

Taller de práctica 3

Prueba de hipótesis para la diferencia de medias con
varianzas poblacionales conocidas, y varianzas poblacionales desconocidas e iguales
(secciones 2.3.1 y 2.3.2)

1. Las base de datos `Imacec+Comercio+Servicio.csv` (link) contiene datos acerca del Imacec de distintos sectores, en particular las columnas son las siguientes:
 - **Imacec**: Corresponde al Imacec general trimestral (nos referiremos a esta variable como **Imacec general**).
 - **Comercio**: Corresponde al Imacec trimestral del área de Comercio.
 - **Servicios**: Corresponde al Imacec trimestral del área de Servicios.
 - **Imacecnominero**: Corresponde al Imacec trimestral fuera del área de minería (y que no recae en las otras categorías).
 - **Trimestre**: Corresponde al número del trimestre en el que se realiza la medición.
 - **Ano**: Corresponde al año de medición.
 - a) Estudie si, en el periodo de años 2015-2022, existe diferencia real entre el la media del Imacec de Comercio y la media del Imacec no Minero, con varianzas poblacionales de 117 y 56 respectivamente. Utilice una confianza del 97.31 %, además, asuma que los datos tienen distribución normal. Para concluir, utilice el método del valor-p y el método del valor crítico.
 - b) Estudie si, para el primer trimestre de todos los años, el promedio del Imacec de Servicios es menor al promedio del Imacec no Minero. Utilice una confianza del 94.8 %, y considere varianzas poblacionales desconocidas e iguales, además, asuma que los datos tienen distribución normal. Para concluir, utilice el método del valor crítico.
 - c) Estudie si, durante el segundo y tercer trimestre, desde el año 2013 hasta el año 2019, el promedio del Imacec no Minero es mayor al promedio del Imacec de Servicios por más de 0.05 puntos, considerando varianzas poblacionales desconocidas e iguales. Utilice una confianza del 89 %, y asuma que los datos tienen una distribución normal. Para concluir, utilice el método del valor-p y el método del intervalo de confianza.
2. Las base de datos `monedas+billetes+circulacion2.csv` (link) contiene datos acerca del total de dinero circulante, diferenciando por billete. Las columnas en la base de datos son las siguientes.
 - **Ano**: Corresponde al año de medición.
 - **Mes**: Corresponde al número del mes en el que se realiza la medición.
 - **Total**: Corresponde al total de billetes circulando, en millones de pesos.
 - **VeinteMil**: Corresponde al total de billetes de \$20000 circulando, en millones de pesos.
 - **DiezMil**: Corresponde al total de billetes de \$10000 circulando, en millones de pesos.
 - **CincoMil**: Corresponde al total de billetes de \$5000 circulando, en millones de pesos.
 - **DosMil**: Corresponde al total de billetes de \$2000 circulando, en millones de pesos.
 - a) Estudie si, para los meses del primer semestre, el promedio de la cantidad de billetes de \$20000 es mayor al promedio de la cantidad de billetes de \$10000 por más de \$21300 (millones de pesos), con desviaciones estándar poblacionales de 2540462 y 1159586 respectivamente. Utilice una confianza del 97 %, además, asuma que los datos tienen distribución normal. Para concluir, utilice el método del valor-p y el método del intervalo de confianza.
 - b) Estudie si, durante los meses del segundo semestre, desde el año 2018 hasta el año 2022, la diferencia entre el promedio de la cantidad de billetes de \$5000 y el promedio de la cantidad de billetes de \$2000 es menor a \$503000 (millones de pesos). Utilice una confianza del 94 %, y considere varianzas poblacionales desconocidas e iguales, además, asuma que los datos tienen distribución normal. Para concluir, utilice el método del valor-p y el método del intervalo de confianza.