

Question 1 Homocedasticidad

Homocedasticidad es una propiedad de la estadística.

Es el fenómeno que ocurre cuando la varianza de errores de un modelo de predicción RLM, es "homogéneo". Es decir, es el mismo a través de todas las variables del modelo.

Esta información puede ser encontrada en:

- [Wikipedia](#)
- [Economipedia](#)

Question 2 Correlación y covarianza

Covarianza

Este es un valor que indica el grado de variación entre 2 variables. Es un dato preliminar que se necesita para sacar la correlación, expuesta más adelante.

Sirve para indicar la dirección de la relación entre 2 variables.

$$s_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

Correlación

Este es un valor que necesita de la covarianza de 2 variables. Indica la fuerza con la que se relacionan estas 2 variables, cuanto más se acercan los valores a 1, mayor la fuerza es.

$$r_{xy} = \frac{s_{xy}}{s_x s_y} \quad r_k = \frac{\sum_{i=1}^{N-k} (x_i - \bar{x}) \cdot (y_{i+k} - \bar{y})}{\sqrt{\sum_{i=1}^{N-k} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=k+1}^N (y_i - \bar{y})^2}}$$

La información de la pregunta 2 se encontró en los siguientes links

- [Covarianza](#)
- [Correlación](#)
- [Covarianza y correlación](#)

Question 3

```
In [1]: import pandas as pd
import numpy as np
import math
```

```
In [2]: df = pd.DataFrame({
    "RAM" : [4,8,12,16],
    "Price" : np.array([7,9,12,16])*1000
})
df.head()
```

```
Out[2]:
```

	RAM	Price
0	4	7000
1	8	9000
2	12	12000
3	16	16000

```
In [3]: df.describe()
```

```
Out[3]:
```

	RAM	Price
count	4.000000	4.000000
mean	10.000000	11000.000000
std	5.163978	3915.780041
min	4.000000	7000.000000
25%	7.000000	8500.000000
50%	10.000000	10500.000000
75%	13.000000	13000.000000
max	16.000000	16000.000000

```
In [4]: vRam = df["RAM"].to_numpy()
vPrice = df["Price"].to_numpy()
sumRam = vRam.sum()
sumPrice = vPrice.sum()
avgRam = sumRam/len(vRam)
avgPrice = sumPrice/len(vPrice)
devRam = vRam - avgRam
devPrice = vPrice - avgPrice
sumDev = np.dot(devRam.T,devPrice)
coVar = sumDev/(len(vRam) - 1)
print(f"La covarianza es >> {coVar}")
```

La covarianza es >> 20000.0

```
In [5]: sumRam = math.sqrt(np.dot(devRam.T,devRam))
sumPrice = math.sqrt(np.dot(devPrice.T,devPrice))
print(f"La correlacion es >> {coVar/(sumRam*sumPrice)}")
```

La correlacion es >> 0.3296902366978935

Question 4 Colinealidad y multicolinealidad

La **multicolinealidad** es la correlación alta entre más de dos variables explicativas. Esta correlación debe de ser fuerte, más no perfecta.

La **colinealidad** es la super categoria de la multicolinealidad, es decir hay varios tipos de multicolinealidad, exacta y aproximada, y la colinealidad es la categoria que encapsula ambas.

Cómo identificar la multicolinealidad

Usando una matriz de correlación se puede identificar el tipo de colinealidad.

Dependiendo del valor de la determinante de la matriz, se puede identificar qué tipo de colinealidad es.

Si la colinealidad es = 0, eso quiere decir que es una **multicolinealidad exacta**.

En caso de que sea cercana a 0, es una **multicolinealidad aproximada**.

Esta información fue encontrada en:

- [Economipedia](#)
- [Wikipedia](#)