES.               *****}

      IF VT1=1 THEN
        BEGIN
         FOR L:=1 TO P DO
             BEGIN
               FOR M:=L TO P DO
                 BEGIN
                   XM[L+1,M+1]:=MXTX[L,M];
                   XM[M+1,L+1]:=MXTX[L,M];
                 END;
               XM[1,L+1]:=MXTX[VT2,L];
               YL[L+1]:=MXTY[L];
               XM[L+1,1]:=MXTX[VT2,L];
             END;
        XM[1,1]:=N;
        YL[1]:=MXTY[VT2];
        SOLUCIONAR:=TRUE;
        INVERTIR:=TRUE;
        INVERSA(XM,INVERSION,SOLUCIONAR,INVERTIR,H+VT1,YS,YL);
       END;
     FOR L:=2-VT1 TO H DO
       BEGIN
         FOR M:=L TO H DO
           IF VT1=0 THEN
             TMI[L-1+VT1,M-1]:=MXTX[M+1,L]
           ELSE
             TMI[L,M]:=XM[M+2,L+1];
           IF VT1=0 THEN
             TMI[L-1+VT1,H+VT1]:=MBEST[L]
           ELSE
             TMI[L,H+VT1]:=YS[L+1];
       END;
     IF VT1=0 THEN
```

```
      BEGIN
        TMI[H+VT1,H+VT1]:=-SCRES;
        CPS:=SCRES;
      END
    ELSE
      BEGIN
        FOR L:=1 TO H+VT1 DO
          CPS:=CPS+YS[L]*YL[L];
        TMI[H+VT1,H+VT1]:=CPS-SCT;
        CPS:=SCT-CPS;
      END;
    FOR L:=1 TO P DO
      BEGIN
        SWEEP(1,L*14+1,L,TMI,H+VT1);
        YM[L]:=-TMI[L*14+1+H+VT1,H+VT1];
      END;
    QUICK(YM,ORE,P);

(*****                CREACION DE LAS MATRICES TMI Y TXM.              *****)

    FOR L:=1 TO P DO
      BEGIN
        FOR M:=L TO P DO
        IF ORE[H+VT1-M]<=ORE[H+VT1-L] THEN
         BEGIN
          TXM[L,M]:=MXTX[ORE[H+VT1-M]+1-VT1,ORE[H+VT1-L]+1-VT1] -MXTX[VT2,
                    ORE[H+VT1-M]+1-VT1]*MXTX[VT2,ORE[H+VT1-L]+1-VT1]/N;
          IF VT1=0 THEN TMI[L,M]:=MXTX[ORE[H-L]+2,ORE[H-M]+1]
          ELSE TMI[L,M]:=XM[ORE[H+VT1-L]+2,ORE[H+VT1-M]+1];
         END
        ELSE
         BEGIN
          TXM[L,M]:=MXTX[ORE[H+VT1-L]+1-VT1,ORE[H+VT1-M]+1-VT1] -MXTX[VT2,
                    ORE[H+VT1-M]+1-VT1]*MXTX[VT2,ORE[H+VT1-L]+1-VT1]/N;
          IF VT1=0 THEN TMI[L,M]:=MXTX[ORE[H-M]+2,ORE[H-L]+1]
          ELSE TMI[L,M]:=XM[ORE[H+1-M]+2,ORE[H+1-L]+1];
         END;
        IF VT1=0 THEN TMI[L,H+VT1]:=MBEST[ORE[H-L]+1]
        ELSE TMI[L,H+VT1]:=YS[ORE[H+1-L]+1];
        TXM[L,P]:=MXTY[ORE[H+VT1-L]+1-VT1]-MXTX[VT2,ORE[H+VT1-L]+1-VT1]*MXTY[VT2]/N;
      END;
      IF VT1=1 THEN TXM[P,P]:=SCT-(SQR(MXTY[VT2])/N)
      ELSE TXM[P,P]:=SCT;
      IF VT1=0 THEN TMI[H,H]:=-SCRES
      ELSE TMI[H+1,H+VT1]:=-CPS;
    ( WRITELN(LST,'PIVOTE  ','VARIABLE  ','RSS PRODUCTO','VARIABLE  ','RSS INVERSA')
    WRITELN(' RSS ORIGEN');)
    IND[1]:=0;
    M:=1;
    TER:=1;
    AVANCE:=TRUE;
    MN:=H+VT1;

