UNIVERSIDAD AUTÓNOMA DEL ESTADO DE HIDALGO

ESCUELA SUPERIOR DE HUEJUTLA

LICENCIATURA EN ADMINISTRACIÓN

REGRESIÓN LINEAL Y CORRELACIÓN

MTRA. ROXANA SIFUENTES CARRILLO

JULIO-DICIEMBRE 2017

LINEAR REGRESSION AND CORRELATION

Abstract

Linear regression is a technique that includes a form of estimation and analysis of the sample data to know if and how 2 or more variables in a population are related to each other. It serves to make possible predictions by approximations.

The term Correlation means mutual relationship, since it indicates the degree to which the values of one variable are related to the values of the other.

Key words: regression, predictions, approximations, correlation, variable.

REGRESIÓN LINEAL

La regresión lineal es una técnica que comprende una forma de estimación y análisis de los datos muestrales para saber sí y cómo se relacionan entre sí 2 o más variables en una población.

Sirve para hacer posible predicciones mediante aproximaciones; puede que no sea exacta y se pueden visualizar mediante gráficas.

Ejemplos:

- Predecir el desempeño de un alumno en la universidad con base en los resultados obtenidos en la preparatoria.

- Distancia de un automóvil precisa para detenerse a partir de su velocidad.

CORRELACIÓN

El objetivo de estudio de correlación es determinar la consistencia de una relación entre observaciones por pares.

El término Correlación significa relación mutua, ya que indica el grado en el que los valores de una variable se relacionan con los valores de la otra.

Ejemplos:

- Relación entre la edad y la resistencia física.
- Tienden a tener mayor escolaridad las personas con altos ingresos en comparación con los de bajos ingresos.
- Puede el éxito en el trabajo predecirse a partir de calificaciones obtenidas en las pruebas de selección.

CORRELACIÓN

La correlación entre dos variables de obtiene mediante la siguiente fórmula:

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

Donde:

x=Variable independiente

y=Variable dependiente

Correlación Positiva: Las 2 variables crecen

Correlación negativa: Una variable crece y la otra decrece.

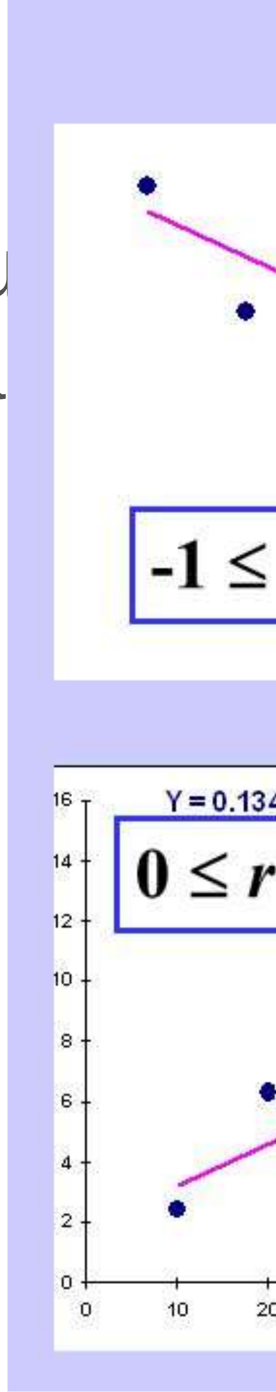
TIPOS DE CORRELACIÓN

La siguiente tabla indica el tipo de correlación entre dos variables de acuerdo al resultado obtenido en la fórmula.

