

**COLEGIO DE INGENIEROS DE VENEZUELA**

**SOCIEDAD DE INGENIERIA DE TASACION DE VENEZUELA**

# **Apéndice 1**

**Aplicación de la Hoja de Cálculo MS-Excel a problemas  
de Regresión Lineal Múltiple**

Ing. Roberto Piol Puppio  
E-Mail: [rpiol@yahoo.com](mailto:rpiol@yahoo.com)  
[www.rpiol.com](http://www.rpiol.com)

## REGRESION MULTIPLE LINEAL CON EXCEL 6.0 (o superior)

### Instrucciones:

- Escoger en fx la función: ESTIMACION.LINEAL
- En la Caja de diálogo marcar la Columna de la Variable Dependiente (y) con el ratón
- En la Caja de diálogo marcar las Columnas de las Variables Independientes (x) con el ratón
- Indicar en la ventanilla "CONSTANTE" el argumento: VERDADERO
- Indicar en la ventanilla "ESTADISTICA" el argumento: VERDADERO
- Marcar con el Ratón el Rango de Salida (\*) de los elementos de la Regresión
- Iluminar con el ratón la "Barra de Fórmulas" la ventanilla donde aparece la fórmula de la regresión
- Apretar simultáneamente las teclas: "CONTROL", "SHIFT" y "ENTER"

(\*) el Rango de Salida tiene en tamaño de: 5 líneas X el # de Variables de columnas

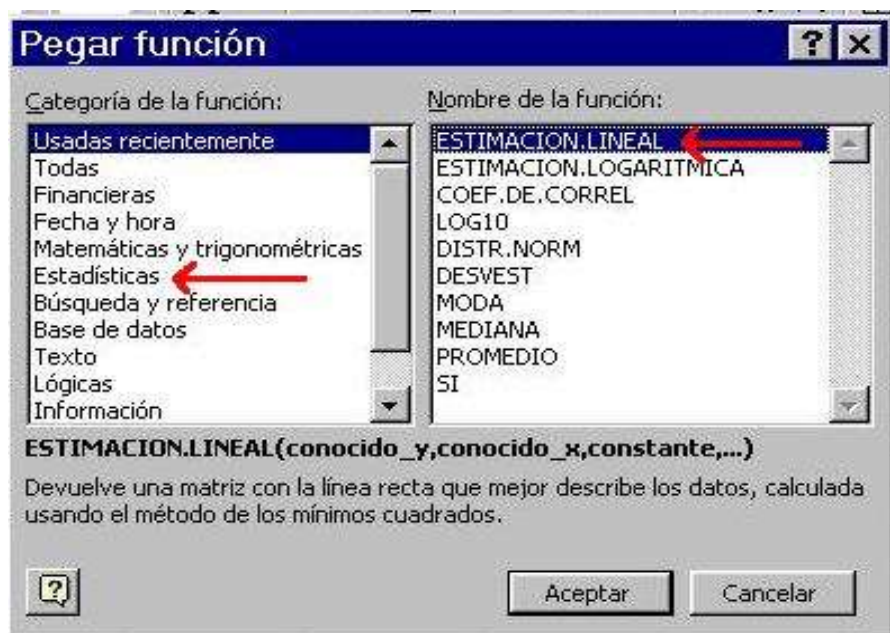
## Galpones Industriales en la Urbanización San Martín. Caracas

