



Regresion Lineal

Materia:

Simulación

Docente:

Ing. Diego Quisi

Estudiante:

Ricardo Vinicio Jara Jara

Definición:

El análisis de regresión involucra el estudio la relación entre dos variables CUANTITATIVAS. En general interesa:

- Investigar si existe una asociación entre las dos variables testeando la hipótesis de independencia estadística.
- Estudiar la fuerza de la asociación, a través de una medida de asociación denominada coeficiente de correlación.
- Estudiar la forma de la relación. Usando los datos propondremos un modelo para la relación y a partir de ella será posible predecir el valor de una variable a partir de la otra.

Para ello proponemos un MODELO que relaciona una variable dependiente (Y) con una variable independiente (X). La decisión sobre qué análisis usar en una situación particular, depende de la naturaleza del variable dependiente y del tipo de función que se propone para relacionar el variable dependiente y la variable independiente.

Llamaremos MODELO MATEMÁTICO a la función matemática que proponemos como forma de relación entre la variable dependiente (Y) y la o las variables independientes.

La función más simple para la relación entre dos variables es la FUNCIÓN LINEAL $Y = a + b X$

- Esta expresión es una aproximación de la verdadera relación entre X e Y.
- Para un dado valor de X el modelo predice un cierto valor para Y.
- Mientras mejor sea la predicción, mejor es el modelo para explicar el fenómeno. Por ejemplo, $Y = 2 X + 3$

Ejercicio en python:

```
In [38]: import numpy as np
import random
import matplotlib.pyplot as plt
#Generacion de Puntos
x = random.sample(range(9), 5)
y = random.sample(range(9), 5)
n = len(x)
x = np.array(x)
y = np.array(y)

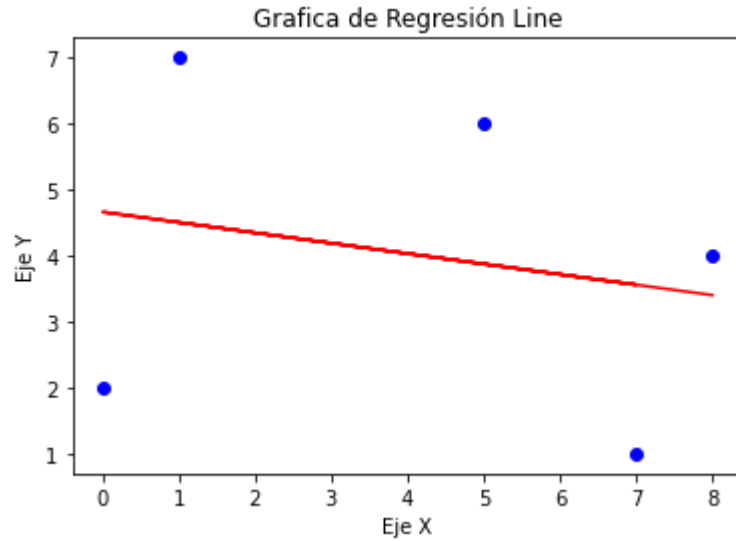
#Sumatoria de los arreglos
sumarx = sum(x)
sumary = sum(y)
sumarx2 = sum(x*x)
sumary2 = sum(y*y)
sumarxy = sum(x*y)

#Promedio
promediox = sumarx/n
promedioy = sumary/n
m = (sumarx*sumary - n*sumarxy)/(sumarx**2 - n*sumarx2)
b = promedioy - m*promediox
print ('Promedios: ',m,b);

#Calculo de operaciones [Predecir]
resultadomax = np.sqrt(sumarx2/n - promediox**2)
resultadomay = np.sqrt(sumary2/n - promedioy**2)
resultadomaxy = sumarxy/n - promediox*promedioy
prediccion = (resultadomaxy/(resultadomax*resultadomay))**2
print ('Predicción:', prediccion);

#Grafica
plt.title('Grafica de Regresión Line')
plt.plot(x,y,'o',color='blue')
plt.plot(x, m*x + b,color='red')
plt.xlabel('Eje X')
plt.ylabel('Eje Y')
plt.show()
```

Promedios: -0.15748031496062992 4.661417322834645
Predicción: 0.04845548152634777



Conclusiones

La regresión lineal es un calculo matematico que nos permite aproximar la relacion de dependencia entre una variable dependiente Y y un variable independiente X, obteniendo de esta manera un resultado o una predicción, con su respectiva grafica para observar los puntos para el analisis.

- El cálculo de la regresión lineal no permite de alguna manera predecir cual seria el futuro de una variable de acuerdo con otra variable.
- Con la ayuda de Python y la librería de numpy y matplotlib lo podemos aplicar la regresión línea y demostrar de morfa grafica los puntos.