

Instrucciones

1. Resuelve las dos partes del problema "El misterioso Helio". Las dos partes del problema se encuentran al final de la presentación Intervalos de confianza con N. Concluye en el contexto del 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. -a Haga una estimación por intervalo con una confianza del 97% para el promedio de porosidad para evaluar si ha disminuido. -b 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 -c ¿Podemos afirmar que la porosidad del helio ha disminuido? El misterioso Helio

Guillermo Romeo Cepeda Medina A01284015

```
import numpy as np
import scipy.stats as stats
import matplotlib.pyplot as plt
#Distribución normal, media 5.3 y desviación estándar 0.75
mu = 5.3
sigma = 0.75
#Nuestra x va de mu - 4*sigma a mu + 4*sigma con 100 puntos
x = np.linspace(mu - 4*sigma, mu + 4*sigma, 100)
n_dist_carbon = stats.norm(mu, sigma)
plt.plot(x, n_dist_carbon.pdf(x), color='black')
plt.show()

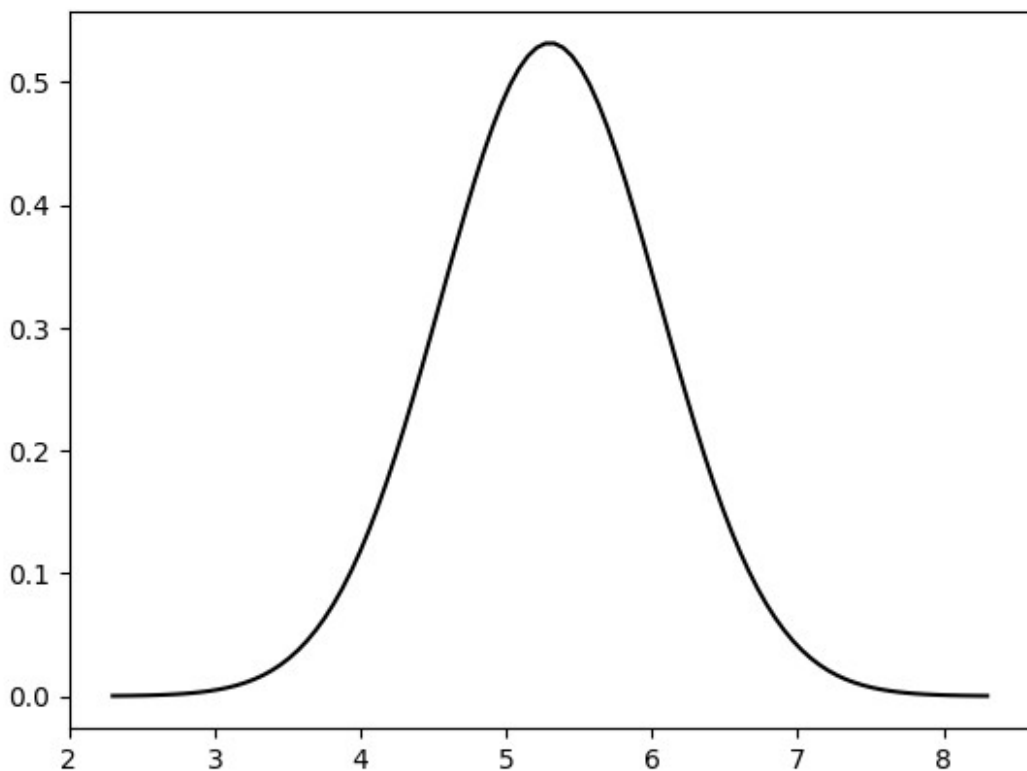
#Se toma una muestra de 20 elementos y se obtiene una media de 4.85
#Estimación de intervalo de confianza de 97% para la media
#Se calcula el error estándar
n = 20
error = sigma/np.sqrt(n)
print(error, "error")
#Se calcula el intervalo de confianza
mu2 = 4.85
intervalo = stats.norm.interval(0.97, loc=mu2, scale=error)
print(intervalo)
#Se grafica el intervalo de confianza
plt.plot(x, n_dist_carbon.pdf(x), color='black')
plt.axvline(5.3, color='blue')
plt.axvline(intervalo[0], color='red')
plt.axvline(intervalo[1], color='red')
plt.show()

#Podemos observar que la media anterior queda fuera del intervalo de
confianza de 97%, por lo que podemos afirmar que la media ha
disminuido.
```

