

Taller de práctica Prueba 1

Parte 1: Intervalos de confianza mediante código de R

La siguiente base de datos contiene datos sobre el PIB (expresado en miles de millones de pesos) de las distintas regiones de Chile en el periodo de años 2013 - 2021. El detalle de las columnas es el siguiente:

- **Reg:** numeración de la región.
- **Region:** nombre de la región.
- **X...:** PIB regional en el año determinado (ejemplo: X2013 corresponde al PIB regional del año 2013).

```
datos = read.csv("https://raw.githubusercontent.com/Dfranzani/Bases-de-datos-para-cursos/main/2023-1/PIB%202013-2021.csv")
str(datos)

## 'data.frame': 16 obs. of 11 variables:
## $ Reg : int 1 2 3 4 5 6 7 8 9 10 ...
## $ Region: chr "Región de Arica y Parinacota" "Región de Tarapacá" "Región de Antofagasta" "Región de Coquimbo" "Región de Valparaíso" "Región de Santiago" "Región de Biobío" "Región de Los Ríos" "Región de Magallanes" "Región de Aysén" "Región de Los Lagos" "Región de Los Andes"
## $ X2013 : num 1303 3943 15239 3744 5504 ...
## $ X2014 : num 1328 4066 15966 3753 5379 ...
## $ X2015 : num 1363 4048 16003 3675 5347 ...
## $ X2016 : num 1397 3962 15334 3824 5453 ...
## $ X2017 : num 1438 4082 14915 3791 5552 ...
## $ X2018 : num 1458 4346 15984 3698 5780 ...
## $ X2019 : num 1457 4544 16026 3539 5892 ...
## $ X2020 : num 1375 4450 15977 3456 5675 ...
## $ X2021 : num 1561 4811 16203 3913 6160 ...
```

1. Para los siguientes códigos en R realice lo siguiente:

- Señale el tipo de intervalo de confianza y exprese la probabilidad asociada a este.
- Explique (y justifique cuando corresponda) cada uno de los valores para los comandos de R utilizados para construir el intervalo de confianza (incluyendo los filtros).
- Interprete la salida de cada uno de los intervalos de confianza, dando una conclusión según la probabilidad planteada anteriormente.

a) Intervalo de confianza:

```
t.test(x = datos$X2013, alternative = "two.sided", conf.level = 0.971)

##
## One Sample t-test
##
## data: datos$X2013
## t = 2.2598, df = 15, p-value = 0.03914
## alternative hypothesis: true mean is not equal to 0
## 97.1 percent confidence interval:
## -653.6879 19774.0979
## sample estimates:
## mean of x
## 9560.205
```

Nota: considere que el punto de comparación en la probabilidad es 0.

b) Intervalo de confianza: (Trabajo independiente)

```

t.test(
  x = datos$X2017,
  y = datos$X2020,
  alternative = "less",
  conf.level = 0.97,
  var.equal = T
)

##
## Two Sample t-test
##
## data:  datos$X2017 and datos$X2020
## t = 0.020873, df = 30, p-value = 0.5083
## alternative hypothesis: true difference in means is less than 0
## 97 percent confidence interval:
##      -Inf 12380.87
## sample estimates:
## mean of x mean of y
## 10212.59 10081.77

```

Nota: considere que el punto de comparación en la probabilidad es 11430.

c) Intervalos de confianza:

```

var.test(
  x = datos$X2015[datos$Reg %in% c(1:5,8:10)],
  y = datos$X2019[datos$Reg %in% c(1:5,8:10)],
  conf.level = 0.81
)

##
## F test to compare two variances
##
## data:  datos$X2015[datos$Reg %in% c(1:5, 8:10)] and datos$X2019[datos$Reg %in% c(1:5, 8:10)]
## F = 1.0094, num df = 7, denom df = 7, p-value = 0.9905
## alternative hypothesis: true ratio of variances is not equal to 1
## 81 percent confidence interval:
##  0.3537856 2.8797225
## sample estimates:
## ratio of variances
##          1.009358

```

```

t.test(
  x = datos$X2015[datos$Reg %in% c(1:5,8:10)],
  y = datos$X2019[datos$Reg %in% c(1:5,8:10)],
  alternative = "less",
  conf.level = 0.81,
  var.equal = T
)

##
## Two Sample t-test
##
## data:  datos$X2015[datos$Reg %in% c(1:5, 8:10)] and datos$X2019[datos$Reg %in% c(1:5, 8:10)]
## t = -0.12606, df = 14, p-value = 0.4507
## alternative hypothesis: true difference in means is less than 0
## 81 percent confidence interval:
##      -Inf 1792.44

```

```
## sample estimates:
## mean of x mean of y
## 5792.401 6081.915
```

Nota: considere que el punto de comparación en la probabilidad es 2000.

