ANÁLISIS DE REGRESIÓN EMPLEANDO EXCEL Y GRAPH

Los primeros y más importantes estudios al respecto se deben a los científicos Francis Galton (1822-1911) y Karl Pearson (1857-1936). Fue Galton quien utilizó por primera vez el término regresión para indicar que, aunque influida por la estatura de sus padres, la estatura de los hijos "regresaba" a la media general.

La regresión examina la relación entre dos variables, pero restringiendo una de ellas con el objeto de estudiar las variaciones de una variable cuando la otra permanece constante. En otras palabras, la regresión es un método que se emplea para predecir el valor de una variable en función de valores dados a la otra variable. En estadística la palabra predecir no se utiliza en el sentido empleado por los astrólogos, futurólogos y mentalistas, sino mas bien en un sentido lógico como es el de utilizar el conocimiento del comportamiento de una variable para obtener información sobre otra variable. Por ejemplo, puede predecirse el resultado que obtendrá un estudiante en su examen final, basados en el conocimiento de las calificaciones promedio de sus exámenes parciales, o predecir la preferencia de los estudiantes por profesiones científicas, conociendo los promedios de sus calificaciones en los estudios escolares.

En todos los casos de regresión existe una dependencia funcional entre las variables. En el caso de dos variables, siendo una de ellas (X) variable independiente y la otra (Y) la dependiente, se habla de regresión de Y sobre X; Por ejemplo, los ingenieros forestales utilizan la regresión de la altura de los árboles sobre su diámetro, lo cual significa que midiendo el diámetro (variable independiente) y reemplazando su valor en una relación definida según la clase de árbol se obtiene la altura, y aun sin necesidad de cálculos aprecian la altura utilizando gráficas de la función de dependencia, altura = función del diámetro.

1) LA RECTA DE LOS MÍNIMOS CUADRADOS

Se llama línea de mejor ajuste y se define como la línea que hace mínima la suma de los cuadrados de las desviaciones respecto a ella de todos los puntos que corresponden a la información recogida.

La recta de los mínimos cuadrados que aproxima el conjunto de puntos $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \dots \dots (X_N, Y_N)$ tomando en cuenta a Y como variable dependiente tiene por ecuación

$$Y = a_0 + a_1 X$$

A esta ecuación suele llamarse recta de regresión de Y sobre X, y se usa para estimar los valores de Y para valores dados de X

Si a la recta de regresión $Y = a_0 + a_1 X$ se le suma en ambos lados $\sum Y = \sum(a_0 + a_1 X)$ se obtiene $\sum Y = a_0 N + a_1 \sum X$

Si a la recta de regresión $Y = a_0 + a_1 X$ se multiplica por X a ambos lados y luego se suma $\sum X\, Y = \sum X(a_0 + a_1 X)$ se obtiene $\sum X\, Y = a_0 \sum X + a_1 \sum X^2$

Las constantes $a_0 \; y \; a_1$ quedan fijadas al resolver simultáneamente las ecuaciones anteriormente encontradas, es decir, al resolver el siguiente sistema de ecuaciones:

$$\begin{cases} \Sigma Y = a_0 N + a_1 \Sigma X \\ \Sigma XY = a_0 \Sigma X + a_1 \Sigma X^2 \end{cases}$$

Que se llaman las ecuaciones normales para la recta de mínimos cuadrados.

Las constantes a_0 y a_1 de las anteriores ecuaciones también se pueden calcular empleando las siguientes fórmulas:

$$a_0 = \frac{\sum Y \cdot \sum X^2 - \sum X \cdot \sum XY}{N\sum X^2 - (\sum X)^2} \qquad a_1 = \frac{N\sum XY - \sum X \cdot \sum Y}{N\sum X^2 - (\sum X)^2}$$

Otra ecuación para los mínimos cuadrados para $x = X - \bar{X}$, $y = Y - \bar{Y}$ de la recta de regresión de Y sobre X es:

$$y = \left(\frac{\sum xy}{\sum \mathrm{x}^2}\right)x$$

La recta de los mínimos cuadrados que aproxima el conjunto de puntos $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \ldots \ldots (X_N, Y_N)$ tomando en cuenta a X como variable dependiente tiene por ecuación:

$$X = b_0 + b_1 Y$$

A esta ecuación suele llamarse recta de regresión de X sobre Y, y se usa para estimar los valores de X para valores dados de Y. Las constantes b_0 y b_1 quedan fijadas al resolver el siguiente sistema de ecuaciones:

$$\begin{cases} \Sigma X = b_0 N + b_1 \Sigma Y \\ \Sigma XY = b_0 \Sigma Y + b_1 \Sigma Y^2 \end{cases}$$

Las constantes b_0 y b_1 del sistema de ecuaciones anterior se pueden calcular empleando las siguientes fórmulas:

$$b_0 = \frac{\sum X \cdot \sum Y^2 - \sum Y \cdot \sum XY}{N\sum Y^2 - (\sum Y)^2} \qquad b_1 = \frac{N\sum XY - \sum X \cdot \sum Y}{N\sum Y^2 - (\sum Y)^2}$$

Otra ecuación para los mínimos cuadrados para $x = X - \bar{X}$, $y = Y - \bar{Y}$ es:

$$x = \left(\frac{\sum xy}{\sum \mathrm{y}^2}\right)y$$

El punto de intersección entre las rectas $Y = a_0 + a_1 X$ con $X = b_0 + b_1 Y$ se simboliza $(\bar{X}, \bar{Y})$ y se llama centroide o centro de gravedad

Ejemplo ilustrativo

Con los datos de la siguiente tabla sobre la altura en centímetros (X) y los pesos en kilogramos (Y) de una muestra de 8 estudiantes varones tomada al azar del segundo semestre de una universidad.

X	152	157	162	167	173	178	182	188
Y	56	61	67	72	70	72	83	92

1) Ajustar la recta de mínimos cuadrados para Y como variable dependiente resolviendo el sistema:

$$\begin{cases} \Sigma Y = a_0 N + a_1 \Sigma X \\ \Sigma XY = a_0 \Sigma X + a_1 \Sigma X^2 \end{cases}$$

2) Ajustar la recta de mínimos cuadrados para Y como variable dependiente empleando las fórmulas:

$$a_0 = \frac{\sum Y \cdot \sum X^2 - \sum X \cdot \sum XY}{N \sum X^2 - (\sum X)^2} \qquad\qquad a_1 = \frac{N \sum XY - \sum X \cdot \sum Y}{N \sum X^2 - (\sum X)^2}$$

3) Ajustar la recta de mínimos cuadrados para Y como variable dependiente empleando la fórmula:

$$y = \left(\frac{\sum xy}{\sum \mathrm{x}^2}\right) x$$

4) Ajustar la recta de mínimos cuadrados para X como variable dependiente resolviendo el sistema:

$$\begin{cases} \Sigma X = b_0 N + b_1 \Sigma Y \\ \Sigma XY = b_0 \Sigma Y + b_1 \Sigma Y^2 \end{cases}$$

5) Calcular el punto centroide.

6) Calcular el coeficiente de determinación.

7) Elaborar el diagrama de dispersión. Y en el mismo diagrama graficar las dos rectas de mínimos cuadrados obtenidas en los pasos anteriores.

8) Estimar el valor de Y cuando X = 200 en el diagrama de dispersión de Y como variable dependiente.

R: 8,2

9) Estimar el valor de X cuando Y= 100 en el diagrama de dispersión X como variable dependiente.

Solución:

Para comenzar a resolver el ejercicio se llena la siguiente tabla:

X	Y	XY	X^2	Y^2
152	56	8512	23104	3136
157	61	9577	24649	3721
162	67	10854	26244	4489
167	72	12024	27889	5184
173	70	12110	29929	4900
178	72	12816	31684	5184
182	83	15106	33124	6889
188	92	17296	35344	8464
$\Sigma X = 1359$	$\Sigma Y = 573$	$\Sigma XY = 98295$	$\Sigma X^2 = 231967$	$\Sigma Y^2 = 41967$

1) Reemplazando valores en el sistema se tiene:

$$\begin{cases} \Sigma Y = a_0 N + a_1 \Sigma X \\ \Sigma XY = a_0 \Sigma X + a_1 \Sigma X^2 \end{cases}$$

$$\begin{cases} 573 = a_0 \cdot 8 + a_1 \cdot 1359 \\ 98295 = a_0 \cdot 1359 + a_1 \cdot 231967 \end{cases} \Rightarrow \begin{cases} 8a_0 + 1359a_1 = 573 \\ 1359a_0 + 231967a_1 = 98295 \end{cases}$$

Resolviendo el sistema por determinantes (regla de Cramer) se obtiene:

$$a_0 = \frac{\Delta a_0}{\Delta} = \frac{\begin{vmatrix} 573 & 1359 \\ 98295 & 231967 \end{vmatrix}}{\begin{vmatrix} 8 & 1359 \\ 1359 & 231967 \end{vmatrix}} = \frac{573 \cdot 231967 - 98295 \cdot 1359}{8 \cdot 231967 - 1359 \cdot 1359} = \frac{-665814}{8855} = -75{,}191$$

$$a_1 = \frac{\Delta a_1}{\Delta} = \frac{\begin{vmatrix} 8 & 573 \\ 1359 & 98295 \end{vmatrix}}{8855} = \frac{8 \cdot 98295 - 1359 \cdot 573}{8855} = \frac{7653}{8855} = 0{,}864$$

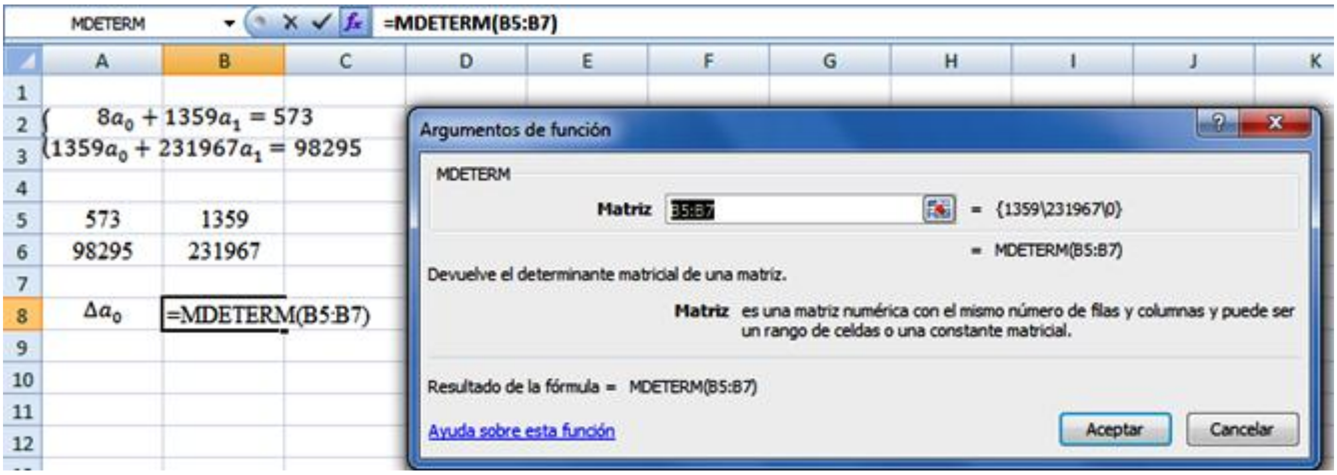
Interpretación:

- El valor $a_1 = 0{,}864$ indica que la recta tiene una pendiente positiva aumentando a razón de 0,864

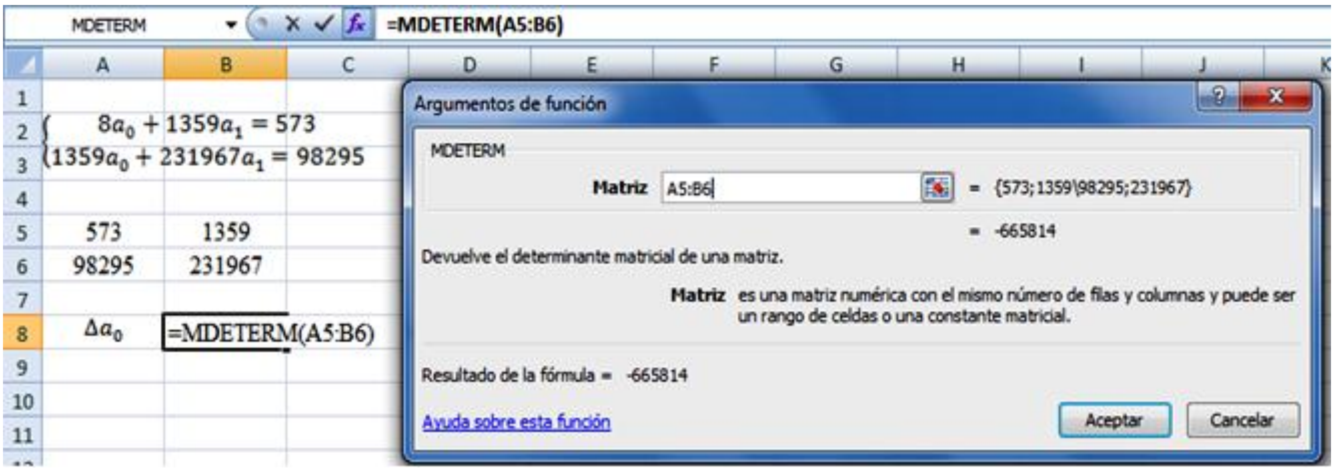
- El valor de $a_0 = -75{,}191$ indica el punto en donde la recta interseca al eje Y cuanto X = 0

En Excel el sistema se resuelve de la siguiente manera:

a) Insertar la función MDETERM.



b) Seleccionar las celdas del Δa_0.



c) Pulsar en Aceptar para culminar con el cálculo del Δa_0.

