

#### Tarea 4 de Probabilidad y Estadística aplicada

El 15 de abril de 1912, durante su viaje inaugural, el Titanic se hundió tras chocar con un iceberg, matando a 1.502 de los 2.224 pasajeros y tripulantes. Se necesitaron alrededor de 7,5 millones de dólares para construir el Titanic y se hundió bajo el océano debido a una colisión. Nosotros procesaremos algunos datos recreados para poder encontrar evidencia estadística interesante sobre las personas que iban a bordo durante la tragedia, tanto los que tuvieron la suerte de salvarse como las que no contaron con ella.

Para esto se subirán los datos a websignatura en un archivo llamado `titanik.csv` que deberá ser procesado en Python.

Además para este trabajo, se recomienda investigar en la librería `scipy.stats` los métodos

- `ttest_ind`
- `ttest_1samp`

Para la realización de pruebas T sobre las muestras que se van a analizar.

#### Parte 1: Estadística Descriptiva

1. Obtener los datos proporcionados en el archivo `titanik.csv` en un dataframe en Python.
2. Corregir los valores vacíos de la columna “age” usando la media de las edades según el género de los pasajeros. Justificar por qué esta es una estrategia razonable para la corrección de los datos faltantes.
3. Calcular la media, mediana, moda, rango, varianza y desviación estándar de las edades.
4. ¿Cuál es la tasa de supervivencia general?
5. ¿Cuál es la tasa de supervivencia por género?
6. Realizar un histograma de las edades de los pasajeros por clase (primera, segunda y tercera). Proponga un modelo para la distribución de la variable edad en el barco.
7. Realizar un diagrama de cajas para las edades de los supervivientes, y otro para las edades de los no supervivientes. ¿Es posible extraer alguna conclusión?

#### Parte 2: Inferencia Estadística

1. Construir un intervalo de confianza, con confianza 95 %, para la edad promedio de las personas en el barco.
2. A partir de los datos de la muestra, con una certeza del 95 % ¿Es posible afirmar que el promedio de edad de las mujeres interesadas en abordar el Titanic es mayor a 56 años? ¿Es posible afirmar lo mismo para los hombres?
3. A partir de los datos de la muestra, con una certeza del 99 % ¿Existe una diferencia significativa en la tasa de supervivencia entre hombres y mujeres? ¿Existe una diferencia significativa en la tasa de supervivencia en las distintas clases?

4. A partir de los datos de la muestra, con una certeza del 95 % ¿Es posible afirmar que en promedio las mujeres eran más jóvenes que los hombres en el barco?

**Sobre el informe:**

- El tiempo para entregar el informe es hasta el sábado 6 de julio inclusive. La entrega se realizará por webasignatura.
- El informe deberá estar en formato pdf, la entrega también deberá incluir los scripts utilizados.
- El informe deberá contener título, fecha, nombre y cédula de los estudiantes.
- Se evaluará: prolijidad del informe, utilización correcta del idioma español, redacción, prolijidad del código presentado en los scripts, conclusiones.