

Tarea semana 4: Intervalo de confianza y prueba de hipótesis para la diferencia de medias

Problema: Stats Health

- La empresa Stats Health está desarrollando un análisis de usabilidad de un nuevo sistema informático de tele-rehabilitación para pacientes de artroplastia de cadera.
 - Vea: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8606428
- Stats Health ha recopilado un dataset acerca de la usabilidad de la interfaz del sistema que contiene:
 - Variables de uso de la aplicación: Time, y la escala de usabilidad dada por el usuario: SUS.
 - Variables demográficas del usuario, Género (Gender), Edad (Age), entre otras.

Archivo de datos:

- DL_data.csv
- Lea el archivo csv y ordene las variables categóricas de una manera adecuada.

■ Encabezado de nuestro datasest.

	SUS	Time	ID	Gender	Age	Kinect	NUI	Games
0	42.5	179.3	student	М	twenty	high	average	high
1	92.5	153.3	student	М	twenty	high	high	high
2	95	135.7	student	М	twenty	average	average	high
3	92.5	148.3	student	F	twenty	average	high	high
4	80	148	student	М	twenty	average	average	high

Archivo de datos:

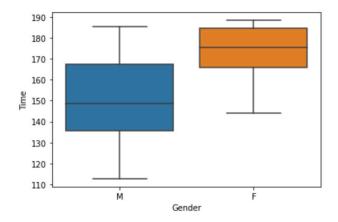
- SUS (System Usability Scale): Escala de Usabilidad del Sistema.
 - Un valor entre 0 y 100, dado por el usuario al evaluar el sistema informático.
 - Ver: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html
- Time: Tiempo en que el usuario realizó la prueba del sistema.
- Gender: M (Male, masculino), F (Female, Femenino).
- Age: twenty (veinte añeros), thirty (treinta añeros), forty (cuarenta añeros).

Análisis a realizar:

- Quiere determinarse si existe alguna diferencia entre la ejecución de las tareas en el sistema de telerehabilitación entre hombres y mujeres.
 - De existir alguna diferencia deben identificarse y realizarse las modificaciones necesarias en el sistema o en la selección de tareas (ejercicios) a realizar para que no exista tal diferencia.

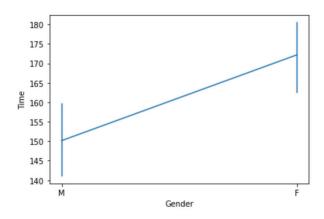
Descripción de los datos

- 1. **.(2 pt)** Escriba el código Python para construir un diagrama de cajas que muestre el género (**Gender**) en x y el tiempo de ejecución de la tarea (**Time**) en y.
- **Discuta:** Observa diferencia en el tiempo de ejecución de la tarea en sistema (Time) realizada por los usuarios de acuerdo al género (Gender). Realice el análisis con respecto a la información mostrada en el diagrama de cajas (resumen de 5 puntos, posición (de los cuartiles), dispersión (IQR, bigotes), etc.).



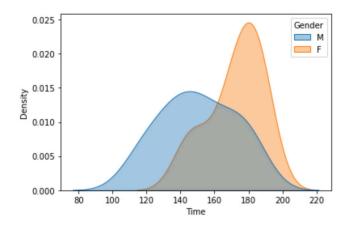
2. **(2 pt)** Escriba el código Python para construir un diagrama de líneas con barras de error (CI: 0.95) que muestre el género (**Gender**) en x y el tiempo de ejecución de la tarea (**Time**) en y.

■ **Discuta:** Observa diferencia en el tiempo de ejecución de la tarea en sistema (Time) por los usuarios de acuerdo al género (Gender).



3. **(2 pt)** Escriba el código Python para construir un diagrama de densidad que muestre el género (**Gender**) en *x* y el tiempo de ejecución de la tarea (**Time**) en *y*.

■ **Discuta:** Observa diferencia en el tiempo de ejecución de la tarea en sistema (Time) por los usuarios de acuerdo al género (Gender).



Prueba de hipótesis para la diferencia de medias

4. (2 pt) Realice una prueba de dos varianzas para los grupos masculino y femenino (Gender).

- Escriba (reutilce) el código Python.
- **Discuta** los resultados.

5.(2 pt) A partir de la prueba anterior realice una prueba de diferencias de medias para los grupos masculino y femenino (Gender).

- Escriba (reutilce) el código Python.
- Discuta los resultados.