

Ejercicio de Evaluación. Descriptiva e Inferencia.

Ejercicio 1. El fichero Excel adjunto contiene, dos variables: la primera es dicotómica con valores 1 (predinástico temprano) y 2 (predinástico tardío) y la segunda contiene la altura de cráneos (mm.) encontrados en un yacimiento arqueológico. La idea es analizar si existen diferencias en la longitud de la altura de los cráneos egipcios a medida que pasa el tiempo. Creo que mayoritariamente tenemos una idea de que las cabezas egipcias son más alargadas y cuando ya llegamos a los romanos son más redondeadas. El cine se ha encargado de hacer muy gráfico todo esto.

Se pide:

- a) obtener con Python las diferentes medidas de centralización y dispersión, asimetría y curtosis estudiadas. Así mismo, obtener el diagrama de caja y bigotes. Se debe hacer por separado para la sub-muestra de los cráneos del predinástico temprano y para la sub-muestra de los del predinástico tardío. Comentar los resultados obtenidos. Estos comentarios son obligatorios
- b) Determinar si cada una de las dos sub-muestras sigue una distribución normal utilizando el test de Kolmogorov-Smirnov.

Ejercicio 2. a) Con los mismos datos del ejercicio anterior, obtener un intervalo de confianza (de nivel 0.9, de nivel 0.95 y de nivel 0.99) para la diferencia entre las medias de la altura de la cabeza en ambos periodos históricos. Interpretar los resultados obtenidos y discutirlos en función del test de normalidad del ejercicio anterior. La interpretación debe ser rigurosa desde el punto de vista estadístico y también marcada por el story telling, es decir, comprensible desde el punto de vista de las

variables respondiendo a la pregunta ¿en qué época la cabeza era más alta?

b) Utilizar el test t para contrastar la hipótesis de que ambas medias son iguales. Explicar qué condiciones se deben cumplir para poder aplicar ese contraste. **Determinar si se cumplen.** Admitiremos de forma natural la independencia entre ambas muestras, así que esa condición no hace falta comprobarla.

Observación: Quiero insistir en que debéis hacer el test t para la diferencia de medias aunque las condiciones no se cumplan. En ese caso discutir la validez de los resultados obtenidos.

Entonces, por favor, todos me debéis entregar dos ficheros, uno Python con las instrucciones utilizadas en la ejecución del trabajo y otro pdf, html o Word en el que se incluyan los comentarios a los resultados obtenidos y las interpretaciones de estos.