	A	B	C	D
1				
2	$\begin{cases} 8a_0 + 1359a_1 = 573 \\ 1359a_0 + 231967a_1 = 98295 \end{cases}$			
3				
4				
5	573	1359		
6	98295	231967		
7				
8	Δa_0	-665814	=MDETERM(A5:B6)	

d) Repetir los pasos anteriores para calcular el Δ y el Δa_1, para luego calcular a_0 y a_1 como indica la siguiente figura:

	A	B	C	D	E	F	G	H
1								
2	$\begin{cases} 8a_0 + 1359a_1 = 573 \\ 1359a_0 + 231967a_1 = 98295 \end{cases}$							
3								
4								
5	573	1359						
6	98295	231967						
7								
8	Δa_0	-665814	=MDETERM(A5:B6)					
9								
10						$a_0 = \frac{\Delta a_0}{\Delta}$	-75,19074	=B8/B14
11	8	1359						
12	1359	231967						
13								
14	Δ	8855	=MDETERM(A11:B12)			$a_1 = \frac{\Delta a_1}{\Delta}$	0,8642575	=B20/B14
15								
16								
17	8	573						
18	1359	98295						
19								
20	Δb_0	7653	=MDETERM(A17:B18)					

Reemplazando valores en la ecuación respectiva se obtiene:

$Y = a_0 + a_1 X \Rightarrow Y = -75{,}191 + 0{,}864X$

2) Con los datos de la tabla anterior se substituye valores en las siguientes ecuaciones:

$$a_0 = \frac{\sum Y \cdot \sum X^2 - \sum X \cdot \sum XY}{N \sum X^2 - (\sum X)^2} = \frac{573 \cdot 231967 - 1359 \cdot 98295}{8 \cdot 231967 - (1359)^2} = \frac{-665814}{8855} - 75{,}191$$

$$a_1 = \frac{N \sum XY - \sum X \cdot \sum Y}{N \sum X^2 - (\sum X)^2} = \frac{8 \cdot 98295 - 1359 \cdot 573}{8 \cdot 231967 - (1359)^2} = \frac{7653}{8855} = 0{,}864$$

Reemplazando valores en la ecuación respectiva se obtiene:

$Y = a_0 + a_1 X \Rightarrow Y = -75{,}191 + 0{,}864X$

3) Se calcula las medias aritméticas de X y Y para llenar la siguiente tabla:

$$\bar{x} = \frac{\sum x_i}{n}$$
$$\bar{X} = \frac{1359}{8} = 169{,}875$$

$$\bar{Y} = \frac{573}{8} = 71{,}625$$

X	Y	$x = X - \bar{X}$	$y = Y - \bar{Y}$	xy	x^2	y^2
152	56	-17,88	-15,625	279,297	319,516	244,141
157	61	-12,88	-10,625	136,797	165,766	112,891
162	67	-7,875	-4,625	36,422	62,016	21,391
167	72	-2,875	0,375	-1,078	8,266	0,141
173	70	3,125	-1,625	-5,078	9,766	2,641
178	72	8,125	0,375	3,047	66,016	0,141
182	83	12,125	11,375	137,922	147,016	129,391
188	92	18,125	20,375	369,297	328,516	415,141
$\Sigma X = 1359$	$\Sigma Y = 573$			$\Sigma xy = 956{,}625$	$\Sigma x^2 = 1106{,}875$	$\Sigma y^2 = 925{,}875$

Reemplazando valores en la fórmula respectiva se obtiene:

$$y = \left(\frac{\sum xy}{\sum \mathrm{x}^2}\right) x \Rightarrow y = \frac{956{,}625}{1106{,}875} x \Rightarrow Y - \bar{Y} = \frac{956{,}625}{1106{,}875} (X - \bar{X})$$

$$Y - 71{,}625 = \frac{956{,}625}{1106{,}875} (X - 169{,}875) \Rightarrow 1106{,}875(Y - 71{,}625) = 956{,}625(X - 169{,}875)$$

$$1106{,}875Y - 79280{,}20838 = 956{,}625X - 162510{,}4984$$

$$1106{,}875Y = 956{,}625X - 162510{,}4984 + 79280{,}20838$$

$$1106{,}875Y = 956{,}625X - 83230{,}29$$

$$Y = \frac{956{,}625X - 83230{,}29}{1106{,}875} \Rightarrow Y = \frac{956{,}625X}{1106{,}875} - \frac{83230{,}29}{1106{,}875} \Rightarrow Y = 0{,}864X - 75{,}19$$

$$Y = -75{,}19 + 0{,}864X$$

4) Reemplazando valores en sistema respectivo se obtiene:

$$\begin{cases} \Sigma X = b_0 N + b_1 \Sigma Y \\ \Sigma XY = b_0 \Sigma Y + b_1 \Sigma Y^2 \end{cases}$$

$$\begin{cases} 1359 = b_0 \cdot 8 + b_1 \cdot 573 \\ 98295 = b_0 \cdot 573 + b_1 \cdot 41967 \end{cases} \Rightarrow \begin{cases} 8b_0 + 573b_1 = 1359 \\ 573b_0 + 41967b_1 = 98295 \end{cases}$$

Resolviendo el sistema se obtiene:
$b_0 = 95{,}871$
$b_1 = 1{,}033$

Reemplazando valores en la ecuación de la recta de mínimos cuadrados se obtiene:

$X = b_0 + b_1 Y$

$X = 95{,}871 + 1{,}033Y$

Interpretación:

- El valor $b_1 = 1{,}033$ indica que la recta tiene una pendiente positiva aumentando a razón de 1,033

- El valor de $b_0 = 95{,}871$ indica el punto en donde la recta interseca al eje X cuanto Y = 0

5) Para calcular el centroide $(\bar{X}, \bar{Y})$ se resuelve el sistema formado por las dos rectas de los mínimos cuadrados en donde X es $\bar{X}$ y Y es $\bar{Y}$.

$$\begin{cases} Y = -75{,}191 + 0{,}864X \\ X = 95{,}871 + 1{,}033Y \end{cases}$$

Al resolver el sistema se obtiene el centroide: X = 169,3 y Y = 71,092

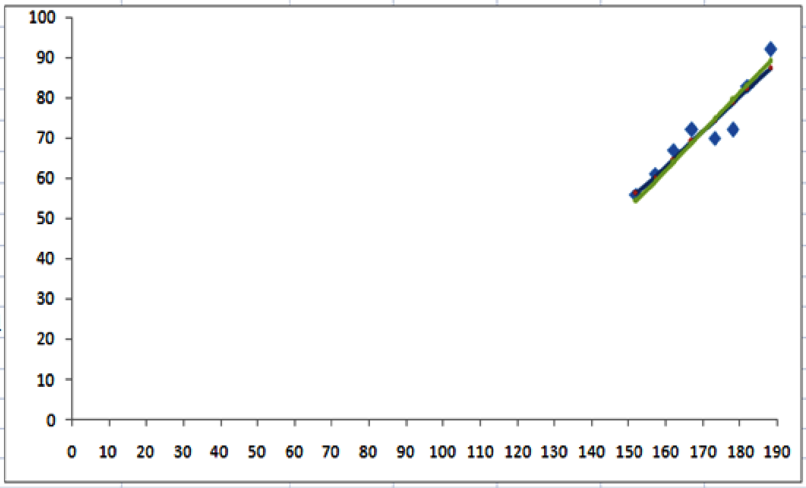
6) Se aplica la ecuación para calcular el coeficiente de Pearson.

$$r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}} = \frac{8 \cdot 98295 - 1359 \cdot 573}{\sqrt{[8 \cdot 231967 - (1359)^2][8 \cdot 41967 - (573)^2]}}$$
$r = 0{,}94497$

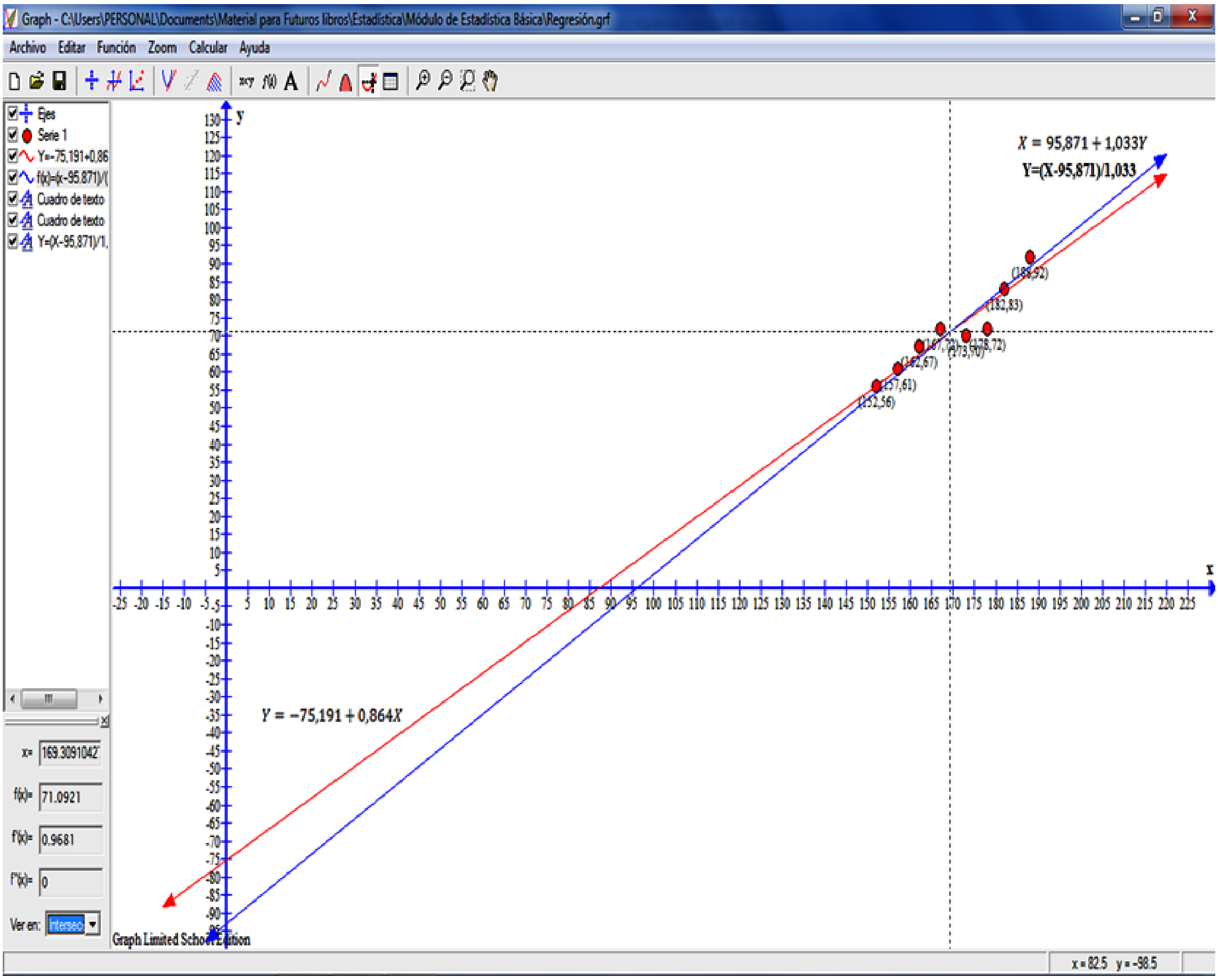
Elevando al cuadrado coeficiente de Pearson queda calculado el coeficiente de determinación.
Coeficiente de determinación = $r^2 = (0{,}94497)^2 = 0{,}893$

7) *En Excel, insertando gráfico de dispersión se obtiene la siguiente figura:*

	A	B	C	D	E
1	X	Y	Y=-75,191+0,864X	y=(X- 95,871)/1,033	
2	152	56	56,137	54,336	
3	157	61	60,457	59,176	
4	162	67	64,777	64,016	
5	167	72	69,097	68,857	
6	173	70	74,281	74,665	
7	178	72	78,601	79,505	
8	182	83	82,057	83,378	
9	188	92	87,241	89,186	
10					
11		a_0	-75,191	b_0	95,871
12		a_1	0,864	b_1	1,033
13					
14					
15					
16					



Empleando el programa Graph se obtiene la siguiente figura:



8) Reemplazando X = 200 en la ecuación solicitada se obtiene:

$Y = -75{,}191 + 0{,}864X = -75{,}191 + 0{,}864 \cdot 200 = -75{,}191 + 172{,}8 = 97{,}609$

9) Reemplazando Y = 100 en la ecuación solicitada se obtiene:

$X = 95{,}871 + 1{,}033Y = X = 95{,}871 + 1{,}033 \cdot 100 = X = 95{,}871 + 103{,}3 = 199{,}171$

TAREA DE INTERAPRENDIZAJE

1) Consulte sobre la biografía de Francis Galton y de Cramer, y realice un organizador gráfico de cada una.