1.- Enterar en la Hoja de Cálculo los referenciales seleccionados

**Nota: Cersiórese que todas las Variables Independientes estén agrupadas en columnas adyacentes**

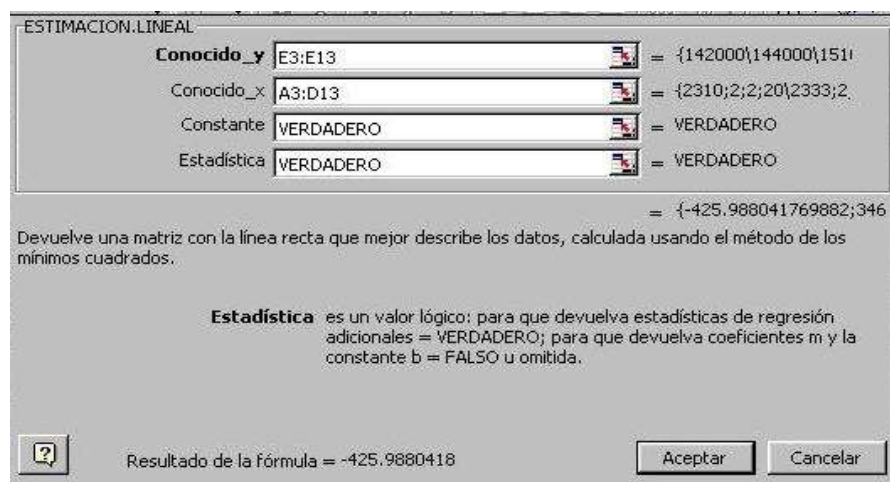
	A	B	C	D	E
1	AREA (M2)	# OFICINAS	# BANOS	EDAD (AÑOS)	P.UNIT
2	X1	X2	X3	X4	Y
3	2,310.00	2	2.0	20	142,000
4	2,333.00	2	2.0	12	144,000
5	2,356.00	3	1.5	33	151,000
6	2,379.00	3	2.0	43	150,000
7	2,402.00	2	3.0	43	139,000
8	2,425.00	4	2.0	23	169,000
9	2,448.00	2	1.5	45	126,000
10	2,471.00	2	2.0	34	142,900
11	2,494.00	3	3.0	23	163,000
12	2,517.00	4	4.0	35	169,000
13	2,540.00	2	3.0	22	149,000

2.- Click en **fx**: Mostrará una caja de diálogo con Dos Columnas



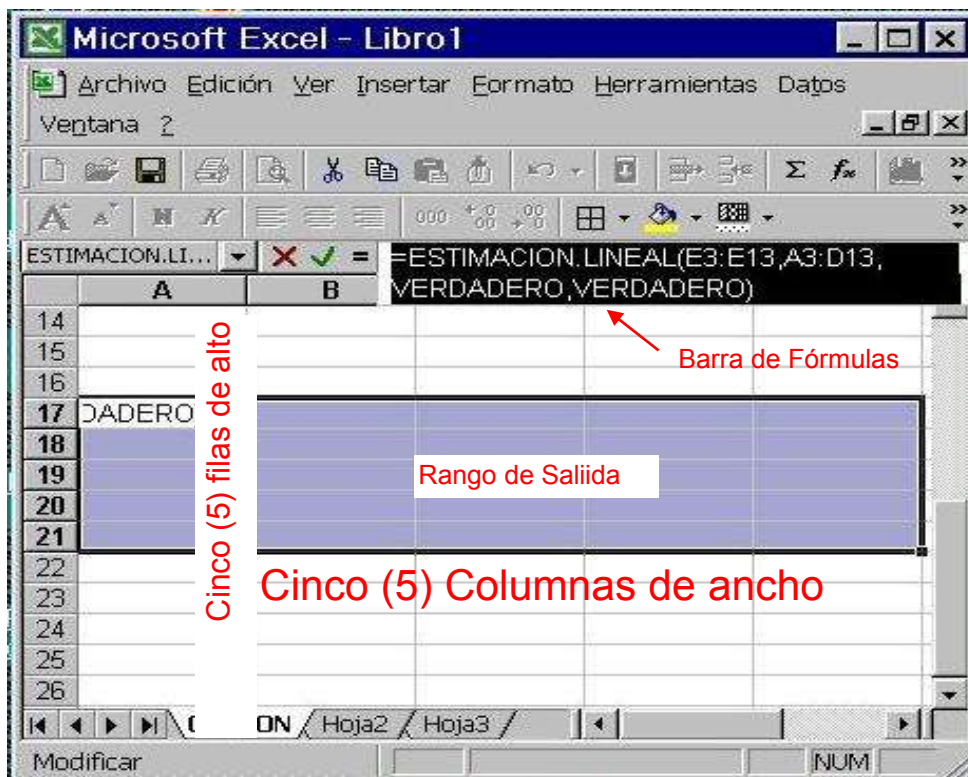
3.- Click en **Estadísticas** (Columna izquierda), e inmediatamente Click en **ESTIMACION.LINEAL** (Columna derecha). La Subrutina lo enviará la caja de diálogo donde:

- a) Deberá iluminar con el ratón la columna de la Variable Independiente (Y)
- b) Deberá iluminar con el ratón conjuntamente las columnas de las Variable Dependientes
- c) Escribir en la ventanilla "**Constante**" el argumento VERDADERO
- d) Escribir en la ventanilla "**Estadística**" el argumento VERDADERO
- e) Click en el boton ACEPTAR; esto lo devolverá a la Hoja de Cálculo

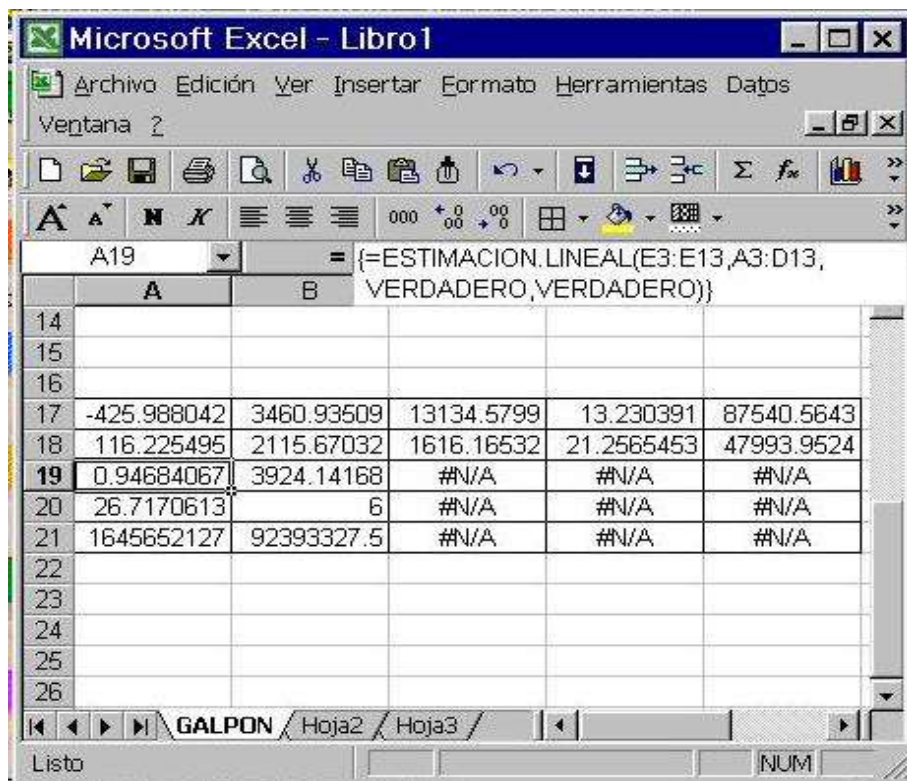


4.- Partiendo de la celda donde Excel indico la salida numérica de la Subrutina ESTIMACION.LINEAL (0.99704807):

- a) Iluminar con el ratón el RANGO DE SALIDA de los elementos de la Regresión  
*Nota: El rango de salida, tiene un formato específico que es el de Cinco (5) líneas de alto por "n" columnas de ancho, donde "n" es igual al número de Variables (Dependientes e Independientes) que tengan los datos que vamos a modelar. En el caso de este ejemplo, el rango de salida sera de Cinco (5) líneas de alto por Cinco (5) columnas de ancho.*
- b) Iluminar con el ratón la "Barra de Fórmulas" (ventanilla donde Excel indica la fórmula de la regresión)
- c) Apretar simultáneamente las teclas "CONTROL", "SHIFT" y "ENTER" (en los teclados en español "CTRL", "Mayus" e "Intro")
- d) Esta acción, permite que en el Rango de Salida, que ya habíamos iluminado, se descomprima permitiendo ver los diferentes elementos de la regresión.



