

# Tests estadísticos con SciPy

Iker Martín, Maribel Castillo y Vicente R. Tomás  
Noviembre 2023

# ÍNDICE

- Funciones estadísticas de Scipy.
- Comprendiendo los resultados.
- Casos prácticos.

# Funciones estadísticas de Scipy

- Veremos 4 tests diferentes

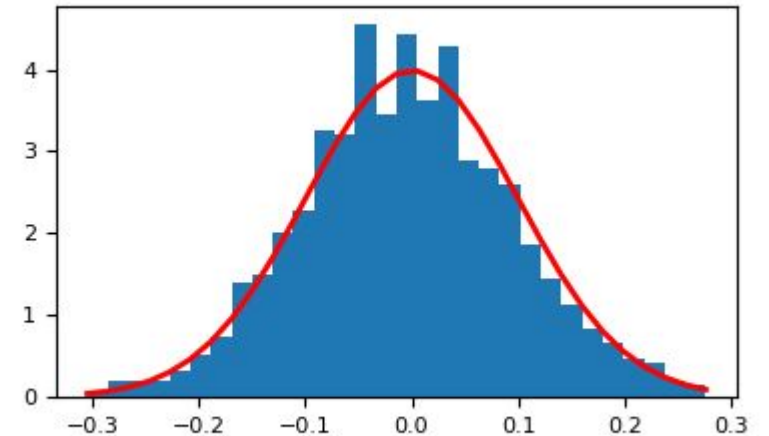
```
stats.shapiro(datos)

stats.levene(datos1, datos2, center='mean') # Distribuciones normales
stats.ttest_ind(datos1, datos2)

stats.levene(datos1, datos2, center='median') # Distribuciones no normales
stats.mannwhitneyu(datos1, datos2)
```

Tests para distribuciones normales

Tests para distribuciones no normales



# Comprendiendo los resultados

- Los anteriores tests estadísticos devuelven un valor P para comprobar que decisión tomar.
- Además es necesario elegir un valor de significancia ( $\alpha$ ). Usualmente 0,05.
- Si  $P > \alpha$ : se acepta la hipótesis nula ( $H_0$ ). En caso contrario se rechaza y se usa la hipótesis alternativa ( $H_1$ ).

Ejemplo:

$$P_x = 0.092$$

$P_x > \alpha \rightarrow$  Aceptar  $H_0 \rightarrow$  En Shapiro indica que la distribución es normal.

$$P_y = 0.026$$

$P_y > \alpha \rightarrow$  Rechazar  $H_0 \rightarrow$  En Shapiro indica que la distribución **no** es normal.

# Casos prácticos

```
[15]: import matplotlib.pyplot as plt
from scipy import stats
import numpy as np

def check_is_H0(p_value, alpha_value=0.05):
    if p_value > alpha_value:
        return True # Aceptar H0
    return False # Rechazar H0

def plot_data(data):
    fig = plt.figure() # Crear una figura
    fig.scatter(data) # Dibujar datos

# Valor de significancia
# Usualmente vale 0.01, 0.05, 0.10. Valores más bajos reducen el error.
alpha_value = 0.05
```

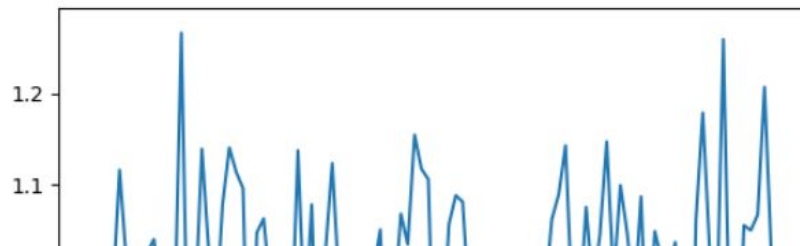
A continuación vamos a realizar una prueba para comprobar si los datos son normales

```
[10]: # Crear una lista con distribución normal
datos = np.random.normal(1, 0.1, 100)
total_datos = range(len(datos))

# Realiza el test de Shapiro-Wilk sobre los datos.
# Devuelve el test estadístico y el valor P.
st,p_value = stats.shapiro(datos)
normal_data = check_is_H0(p_value, alpha_value)
print("Los datos son normales: " + str(normal_data))
plt.plot(total_datos, datos)
```

Los datos son normales: True

```
[10]: [<matplotlib.lines.Line2D at 0x7fa64cd9f520>]
```



Iker Martín, Maribel Castillo y Vicente R. Tomás  
martini@uji.es; castillo@uji.es; vtomas@uji.es