(*****             SELECCION DE LOS SUBCONJUNTOS DE VARIABLES           *****)
```

```
REPEAT
  AVANCE:=TRUE;
  WHILE AVANCE=TRUE DO
    BEGIN
      IF IND[M]<P-1 THEN
        BEGIN
          M:=M+1;
          IND[M]:=IND[M-1]+1;
        END
      ELSE
        BEGIN
          M:=M-1;
          IND[M]:=IND[M]+1;
          TER:=IND[M];
        END;
      IF M<MN  THEN AVANCE:=FALSE;
      IF M=1  THEN AVANCE:=FALSE;
    END;
  IF M<>1 THEN
    BEGIN
      SWEEP(IND[M-1]*14+1,IND[M]*14+1,IND[M],TXM,P);
      SWEEP((TER-1)*14+1,(IND[M]-1)*14+1,IND[M],TMI,H+VT1);
      DIF:=P-1-IND[M];
    (  WRITELN(LST);
      WRITE(LST,IND[M]:2,' ':9);
      FOR L:=2 TO M DO WRITE(LST,' ',IND[M]:2);
      WRITE(LST,TXM[IND[M]*14+1+P,P]:8:4,' ':4);
      FOR L:=2 TO M-1 DO WRITE(LST,' ',IND[L]:2);
      FOR L:=1 TO DIF DO WRITE(LST,' ',IND[M]+L:2);
      WRITE(LST,' ',P:2);
      WRITE(LST,' ':4,-TMI[IND[M]*14+1+H+VT1,H+VT1]:8:4);
      WRITELN(LST);WRITE(LST,-TMI[TER+H+VT1,H+VT1]:8:4);WRITELN(LST);)
      IF IND[M]=0 THEN COL:=0 ELSE COL:=IND[M]*14+1;
      IF TER=1 THEN
        BEGIN
          RSS[M-1]:=TXM[COL+P,P];
          FOR L:=2 TO M DO
          VREG[M-1,L-1]:=IND[L];
        END
      ELSE
        BEGIN
          NVF:=M+DIF-1;
          IF RSS[M-1]>TXM[COL+P,P] THEN
            BEGIN
              RSS[M-1]:=TXM[COL+P,P];
              FOR L:=2 TO M DO
               VREG[M-1,L-1]:=IND[L];
            END;
            JKL:=TMI[COL+H+VT1,H+VT1]* (-1);
          IF RSS[NVF]>JKL THEN
            BEGIN
              RSS[NVF]:=JKL;
              FOR L:=2 TO M-1 DO
```

```
              VREG[NVF,L-1]:=IND[L];
              FOR L:=1 TO DIF DO
              VREG[NVF,M-2+L]:=IND[M]+L;
              VREG[NVF,NVF]:=P;
            END;
            IF TER=1 THEN COL:=0 ELSE COL:=(TER-1)*14+1;
            JKL:=TM*[COL+H+VT1,H+VT1]*(-1);
           IF RSS[M-1]<=JKL THEN
            BEGIN
              AVANCE:=TRUE;
             { WRITELN(LST,'AVANCE');}
              MN:=M;
            END
          ELSE MN:=P;
        END;
     END;
  UNTIL M<=1;