5.- La Salida de la Subrutina ESTIMACION.LINEAL, es presentada por EXCEL de la siguiente forma:



6.- La interpretación de esta Salida es la siguiente:

E	D	C	B	A
Se4	Se3	Se2	Se1	Se0
R^2	SeY			
F	df			
SCR	SCE			

Donde:

<b>A:</b>	Coeficiente del término independiente
<b>B:</b>	Coeficiente que acompaña a X1 (Area)
<b>C:</b>	Coeficiente que acompaña a X2 (Nro. De Oficinas)
<b>D:</b>	Coeficiente que acompaña a X3 (Nro. De Baños)
<b>E:</b>	Coeficiente que acompaña a X4 (Edad del Galpón)
<b>Se0:</b>	Error estándar de la constante A
<b>Se1 al Se4:</b>	Error estándar de cada uno de los coeficientes de las Variables Independientes
<b>SeY:</b>	Error estándar de la correlación
<b>R^2:</b>	Coeficiente de Determinación
<b>F:</b>	Estadístico F
<b>df:</b>	Grados de libertad [ n - (k+1) ]
<b>SCR:</b>	Suma del cuadrado de la regresión
<b>SCE:</b>	Suma del cuadrado del error

7.- El Resultado de la Regresión Lineal Múltiple será:

*Modelo de Regresión Lineal:*  $Y = A + B*X1 + C*X2 + D*X3 + E*X4$

$Y = 87540.56 + 13.23*X1 + 13134.58*X2 + 3460.94*X3 - 425.99*X4$

$R^2 = 94.68\%$

$F = 26.71$

$F_0 = 4.53$

## REGRESION MULTIPLE LOGARITMICA CON EXCEL 6.0 (o superior)

La Subrutina para el Cálculo de la Regresión Múltiple Logarítmica de Excel es similar a la Lineal, ya estudiada. El comando para activarla es :

**ESTIMACION.LOGARITMICA.**

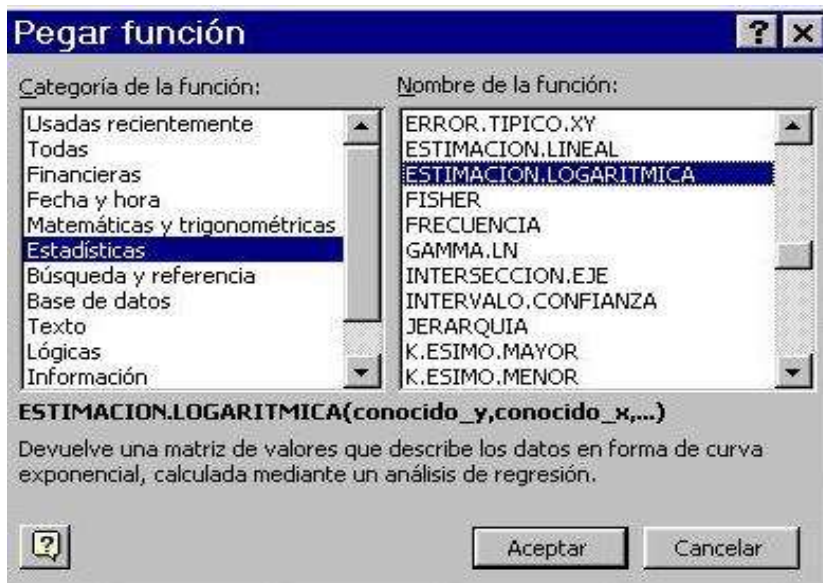
1.- Se expondrá brevemente las Instrucciones:

### **Instrucciones:**

- a) Escoger en fx la función: ESTIMACION.LOGARITMICA
- b) En la Caja de diálogo marcar la Columna de la Variable Dependiente (y) con el ratón
- c) En la Caja de diálogo marcar las Columnas de las Variables Independientes (x) con el ratón
- d) Indicar en la ventanilla "CONSTANTE" el argumento: VERDADERO
- e) Indicar en la ventanilla "ESTADISTICA" el argumento: VERDADERO
- f) Marcar con el Ratón el Rango de Salida (\*) de los elementos de la Regresión
- g) Iluminar con el ratón la "Barra de Fórmulas" la caja donde aparece la fórmula de la regresión
- h) Apretar simultáneamente las teclas: "CONTROL", "SHIFT" y "ENTER"

