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 www.rpiol.com

REGRESION MULTIPLE LINEAL CON EXCEL 6.0 (o superior)

Instrucciones:

a) Escoger en fx la función: ESTIMACION.LINEAL

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

c) En la Caja de diálogo marcas 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) lluminar con el ratón la "Barra de Fórmulas" la ventanilla 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

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

	Microsoft	Excel - Lib	ro1		_ 🗆 ×
1	Archivo Edici	ión <u>V</u> er <u>I</u> nser	tar <u>F</u> ormato	Herramientas	Datos
Ve	<u>n</u> tana <u>?</u>				_ 8 ×
1n			a A D.	- □ ⇒ 3+	Σf_{*}
1.**			000 + 0 00		
JA	АЛЛ		000 00 + 0		• .
	G15				
	A	В	C	D	E
1	AREA (MZ)	# OFICINAS	# BANOS	EDAD (ANOS)	P.UNIT -
2	X1	X2	X 3	X4	<u> </u>
3	2,310.00	0 2 2.0		20	142,000
4	2,333.00	2	2.0	12	144,000
5	2,356.00	2,356.00 3		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 💂
	GAL	PON / Hoja2 /	Hoja3 /	Í∙	
List	o	1		V 13	NUM

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

ategoria de la funcion:	
Todas Financieras Fecha y hora Matemáticas y trigonométricas Estadísticas Búsqueda y referencia Base de datos Texto Lógicas Información	ESTIMACION.LOGARITMICA COEF.DE.CORREL LOG10 DISTR.NORM DESVEST MODA MEDIANA PROMEDIO SI
STIMACION.LINEAL(conocido	y,conocido_x,constante,)
evuelve una matriz con la línea r Isando el método de los mínimos (ecta que mejor describe los datos, calculad uadrados.

3.- Click en **Estadísticas** (Columna izquierda), e inmeditamente 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

ESTIMACION.LINEAL		
Conocido_y	E3:E13	142000\144000\151
Conocido_x	A3:D13	1 = {2310;2;2;20\2333;2
Constante	VERDADERO	VERDADERO
Estadística	VERDADERO	VERDADERO
Devuelve una matriz con la líne: mínimos cuadrados. Estadí :	a recta que mejor describe los stica es un valor lógico: para adicionales = VERDADI constante b = FALSO (datos, calculada usando el método de los a que devuelva estadísticas de regresión ERO; para que devuelva coeficientes m y la u omitida.
লা		

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

a) lluminar 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) lluminar 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.

🔀 Microsoft Excel - Libro1 📃 🗖 🗙
Archivo Edición Ver Insertar Eormato Herramientas Datos
D 🛩 🖬 🗇 🖪 👗 🛍 🖻 🗗 🗠 - 🖬 🗁 🚈 🗶 🖍 🏙 😤
🗛 🔺 🖪 🖉 😳 *8 .98 🖽 • 🦄 • 🕅 • 👌
ESTIMACION.LI 🛃 🗙 🗸 = ESTIMACION.LINEAL(E3:E13,A3:D13,
A B VERDADERO, VERDADERO)
14 0 15 0 16 0 17 DADERO
18 See 19 E 20 G
21 0
$\frac{22}{23}$ $\stackrel{e}{\odot}$ Cinco (5) Columnas de ancho
25
Modificar NUM

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

	Microsc	oft E	Exc	el -	Lib	ro1								-		×
Ver	<u>A</u> rchivo [otana <u>2</u>	Edicio	ón y	<u>/</u> er	Insei	rtar	Eor	mato	He	rrai	mie	ntas	Da	tos J	_ [8]	×
	😂 🖬 🛛	9	Q.	Ж	6	8	٥	5	- 1		⊒+	3+c	Σ	f.		>> *
A	A N	K	E	=		000	+_0	• 00 • 0	⊞ •	8	• •		*			>> *
	A19	-		3	= {=	ESTI	MA	CION	LIN	EAI	_(E3	3:E1	3,A3	8:D13	3,	
	Α			В		ERD	AD	ERO,	VER	DA	DE	RO)]	<u> </u>			1000
14																
15	N		-			-			-			_	_			
16	105.000	0.10	0.40			1 40	101	5700	1 2	0.0		-	071	10.0	0.00	
17	-425.988	1042	346	50.93	3509	13	134	.5/95		13.4	2303	391	8/5	040.5	643	
18	116.225	495	21	15.6/	1032	16	16.	16532	2	25	12.0	153	475	193.9	524	
19	0.94684		392	24.14	1168		#\\/	A	100	#1	VA VA		7	#WA		
20	26.7170	613	000	0000	07.5		#N/	A	-	#1\ 21b	VA VA		Ĩ		-	
21	1645652	127	923	5933.	27.5	1	₩W	A	-	H)	IΑ	-	7	¶V/A	-	
22						-			-			-				4
20	<u>.</u>		-			-									-	ŝ.
24	-				_	- C-			-			-	_		-	
20										_		_				
14		GAL D	ON	/Hor	ia2 /	Hoi	- 54	1	1.1				1		-	ř
Lint		JALL	0.47	(110)		- i Oje			1.1	-	_	-10	_	6 H (6 A		
LISU	0				3		Ja:					1		LACUAL	1	11.