2) Dada la siguiente tabla sobre la altura en centímetros (X) y los pesos en kilogramos (Y) de una muestra de 8 estudiantes varones tomada al azar del segundo semestre de una universidad.

X	150	155	160	165	170	175	180	185
Y	55	60	63	67	70	74	79	85

2.1) Ajuste la recta de mínimos cuadrados para Y como variable dependiente resolviendo el sistema. El sistema resuelva de manera manual y mediante Excel.

$$\begin{cases} \Sigma Y = a_0 N + a_1 \Sigma X \\ \Sigma XY = a_0 \Sigma X + a_1 \Sigma X^2 \end{cases}$$

$$Y = -66{,}869 + 0{,}812X$$

2.2) Ajuste la recta de mínimos cuadrados para Y como variable dependiente empleando las fórmulas

$$a_0 = \frac{\sum Y \cdot \sum X^2 - \sum X \cdot \sum XY}{N \sum X^2 - (\sum X)^2} \qquad a_1 = \frac{N \sum XY - \sum X \cdot \sum Y}{N \sum X^2 - (\sum X)^2}$$

$$a_0 = -66{,}869\ ; \quad a_1 = 0{,}812$$

2.3) Ajuste la recta de mínimos cuadrados para Y como variable dependiente empleando la fórmula

$$y = \left(\frac{\sum xy}{\sum \mathrm{x}^2}\right) x$$

$$Y = -66{,}869 + 0{,}812X$$

2.4) Ajuste la recta de mínimos cuadrados para X como variable dependiente resolviendo el sistema. El sistema resuelva de manera manual y mediante Excel.

$$\begin{cases} \Sigma X = b_0 N + b_1 \Sigma Y \\ \Sigma XY = b_0 \Sigma Y + b_1 \Sigma Y^2 \end{cases}$$

$$X = 83{,}18 + 1{,}22Y$$

2.5) Ajuste la recta de mínimos cuadrados para X como variable dependiente empleando las fórmulas

$$b_0 = \frac{\sum X \cdot \sum Y^2 - \sum Y \cdot \sum XY}{N \sum Y^2 - (\sum Y)^2} \qquad b_1 = \frac{N \sum XY - \sum X \cdot \sum Y}{N \sum Y^2 - (\sum Y)^2}$$

$$b_0 = 83.18\ ; \quad b_1 = 1{,}22$$

2.6) Ajuste la recta de mínimos cuadrados para X como variable dependiente empleando la fórmula

$$x = \left(\frac{\sum xy}{\sum \mathrm{y}^2}\right) y$$

$$X = 83{,}18 + 1{,}22Y$$

2.7) Calcule el punto centroide de manera manual y empleando Excel.

$$\bar{X} = 170{,}9\ ; \ \bar{Y} = 71{,}9$$

2.8) Calcule el coeficiente de determinación.

0,99

2.9) Elabore el diagrama de dispersión. Y en el mismo diagrama graficar las dos rectas de mínimos cuadrados obtenidas en los pasos anteriores. Elabore de manera manual, empleando Excel y el programa Graph.

2.10) Estime el valor de Y cuando X = 173 en el diagrama de dispersión de Y como variable dependiente.

73,6

2.11) Estime el valor de X cuando Y = 73 en el diagrama de dispersión de Y como variable dependiente.

172,2

3) Cree y resuelva un ejercicio similar al anterior con datos obtenidos de 10 amigas suyas.

2) LA PARÁBOLA DE LOS MÍNIMOS CUADRADOS

La parábola de mínimos cuadrados que aproxima el conjunto de puntos $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \ldots (Y_N, Y_N)$ tiene ecuación dada por $Y = a_0 + a_1X + a_2X^2$, donde las constantes a_0, a_1 y a_2 se determinan al resolver simultáneamente el sistema de ecuaciones que se forma al multiplicar la ecuación $Y = a_0 + a_1X + a_2X^2$ por $1, X, Y$ sucesivamente, y sumando después.

$$\begin{cases} \Sigma Y = a_0N + a_1\Sigma X + a_2\Sigma X^2 \\ \Sigma XY = a_0\Sigma X + a_1\Sigma X^2 + a_2\Sigma X^3 \\ \Sigma X^2Y = a_0\Sigma X^2 + a_1\Sigma X^3 + a_2\Sigma X^4 \end{cases}$$

Ejemplo ilustrativo

La siguiente tabla muestra la población de un país en los años 1960-2010 en intervalos de 5 años.

Año	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010
Población (millones)	4,52	5,18	6,25	7,42	8,16	9,12	10,92	11,62	12,68	13,12	13,97

1) Ajustar una parábola de mínimos cuadrados de la forma $Y = a_0 + a_1X + a_2X^2$

2) Calcular los valores de tendencia para los años dados.

3) Estimar la población para los años 2015 y 2020.

4) Calcular el coeficiente de determinación.

5) Elaborar un diagrama de dispersión, y en el mismo diagrama graficar la parábola de los mínimos cuadrados.

Nota: Se recomienda codificar o cambiar la numeración de los años, eligiendo X de modo que el año central, 1985, corresponda a X= 0, para que se hagan más fáciles los cálculos.

Solución:

1) Para ajustar una parábola de mínimos cuadrados se llena la siguiente tabla:

Año	X	Y	X^2	X^3	X^4	XY	X^2Y
1960	-5	4,52	25	-125	625	-22,6	113
1965	-4	5,18	16	-64	256	-20,72	82,88
1970	-3	6,25	9	-27	81	-18,75	56,25
1975	-2	7,42	4	-8	16	-14,84	29,68
1980	-1	8,16	1	-1	1	-8,16	8,16
1985	0	9,12	0	0	0	0	0
1990	1	10,92	1	1	1	10,92	10,92
1995	2	11,62	4	8	16	23,24	46,48
2000	3	12,68	9	27	81	38,04	114,12
2005	4	13,12	16	64	256	52,48	209,92
2010	5	13,97	25	125	625	69,85	349,25
Σ	0	102,96	110	0	1958	109,46	1020,66

Se reemplaza valores en el sistema y se obtiene:

$$\begin{cases} \Sigma Y = a_0 N + a_1 \Sigma X + a_2 \Sigma X^2 \\ \Sigma XY = a_0 \Sigma X + a_1 \Sigma X^2 + a_2 \Sigma X^3 \\ \Sigma X^2 Y = a_0 \Sigma X^2 + a_1 \Sigma X^3 + a_2 \Sigma X^4 \end{cases}$$

$$\begin{cases} 102{,}96 = \mathrm{a}_0 \cdot 11 + \mathrm{a}_1 \cdot 0 + \mathrm{a}_2 \cdot 110 \\ 109{,}46 = \mathrm{a}_0 \cdot 0 + \mathrm{a}_1 \cdot 110 + \mathrm{a}_2 \cdot 0 \\ 1020{,}66 = \mathrm{a}_0 \cdot 110 + \mathrm{a}_1 \cdot 0 + \mathrm{a}_2 \cdot 1958 \end{cases} \Rightarrow \begin{cases} 11\mathrm{a}_0 + 0\mathrm{a}_1 + 110\mathrm{a}_2 = 102{,}96 \\ 0\mathrm{a}_0 + 110\mathrm{a}_1 + 0\mathrm{a}_2 = 109{,}46 \\ 110\mathrm{a}_0 + 0\mathrm{a}_1 + 1958\mathrm{a}_2 = 1020{,}66 \end{cases}$$

Resolviendo el sistema empleando determinantes (regla de Cramer) se obtiene:

$$a_0 = \frac{\Delta a_0}{\Delta} = \frac{\begin{vmatrix} 102{,}96 & 0 & 110 \\ 109{,}46 & 110 & 0 \\ 1020{,}66 & 0 & 1958 \end{vmatrix}}{\begin{vmatrix} 11 & 0 & 110 \\ 0 & 110 & 0 \\ 110 & 0 & 1958 \end{vmatrix}} = \frac{\begin{vmatrix} 102{,}96 & 0 & 110 \\ 109{,}46 & 110 & 0 \\ 1020{,}66 & 0 & 1958 \\ 102{,}96 & 0 & 110 \\ 109{,}46 & 110 & 0 \end{vmatrix}}{\begin{vmatrix} 11 & 0 & 110 \\ 0 & 110 & 0 \\ 110 & 0 & 1958 \\ 11 & 0 & 110 \\ 0 & 110 & 0 \end{vmatrix}}$$

$$\mathrm{a}_0 = \frac{22175524{,}8 + 0 + 0 - 12349986 - 0 - 0}{2369180 + 0 + 0 - 1331000 - 0 - 0} = \frac{9825538{,}8}{1038180} = 9{,}464$$

$$a_1 = \frac{\Delta a_1}{\Delta} = \frac{\begin{vmatrix} 11 & 102{,}96 & 110 \\ 0 & 109{,}46 & 0 \\ 110 & 1020{,}66 & 1958 \end{vmatrix}}{1038180} = \frac{\begin{vmatrix} 11 & 102{,}96 & 110 \\ 0 & 109{,}46 & 0 \\ 110 & 1020{,}66 & 1958 \\ 11 & 102{,}96 & 110 \\ 0 & 109{,}46 & 0 \end{vmatrix}}{1038180}$$

$$a_1 = \frac{23577549{,}48 + 0 + 0 - 1324466 - 0 - 0}{1038180} = \frac{2357549{,}48}{1038180} = 0{,}995$$

$$a_2 = \frac{\Delta a_2}{\Delta} = \frac{\begin{vmatrix} 11 & 0 & 102{,}96 \\ 0 & 110 & 109{,}46 \\ 110 & 0 & 1020{,}66 \end{vmatrix}}{1038180} = \frac{\begin{vmatrix} 11 & 0 & 102{,}96 \\ 0 & 110 & 109{,}46 \\ 110 & 0 & 1020{,}66 \\ 11 & 0 & 102{,}96 \\ 0 & 110 & 109{,}46 \end{vmatrix}}{1038180}$$

$$a_2 = \frac{1234998{,}6 + 0 + 0 - 1245816 - 0 - 0}{1038180} = \frac{-10817{,}4}{1038180} = -0{,}01$$

El sistema resuelto en Excel se muestra en la siguiente figura:

	A	B	C	D	E	F	G	H
16						a_0	a_1	a_2
17	$11a_0 + 0a_1 + 110a_2 = 102{,}96$				102,96	11	0	110
18	$0a_0 + 110a_1 + 0a_2 = 109{,}46$				109,46	0	110	0
19	$110a_0 + 0a_1 + 1958a_2 = 1020{,}66$				1020,66	110	0	1958
20								
21	102,96	0	110					
22	109,46	110	0		$a_0 = \frac{\Delta a_0}{\Delta}$	9,464	=B24/B40	
23	1020,66	0	1958					
24	Δa_0	9825538,8	=MDETERM(A21:C23)					
25								
26								
27	11	102,96	110					
28	0	109,46	0		$a_1 = \frac{\Delta a_1}{\Delta}$	0,995	=B30/B40	
29	110	1020,66	1958					
30	Δa_1	1033083,5	=MDETERM(A27:C29)					
31								
32	11	0	102,96					
33	0	110	109,46					
34	110	0	1020,66					
35	Δa_2	-10817,4	=MDETERM(A32:C34)					
36								
37	11	0	110					
38	0	110	0		$a_2 = \frac{\Delta a_2}{\Delta}$	-0,010	=B35/B40	
39	110	0	1958					
40	Δ	1038180	=MDETERM(A37:C39)					