{*****                    MUESTRA DE RESULTADOS.                    *****}

  IF IMPRESION THEN
    BEGIN
      WRITELN(LST);WRITELN(LST);WRITELN(LST);WRITELN(LST);
      WRITELN(LST,'          SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES');
    END;
  CLRSCR;
  WRITELN(^J,^M,^M,^M,^M,'          SELECCION DE LOS MEJORES SUBCONJUNTOS DE VARIABLES');
  FOR L:=1 TO P-1 DO
   BEGIN
    IF L MOD 6=0 THEN
      BEGIN
        WRITELN;CONTINUAR;CLRSCR;
         WRITELN(^J,^M,^M,^M,^M,'        . SELECCION DE LOS MEJORES SUBCONJUNNTOS DE VARIABLES')
      END;
    WRITELN(^J,^M,^M,'MEJOR SUBCONJUNTO CON ',L:2,' VARIABLES');
    WRITE('  VARIABLES  :  ');
    IF IMPRESION THEN
       BEGIN  WRITELN(LST);
          WRITELN(LST,'MEJOR SUBCONJUNTO CON ',L:2,' VARIABLES');
          WRITE(LST,'  VARIABLES  :  ');
       END;
    FOR M:=1 TO L DO
       BEGIN
          LL:=VREG[L,M];
          IF M=1 THEN
           BEGIN
             WRITE(ORE[H+VT1-LL]:2);
             IF IMPRESION THEN WRITE(LST,ORE[H+VT1-LL]:2);
           END
          ELSE
           BEGIN
             WRITE(',',ORE[H+VT1-LL]:2);
             IF IMPRESION THEN WRITE(LST,',',ORE[H+VT1-LL]:2);
           END;
```