**(\*) el Rango de Salida tiene en tamaño de: 5 líneas X el # de Variables de columnas**

a) Escoger en **fx** la función: ESTIMACION.LOGARITMICA

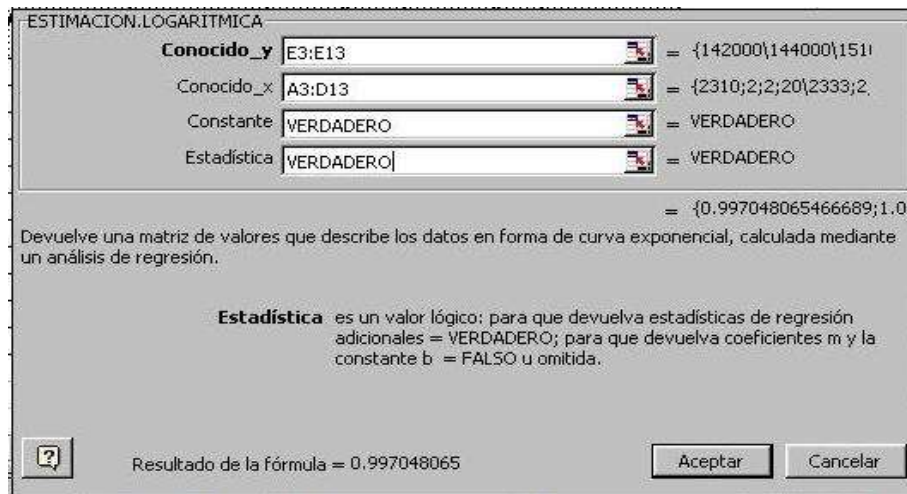


b) En la Caja de diálogo marcar la Columna de la Variable Dependiente (y) con el ratón

c) En la Caja de diálogo marcar las Columnas de las Variables Independientes (x) con el ratón

d) Indicar en la ventanilla "CONSTANTE" el argumento: VERDADERO

e) Indicar en la ventanilla "ESTADISTICA" el argumento: VERDADERO



- f) Marcar con el Ratón el Rango de Salida de los elementos de la Regresión
- g) Iluminar con el ratón la "Barra de Fórmulas", ventanilla donde aparece la fórmula de la regresión
- h) Apretar simultáneamente las teclas: "CONTROL", "SHIFT" y "ENTER"



**Nota:** el Rango de Salida tiene en tamaño de: 5 líneas por el # de Variables de columnas

3.- El Resultado de la Regresión Logarítmica Múltiple sera:

$$\begin{aligned}
 & \text{Modelo de Regresión: } Y = A * B^{X1} * C^{X2} * D^{X3} * E^{X4} \\
 & Y = 104,481.9964 * 1.0001^{X1} * 1.0901^{X2} * 1.0252^{X3} * 0.997^{X4} \\
 & R^2 = 92.82\% \quad F = 19.3953 \quad Fo = 4.54
 \end{aligned}$$

4.- ¿Que Modelo utilizar Lineal o Logarítmico?

El modelo que mejor explica el Fenómeno: "Precios Unitarios de Galpones Industriales en la Urbanización Industrial San Martín. Caracas"; sera aq ue cuyo Coeficiente de Determinación R^2 sea Mayor.

**CONCLUSION:** En este caso, el Modelo LINEAL explica mejor que el EXPONENCIAL



## CALCULO DE LA MATRIZ DE CORRELACION CON EXCEL 6.0 (o superior)

Para calcular la Matriz de Correlación, utilizando las Funciones (**fx**) Estadísticas de la hoja de cálculo Excel; hay que proceder de manera manual. La Matriz de Correlación debe ser calculada factor por factor (procedimiento algo engorroso)

1.- Se expondrán brevemente las instrucciones:

### Instrucciones:

- a) Crear en la hoja la matriz de correlación: Diagonal principal = 1; Ignorar la 1ra. Columna  
b) Utilizando la función (fx) "COEF.DE.CORREL", generar los 6 coeficientes de covarianza:

X1 - X2  
X1 - X3  
X1 - X4  
X2 - X3  
X2 - X4  
X3 - X4

- c) Completar la Matriz

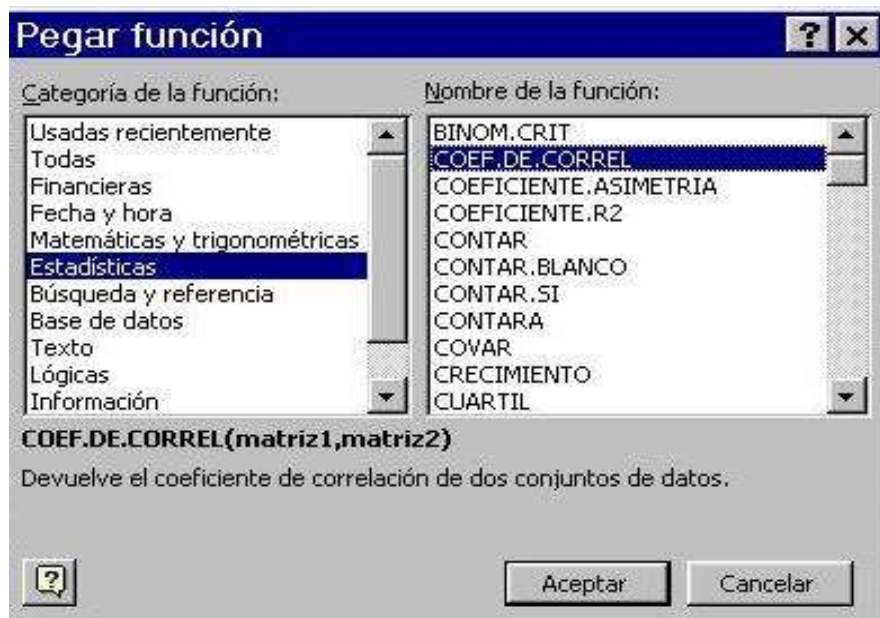
2.- Crear en la hoja de cálculo, en forma manual, el esqueleto de la matriz de correlación:

	CTE.	X1	X2	X3	X4
CTE.					
X1					
X2					
X3					
X4					

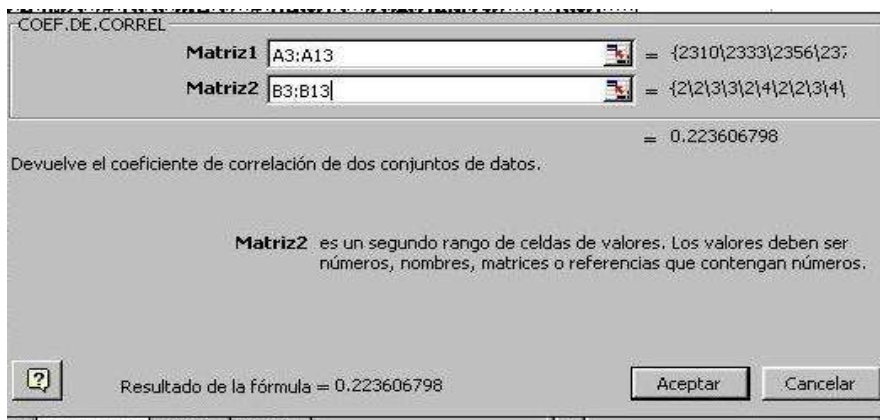
3.- Crear la Diagonal Principal con unos (1) e ignorar la Primera Columna

	CTE.	X1	X2	X3	X4
CTE.	1,00				
X1	-o-	1,00			
X2	-o-		1,00		
X3	-o-			1,00	
X4	-o-				1,00

4.- Click en **fx**: Mostrará una caja de diálogo con Dos Columnas seleccionar en la columna de la izquierda: "Estadísticas" y en la columna de la derecha: "COEF.DE.CORREL."



5.- Dentro del "esqueleto" de la Matriz, ubicarse el la intersección de la Variable Independientes X1-X2. Click en **fx**, se abrirá la caja de diálogo con dos ventanillas (Matriz 1 y Matriz 2).