Reemplazando los valores encontrados se obtiene la ecuación de la parábola de mínimos cuadrados:

$Y = a_0 + a_1X + a_2X^2 \Rightarrow Y = 9{,}464 + 0{,}995X - 0{,}01X^2$

2) Los valores de tendencia se obtienen al reemplazar los valores de X en la ecuación de la parábola de mínimos cuadrados, los cuales se presenta en la siguiente tabla:

Año	X	Y	Valores de tendencia $Y = 9{,}464 + 0{,}995X - 0{,}01X^2$
1960	-5	4,52	4,24
1965	-4	5,18	5,32
1970	-3	6,25	6,39
1975	-2	7,42	7,43
1980	-1	8,16	8,46
1985	0	9,12	9,46
1990	1	10,92	10,45
1995	2	11,62	11,41
2000	3	12,68	12,36
2005	4	13,12	13,28
2010	5	13,97	14,19

3) Para estimar la población de los años 2015 y 2020 se transforma estos años a X siguiendo la secuencia de la tabla anterior, siendo X = 6 para el año 2015 y X= 7 para el 2020

Entonces para el 2015 se tiene:
$Y = 9{,}464 + 0{,}995X - 0{,}01X^2 = 9{,}464 + 0{,}995(6) - 0{,}01(6)^2 = 9{,}464 + 5{,}97 - 0{,}36 = 15{,}074$

Para el 2020 se tiene:
$Y = 9{,}464 + 0{,}995X - 0{,}01X^2 = 9{,}464 + 0{,}995(7) - 0{,}01(7)^2 = 9{,}464 + 6{,}965 - 0{,}49 = 15{,}939$

4) Se llena la siguiente tabla y se aplica la ecuación para calcular el coeficiente de Pearson

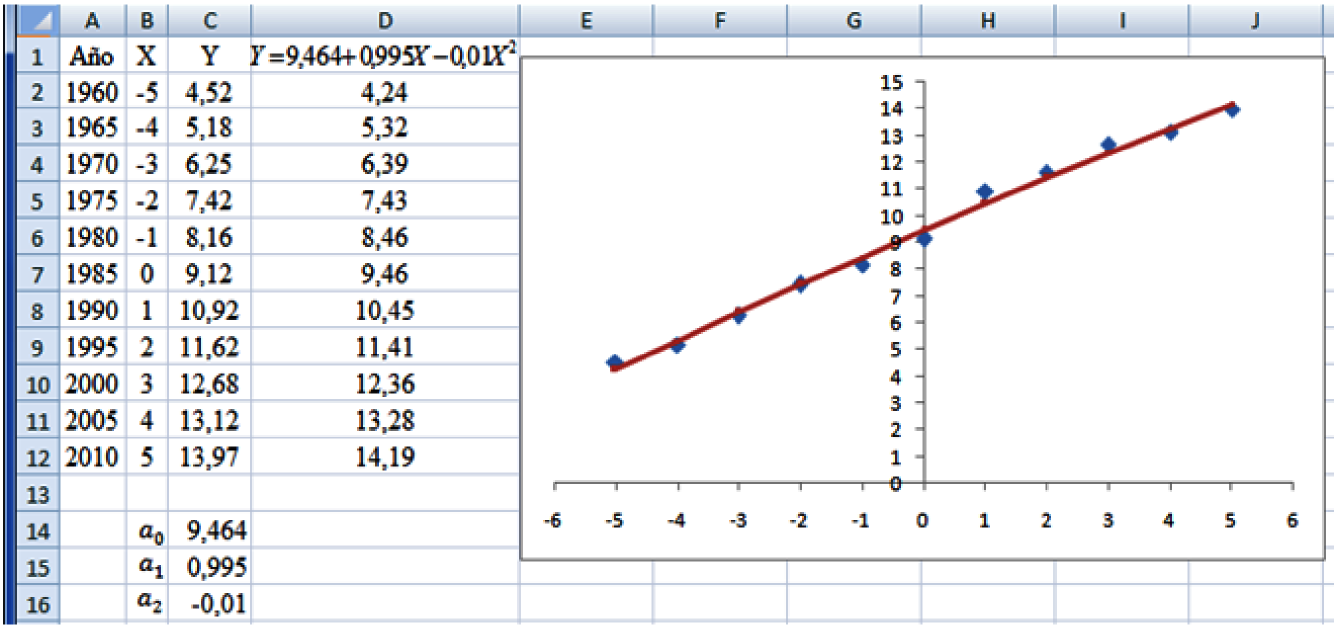
Año	X	Y	X^2	XY	Y^2
1960	-5	4,52	25	-22,6	20,430
1965	-4	5,18	16	-20,72	26,832
1970	-3	6,25	9	-18,75	39,063
1975	-2	7,42	4	-14,84	55,056
1980	-1	8,16	1	-8,16	66,586
1985	0	9,12	0	0	83,174
1990	1	10,92	1	10,92	119,246
1995	2	11,62	4	23,24	135,024
2000	3	12,68	9	38,04	160,782
2005	4	13,12	16	52,48	172,134
2010	5	13,97	25	69,85	195,161
Σ	0	102,96	110	109,46	1073,490

$$r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}} = \frac{11 \cdot 109{,}46 - 0 \cdot 102{,}96}{\sqrt{[11 \cdot 110 - (0)^2][11 \cdot 1073{,}490 - (102{,}96)^2]}}$$
$r = 0{,}996$

Elevando al cuadrado coeficiente de Pearson queda calculado el coeficiente de determinación.
Coeficiente de determinación = $r^2 = (0{,}996)^2 = 0{,}992$
El coeficiente de determinación calculado en Excel se muestra en la siguiente figura:

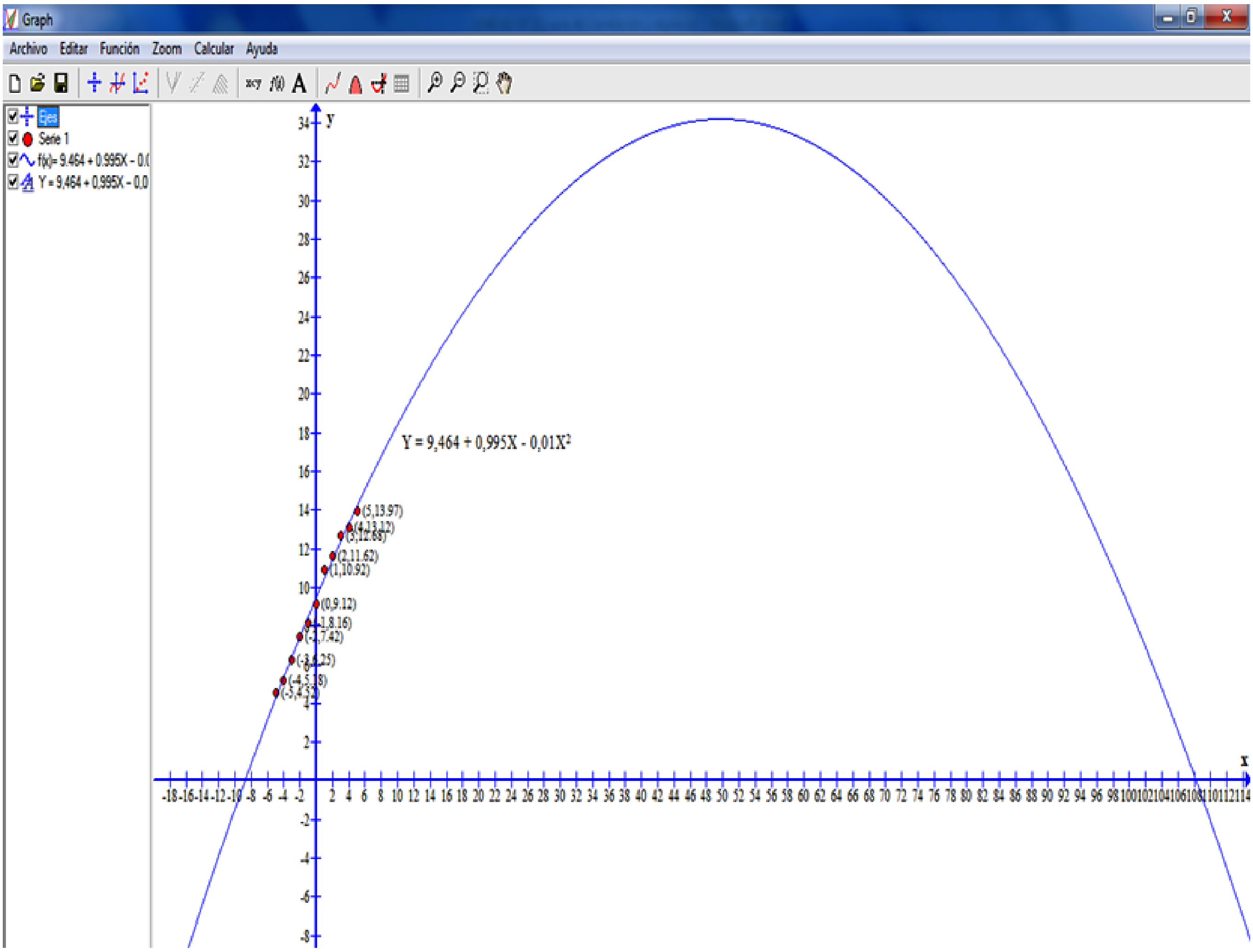
	A	B	C	D	E
1	**Año**	**X**	**Y**		
2	1960	-5	4,52		
3	1965	-4	5,18		
4	1970	-3	6,25		
5	1975	-2	7,42		
6	1980	-1	8,16		
7	1985	0	9,12		
8	1990	1	10,92		
9	1995	2	11,62		
10	2000	3	12,68		
11	2005	4	13,12		
12	2010	5	13,97		
13					
14	r	0,9960666	=COEF.DE.CORREL(B2:B12;C2:C12)		
15		0,9960666	=COEF.DE.CORREL(A2:A12;C2:C12)		
16	r^2	0,9921487	=E15^2		
17		0,9921487	=COEFICIENTE.R2(C2:C12;B2:B12)		

5) *El diagrama de dispersión y la parábola de los mínimos cuadrados mediante Excel se muestra en la siguiente figura:*



Año	X	Y	$Y=9{,}464+0{,}995X-0{,}01X^2$
1960	-5	4,52	4,24
1965	-4	5,18	5,32
1970	-3	6,25	6,39
1975	-2	7,42	7,43
1980	-1	8,16	8,46
1985	0	9,12	9,46
1990	1	10,92	10,45
1995	2	11,62	11,41
2000	3	12,68	12,36
2005	4	13,12	13,28
2010	5	13,97	14,19
	a_0	9,464	
	a_1	0,995	
	a_2	-0,01	

Empleando el programa Graph se obtiene la siguiente figura:



TAREA DE INTERAPRENDIZAJE

1) La siguiente tabla muestra la población aproximada de la Provincia de Imbabura en los años 1960-2010 en intervalos de 5 años.

Año	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010
Población (miles)	123	140	170	201	221	247	296	315	344	356	379

1.1) Ajuste una parábola de mínimos cuadrados de la forma $Y = a_0 + a_1X + a_2X^2$

$Y = 256{,}464 + 26{,}991X - 0{,}265X^2$

1.2) Calcule los valores de tendencia para los años dados de manera manual y empleando Excel.

Año	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010
Valor de tendencia	114,88	144,26	173,11	201,42	229,21	256,46	283,19	309,39	335,05	360,19	384,79

1.3) Estime la población para los años 2015 y 2020

Año 2015 = 408,87 miles de habitantes
Año 2020 = 432,42 miles de habitantes

1.4) Calcule el coeficiente de determinación de manera manual y empleando Excel.

0,992

1.5) Elabore un diagrama de dispersión, y en el mismo diagrama graficar la parábola de los mínimos cuadrados de manera manual, empleando Excel y empleando Graph.

2) Cree y resuelva un ejercicio de aplicación de la parábola de los mínimos cuadrados con datos de la población del Ecuador o de cualquier otro país de manera manual, empleando Excel y Graph.