```
            END;
            CP_MALLOWS(RSS[L],L,CPS1);
            WRITELN;
            WRITE('  SCRES=  ',RSS[L]:8:4,'   Cp_MALLOWS=  ',CPS1:8:4);WRITELN;
            IF IMPRESION THEN
              BEGIN
                WRITELN(LST);
                WRITE(LST,'  SCRES=  ',RSS[L]:8:4,'   Cp_MALLOWS=  ',CPS1:8:4);
                WRITELN(LST);
              END;
        END;
      END
      ELSE
         BEGIN
           GOTOXY(5,22); WRITELN('> SOLO SE TIENE UNA VARIABLE');
         END;
      WRITELN;
      CONTINUAR;
   END;
   END.
```

A N E X O B

DEMOSTRACION DE QUE X´X ES DEFINIDA POSITIVA

En la solución del sistema

$$(X'X)\hat{\beta} = X'Y$$

utilizamos el método de la descomposición de Cholesky por las ventajas mencionadas en el Capítulo I.

El método de la descomposición de Cholesky requiere que la matriz $X'X$ sea definida positiva.

El siguiente teorema ayuda a demostrar que $X'X$ es definida positiva.

Teorema. *Si X es una matriz de (n x p) de rango p, entonces X'X es positiva definida.*

Prueba: $xX'Xx = y'y > 0$. La igualdad se cumple $\Leftrightarrow$ $Xx = 0$ $\Leftrightarrow$ $x = 0$, ésto porque las columnas de X son linealmente independientes. ♦

(El rango de X es (n x p) por hipótesis).

A N E X O C

MANEJO Y UTILIZACION DE ARCHIVOS

En el paquete se utilizan los archivosde tipo texto o también llamado "Standard Data File, (SDF)", este tipo de archivo puede ser creado en el editor del TURBO Pascal o en cualquier otro editor de textos, como el archivo no-documento de WORDSTAR.

Cada renglón del archivo de texto es considerado como un registro y éstos deben estar en formato libre, es decir, los números separados por comas y sin espacios entre estos. Al crear un registro el valor de la primera componente debe ser el de Y (aúnque solamente se haya escogido la opción de LEER X) y las siguientes componentes serán los valores de las X's. Como cada renglón es un registro, su longitud máxima es de 255 columnas. Si en algún renglón existe un espacio, ya sea al principio o entre los valores, o algún caracter no numérico (excepto las comas y puntos), entonces el programa indicará error en dicho renglón y lo descartará de los datos.

Al crearse un archivo de datos (no importa el tipo) no se debe incluír la columna de unos para el término independiente, ya que los procedimientos están diseñados para considerarlo.

El nombre de los archivos está compuesto por dos partes, la primera parte consta de 1 a 8 caracteres, la segunda parte se le llama Extensión y consiste de un punto (.) seguido de 1 a 3 caracteres y es opcional. Los caracteres legales para el nombre son:

A-Z 0-9 $ & # % ' ( ) - _ @ ^ { } ~ !

Cuando el programa pregunta por el nombre de un archivo, puede ser debido a dos situaciones: 1) Encontrar un archivo en el disco para leerlo, sí no existe, se volverá a pedir el nombre. 2) Para crear un archivo con ese nombre, si ya existe el archivo o es ilegal el nombre, se volverá a pedir.

En la lectura de archivos se pregunta cuantos registros se leerán (n), pero no se podrá leer más registros de los que existen, debido a que la lectura se suspende cuando se encuentra el fin del archivo, y si n es mayor al numero de registros, entonces, se actualizará al número de registros leídos.

A N E X O   D

POLINOMIOS ORTOGONALES DE HERMITE[1]

Dáda una función de distribución $F(x)$ con momento finitos

$$\alpha_v = \int_{-\infty}^{\infty} X^{v} dF(X)$$

de todos los ordenes, existe una sucesión de polinomios $P_0(x), P_1(x), \ldots,$ determinados por las siguientes condiciones:

i) $P_n(x)$ es de grado n y el coeficiente de X^n en $P_n(x)$ es positivo.

ii) Los $P_n(x)$ satisfacen las condiciones de ortogonalidad:

$$\int_{-\infty}^{\infty} P_m(X)\ P_n(X)\ dF(X) = \begin{cases} 1 & \text{para } m=n \\ 0 & \text{para } m\neq n \end{cases}$$

$P_0(x)=1$ y $P_n(x)= u_0+u_1 X + \ldots + u_n X^n$ con $n>0$, así obtendremos los coeficientes u_i mediante i) y ii).

1) Cramer, H., "Metodos Matematicos de Estadistica", pp.

$\int_{-\infty}^{\infty} X^i P_n(X)\ dF(X)=0$ para $i=0,1,2,\ldots,n-1$, efectuando las integrales se obtienen n ecuaciones lineales homogeneas entre las n+1 incógnitas ($u_o,\ldots,u_n$). Además se sabe que $\int_{-\infty}^{\infty} P_n^2(X)\ dF(X)=1$.

Consideremos el caso particular de la función de densidad continua $f(x)=F'(x)$, y sean $P_o(X),P_1(x)\ldots$ los polinomios ortogonales correspondientes. Si $g(x)$ es otra función de densidad, podemos desarrollar $g(x)$ en una serie de la forma:

$$g(x) = b_oP_o(x)f(x) + b_1P_1(x)\ f(x) + \ldots$$

$$= f(x)\ \{\ b_oP_o(x) + b_1P_1(x) + \ldots\ \},$$

multiplicando por $P_n(X)$ e integrando término a término y por i) y ii) obtenemos que:

$$b_n = \int_{-\infty}^{\infty} P_n(X)\ g(X)\ d(x)$$

y en particular $b_o=1$.

Los polinomios de Hermite $H_n(x)$ están definidos por las relaciones:

$$\left(\frac{d}{dx}\right)^n e^{-\frac{x^2}{2}} = (-1)^n\ H_n(X)\ e^{-\frac{x^2}{2}}$$

$$(n=0,1,2,\ldots)$$

$H_n(x)$ es un polinomio de grado n, y se tiene:

$$H_o(x) = 1,\ \ H_1(x) = X,\ H_2(x) = X^2-1,\ \ \text{etc.}$$

Así,

$$\int_{-\infty}^{\infty} H_m(X)\ H_n(X)\ d\Phi(X) = \int_{-\infty}^{\infty} H_m(X)\ H_n(X) e^{-\frac{x^2}{2}}\ dx = \begin{cases} n! & \text{para } m=n \\ 0 & \text{para } m\neq n \end{cases}$$