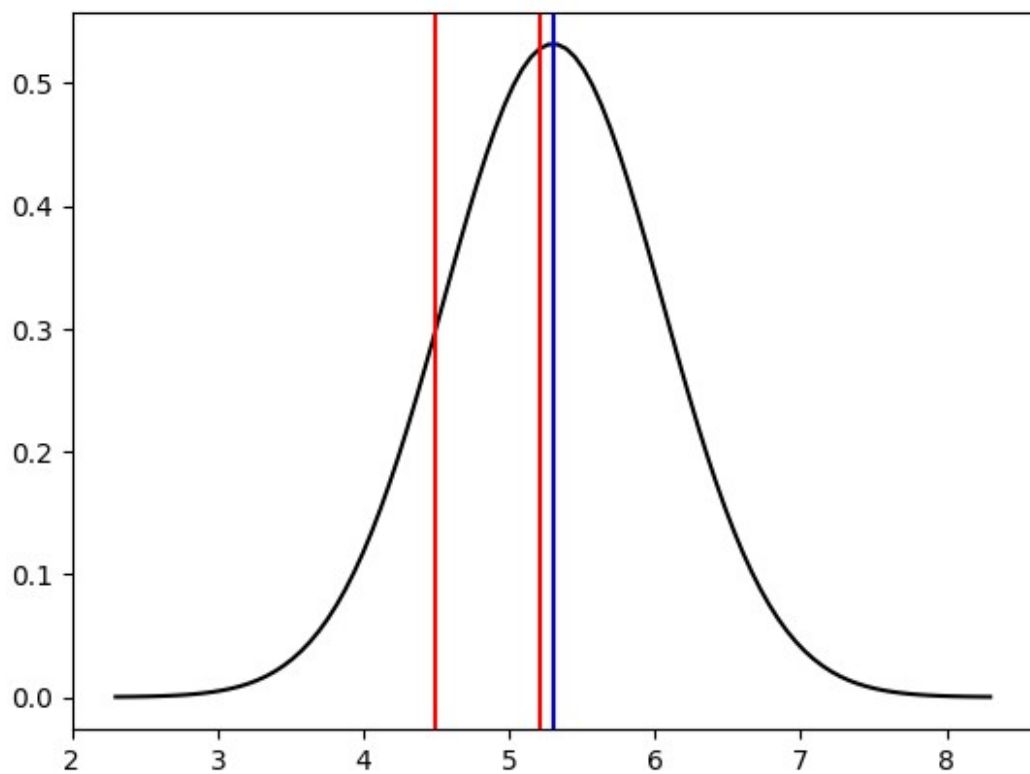
```
#Para una media de 4.56 con una muestra de 16 elementos
mu2 = 4.56
error2 = sigma/np.sqrt(16)
print(error2,"error2")
intervalo2 = stats.norm.interval(0.97, loc=mu2, scale=error2)
print(intervalo2)
```

```
plt.plot(x, n_dist_carbon.pdf(x), color='black')
plt.axvline(5.3, color='blue')
plt.axvline(intervalo2[0], color='red')
plt.axvline(intervalo2[1], color='red')
plt.show()
```

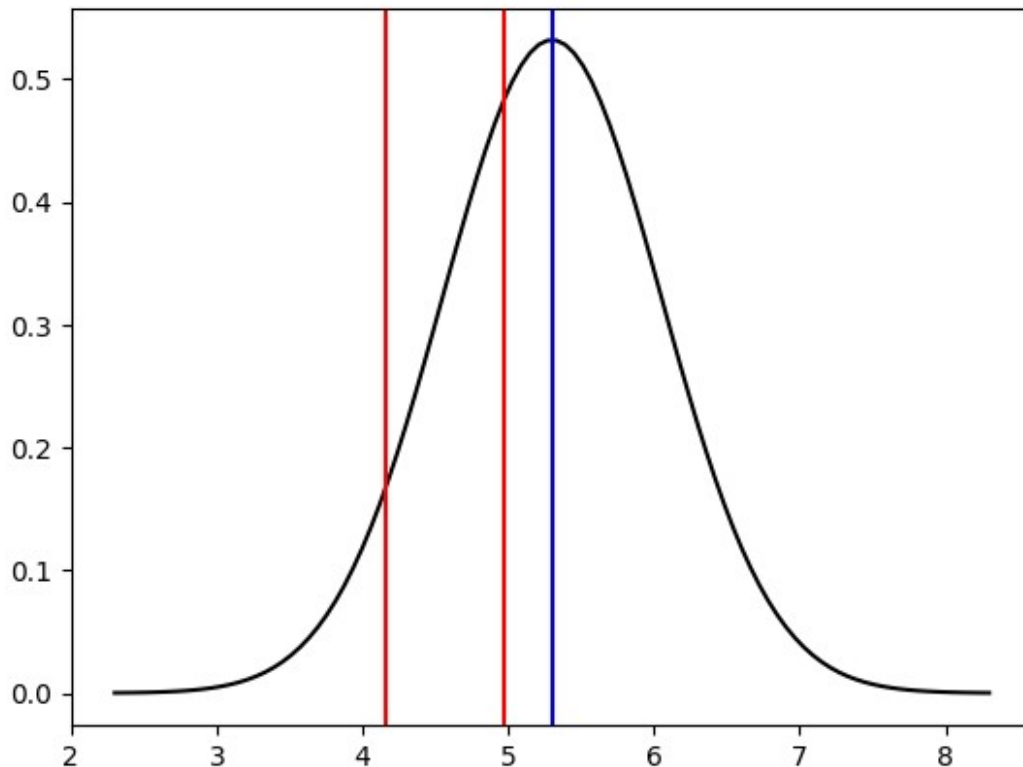
#Si, podemos afirmar que la media ha disminuido, ya que la media anterior queda fuera del intervalo de confianza de 97% y la media actual queda dentro del intervalo de confianza de 97%.



```
0.16770509831248423 error
(4.486064779880205, 5.213935220119795)
```