3) REGRESIÓN EXPONENCIAL

Cuando la curva de regresión de y sobre x es exponencial, es decir para cualquier x considerada, la media de la distribución está dada por la siguiente ecuación predictora:

$$Y = \alpha \cdot \beta^X$$

Tomando logaritmos en ambos miembros:

$$logY = log\alpha + X \cdot log\beta$$

Y se puede estimar ahora log Y y log β, y de ahí obtener α y β, aplicando *los métodos de los mínimos cuadrados*.

Donde las constantes α y β quedan fijadas al resolver simultáneamente las ecuaciones:

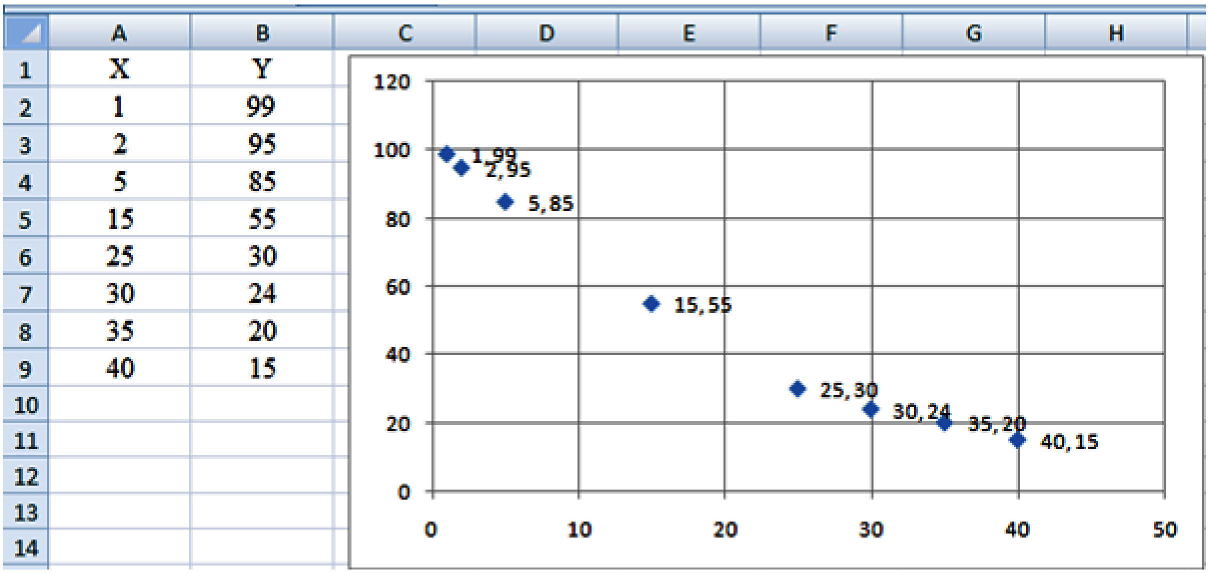
$$\begin{cases} \Sigma logY = log\alpha \cdot N + log\beta \cdot \Sigma \mathrm{X} \\ \Sigma \mathrm{X} \cdot \mathrm{logY} = \mathrm{log\alpha} \cdot \Sigma \mathrm{X} + \mathrm{log\beta} \cdot \Sigma X^2 \end{cases}$$

Ejemplo ilustrativo: Las cifras siguientes son datos sobre el porcentaje de llantas radiales producidas por cierto fabricante que aún pueden usarse después de recorrer cierto número de millas:

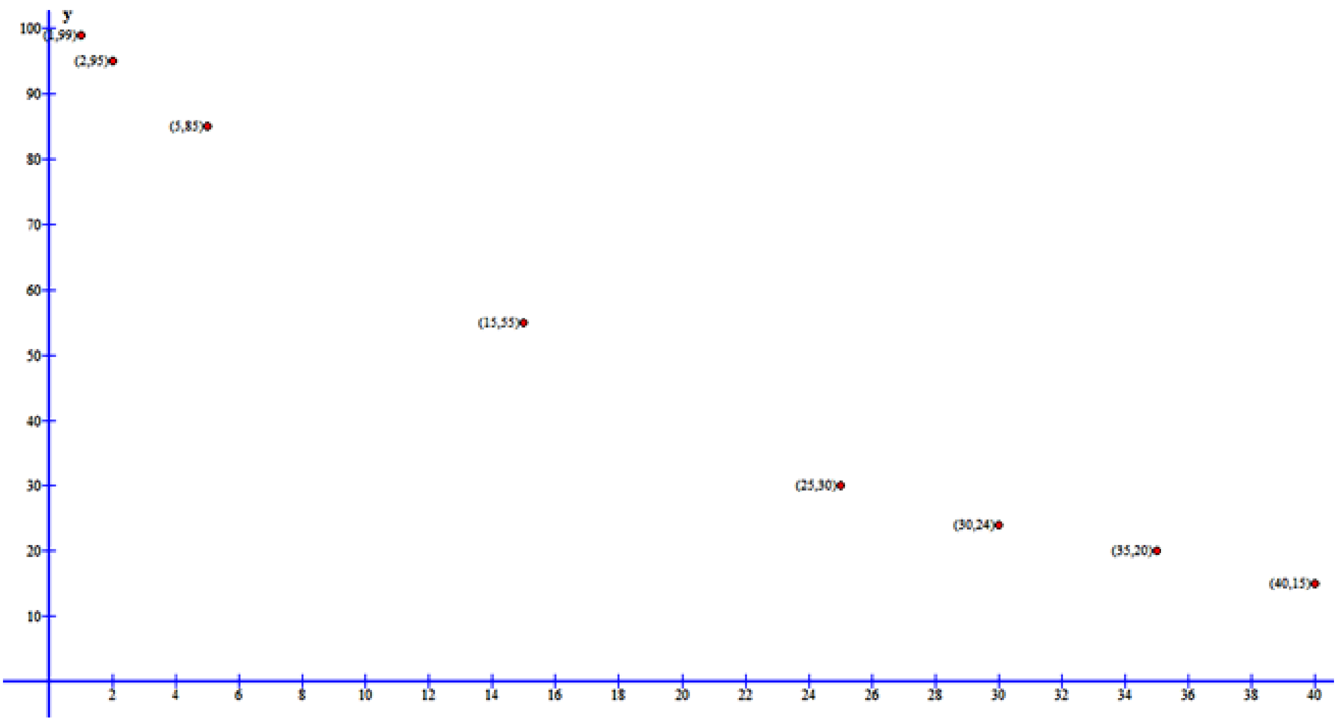
Miles de Millas recorridas (X)	1	2	5	15	25	30	35	40
Porcentaje útil (Y)	99	95	85	55	30	24	20	15

1) Elaborar el diagrama de dispersión.
2) Ajustar una curva exponencial aplicando el método de mínimos cuadrados.
3) Calcular la ecuación predictora.
4) Graficar la ecuación predictora.
5) Estimar qué porcentaje de las llantas radiales del fabricante durarán 50000 millas.

Solución:
1) *Elaborando el diagrama de dispersión empleando Excel se obtiene la siguiente figura:*



Empleando el programa Graph se obtiene la siguiente figura:



2) Se llena la siguiente tabla:

X	Y	$logY$	X^2	$X \cdot logY$
1	99	1,996	1	1,996
2	95	1,978	4	3,955
5	85	1,929	25	9,647
15	55	1,740	225	26,105
25	30	1,477	625	36,928
30	24	1,380	900	41,406
35	20	1,301	1225	45,536
40	15	1,176	1600	47,044
$\Sigma X = 153$		Σ log Y=12,97759	Σ X^2 = 4605	Σ X· logY = 212,61769

Resolviendo empleando Excel se muestra en la siguiente figura:

	A	B	C	D	E	F	G	H	I
1		X	Y	log Y		X^2		X logY	
2		1	99	1,99564	=LOG10(C2)	1	=B2^2	1,99564	=B2*D2
3		2	95	1,97772	=LOG10(C3)	4	=B3^2	3,95545	=B3*D3
4		5	85	1,92942	=LOG10(C4)	25	=B4^2	9,64709	=B4*D4
5		15	55	1,74036	=LOG10(C5)	225	=B5^2	26,10544	=B5*D5
6		25	30	1,47712	=LOG10(C6)	625	=B6^2	36,92803	=B6*D6
7		30	24	1,38021	=LOG10(C7)	900	=B7^2	41,40634	=B7*D7
8		35	20	1,30103	=LOG10(C8)	1225	=B8^2	45,53605	=B8*D8
9		40	15	1,17609	=LOG10(C9)	1600	=B9^2	47,04365	=B9*D9
10	Σ	153		12,97759		4605		212,61769	
11				=SUMA(D2:D9)		=SUMA(E2:E9)		=SUMA(F2:F9)	
12	N	8	=CONTAR(A2:A9)						

Reemplazando valores en el sistema se obtiene:

$$\begin{cases} \Sigma logY = log\alpha \cdot N + log\beta \cdot \Sigma \mathrm{X} \\ \Sigma \mathrm{X} \cdot \log\mathrm{Y} = \log\alpha \cdot \Sigma \mathrm{X} + \log\beta \cdot \Sigma X^2 \end{cases}$$

$$\begin{cases} 12{,}97759 = \log\alpha \cdot 8 + \log\beta \cdot 153 \\ 212{,}61769 = \log\alpha \cdot 153 + \log\beta \cdot 4605 \end{cases} \Rightarrow \begin{cases} 8\log\alpha + 153\log\beta = 12{,}97759 \\ 153\log\alpha + 4605\log\beta = 212{,}61769 \end{cases}$$

Al resolver el sistema se obtiene:

$$log\alpha = \frac{\Delta_\alpha}{\Delta} = \frac{\begin{vmatrix} 12{,}97759 & 153 \\ 212{,}61769 & 4605 \end{vmatrix}}{\begin{vmatrix} 8 & 153 \\ 153 & 4605 \end{vmatrix}} = \frac{59761{,}80195 - 32530{,}50657}{36840 - 23409} = \frac{27231{,}295389}{13431} = 2{,}027495747$$

$$log\beta = \frac{\Delta_\beta}{\Delta} = \frac{\begin{vmatrix} 8 & 12{,}97759 \\ 153 & 212{,}61769 \end{vmatrix}}{13431} = \frac{1700{,}944152 - 1985{,}57127}{13431} = \frac{-284{,}627118}{13431} = -0{,}02119180389$$

Reemplazando valores se obtiene:

$$logY = log\alpha + X \cdot log\beta$$

$$logY = 2{,}027496 - 0{,}02119X$$

Aplicando el antilogaritmo se obtiene:

$$\alpha = \text{anti log } 2{,}027495747 = 106{,}536$$

$$\beta = \text{anti log}(-0{,}02119180389) = 0{,}952$$

Resolviendo empleando Excel se muestra en la siguiente figura:

	A	B	C	D	E	F
13						
14	$\begin{cases} 8\log\alpha + 153\log\beta = 12{,}97759 \\ 153\log\alpha + 4605\log\beta = 212{,}61769 \end{cases}$					
15						
16						
17	8,000	153	12,9776			
18	153	4605	212,6177			
19						
20	12,97759	153				
21	212,61769	4605				
22			27231,3151460	=MDETERM(A20:B21)		
23						
24	8	12,97759				
25	153	212,61769				
26			-284,6304178	=MDETERM(A24:B25)		
27						
28	8	153				
29	153	4605				
30			13431	=MDETERM(A28:B29)		
31						
32	log α	2,02749722	=C22/C30	α	106,536	=10^G21
33						
34	log β	-0,02119205	=C26/C30	β	0,952	=10 ^G23