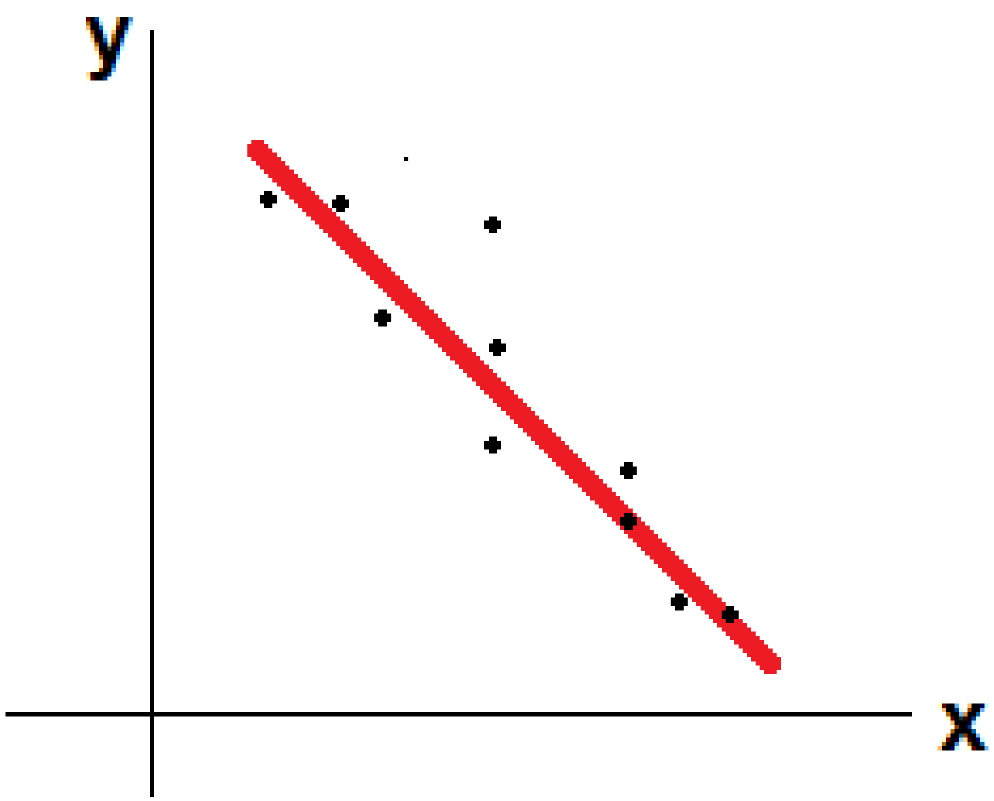
Tipo de Correlación	Rango
Perfecta Relación	±0.85 a ±1
Fuerte Relación	±0.70 a ±0.85
Mediana Relación	±0.40 a ±0.70
Poca Relación	±0.20 a ±0.40
Nula Relación	± 0 a ±0.20