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

E	D	С	В	Α
Se4	Se3	Se2	Se1	Se0
R^2	SeY			
F	df			
SCR	SCE			

Donde:

A :	Coeficiente del término independiente
B:	Coeficiente que acompaña a X1 (Area)
C:	Coeficiente que acompaña a X2 (Nro. De Oficinas)
D:	Coeficiente que acompaña a X3 (Nro. De Baños)
E:	Coeficiente que acompaña a X4 (Edad del Galpón)
Se0:	Error estándar de la constante A
Se1 al Se4:	Error estándar de cada uno de los coeficientes de las Variables Independientes
SeY:	Error estándar de la correlación
R^2:	Coeficiente de Determinación
F:	Estadístico F
df:	Grados de libertad [n - (k+1)]
SCR:	Suma del cuadrado de la regresión
SCE:	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*X4Y = 87540.56 + 13.23*X1 + 13134.58*X2 + 3460.94*X3 - 425.99*X4 $R^2 = 94.68\%$ F = 26.71F = 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 al 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 marcas 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) lluminar 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

Isadas recientemente	The second s
isauas reciencemence	ERROR.TIPICO.XY
odas	ESTIMACION, LINEAL
inancieras	ESISIMACIONA LOIGARISIMICA
echa y hora	FISHER
latematicas y trigonometricas	FRECUENCIA
stadisticas	GAMMA.LN
usqueda y rererencia	INTERSECCION.EJE
ase de datos	
éxico	LESTINO MAYOD
	K ESTMO MENOR
5TIMACION.LOGARITMICA(cor	iocido_y,conocido_x,)
evuelve una matriz de valores que	describe los datos en forma de curva
	76

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

ESTIMAC	ION.LOGARITMICA	
	Conocido_y E3:E13	142000\144000\151
	Conocido_x A3:D13	1 = {2310;2;2;20\2333;2
	Constante VERDADERO	VERDADERO
	Estadística VERDADERO	VERDADERO
Devuelve un análisi	una matriz de valores que describe los datos en fo s de regresión.	= -{0.997048065466689;1.0 prma de curva exponencial, calculada mediante
	Estadística es un valor lógico: par adicionales = VERDAD constante b = FALSC	ra que devuelva estadísticas de regresión JERO; para que devuelva coeficientes m y la) u omitida.
2		

f) Marcar con el Ratón el Rango de Salida de los elementos de la Regresión

g) lluminar con el ratón la "Barra de Fórmulas", ventanilla donde aparece la fórmula de la regresión

h) Apretar simultáneamente las teclas: "CONTROL", "SHIFT" y "ENTER"

	Microsoft E	Excel - Libi	ro1			
2	<u>A</u> rchivo <u>E</u> dició	ón ⊻er Inser	tar <u>E</u> ormato	<u>H</u> erramientas	Da <u>t</u> os Ve <u>n</u> t	ana <u>8 ×</u>
	288	Q. % Pa (8 🗄 🗠 -	₫ ⇒ 3~	Σ f*	,昌 ?
A	A H K		000 *08 +08 E	8 - 🕭 - 🕅	-	>> *
	A34 <u>*</u>	= {=E B VE	STIMACION.I RDADERO,V	_OGARITMICA (ERDADERO))	A(E3:E13,A3:I }	D13,
31						-
32	0.99704807	1.02518297	1.09012376	1.00006524	104481.996	
33	0.00090584	0.01648919	0.01259613	0.00016567	0.37405703	
34	0.92821352	0.03058412	#N/A	#N/A	#N/A	
35	19.3953003	6	#N/A	#N/A	#N/A	
36	0.07256855	0.00561233	#N/A	#N/A	#N/A	
37						
38		06/ N		10 M		. .
4 4	GALP	ON / Hoja2 /	Hoja3 /	•		•
List	5				NUM	- /

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:

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

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 aque 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 funcion (fx) "COEF.DE.CORREL", generar los 6 coeficientes de covarianza:
X1 - X2
X1 - X3
X1 - X4
X2 - X3
X2 - X4
X3 - X4
c) Complatar 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	-0-	1,00			
X2	-0-		1,00		
X3	-0-			1,00	
X4	-0-				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).

as de valores. Los valores deben ser
o referencias que contengan números

6). Ilumine con el ratón la variable idependiente 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-	0,22	1,00		
X3	-0-			1,00	
X4	-0-				1,00

7.- Utilizando la funcion (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.

