

06_Actividad_correlacion.R

Usuario

2023-10-02

```
# Correlacion
# 02/10/2023
# Leobardo Estrella
library(repmis)
erupciones <-
source_data("https://www.dropbox.com/s/liir6sil7hkqlxs/erupciones.csv?dl=
1")

## Downloading data from:
https://www.dropbox.com/s/liir6sil7hkqlxs/erupciones.csv?dl=1

## SHA-1 hash of the downloaded data file is:
## b07708389ddf62ee20d19c759c88d7dc2d0da3ac

# Estadísticas descriptivas -----
--
mean(erupciones$waiting)

## [1] 70.89706

median(erupciones$waiting)

## [1] 76

var(erupciones$waiting)

## [1] 184.8233

sd(erupciones$waiting)

## [1] 13.59497

mean(erupciones$eruptions)

## [1] 3.487783

median(erupciones$eruptions)

## [1] 4

var(erupciones$eruptions)

## [1] 1.302728

sd(erupciones$eruptions)
```

```
## [1] 1.141371

# Hipotesis -----
--
t.test(erupciones$eruptions, erupciones$waiting)

##
## Welch Two Sample t-test
##
## data: erupciones$eruptions and erupciones$waiting
## t = -81.489, df = 274.82, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -69.03776 -65.78079
## sample estimates:
## mean of x mean of y
## 3.487783 70.897059

# Hipotesis Alternativa H1.

# Si hay una diferencia, porque entre mas tiempo de espera entre
erupciones las duraciones de las erupciones son mayores.

# Hipotesis Nula H0.

#No hay diferencia, entre mas tiempo de espera entre erupciones las
duraciones son iguales.

# Hipotesis aceptada: # Alternativa H1 ya que el p value es mayor que
cero.

# Correlacion -----
--
plot(erupciones$waiting, erupciones$eruptions,
      xlab = "Tiempo de espera entre erupciones (min)",
      ylab = "Duracion de las erupciones (min)",
      col = "red",
      pch = 19,
      main = "Leobardo Estrella")

cor.test(erupciones$eruptions, erupciones$waiting)

##
## Pearson's product-moment correlation
##
## data: erupciones$eruptions and erupciones$waiting
## t = 34.089, df = 270, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
```

```
## 95 percent confidence interval:  
## 0.8756964 0.9210652  
## sample estimates:  
## cor  
## 0.9008112  
text(50,4.5, "r = 0.9008112***")
```