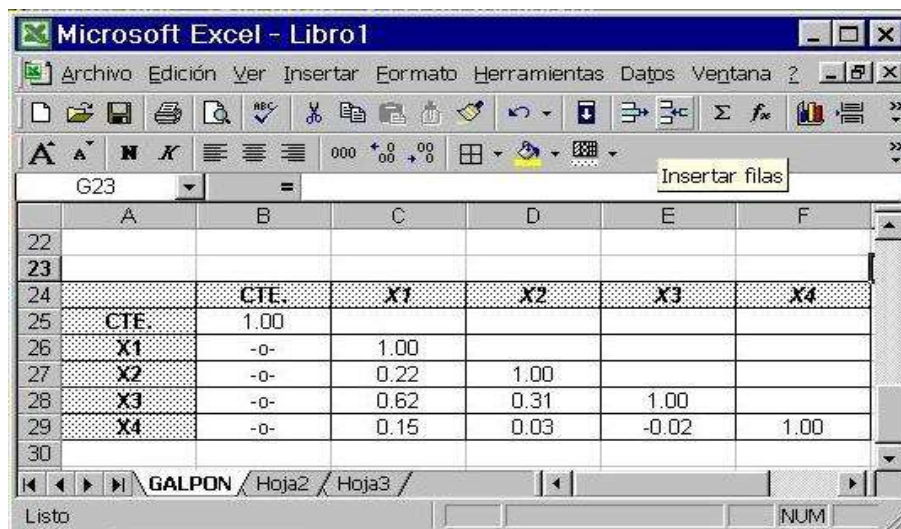
6). Ilumine con el ratón la variable independiente X1 para una de las ventanillas y a la variable independiente X2 para la otra. Click en Aceptar. Esta operación generó el Coeficiente de Correlación entre las Variables independientes X1 y X2

	CTE.	X1	X2	X3	X4
CTE.	1,00				
X1	-0-	1,00			
X2	-0-	<b>0,22</b>	1,00		
X3	-0-			1,00	
X4	-0-				1,00

7.- Utilizando la función (fx) / "COEF.DE.CORREL", generar los Cinco (5) Coeficientes de correlación faltantes para los pares:

- X1 - X3
- X1 - X4
- X2 - X3
- X2 - X4
- X3 - X4

y completar la Matriz de Correlación.



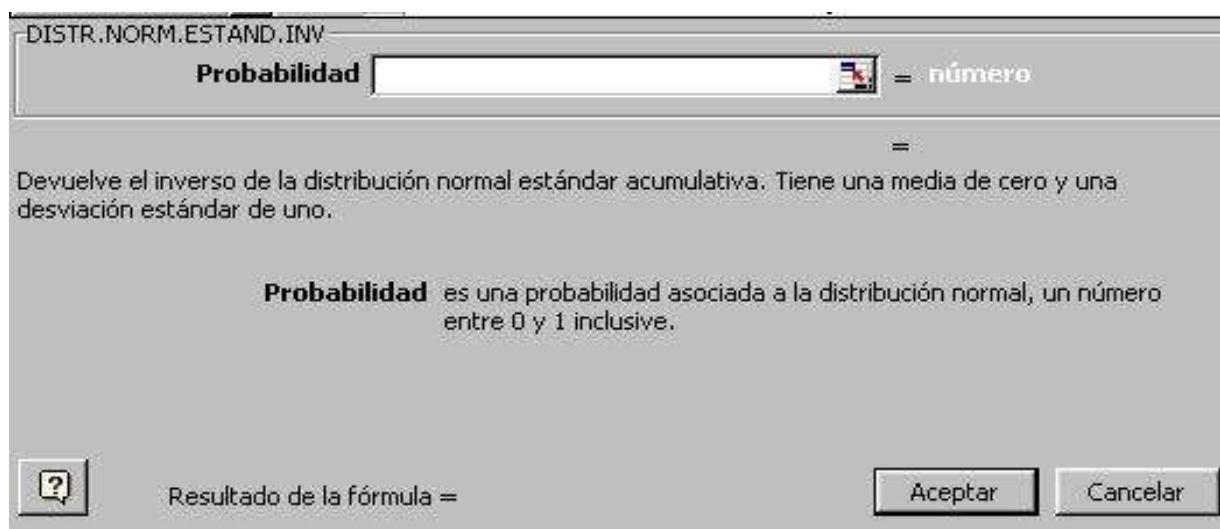
Observese, que no hay problemas de Multicolinealidad entre las Variables Independientes

## DISTRIBUCION NORMAL ESTANDAR INVERSA (IDF)

Para calcular el coeficiente "k", que genera la Distribución Normal Estandar Inversa:

1.- Escoger en **fx** la función: DISTR.NORM.ESTAND.INV (IDF en Inglés) de las funciones estadísticas de la Hoja de Cálculo Excel.

