

# Trabajo Práctico 4 en R

## Test de hipótesis

En este link podrán acceder a los datos del test AB sobre lo que trabajemos en esta actividad.

1. Recordemos el problema del test AB y supongamos que tomaremos la decisión (si implementar o no la nueva versión) en base al tiempo de permanencia medio en la página. Supongamos, también, que el tiempo de permanencia medio con la versión actual es conocido e igual a 60 segundos.

Es decir, el responsable de la página de “DormiTown” quiere determinar si el tiempo de permanencia medio en la página por sesión con la nueva versión es mayor a 60 segundos, en cuyo caso cambiará la versión.

Para ello, selecciona una muestra de visitantes a los que les enseña la nueva versión y registra el tiempo de permanencia en la página de cada uno de ellos. En base a estos datos realizará un test para decidir si cambia o no la versión de la página. Suponga que la tienda está dispuesta a correr un riesgo del 1 % de cambiar la versión actual cuando en realidad el tiempo de permanencia medio con la nueva versión es igual al de la actual.

- a) Defina (con palabras) el parámetro de interés.
  - b) Plantee las hipótesis nula y alternativa del test que debe realizar el responsable de la página.
  - c) ¿Cuál es el nivel del test?
  - d) Describa las variables de la muestra aleatoria, es decir, defínalas con palabras y especifique todo lo que sabe de su distribución.
2. Para simular el experimento que haría el responsable de la página, baje un conjunto con 30 datos de la versión nueva (versión B) usando su DNI, impórtelo a R en un data frame llamado `datos` y luego guarde la variable `tiempo` de dicho data frame en un vector llamado `tiempos`.
    - a) Calcule la media muestral del tiempo de permanencia en la muestra. ¿Es mayor que 60 segundos?
    - b) Realice un boxplot de los datos, ¿qué observa?
    - c) Realice un histograma de `tiempos` y superponga la curva normal con los parámetros que considere adecuados para evaluar si es razonable asumir que los datos son normales.

3. En base al ítem anterior, ¿diría que la distribución de los datos es normal?
4. En base a la respuesta del ítem anterior, ¿cuál de los test vistos hasta ahora realizaría? Para responder esta pregunta...
  - a) Indique cuál es el estadístico del test, qué distribución tiene bajo  $H_0$  y si ésta es exacta o asintótica. Justifique.
  - b) Plantee la región de rechazo.
5. En base a la muestra que guardó en **tiempos**, indique la conclusión del test en el contexto del problema y diga qué decisión tomaría respecto de cambiar o no la versión de la página.
6. Calcule el p-valor de la muestra.
7. Sin hacer más cuentas, indique qué decisión tomaría a nivel 5 %. Justifique.