TIPOS DE CORRELACIÓN

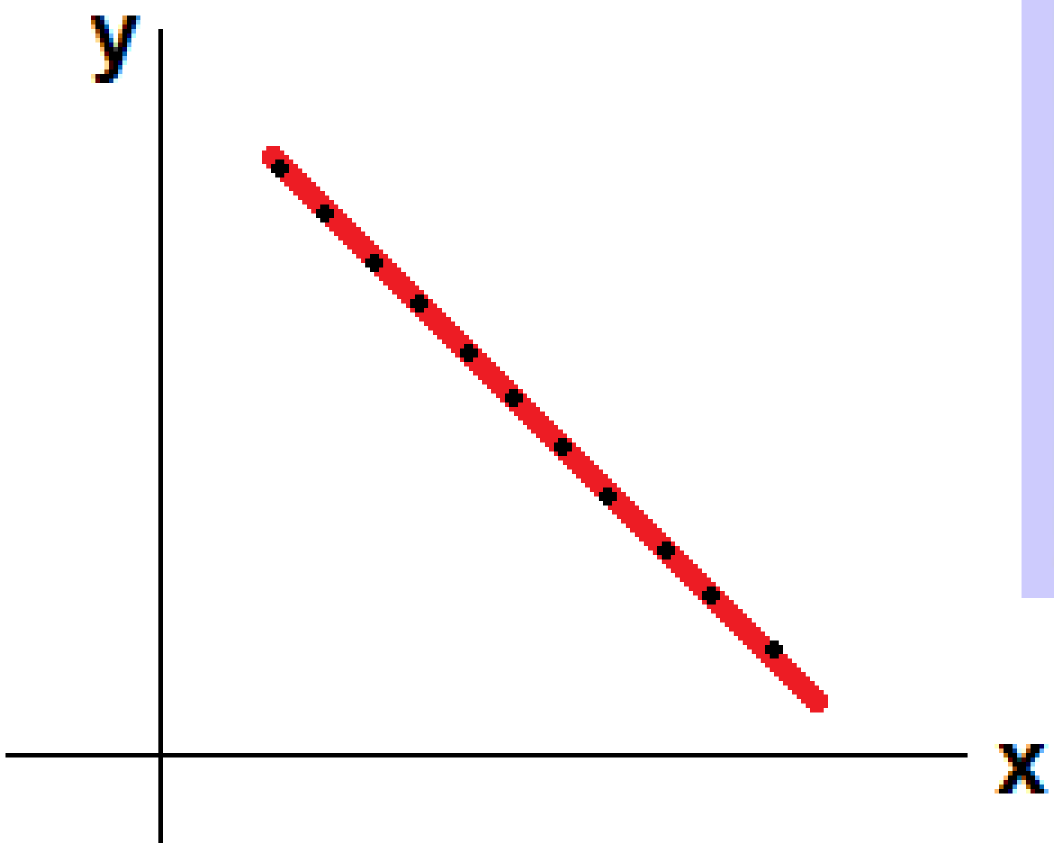
El tipo de correlación entre dos variables también pu observarse al graficar el valor de cada una de las varia observadas.