🔀 M	icrosoft I	Excel - Lib	ro1				×
Ar	rchivo <u>E</u> dicio	ón ⊻er Insei	rtar <u>F</u> ormato	Herramientas	Da <u>t</u> os Ve <u>n</u>	tana <u>?</u>	<u>N</u> ×
	F 🖬 🍯	🗟 🚏 🐰	B B &	🝼 🗠 🗖	⊒⇒ ⊒∞ Σ	f* 🚺 🖶	u » 1 ≠
A A	A N K	EEE	000 * 00 + 00	🆽 + 🖄 + 🕮	-		» *
0	G23 💌	=			Inserta	r filas	
	A	В	C	D	E	F	
22							
23							1
24		CTE,	X1	X2	X 3	X4	Ĩ
25	CTE.	1.00		()			
26	X1	-0-	1.00				
27	X2	-0-	0.22	1.00			
28	XJ	-0-	0.62	0.31	1.00		
29 🔛	X4	-0-	0.15	0.03	-0.02	1.00	
30							-
14 4	► ► GALF	ON / Hoja2 /	Hoja3 /	•		•	
Listo			<u> </u>			NUM	- //

Observese, que no hay problemas de Multicolinealida entre la Variables Independientes

DISTRIBUCION NORMAL ESTANDAR INVERSA (IDF)

Para calcular el coeficiente "k", que genera la Distibución 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 presenteda por Excel, la Probabilidad "p", bien como número, fracción o indicando la celda donde ya se haya calculado

DISTR.NORM.ESTAND.INV Probabilidad	💽 = número
	=
Devuelve el inverso de la distribución normal están desviación estándar de uno.	dar acumulativa. Tiene una media de cero y una
Probabilidad es una proba entre 0 y 1 i	abilidad asociada a la distribución normal, un número nclusive.
Resultado de la fórmula =	Aceptar Cancelar

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

NOTA IMPORTANTE:

de instalar el conjunto de r con los comandos:	nacros : "Análisis de Datos", des	de el CD de in	nstalación y proceder
<u>H</u> erramientas —	Aná <u>l</u> isis de datos]	Coeficiente de Correlación
una vez abierta la cája de o (sub-caja: "Rango de Entra la salida de la matriz (Sub-	liálogo, iluminar las Variables Ind Ida") e indicarle en que celda de l caja: "Rango de Salida)	lependientes (∣a hoja de cálc	(X) con el ratón culo se requiere
Esta macro, es especialme muchas variables indepen	nte útil cuando se va a analizar n dientes.	nodelos de reç	gresión con

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%
В	987	100	7	0,515%
С	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%
Н	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
0,284908723	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
0,304787497	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	
В	987	1.145	-158	
С	1.195	1.259	-64	
D	1.500	972	528	ATIPICO
E	975	976	-1	
F	705	878	-173	
G	872	1.005	-133	
Н	1.007	861	146	
	1.450	1.323	127	
J	1.332	1.156	176	ATIPICO
K	1.082	1.152	-70	
L	750	1.029	-279	ATIPICO

NUEVA SERIE

REFER.	P.UNIT	AREA	EDAD	% DE
	Y	M2 X1	X2	X3
A	1.245	75	1	0,942%
В	987	100	7	0,515%
С	1.195	80	3	0,363%
E	975	100	13	0,561%
F	705	125	16	0,874%
G	872	110	12	0,623%
Н	1.007	90	19	0,409%
-	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
0,940585712	0,063720358	#N/A	#N/A
31,66193639	6	#N/A	#N/A
0,385669365	0,024361704	#N/A	#N/A

94,06%	R^2 =
31,6619	F =
3	k =
10	n =
6	n - (k+1) =
4,76	Fo =
F >> Fo	

Modelo de Correlación:

Y = 2673,544 * 0,99131359^X1*0,99426175^X2*0,00000069683^X3

Rev.: Febrero-2011