

## ▼ Actividad 4 - Intervalos de confianza

Frida Cano Falcón - A01752953

## ▼ Extracción de datos

```
1 from google.colab import drive
2 drive.mount('/content/drive')
3 %cd "/content/drive/MyDrive/7mo Semestre/Estadistica"
4 !ls

Mounted at /content/drive
/content/drive/MyDrive/Semestres/7mo Semestre/Estadistica
Act1_Distribuciones_FridaCano_A01752953.ipynb
Act2_ExplorandoBases_FridaCano_A01752953.ipynb
Act3_Transformaciones_FridaCanoFalcon_A01752953.ipynb
Act4_IntervalosConfianza_FridaCanoFalcon_A01752953.ipynb
'El marcapasos.csv'
mc-donalds-menu-1.csv
precios_autos.csv
```

Importamos las librerías necesarias

```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3 import numpy as np
4 from scipy import stats
5 # Importing library
6 from scipy.stats import norm
7 import random
8 import math
```

## ▼ Primer Problema

Suponga que la porosidad al helio (en porcentaje) de muestras de carbón, tomadas de cualquier veta en particular, está normalmente distribuida con una desviación estándar verdadera de 0.75. Se sabe que 10 años atrás la porosidad media de helio en la veta era de 5.3 y se tiene interés en saber si actualmente ha disminuido. Se toma una muestra al azar de 20 especímenes y su promedio resulta de 4.85.

1. Haga una estimación por intervalo con una confianza del 97% para el promedio de porosidad para evaluar si ha disminuido.

```
1 # En este caso tenemos una varianza conocida
2 alpha = 1 - 0.97
3 z = abs(norm.ppf(alpha/2))
4 x_media = 4.85
5 sigma = 0.75
6 n = 20
7
8 low_int = x_media - (z*(sigma/math.sqrt(n)))
9 up_int = x_media + (z*(sigma/math.sqrt(n)))
10 print(low_int, up_int)

4.486064779880205 5.213935220119795
```

2. Se toma otra muestra de tamaño 16. El promedio de la muestra fue de 4.56. Calcule el intervalo de confianza al 97% de confianza.

```
1 alpha = 1 - 0.97
2 z = abs(norm.ppf(alpha/2))
3 x_media = 4.56
4 sigma = 0.75
5 n = 16
6
7 low_int = x_media - (z*(sigma/math.sqrt(n)))
8 up_int = x_media + (z*(sigma/math.sqrt(n)))
9 print(low_int, up_int)
```

4.153108054202894 4.966891945797105

3. ¿Podemos afirmar que la porosidad del helio ha disminuido?

## ▼ Segundo problema

Suponga que la porosidad al helio (en porcentaje) de muestras de carbón, tomadas de cualquier veta en particular, está normalmente distribuida con una desviación estándar verdadera de 0.75.

1. ¿Qué tan grande tiene que ser el tamaño de la muestra si se desea que el ancho del intervalo con un 95% de confianza no sobrepase de 0.4?

Sabemos que:

$$n = \left( \frac{Z * \sigma}{E} \right)^2$$

Donde:

- $n$  : el tamaño de muestra
- $Z$  : el valor crítico de la distribución normal estándar para el nivel de confianza deseado. Para un 95% de confianza,  $Z$  es aproximadamente igual a 1.96 (este valor se encuentra en las tablas de distribución normal estándar).
- $\sigma$  : la desviación estándar verdadera del atributo que estamos midiendo. En este caso 0.75.

```
1 n = ((1.96*0.75)/0.2)**2
2 print("El tamaño de la muestra debe ser aproximadamente de : " + str(n))
```

El tamaño de la muestra debe ser aproximadamente de : 54.022499999999994

2. ¿Qué tamaño de muestra necesita para estimar la porosidad promedio verdadera dentro de 0.2 unidades alrededor de la media muestral con una confianza de 99%?

Sabemos que:

$$n = \left( \frac{Z * \sigma}{E} \right)^2$$

Donde:

- $n$  : el tamaño de muestra
- $Z$  : el valor crítico de la distribución normal estándar para el nivel de confianza deseado. Para un 99% de confianza,  $Z$  es aproximadamente igual a 2.576 (este valor se encuentra en las tablas de distribución normal estándar).
- $\sigma$  : la desviación estándar verdadera del atributo que estamos midiendo. En este caso 0.75.

```
1 n = ((2.576*0.75)/0.2)**2
2 print("El tamaño de la muestra debe ser aproximadamente de : " + str(n))

El tamaño de la muestra debe ser aproximadamente de : 93.31559999999998
```



## ▼ Marcapasos

Con el archivo de datos de El Marcapasos haz los intervalos de confianza para la media de dos de las siguientes variables:



- Intensidad de pulsos con y sin Marcapasos
- Periodo entre pulso con y sin Marcapasos

Cargamos la base de datos

```
1 df = pd.read_csv('El marcapasos.csv')
2 df.head()
```

	Periodo entre pulsos	Intensidad de pulso	Marcapasos	
0	1.2	0.131	Sin MP	
1	0.9	0.303	Sin MP	
2	0.9	0.297	Sin MP	
3	0.8	0.416	Sin MP	
4	0.7	0.585	Sin MP	

```
1 df.describe()
```

	Periodo entre pulsos	Intensidad de pulso	
count	102.000000	102.000000	
mean	1.001471	0.201520	
std	0.301936	0.123047	
min	0.690000	0.005000	
25%	0.830000	0.109000	
50%	0.935000	0.191000	
75%	1.100000	0.292250	
max	2.800000	0.585000	

Intensidad de pulso

```

1 # En este caso tenemos una variancia conocida
2 alpha = 1 - 0.95
3 z = abs(norm.ppf(alpha/2))
4 # Media
5 m = 0.201520
6 # Desviacion estandar
7 sigma = 0.123047
8 # Número de muestras
9 n = 102
10
11 low_int = m - (z*(sigma/math.sqrt(n)))
12 up_int = m + (z*(sigma/math.sqrt(n)))
13 print("Intervalo de confianza de Intensidad de pulso \n")
14 print(low_int, up_int)

Intervalo de confianza de Intensidad de pulso

0.1776408405879732 0.22539915941202682

```

## Periodo entre pulsos

```

1 # En este caso tenemos una variancia conocida
2 alpha = 1 - 0.95
3 z = abs(norm.ppf(alpha/2))
4 # Media
5 m = 1.001471
6 # Desviacion estandar
7 sigma = 0.301936
8 # Número de muestras
9 n = 102
10
11 low_int = m - (z*(sigma/math.sqrt(n)))
12 up_int = m + (z*(sigma/math.sqrt(n)))
13 print("Intervalo de confianza de Intensidad de pulso \n")
14 print(low_int, up_int)

Intervalo de confianza de Intensidad de pulso

0.9428756837693749 1.0600663162306252

```

Grafica los intervalos. En un gráfico la intensidad de pulso con y sin marcapasos y en otro gráfico el periodo entre pulso con y sin marcapasos. Interpreta el resultado.

```

1 x = np.linspace(np.min(df['Intensidad de pulso']),np.max(df['Intensidad de pulso']))
2

```

---

✓ 0 s completado a las 15:02