Correlación Negativa



-1≤ r ≤ 0

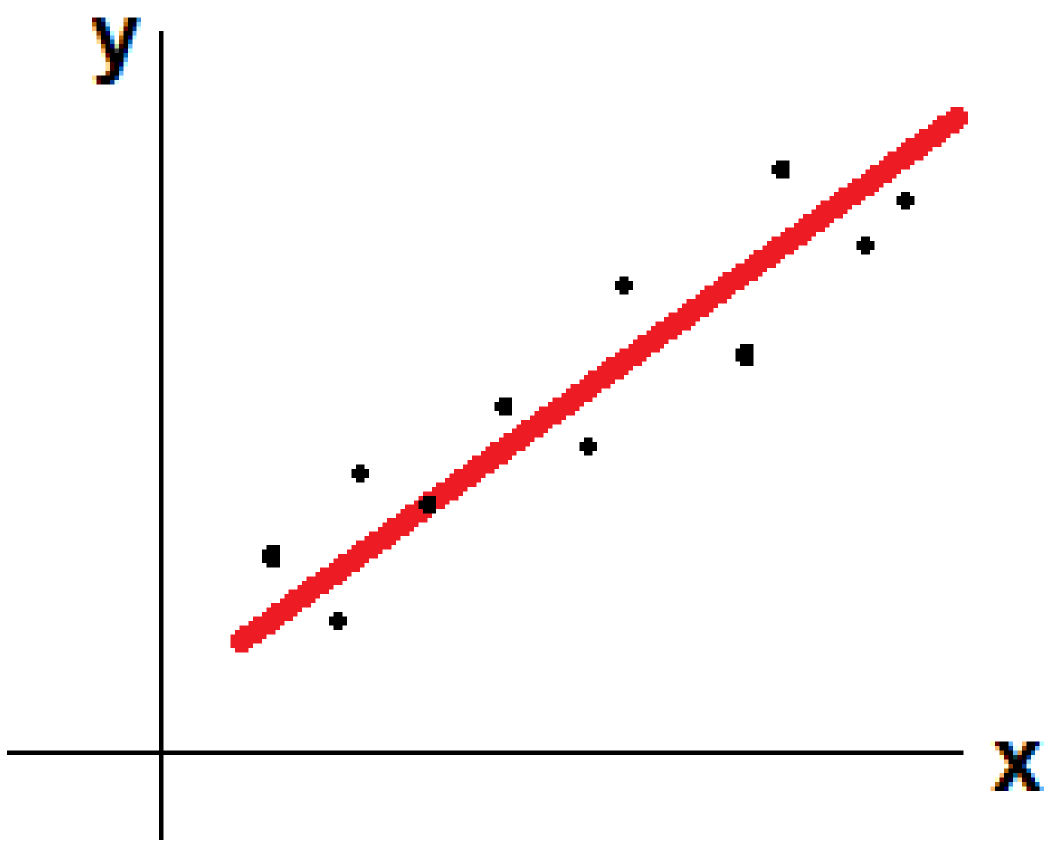


r=-1

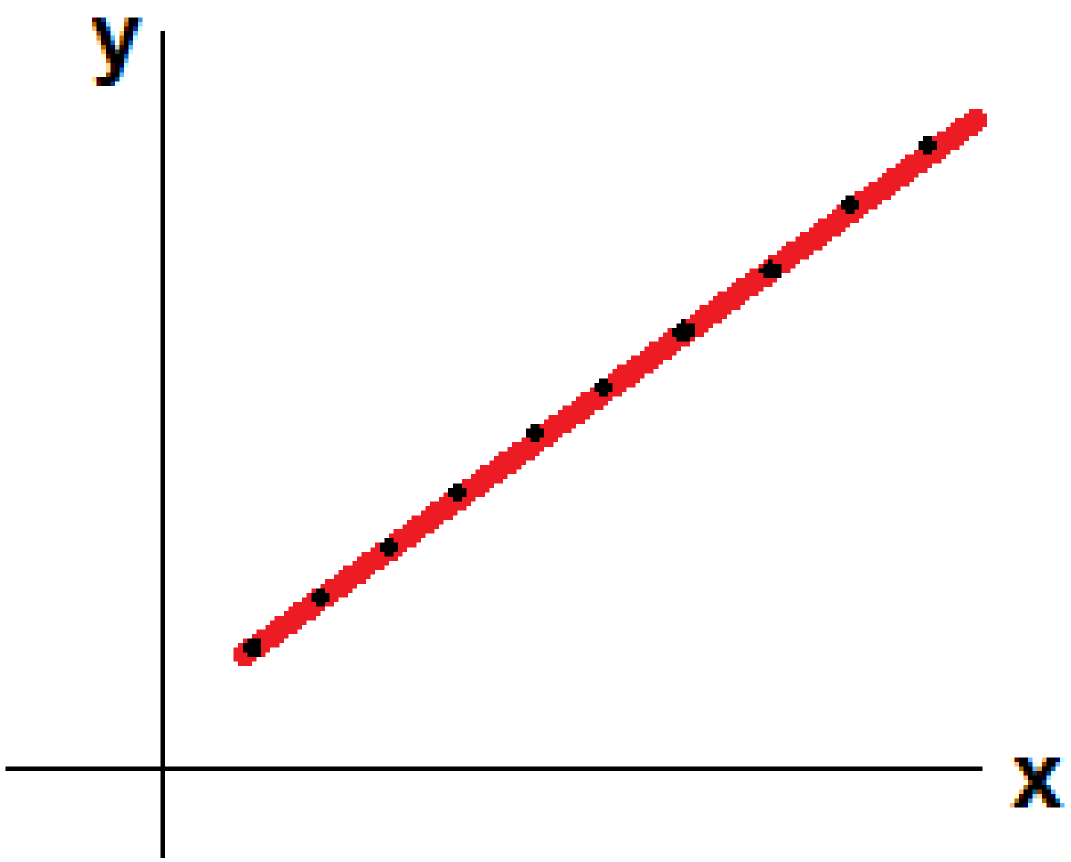
Imagen 1

TIPOS DE CORRELACIÓN

Correlación Positiva



0≤ r ≤ +1



r=+1

Imagen 1

ECUACION LINEAL

Dos características importantes de una ecuación lineal son:

a) La pendiente de la recta.

b) La localización de la recta en algún punto.

