

Desafío - Inferencia e hipótesis

En este desafío validaremos nuestros conocimientos de estadística inferencial. Para lograrlo, realizarás inferencias y pruebas de hipótesis a partir de la base de datos **earnings.csv**.

Lee todo el documento antes de comenzar el desarrollo **individual**, para asegurarte de tener el máximo de puntaje y enfocar bien los esfuerzos.

Tiempo asociado: 2 horas cronológicas

Descripción

Podemos registrar muchas características de una persona, y buscar ver si algunas influyen o no en otras. Realiza las siguientes actividades para verificarlo.

1. Considerando las variables **earn**, **height** y **age**, vamos a suponer que nuestro dataset fuera una población completa. (Prepara los datos adecuadamente).
 - a. Realiza 5 pruebas con muestras de tamaño $n = 25$, en cada caso, con niveles de significancia diferentes pero inferiores a 0,1, para confirmar o rechazar las siguientes hipótesis considerando las alternativas. Crea para ello una fórmula que reciba los parámetros adecuados y responda "con nivel de confianza del ...%, se acepta la hipótesis nula/alternativa".

earn	height	age
$H_0: \bar{x} = 20000$ $H_1: \bar{x} \neq 20000$	$H_0: \bar{x} = 55$ $H_1: \bar{x} \neq 55$	$H_0: \bar{x} = 44$ $H_1: \bar{x} \neq 44$
$H_0: \bar{x} = 35000$ $H_1: \bar{x} \neq 35000$	$H_0: \bar{x} = 68$ $H_1: \bar{x} \neq 68$	$H_0: \bar{x} = 56$ $H_1: \bar{x} \neq 56$
$H_0: \bar{x} = 25000$ $H_1: \bar{x} > 25000$	$H_0: \bar{x} = 63$ $H_1: \bar{x} > 63$	$H_0: \bar{x} = 51$ $H_1: \bar{x} > 51$
$H_0: \bar{x} = 32000$ $H_1: \bar{x} < 32000$	$H_0: \bar{x} = 70$ $H_1: \bar{x} < 70$	$H_0: \bar{x} = 62$ $H_1: \bar{x} < 62$

- b. Calcula la media poblacional para cada variable. ¿Son correctos los resultados obtenidos por tus pruebas?

2. Crea una función que estime la proporción de hombres en el dataset, considerando una muestra de 50 individuos, con niveles de significancia de 0,05 y 0,01. Escribe con palabras tus resultados ejecutar la función e interpreta. Compara con la proporción real.
3. Considerando el dataset como una muestra, verifica si el género (male=1 significa 'hombre') influye sobre el sueldo "earn" de las personas. Explica y justifica tu procedimiento.

Requerimientos

1. Realiza pruebas de hipótesis, considerando el enunciado de cada una y las implementa en Python mediante funciones **(5 Puntos)**
2. Infiere sobre medias poblacionales y proporciones, interpretando error y significancia. **(2 Puntos)**
3. Plantea, interpreta e implementa hipótesis de prueba para muestras independientes. **(3 Puntos)**



¡Mucho éxito!

Consideraciones y recomendaciones

Debes entregar tu trabajo en un archivo de Jupyter Notebook, con el desarrollo de tu trabajo y las explicaciones necesarias de tu procedimiento.

Recuerda explicar y explicitar los pasos realizados en tu procedimiento.