

Desafío - Pruebas de hipótesis

- Para poder realizar este desafío debes haber revisado la lectura y videos correspondiente a la unidad.
- Crea una carpeta de trabajo y guarda todos los archivos correspondientes (notebook y csv).
- Una vez terminado el desafío, comprime la carpeta y sube el `.zip`

1. Evaluar juicios de hipótesis

A continuación se presenta una serie de enunciados de hipótesis. Usted debe discernir si es posible rechazar la hipótesis nula.

- 75 individuos elegidos al azar fueron alcanzados por activistas pro-LGTB que se identificaron como homosexuales, y 75 alcanzados por activistas pro-LGTB que se identificaron como heterosexuales. El objetivo era medir actitudes frente a la adopción homoparentales. La organización encargada de procesar los datos obtuvo los siguientes resultados:
 - El 67% de los encuestados por heterosexuales se mostró a favor de la adopción homoparental, mientras que un 72% de los encuestados por activistas que se identificaron como homosexuales se mostraron a favor de la adopción homoparental.
 - Asumiendo que la hipótesis nula es que ambos porcentajes no son diferentes, y la hipótesis alternativa es que existe una diferencia substancial entre ambos.
 - Con un valor de prueba $p = 11.83$ y un puntaje de corte de 2.58, ¿qué se debe concluir?
- Un investigador de la Sociedad de Abstemios de Chile sugiere que los hombres tienen mayores niveles de consumo de alcohol que las mujeres, siendo esta diferencia estadísticamente significativa. Esto en base a su estudio realizado con 48 individuos, donde contrastó su puntaje de prueba a una distribución asintóticamente normal. ¿Qué se puede decir sobre su estudio?

2. Importe la base de datos utilizada la sesión presencial anterior

- Agregue una serie de variables binarias por cada continente de la variable `region`.
 - *tip*: Utilice `np.where` para ello.
- De manera similar a la sesión anterior, enfóquese en las siguientes variables:
 - Apellidos desde la A hasta la N: Enfocarse en las variables `chldmort`, `adfert` y `life`.
 - Apellidos desde la N hasta la Z: Enfocarse en las variables `femlab`, `literacy` y `school`.

3. Implemente una función de prueba de hipótesis a mano

- La función debe ingresar los siguientes argumentos:
 - `df` : La tabla de datos.
 - `variable` : La variable a analizar.
 - `binarize` : El indicador binario a utilizar.
- *tips*:
 - Separe la variable en dos, utilizando el indicador binario. Recuerde eliminar los perdidos con `dropna()`.
 - Implemente `ttest_ind` de `scipy.stats` y guarde el valor `t` y `pval`.
 - Reporte las medias para cada grupo (0 y 1).
 - Reporte la diferencia de entre las medias.
 - Reporte los valores `t` y `p`

4. Implemente una función que grafique los histogramas para ambas muestras.

- Genere una función que devuelva un gráfico donde visualice las dos histogramas cuando la variable es 1 y 0, respectivamente.
 - *tips*: Refactorize la función incluyendo el método `hist` de `matplotlib.pyplot`. Incluya los argumentos `alpha` y `label`.
- Para las tres variables de interés acorde a su grupo, analice las diferencias de medias por cada continente, y posteriormente grafique. Concluya con los principales resultados al respecto.