Fórmula:

$$y = (a + bx) \pm \quad Se$$

Donde:

y= Variable dependiente

x= Variable independiente

a y b= Mínimos cuadrados

Se= Margen de error

MÉTODO DE MÍNIMOS CUADRADOS

Es el procedimiento que más se utiliza para adaptar una recta a un conjunto de puntos, mediante las siguientes fórmulas:

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

$$a = \frac{\sum y - b\left(\sum x\right)}{n}$$

MARGEN DE ERROR

La determinante primaria de la exactitud es el grado de dispersión de la población; cuánto más disperso esté, menor será la exactitud de las estimaciones.

El grado de dispersión en la población se puede estimar a partir del grado de dispersión en las observaciones de las muestras con respecto a la línea de regresión calculada, utilizando la siguiente fórmula:

$$Se = \sqrt{\frac{\sum y^2 - a\left(\sum y\right) - b\left(\sum xy\right)}{n-2}}$$

EJEMPLO

Los siguientes datos indican el Monto del seguro en función del Ingreso anual de cada trabajador; Determine la ecuación de la recta y las predicciones.

Trabajador	Monto del seguro (miles de pesos) "Y"	Ingreso Anual (miles de pesos) "X"	xy	x^2	y^2
1	10	20	200	400	100
2	12	25	300	625	144
3	15	26	390	676	225
4	10	18	180	324	100
5	15	16	240	256	225
6	20	17	340	289	400
7	30	31	930	961	900
8	5	13	65	169	25
9	40	38	1520	1444	1600
10	50	41	2050	1681	2500
11	40	42	1680	1764	1600
12	55	45	2475	2025	3025
N=12	$\sum Y = 302$	$\sum X = 332$	$\sum XY = 10370$	$\sum x^2 = 10614$	$\sum y^2 = 10844$