0.1875 error2
(4.153108054202894, 4.966891945797105)



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.

-a ¿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?

-b ¿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%?

El misterioso Helio Respuesta: 55, 94

```
sigma = 0.75
intervalo = 0.95
ancho_intervalo = 0.4
#Resolver para n
n = (2*sigma*stats.norm.ppf((intervalo+1)/2)/ancho_intervalo)**2
#menor a 0.4
print("n = ",np.ceil(n))

n = 55.0
```

```

sigma = 0.75
intervalo = 0.99
ancho = 0.2
n = (2*sigma*stats.norm.ppf((intervalo+1)/2)/ancho_intervalo)**2
print("n =", np.ceil(n))

n = 94.0

```

1. Con el archivo de datos de El Marcapasos Download 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

1. 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.

```

import pandas as pd
df =
pd.read_csv('/Users/guillermocepeda/C:C++/Implementacion_IA_a01284015/
Android/El_marcapasos.csv', sep=',')
print(df.head())

#separar en dos dataframes, ['Marcapasos'] == 1 y ['Marcapasos'] == 0
cm = df[df['Marcapasos'] == "Sin MP"]
sm = df[df['Marcapasos'] == "Con MP"]
print(cm.head())
pmean_cm = cm['Periodo entre pulsos'].mean()
pmean_sm = sm['Periodo entre pulsos'].mean()
print("Periodo Con marcapasos ", pmean_cm, " Periodo Sin marcapasos:
", pmean_sm)
imean_cm = cm['Intensidad de pulso'].mean()
imean_sm = sm['Intensidad de pulso'].mean()
print("Intensidad Con marcapasos ", imean_cm, " Intensidad Sin
marcapasos: ", imean_sm)

#Intervalos de confianza
#Intervalo de confianza para la media
desv_pcm = cm['Periodo entre pulsos'].std()
desv_psm = sm['Periodo entre pulsos'].std()
print("Desviación estándar Con marcapasos ", desv_pcm, " Desviación
estándar Sin marcapasos: ", desv_psm)
desv_icm = cm['Intensidad de pulso'].std()
desv_ism = sm['Intensidad de pulso'].std()
print("Desviación estándar Con marcapasos ", desv_icm, " Desviación
estándar Sin marcapasos: ", desv_ism)

intervalo = 0.95
n_cm = len(cm)
n_sm = len(sm)
error_pcm = desv_pcm/(n_cm**0.5)