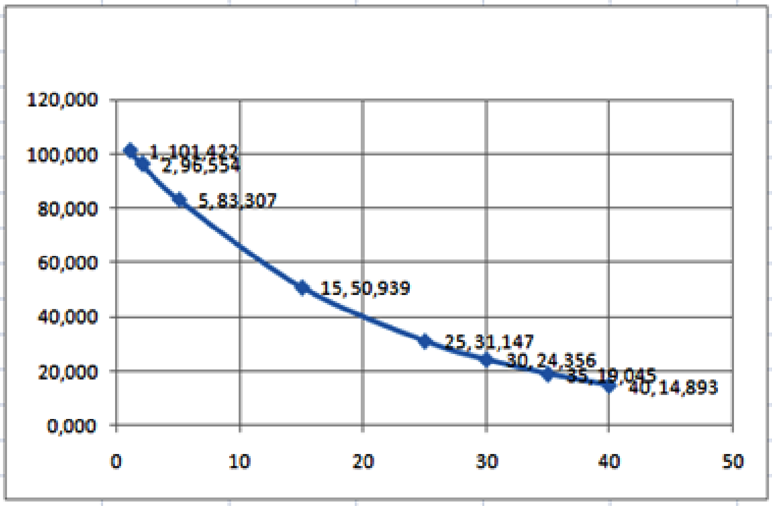
3) Reemplazando en la ecuación predictora se obtiene:

$$Y = \alpha \cdot \beta^X$$

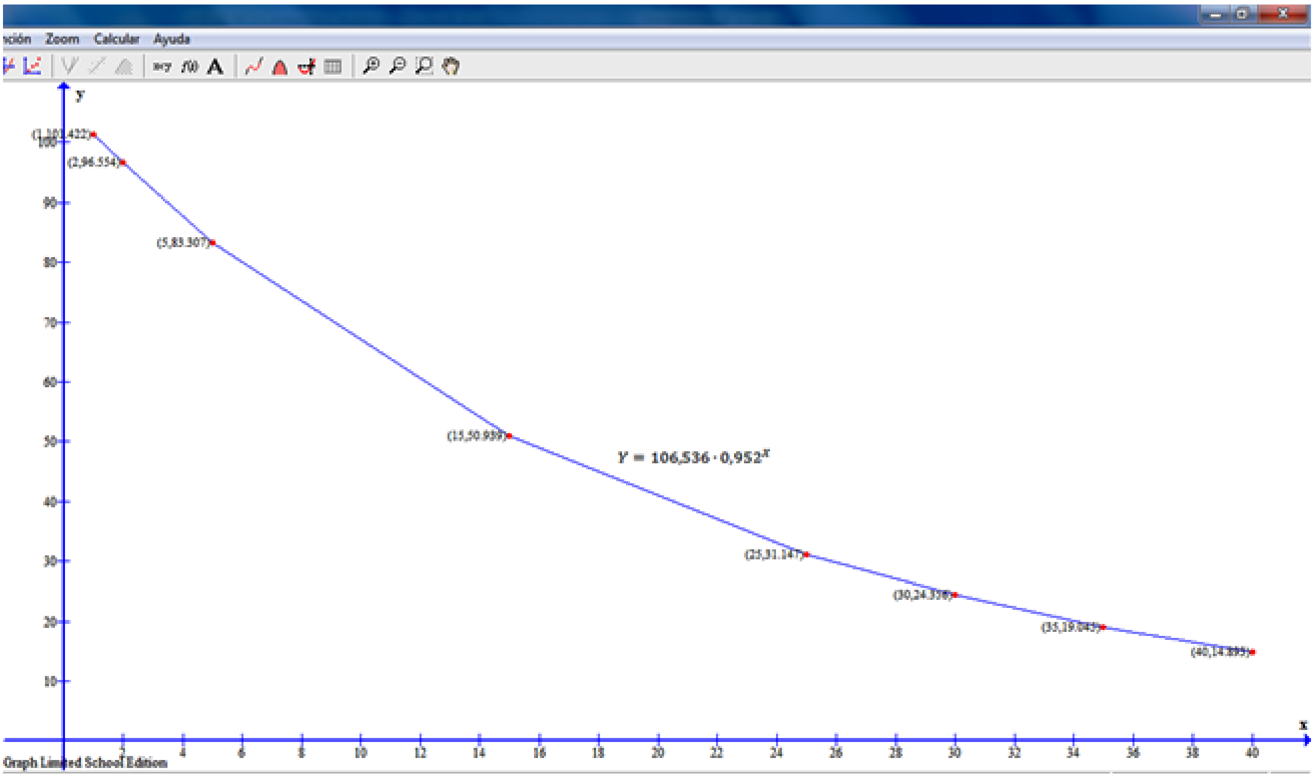
$$Y = 106{,}536 \cdot 0{,}952^X$$

4) *Graficando la ecuación predictora empleando Excel se obtiene la siguiente figura:*

	A	B	C
1	X	$Y = 106{,}536 \cdot 0{,}952^X$	
2	1	101,422	=$B$12*$B$13^A2
3	2	96,554	=$B$12*$B$13^A3
4	5	83,307	=$B$12*$B$13^A4
5	15	50,939	=$B$12*$B$13^A5
6	25	31,147	=$B$12*$B$13^A6
7	30	24,356	=$B$12*$B$13^A7
8	35	19,045	=$B$12*$B$13^A8
9	40	14,893	=$B$12*$B$13^A9
10			
11			
12		106,536	
13		0,952	
14			
15			



En Graph se obtiene la siguiente figura:



5) La estimación del porcentaje de llantas radiales que durarán 50000 millas se obtiene reemplazando en la ecuación predictora el valor de X = 50

$$Y = 106{,}536 \cdot 0{,}952^X$$
$$Y = 106{,}536 \cdot 0{,}952^{50} = 9{,}106$$

Entonces el porcentaje sería de 9,106%

TAREA DE INTERAPRENDIZAJE

1) Elabore un organizador gráfico sobre la regresión exponencial.

2) Las cifras siguientes son datos sobre el porcentaje de llantas radiales producidas por cierto fabricante que aún pueden usarse después de recorrer cierto número de millas:

Miles de Millas recorridas (X)	1	2	5	10	20	30	40	50
Porcentaje útil (Y)	98	92	80	64	36	32	17	11

2.1) Ajuste una curva exponencial aplicando el método de mínimos cuadrados. Resolver manualmente y empleando Excel. Realizar los cálculos empleando la mayor cantidad de decimales.

$$logY = 1{,}9988 - 0{,}0189X$$

2.2) Calcule la ecuación predictora.

$$Y = 99{,}72 \cdot 0{,}9574^{X}$$

2.3) Grafique la ecuación predictora de manera manual, empleando Excel y el programa Graph.

2.4) Estime qué porcentaje de las llantas radiales del fabricante durarán 35000 millas.

21,7%

3) Cree y resuelva un ejercicio empleando los conocimientos de la regresión exponencial de manera manual, empleando Excel y Graph.

4) REGRESIÓN POTENCIAL

La regresión potencial tiene por ecuación predictora:

$$Y = \alpha \cdot X^{\beta}$$

Y la regresión recíproca es:

$$Y = \frac{1}{\alpha + \beta \cdot X}$$

Para el primer caso los valores siguen una ley potencial. Si la ecuación predictora está dada por: $Y = \alpha \cdot X^{\beta}$, tomando logaritmos en ambos miembros, queda:

$$logY = log\alpha + \beta \cdot logX$$

Donde las constantes α y β quedan fijadas al resolver simultáneamente las ecuaciones:

$$\begin{cases} \Sigma logY = log\alpha \cdot N + \beta \cdot \Sigma \mathrm{logX} \\ \Sigma\mathrm{logX} \cdot \mathrm{logY} = \mathrm{log}\alpha \cdot \Sigma\mathrm{logX} + \beta \cdot \Sigma(logX)^2 \end{cases}$$

Para el segundo caso, si la ecuación predictora está dada por $Y = 1/(\alpha + \beta \cdot X)$, entonces invirtiendo, la misma expresión se puede escribir $1/Y = (\alpha + \beta \cdot X)/1$, o sea:

$$Y = \frac{1}{\alpha + \beta \cdot X} \Rightarrow \frac{1}{Y} = \alpha + \beta \cdot X$$

Donde las constantes α y β quedan fijadas al resolver simultáneamente las ecuaciones:

$$\begin{cases} \Sigma\dfrac{1}{Y} = \alpha \cdot N + \beta \cdot \Sigma X \\ \Sigma X \cdot \dfrac{1}{Y} = \alpha \cdot \Sigma X + \beta \cdot \Sigma X^2 \end{cases}$$

Ejemplos ilustrativo N° 1

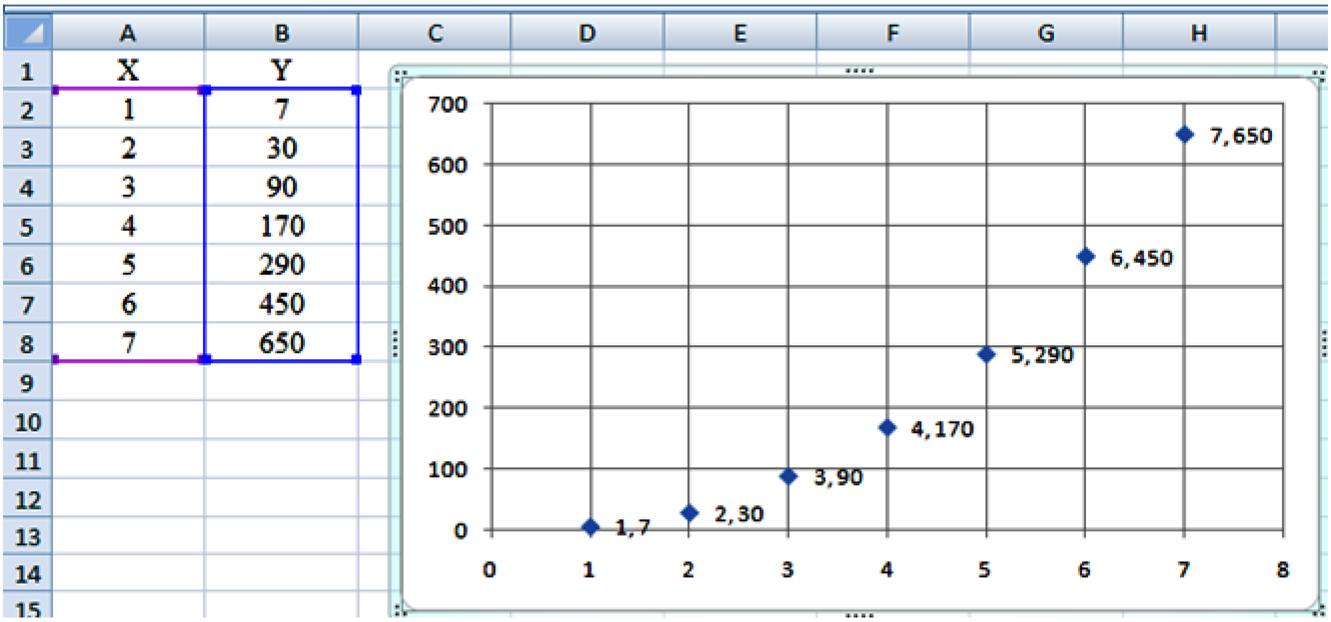
Sea el siguiente conjunto de valores, las lecturas de un experimento donde X es el volumen (variable independiente) e Y es la presión de una masa dada de gas (variable resultante).

X	1	2	3	4	5	6	7
Y	7	30	90	170	290	450	650

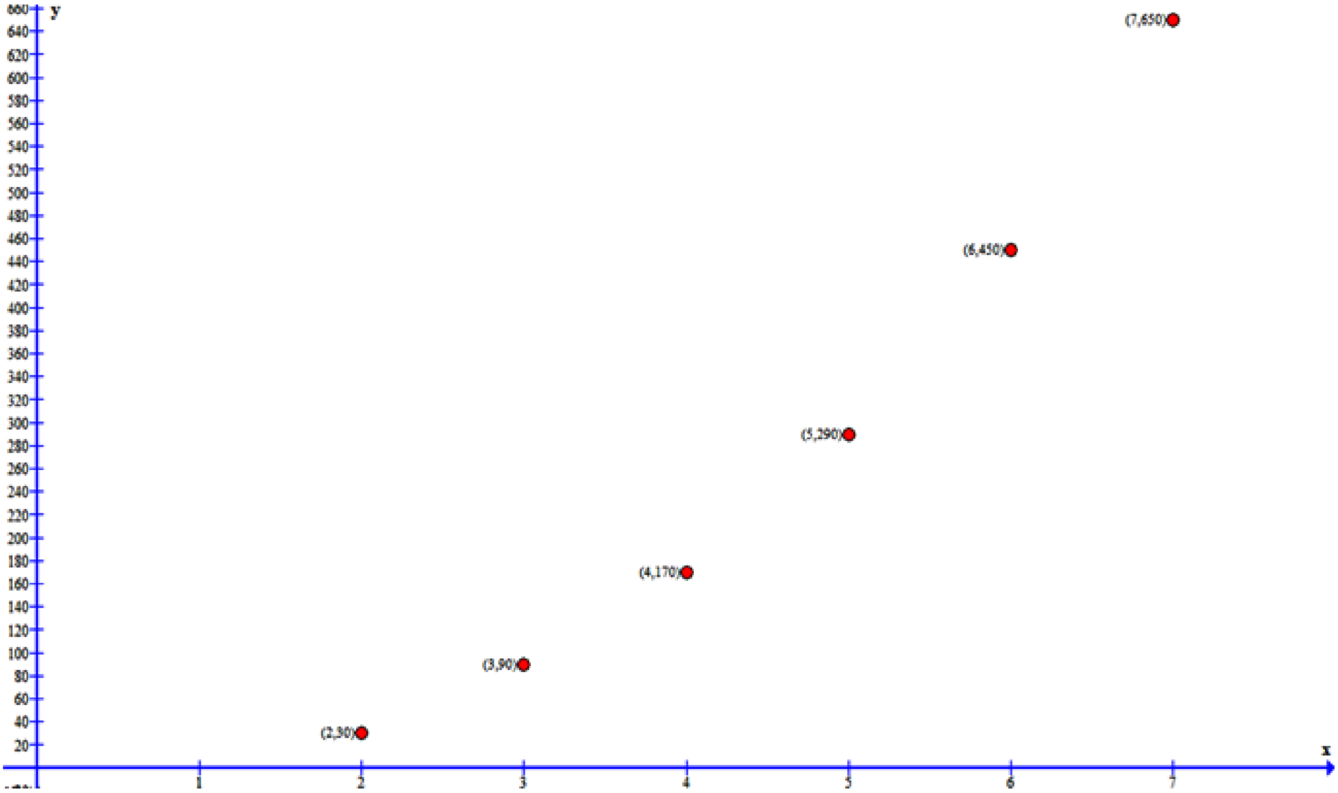
1.1) Elaborar el diagrama de dispersión.
1.2) Ajustar una curva exponencial aplicando el método de mínimos cuadrados.
1.3) Calcular la ecuación predictora.
1.4) Graficar la ecuación predictora.
1.5) Estimar la presión de la masa de gas de volumen 9.