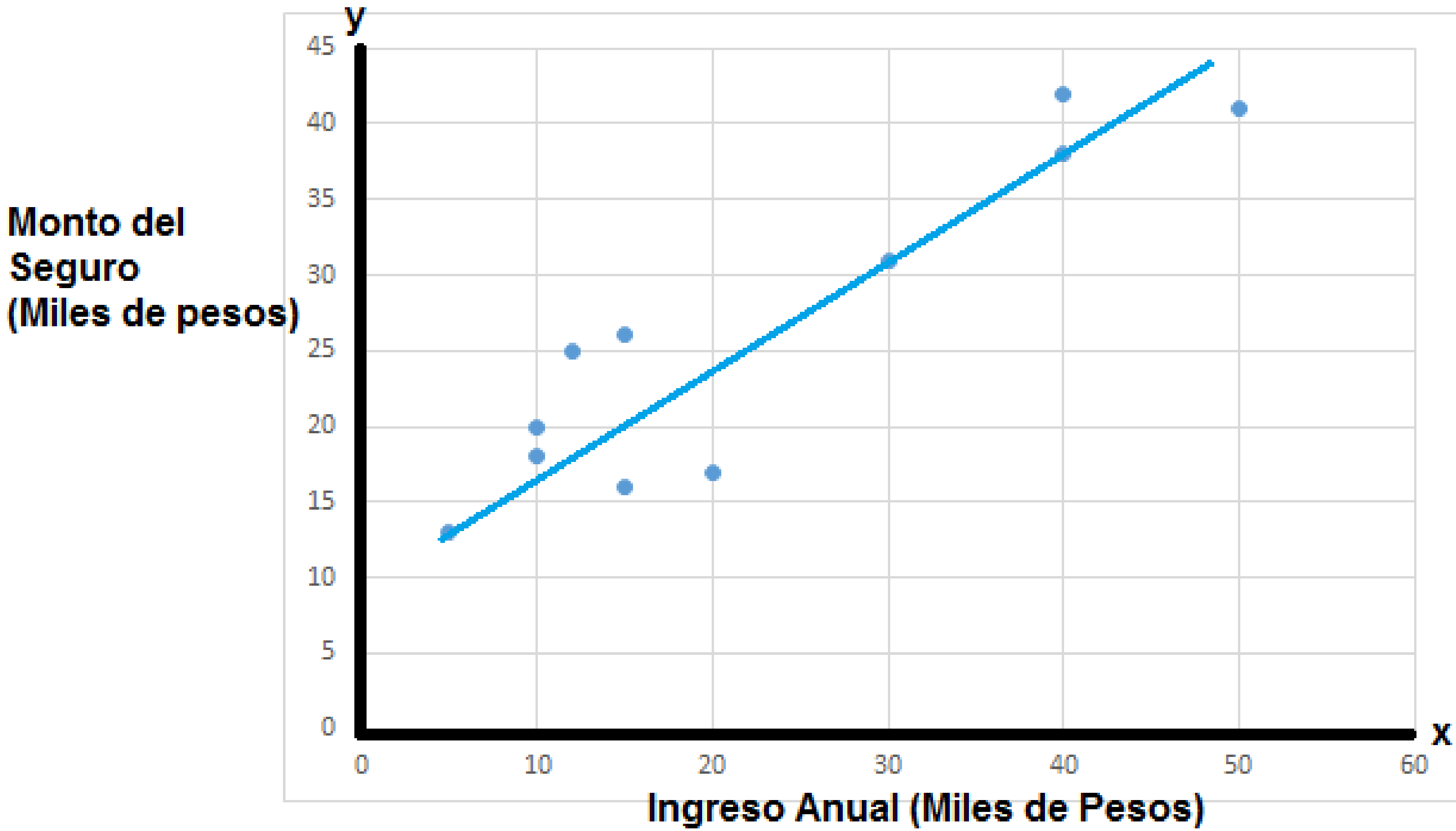
CORRELACIÓN

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

$$r = \frac{12(10370) - (332)(302)}{\sqrt{[12(10614) - (332)^2][12(10844) - (302)^2]}} = 0.9356$$

$$r = 93.56\% \quad \textit{Perfecta correlación positiva}$$

GRÁFICA



$r = 93.56\% \quad Perfecta\ correlación\ positiva$

MINIMOS CUADRADOS

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

$$b = \frac{12(10370) - (332)(302)}{[12(10614) - (332)^2]} = 1.41$$

$$a = \frac{\sum y - b\left(\sum x\right)}{n}$$

$$a = \frac{302 - 1.41(332)}{12} = -13.84$$

MARGEN DE ERROR

$$Se = \sqrt{\frac{\sum y^2 - a\left(\sum y\right) - b\left(\sum xy\right)}{n-2}}$$

$$Se = \sqrt{\frac{10844 - (-13.84)(302) - (1.41)(10370)}{12-2}}$$

$Se =$6.34

ECUACION LINEAL

$$y = \left(a + bx\right) \pm \quad Se$$

$$y = (-13.84 + 1.41x) \pm 6.34$$

El monto del seguro es igual a -13.84 miles de pesos más 1.41 miles de pesos por el ingreso anual del trabajador, con un margen de error de 38.15 miles de pesos

PREDICCIONES

Determinar el monto del seguro de un trabajador, si su ingreso anual es de 240 mil pesos

$$y = (-13.84 + 1.41x) \pm 6.34$$

$$y = (-13.84 + 1.41(29)) \pm 6.34$$

$$y = 27.05 \pm 6.34$$

$$y = 20.71\ a\ 33.39\ Miles\ de\ pesos$$

BIBLIOGRAFIA

Diaz A. (2013). *Estadística aplicada a la administración y a la economía.* México D.F.: Mcgraw-Hill.

Lind D. (2012).*Estadística Aplicada a los negocios y a la economía.* Mcgraw-Hill.