2.- Ingresar en la ventanilla de la caja de diálogo presentada por Excel, la Probabilidad "p", bien como número, fracción o indicando la celda donde ya se haya calculado



3.- Click en "Aceptar", generará la salida de esta subrutina que será el valor del coeficiente "k"

### NOTA IMPORTANTE:

Nota: Para aquellos usuarios de la versión "Professional" de Excel; tienen la oportunidad de instalar el conjunto de macros : "Análisis de Datos", desde el CD de instalación y proceder con los comandos:

Herramientas → Análisis de datos → Coeficiente de Correlación

una vez abierta la caja de diálogo, iluminar las Variables Independientes (X) con el ratón (sub-caja: "Rango de Entrada") e indicarle en que celda de la hoja de cálculo se requiere la salida de la matriz (Sub-caja: "Rango de Salida")

Esta macro, es especialmente útil cuando se va a analizar modelos de regresión con muchas variables independientes.

Ejemplo completo de Aplicación:

Sea la siguiente serie correspondientes a apartamentos en una Urbanización del suroeste de Caracas:

REFER.	P.UNIT Bs.F./M2 Y	AREA M2 X1	EDAD AÑOS X2	% DE CONDominio X3
A	1.245	75	1	0,942%
B	987	100	7	0,515%
C	1.195	80	3	0,363%
D	1.500	100	8	1,000%
E	975	100	13	0,561%
F	705	125	16	0,874%
G	872	110	12	0,623%
H	1.007	90	19	0,409%
I	1.450	60	1	0,201%
J	1.332	85	8	0,277%
K	1.082	90	7	0,414%
L	750	51	3	0,836%

## LINEAL

-25305,4062	-25,9970171	2,264762828	1250,7204
29889,83393	19,56592643	5,694049533	395,98012
<b>0,284908723</b>	256,0419548	#N/A	#N/A
1,062461001	8	#N/A	#N/A
208956,8058	524459,8609	#N/A	#N/A

## LOGARITMICA

3,49749E-13	0,975890266	1,002286345	1252,68
27,7506444	0,01816561	0,005286531	0,3676402
<b>0,304787497</b>	0,237717254	#N/A	#N/A
1,169090969	8	#N/A	#N/A
0,198194213	0,452075942	#N/A	#N/A

El Modelo Logarítmico, explica mejor el fenómeno

MATRIZ DE CORRELACION:

	AREA	EDAD	CONDOM
AREA	1		
EDAD	0,725799624	1	
CONDOM	0,23161363	0,039900051	1

No hay problemas de Multicolinialidad entre las variables independientes

### CALCULO DE LOS VALORES ATIPICOS (Método Empírico)

REF	Y observ.	Y calc.	Residuo	
A	1.245	1.107	138	
B	987	1.145	-158	
C	1.195	1.259	-64	
<b>D</b>	<b>1.500</b>	<b>972</b>	<b>528</b>	ATIPICO
E	975	976	-1	
F	705	878	-173	
G	872	1.005	-133	
H	1.007	861	146	
I	1.450	1.323	127	
<b>J</b>	<b>1.332</b>	<b>1.156</b>	<b>176</b>	ATIPICO
K	1.082	1.152	-70	
<b>L</b>	<b>750</b>	<b>1.029</b>	<b>-279</b>	ATIPICO

### NUEVA SERIE

REFER.	P.UNIT Y	AREA M2 X1	EDAD AÑOS X2	% DE CONDominio X3
A	1.245	75	1	0,942%
B	987	100	7	0,515%
C	1.195	80	3	0,363%
E	975	100	13	0,561%
F	705	125	16	0,874%
G	872	110	12	0,623%
H	1.007	90	19	0,409%
I	1.450	60	1	0,201%
J	1.332	85	8	0,277%
K	1.082	90	7	0,414%

### LOGARITMICA

6,96826E-07	0,994261752	0,991313595	2673,5435
11,00904861	0,005386193	0,002090944	0,1338812
<b>0,940585712</b>	0,063720358	#N/A	#N/A
<b>31,66193639</b>	6	#N/A	#N/A
0,385669365	0,024361704	#N/A	#N/A

R <sup>2</sup> =	94,06%
F =	31,6619
k =	3
n =	10
n - (k+1) =	6
Fo =	4,76
	<b>F &gt;&gt; Fo</b>

Modelo de Correlación:  $Y = 2673,544 * 0,99131359^{X1} * 0,99426175^{X2} * 0,0000069683^{X3}$