

Taller de práctica Prueba 2

Pruebas de hipótesis para la diferencia de medias, y comparación de varianzas

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
## $ 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 ...
```

A continuación, se presentan diversos códigos asociados a pruebas de hipótesis; especifique: la(s) media(s) (y las varianzas cuando corresponda), las hipótesis (H_0 y H_1), el tipo de prueba, la(s) regla(s) de rechazo a utilizar y la confianza. Interprete el resultado de la prueba de hipótesis utilizando tanto método del valor p como el método del intervalo de confianza.

1. Prueba de hipótesis:

```
t.test(x = datos$X2013, y = datos$X2021, alternative = "greater", conf.level = 0.971,
       mu = 1102, var.equal = F)

##
## Welch Two Sample t-test
##
## data: datos$X2013 and datos$X2021
## t = -0.41806, df = 29.354, p-value = 0.6605
## alternative hypothesis: true difference in means is greater than 1102
## 97.1 percent confidence interval:
## -14397.58 Inf
## sample estimates:
## mean of x mean of y
## 9560.205 11168.414
```

2. Prueba de hipótesis::

```
t.test(x = datos[3,-c(1,2)], y = datos[5,-c(1,2)], alternative = "less",
       conf.level = 0.891, mu = 12300, var.equal = T)
```

```
##
## Two Sample t-test
##
## data: datos[3, -c(1, 2)] and datos[5, -c(1, 2)]
## t = -12.585, df = 16, p-value = 5.147e-10
## alternative hypothesis: true difference in means is less than 12300
## 89.1 percent confidence interval:
##      -Inf 10324.69
## sample estimates:
## mean of x mean of y
## 15738.580 5637.999
```

3. Prueba de hipótesis:

```
t.test(
  x = datos$X2017,
  y = datos$X2020,
  alternative = "less",
  conf.level = 0.97,
  mu = 11430,
  var.equal = T
)

##
## Two Sample t-test
##
## data: datos$X2017 and datos$X2020
## t = -1.8029, df = 30, p-value = 0.04073
## alternative hypothesis: true difference in means is less than 11430
## 97 percent confidence interval:
##      -Inf 12380.87
## sample estimates:
## mean of x mean of y
## 10212.59 10081.77
```

4. Prueba de hipótesis:

```
t.test(
  x = datos$X2014[datos$Reg %in% 1:10],
  y = datos$X2018[datos$Reg %in% 1:10],
  alternative = "two.sided",
  conf.level = 0.946,
  mu = 20200,
  var.equal = F
)

##
## Welch Two Sample t-test
##
## data: datos$X2014[datos$Reg %in% 1:10] and datos$X2018[datos$Reg %in% 1:10]
## t = -2.1301, df = 17.864, p-value = 0.04733
## alternative hypothesis: true difference in means is not equal to 20200
## 94.6 percent confidence interval:
## -21609.93 19527.02
## sample estimates:
```

```
## mean of x mean of y
## 13084.95 14126.40
```

5. Prueba de hipótesis:

```
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,
  mu = 1710,
  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 = -0.87062, df = 14, p-value = 0.1993
## alternative hypothesis: true difference in means is less than 1710
## 81 percent confidence interval:
## -Inf 1792.44
## sample estimates:
## mean of x mean of y
## 5792.401 6081.915
```