Solución:

1.1) *El diagrama de dispersión elaborado en Excel se presenta en la siguiente figura*:



El diagrama de dispersión elaborado en Graph se presenta en la siguiente figura:



1.2) Para ajustar una curva exponencial aplicando el método de mínimos cuadrados se llena la siguiente tabla:

X	Y	log X	log Y	log X· log Y	(log X)^2
1	7	0,0000	0,8451	0,0000	0,0000
2	30	0,3010	1,4771	0,4447	0,0906
3	90	0,4771	1,9542	0,9324	0,2276
4	170	0,6021	2,2304	1,3429	0,3625
5	290	0,6990	2,4624	1,7211	0,4886
6	450	0,7782	2,6532	2,0646	0,6055
7	650	0,8451	2,8129	2,3772	0,7142
Σ X=28		Σ logX=3,7024	Σ logY=14,4354	Σ log X· log Y =8,8829	Σ(log X)^2= 2,4890

Reemplazando valores en el sistema de ecuaciones se obtiene:

$$\begin{cases} \Sigma logY = log\alpha \cdot N + \beta \cdot \Sigma \mathrm{logX} \\ \Sigma \mathrm{logX} \cdot \mathrm{logY} = \mathrm{log}\alpha \cdot \Sigma \mathrm{logX} + \beta \cdot \Sigma (logX)^2 \end{cases}$$

$$\begin{cases} 14{,}4354 = log\alpha \cdot 7 + \beta \cdot 3{,}7024 \\ 8{,}8829 = log\alpha \cdot 3{,}7024 + \beta \cdot 2{,}4890 \end{cases} \Rightarrow \begin{cases} 7log\alpha + 3{,}7024\beta = 14{,}4354 \\ 3{,}7024log\alpha + 2{,}4890\beta = 8{,}8829 \end{cases}$$

Al resolver el sistema se obtiene: log α = 0,819; β = 2,351

Reemplazando valores en la ecuación predictora expresada en logaritmos se tiene:

$$logY = log\alpha + \beta \cdot logX$$
$$logY = 0{,}819 + 2{,}351 \cdot logX$$

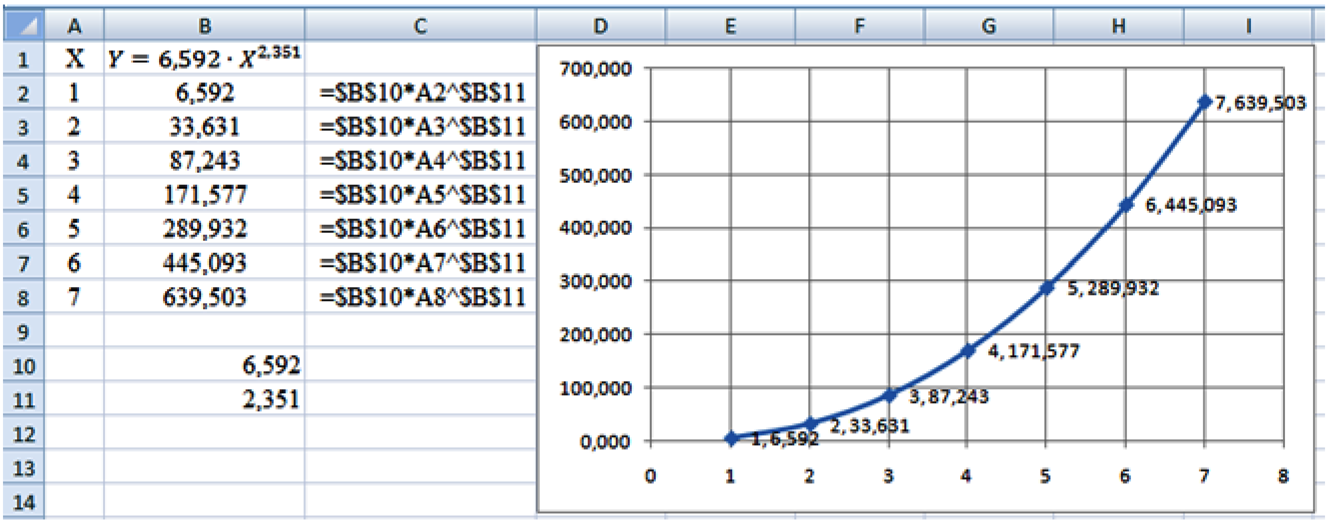
1.3) Para calcular la ecuación predictora, primero se calcula el valor de α de la siguiente manera:

$$log\alpha = 0{,}819 \Rightarrow \alpha = antilog\ 0{,}819 = 6{,}592$$

Reemplazando en la ecuación predictora se obtiene:

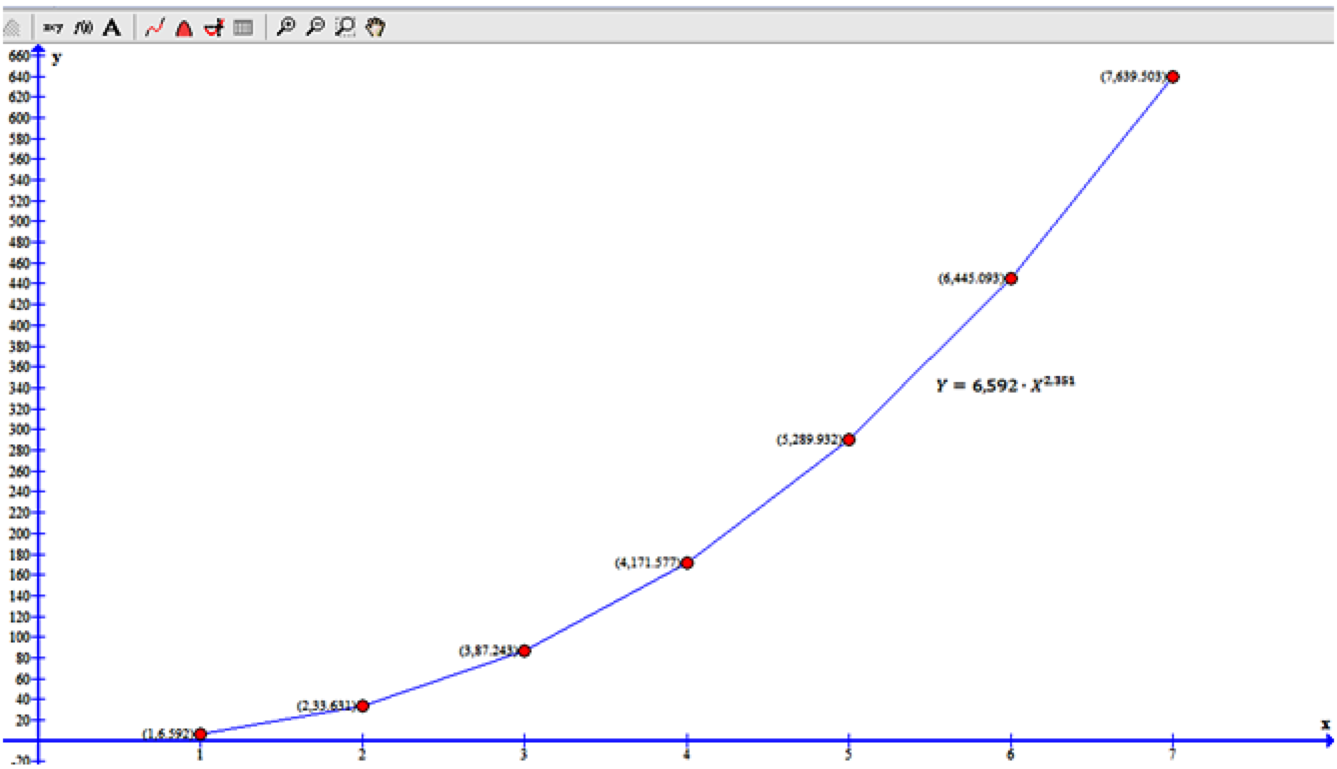
$$Y = \alpha \cdot X^{\beta}$$
$$Y = 6{,}592 \cdot X^{2{,}351}$$

1.4) *Graficando la ecuación predictora mediante Excel se muestra en la siguiente figura:*



	A	B	C
1	X	$Y = 6{,}592 \cdot X^{2{,}351}$	
2	1	6,592	=$B$10*A2^$B$11
3	2	33,631	=$B$10*A3^$B$11
4	3	87,243	=$B$10*A4^$B$11
5	4	171,577	=$B$10*A5^$B$11
6	5	289,932	=$B$10*A6^$B$11
7	6	445,093	=$B$10*A7^$B$11
8	7	639,503	=$B$10*A8^$B$11
9			
10		6,592	
11		2,351	
12			
13			
14			

Empleando Graph se obtiene la siguiente figura:



1.5) Para estimar la presión de la masa de gas de volumen 9 se reemplaza el valor X = 9 en la ecuación predictora

$Y = 6{,}592 \cdot X^{2{,}351}$

$Y = 6{,}592 \cdot 9^{2{,}351} = 875{,}35$

Ejemplo ilustrativo N° 2

Sea el siguiente conjunto de valores, las lecturas de un experimento donde X es la variable independiente e Y la variable resultante.

X	1	2	3	4	5	6	7
Y	1,4	1	0,9	0,7	0,6	0,55	0,5

2.1) Elaborar el diagrama de dispersión.

2.2) Calcular las constantes α y β, aplicando el método de mínimos cuadrados.

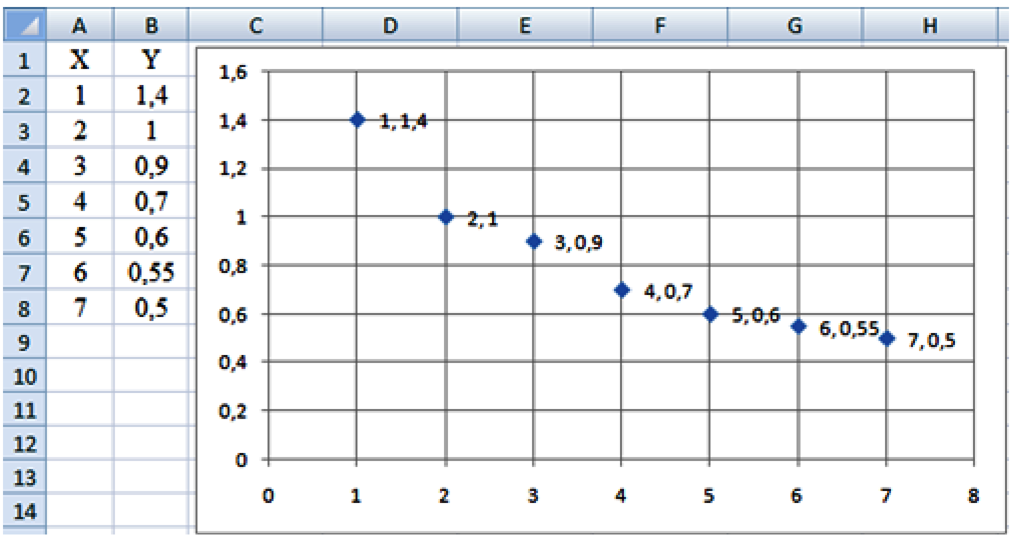
2.3) Calcular la ecuación predictora.