```

```

error_psm = desv_psm/(n_sm**0.5)
intervalo_pcm = [pmean_cm - error_pcm*1.96,pmean_cm + error_pcm*1.96]
intervalo_psm = [pmean_sm - error_psm*1.96,pmean_sm + error_psm*1.96]

error_icm = desv_icm/(n_cm**0.5)
error_ism = desv_ism/(n_sm**0.5)
intervalo_icm = [imean_cm - error_icm*1.96,imean_cm + error_icm*1.96]
intervalo_ism = [imean_sm - error_ism*1.96,imean_sm + error_ism*1.96]

x_pcm = np.linspace(pmean_cm-2*desv_pcm,pmean_cm+2*desv_pcm,1000)
y = stats.norm(pmean_cm,desv_pcm).pdf(x_pcm)
x_psm = np.linspace(pmean_sm-2*desv_psm,pmean_sm+2*desv_psm,1000)
y = stats.norm(pmean_sm,desv_psm).pdf(x_psm)

plt.plot(x_pcm,y)
plt.axvline(intervalo_pcm[0],color='red')
plt.axvline(intervalo_pcm[1],color='red')
plt.axvline(pmean_cm,color='green')
plt.axvline(intervalo_psm[0],color='red')
plt.axvline(intervalo_psm[1],color='red')
plt.axvline(pmean_sm,color='green')
plt.title("Pulsos")
plt.show()
print("Media Sin marcapasos ",pmean_sm," Desviación estándar Sin
marcapasos: ",desv_psm, " Intervalo de confianza para la media:
",intervalo_psm)
print("Media Con marcapasos ",pmean_cm," Desviación estándar Con
marcapasos: ",desv_pcm, " Intervalo de confianza para la media:
",intervalo_pcm)

x_icm = np.linspace(imean_cm-2*desv_icm,imean_cm+2*desv_icm,1000)
y = stats.norm(imean_cm,desv_icm).pdf(x_icm)
x_ism = np.linspace(imean_sm-2*desv_ism,imean_sm+2*desv_ism,1000)
plt.plot(x_icm,y)
plt.axvline(intervalo_icm[0],color='red')
plt.axvline(intervalo_icm[1],color='red')
plt.axvline(imean_cm,color='green')
plt.axvline(intervalo_ism[0],color='red')
plt.axvline(intervalo_ism[1],color='red')
plt.axvline(imean_sm,color='green')
plt.title("Intensidad")
plt.show()
print("Media Sin marcapasos ",imean_sm," Desviación estándar Sin
marcapasos: ",desv_ism, " Intervalo de confianza para la media:
",intervalo_ism)
print("Media Con marcapasos ",imean_cm," Desviación estándar Con

```

```
marcapasos: ",desv_icm, " Intervalo de confianza para la media:
",intervalo_icm)
```

	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

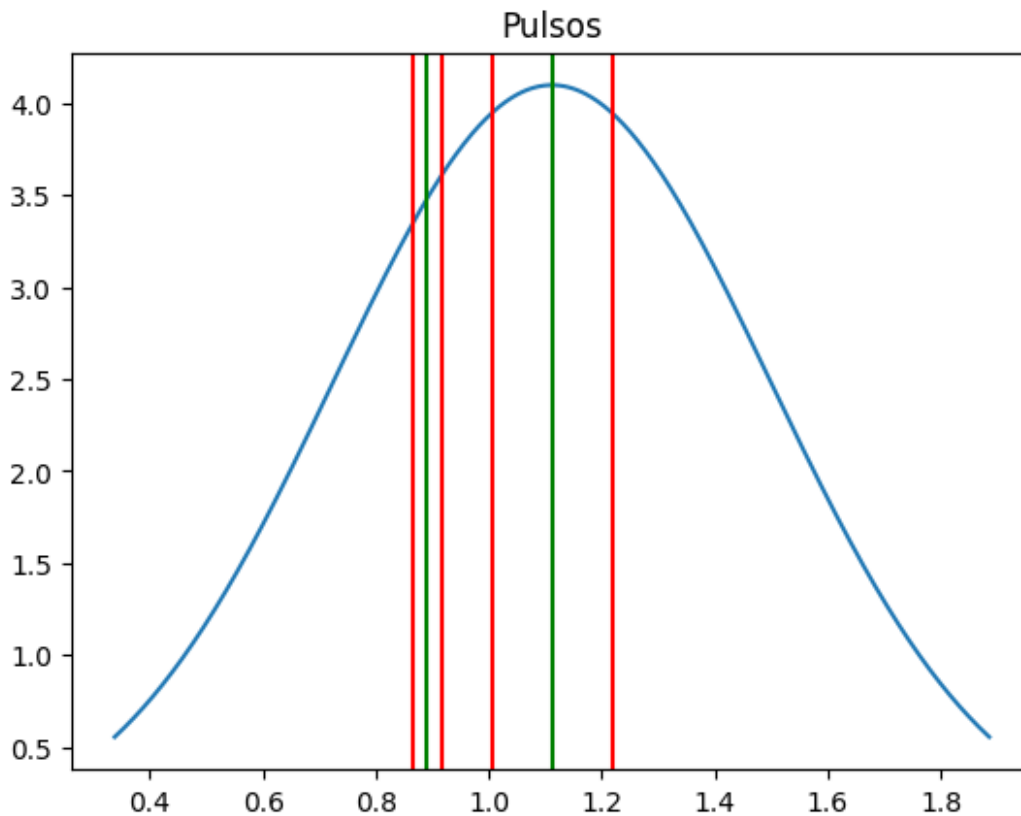
	Periodo entre pulsos	Intensidad de pulso	Marcapasos
0	1.2	0.131	Sin MP
1	0.9	0.303	Sin MP
2	0.9	0.297	Sin MP
3	0.8	0.416	Sin MP
4	0.7	0.585	Sin MP