Parte 2: Cálculo de intervalos de confianza

La siguiente base de datos contiene datos sobre el PIB (expresado en miles de millones de pesos) de las distintas regiones de Chile en el periodo de años 2013 - 2021. El detalle de las columnas es el siguiente:

- **Reg:** numeración de la región.
- **Region:** nombre de la región.
- **X...**: PIB regional en el año determinado (ejemplo: X2013 corresponde al PIB regional del año 2013).

Con la siguiente salida es de 10 valores aleatorios de la base de datos. Asumiendo que las variables tiene distribución normal y son independientes, realice lo siguiente:

```
head(datos[1:10, c(3,4)], 10)
```

```
##      X2013    X2014
## 1  1302.85  1328.39
## 2  3943.47  4065.96
## 3 15238.75 15965.84
## 4   3743.65  3753.41
## 5   5504.03  5379.06
## 6 13572.06 13220.14
## 7  71021.42 72231.05
## 8   6823.17  7012.54
## 9   5904.18  5903.31
## 10 1981.16 1989.79
```

1. Un intervalo de confianza para estudiar si la media del 2014 es menor a 13000 miles de millones de pesos. Utilice un nivel de significancia del 5 %, y considere una desviación estándar poblacional de 17210.14 (considere que $Z_{0.05} = -1.6448$ y $Z_{1-0.05} = 1.6448$).
2. Un intervalo de confianza para estudiar si la media del 2013 es mayor a 9000 miles de millones de pesos. Utilice un nivel de significancia del 10 %, y considerando una varianza poblacional de 453794323 (considere que $Z_{1-0.1} = 1.2815$).
3. (Trabajo independiente) Un intervalo de confianza para estudiar si la media del 2013 es mayor a 1420 miles de millones de pesos. Utilice un nivel de significancia del 12 %, y considere una desviación estándar poblacional de 20929.41 (considere que $Z_{1-0.12} = 1.1749$).
4. (Trabajo independiente) Un intervalo de confianza para estudiar si la media del 2014 es distinta a 28090 miles de millones de pesos. Utilice un nivel de significancia del 8 %, y considerando una desviación estándar poblacional de 21302.45 (considere que $Z_{1-0.04} = 1.7506$).
5. Un intervalo de confianza para estudiar si la media del 2013 es menor a 24100 miles de millones de pesos. Utilice un nivel de significancia del 7 %, y considere una varianza poblacional de 438040402 (considere que $Z_{0.07} = -1.4757$ y $Z_{1-0.07} = 1.4757$).
6. (Trabajo independiente) Un intervalo de confianza para estudiar si la media del 2013 es distinta a 18500 miles de millones de pesos. Utilice un nivel de significancia del 4 %, y considerando una varianza poblacional de 453794323 (considere que $Z_{1-0.02} = 2.053749$ y).
7. (Trabajo independiente) Un intervalo de confianza para estudiar si la media del 2013 es diferente a la media del 2014, considerando desviaciones estándar poblacionales de 16922.05 y 17210.14 respectivamente. Utilice un nivel de significancia del 5 % (considere que $Z_{1-0.05/2} = 1.9599$).

8. (Trabajo independiente) Un intervalo de confianza para estudiar si la media del 2013 es mayor a 9000 miles de millones de pesos. Utilice un nivel de significancia del 10 % (asuma que $t_{1-0.1, k=9} = 1.3830$). Nota: el promedio y varianza de la muestra se deben calcular manualmente con los datos de la base dados en el enunciado.
9. Un intervalo de confianza para estudiar si la diferencia entre media del 2013 y la media del 2014 es menor a 5 miles de millones de pesos. Utilice un nivel de significancia del 3 %. Considere varianzas poblacionales desconocidas e iguales (considere que $t_{0.03, k=18} = -2.0070$ y $t_{1-0.03, k=18} = 2.0070$).
10. (Trabajo independiente) Un intervalo de confianza para estudiar si la diferencia entre media del 2013 y la media del 2014 es mayor a 5 miles de millones de pesos. Utilice un nivel de significancia del 7 %. Considere varianzas poblacionales desconocidas y distintas (considere que $t_{0.07, k=18} = -1.5439$ y $t_{1-0.07, k=18} = 1.5439$).
11. (Trabajo independiente) Un intervalo de confianza para estudiar si la media del 2014 es menor a 25000 miles de millones de pesos. Utilice un nivel de significancia del 5 %, y considere una desviación estándar poblacional de 17210.14 (considere que $Z_{0.05} = -1.6448$ y $Z_{1-0.05} = 1.6448$).
12. (Trabajo independiente) Un intervalo de confianza para estudiar si la media del 2014 es distinta a 28090 miles de millones de pesos. Utilice un nivel de significancia del 8 %, y considerando una desviación estándar poblacional de 21302.45 (considere que $Z_{1-0.04} = 1.7506$).