2.4) Graficar la ecuación predictora.

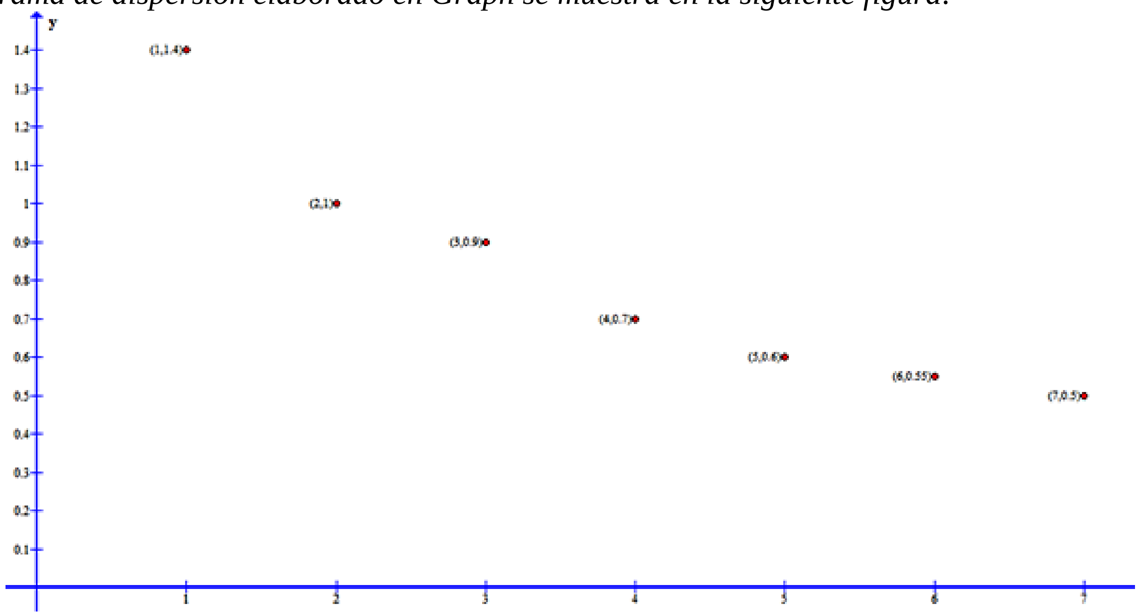
2.5) Estimar el valor de Y para X = 9

Solución:

2.1) *El diagrama de dispersión elaborado en Excel se muestra en la siguiente figura:*



El diagrama de dispersión elaborado en Graph se muestra en la siguiente figura:



2.2) Para calcular las constantes α y β, aplicando el método de mínimos cuadrados se llena la siguiente tabla:

X	Y	$1/Y$	$X(1/Y)$	X^2
1	1,4	0,7143	0,7143	1
2	1	1,0000	2,0000	4
3	0,9	1,1111	3,3333	9
4	0,7	1,4286	5,7143	16
5	0,6	1,6667	8,3333	25
6	0,55	1,8182	10,9091	36
7	0,5	2,0000	14,0000	49
$\Sigma X = 28$		$\Sigma\,(1/Y) = 9{,}7388$	$\Sigma\, X(1/Y) = 45{,}0043$	$\Sigma X^2 = 140$

Reemplazando valores en el siguiente sistema se obtiene:

$$\begin{cases} \Sigma \dfrac{1}{Y} = \alpha \cdot N + \beta \cdot \Sigma \mathrm{X} \\ \Sigma \mathrm{X} \cdot \dfrac{1}{Y} = \alpha \cdot \Sigma \mathrm{X} + \beta \cdot \Sigma X^2 \end{cases}$$

$$\begin{cases} 9{,}7388 = \alpha \cdot 7 + \beta \cdot 28 \\ 45{,}0043 = \alpha \cdot 28 + \beta \cdot 140 \end{cases} \Rightarrow \begin{cases} 7\alpha + 28\beta = 9{,}7388 \\ 28\alpha + 140\beta = 45{,}0043 \end{cases}$$

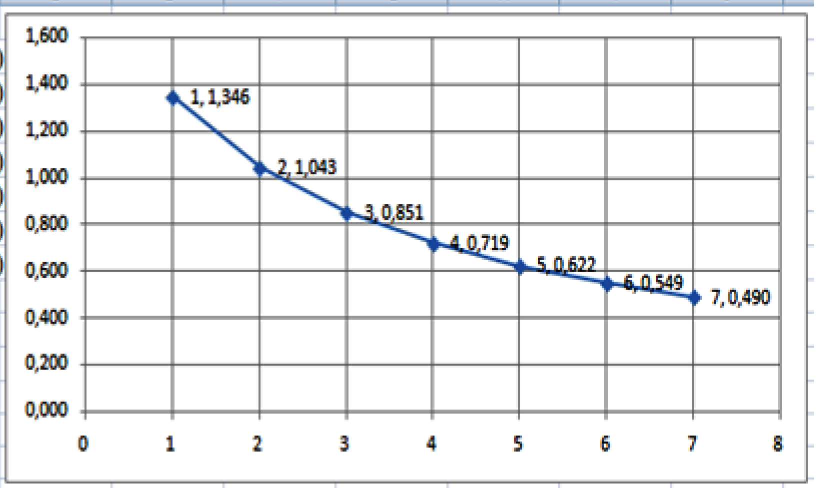
Al resolver el sistema se obtiene:
$\alpha = 0{,}5271$; $\beta = 0{,}2160$

2.3) Para calcular la ecuación predictora se reemplaza los valores encontrados de α y β, y se obtiene:

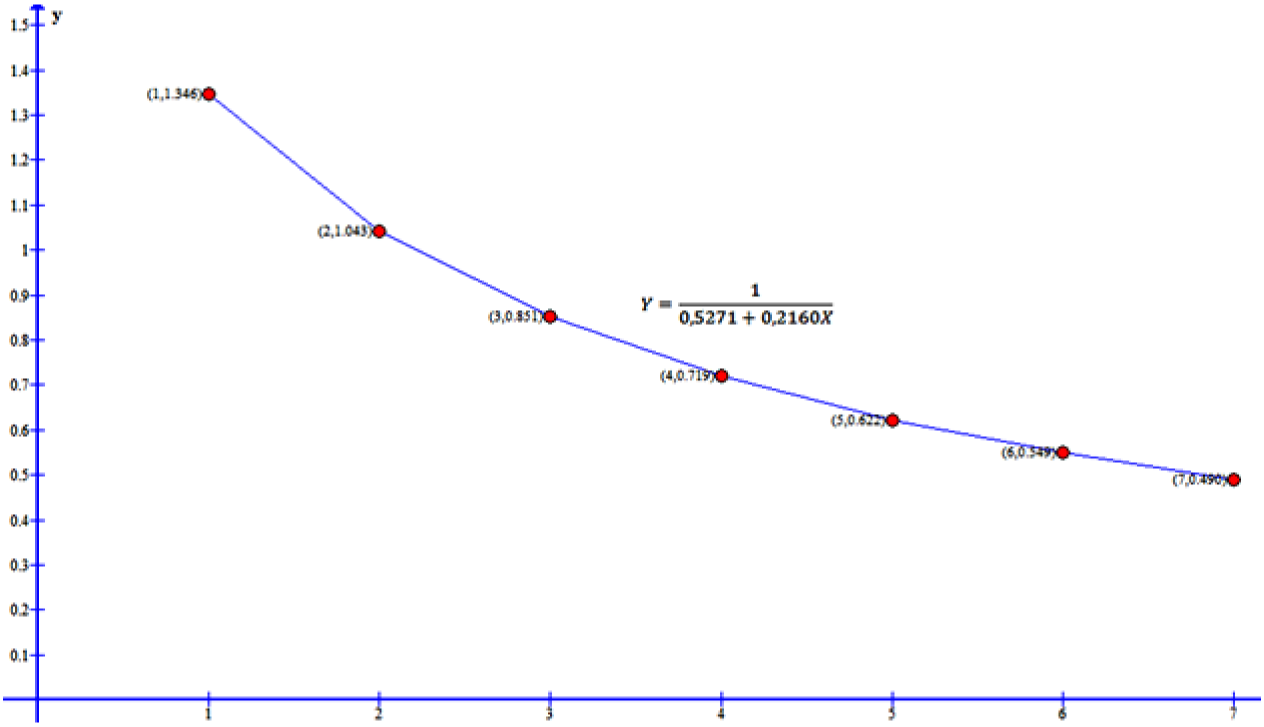
$$Y = \frac{1}{\alpha + \beta \cdot X} \Rightarrow Y = \frac{1}{0{,}5271 + 0{,}2160X}$$

2.4) *La gráfica la ecuación predictora elaborada en Excel se muestra en la siguiente figura:*

	A	B	C
1	X	Y = 1/(0,5271+0,2160X)	
2	1	1,346	=1/($B$10+$B$11*A2)
3	2	1,043	=1/($B$10+$B$11*A3)
4	3	0,851	=1/($B$10+$B$11*A4)
5	4	0,719	=1/($B$10+$B$11*A5)
6	5	0,622	=1/($B$10+$B$11*A6)
7	6	0,549	=1/($B$10+$B$11*A7)
8	7	0,490	=1/($B$10+$B$11*A8)
9			
10		0,5271	
11		0,216	
12			
13			
14			



La gráfica la ecuación predictora elaborada en Graph se muestra en la siguiente figura:



2.5) Para estimar el valor de Y para X = 9 se reemplaza el valor de X en la ecuación predictora.

$$Y = \frac{1}{0,5271 + 0,2160X}$$

$$Y = \frac{1}{0,5271 + 0,2160 \cdot 9} = 0,405$$

TAREA DE INTERAPRENDIZAJE

1) Elabore un organizador gráfico sobre la regresión potencial.

2) Sea el siguiente conjunto de valores, las lecturas de un experimento donde X es el volumen (variable independiente) e Y es la presión de una masa dada de gas (variable resultante).

X	1	2	3	4	5	6	7
Y	5	35	90	180	300	460	670

2.1) Elabore el diagrama de dispersión de manera manual, empleando Excel y Graph

2.2) Ajuste una curva exponencial aplicando el método de mínimos cuadrados empleando por lo menos 4 decimales para los cálculos.

$$logY = 0,7437 + 2,4883 \cdot logX$$

2.3) Calcule la ecuación predictora.

$$Y = 5,5424 \cdot X^{2,4883}$$

2.4) Grafique la ecuación predictora de manera manual, empleando Excel y Graph.

2.5) Estime la presión de la masa de gas de volumen 8.

979,17

3) Cree y resuelva un ejercicio similar al anterior.

4) Sea el siguiente conjunto de valores, las lecturas de un experimento donde X es la variable independiente e Y la variable resultante.

X	1	2	3	4	5	6	7
Y	1,5	1	0,8	0,9	0,5	0,4	0,3

4.1) Elabore el diagrama de dispersión de manera manual, empleando Excel y Graph.

4.2) Calcule las constantes α y β, aplicando el método de mínimos cuadrados de manera manual y empleando Excel.

α = 0,0159; β = 0,4196

4.3) Calcule la ecuación predictora.

$$Y = \frac{1}{0{,}0159 + 0{,}4196X}$$

4.4) Grafique la ecuación predictora de manera manual, empleando Excel y Graph.

4.5) Estime el valor de Y para X = 8

0,2965

5) Cree y resuelva un ejercicio similar al anterior de manera manual, empleando Excel y el programa Graph.

REFERENCIAS BIBLIOGRÁFICAS

SUÁREZ, Mario, (2012), TAPIA, Fausto Interaprendizaje de Estadística Básica, Universidad Técnica del Norte Ibarra, Ecuador.

SUÁREZ, Mario, (2011), La recta de los mínimos cuadrados, www.monografias.com/trabajos85/

SUÁREZ, Mario (2011), Análisis de regresión mediante la parábola de los mínimos cuadrados, www.monografias.com/trabajos86/

SUÁREZ, Mario (2011), Regresión potencial mediante el método de los mínimos cuadrados, www.monografias.com/trabajos89/

SUÁREZ, Mario (2011), Regresión exponencial mediante el método de los mínimos cuadrados, www.monografias.com/trabajos89/