Periodo Con marcapasos 1.111764705882353 Periodo Sin marcapasos: 0.8911764705882353

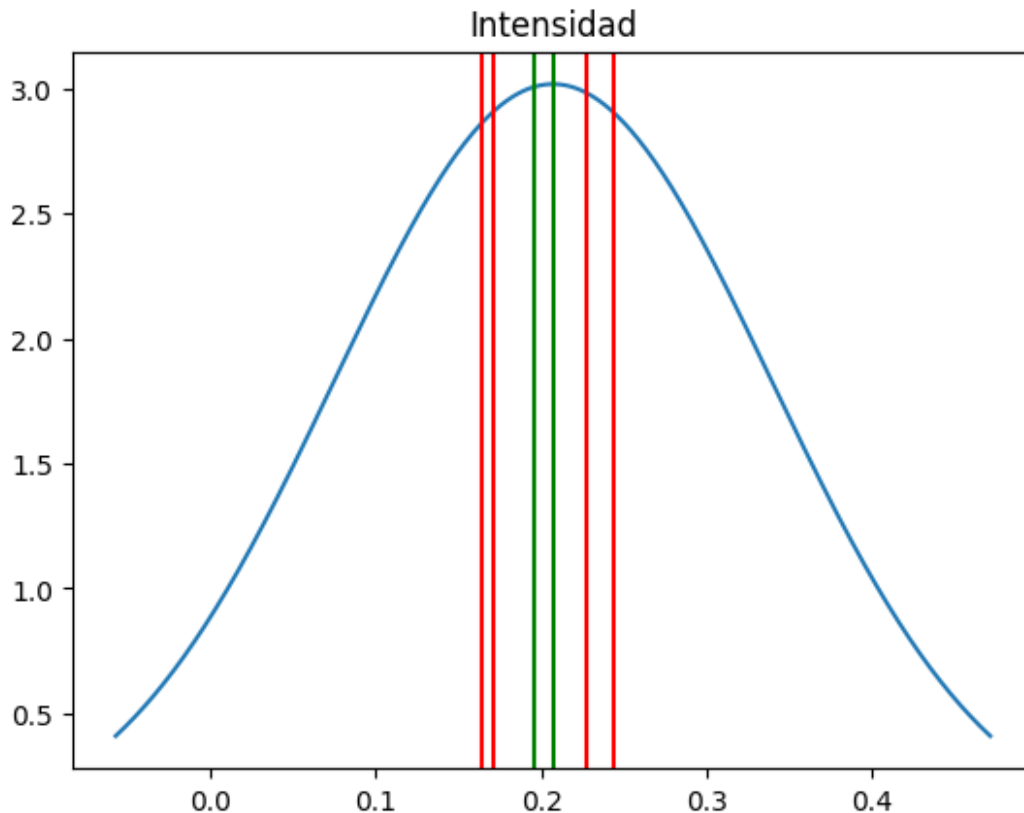
Intensidad Con marcapasos 0.20709803921568626 Intensidad Sin marcapasos: 0.19594117647058829

Desviación estándar Con marcapasos 0.3871160336764828 Desviación estándar Sin marcapasos: 0.09735804145161364

Desviación estándar Con marcapasos 0.13215101284545053 Desviación estándar Sin marcapasos: 0.11426537739222775



Media Sin marcapasos 0.8911764705882353 Desviación estándar Sin marcapasos: 0.09735804145161364 Intervalo de confianza para la media: [0.8644560794011604, 0.9178968617753103]
 Media Con marcapasos 1.111764705882353 Desviación estándar Con marcapasos: 0.3871160336764828 Intervalo de confianza para la media: [1.0055188149379943, 1.2180105968267116]



Media Sin marcapasos 0.19594117647058829 Desviación estándar Sin marcapasos: 0.11426537739222775 Intervalo de confianza para la media: [0.16458048414835724, 0.22730186879281933]
 Media Con marcapasos 0.20709803921568626 Desviación estándar Con marcapasos: 0.13215101284545053 Intervalo de confianza para la media: [0.1708285466668499, 0.2433675317645226]

El resultado que interpreto, es que mientras que los pulsos sin marcapasos, tienen un promedio, así como un intervalo de confianza fuera del de los que tienen Marca Pasos, la intensidad no se sale del intervalo del intervalo de confianza de los que sí tienen, por lo que se encuentra un cambio bastante grande en la población sin marcapasos cuando se habla de pulsos.