

Asignacion_6.R

jryma

2020-10-26

```
#Carolina Guadalupe Hernández García.  
#Matricula 2049747
```

```
#Ejercicio 1.
```

```
erupciones <- read.csv("erupciones.csv")  
plot(erupciones, xlab= "Tiempo de espera entre erupciones (min)",  
      ylab= "Duración de las erupciones (min)", pch = 19)
```

```
mean(erupciones$eruptions)
```

```
## [1] 3.487783
```

```
mean(erupciones$waiting)
```

```
## [1] 70.89706
```

```
sd(erupciones$eruptions)
```

```
## [1] 1.141371
```

```
sd(erupciones$waiting)
```

```
## [1] 13.59497
```

```
var(erupciones$eruptions)
```

```
## [1] 1.302728
```

```
var(erupciones$waiting)
```

```
## [1] 184.8233
```

```
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

#La prueba de correlación muestra un coeficiente de relación = 0.90 y un
valor de  $p = 2.2e-16$ .
#Por lo tanto, si existe relación significativa entre el tiempo de espera
entre erupciones y la duración de las erupciones.

#Regresión lineal.

#H0= La relación entre el tiempo de espera entre erupciones y la duración
de las erupciones es igual.
#H1= La relación entre el tiempo de espera entre erupciones y la duración
de las erupciones es diferente.

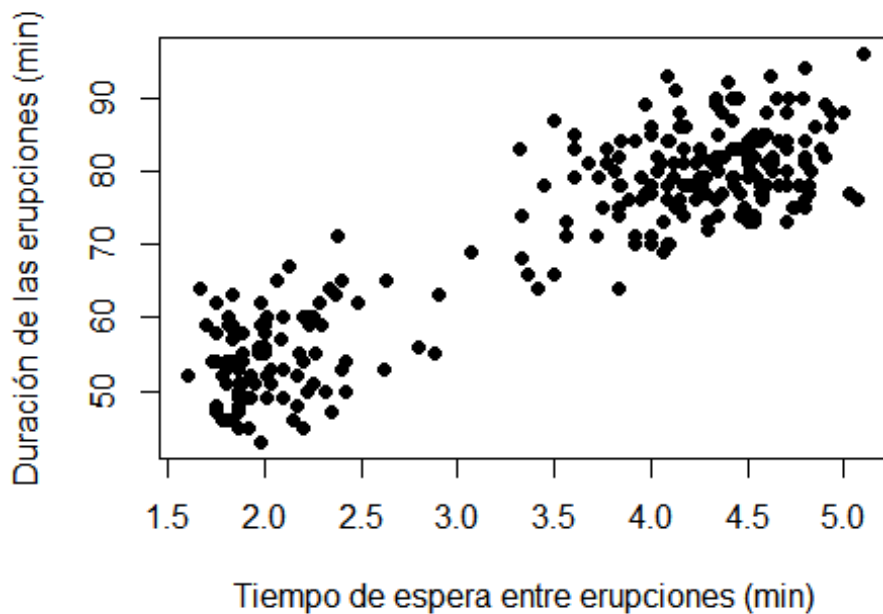
lm.erup <- lm(erupciones$eruptions ~ erupciones$waiting)
lm.erup

##
## Call:
## lm(formula = erupciones$eruptions ~ erupciones$waiting)
##
## Coefficients:
##      (Intercept)  erupciones$waiting
##      -1.87402      0.07563

summary(lm.erup)

##
## Call:
## lm(formula = erupciones$eruptions ~ erupciones$waiting)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29917 -0.37689  0.03508  0.34909  1.19329
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -1.874016    0.160143  -11.70  <2e-16 ***
## erupciones$waiting  0.075628    0.002219   34.09  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4965 on 270 degrees of freedom
## Multiple R-squared:  0.8115, Adjusted R-squared:  0.8108
## F-statistic: 1162 on 1 and 270 DF,  p-value: < 2.2e-16

abline(lm.erup, col= "red")
```



#Valor intercepto (alfa)= -1.87
#Valor de la pendiente (beta) = 0.075
#Valor p= 2.2e-16 Por lo tanto si son significativas

```
residual2 <- resid(lm.erup)
predichos2 <- fitted(lm.erup)
residual.3 <- residual2^2
```

```
Cuadro1 <- round(data.frame(erupciones$eruptions, erupciones$waiting,
residual2, predichos2, residual.3), 2)
head(Cuadro1)
```

```
##   erupciones.eruptions erupciones.waiting residual2 predichos2
residual.3
## 1                3.60                79      -0.50        4.10
0.25
## 2                1.80                54      -0.41        2.21
0.17
## 3                3.33                74      -0.39        3.72
0.15
## 4                2.28                62      -0.53        2.81
0.28
## 5                4.53                85      -0.02        4.55
0.00
## 6                2.88                55       0.60        2.29
0.36
```

```

sum(Cuadro1$residual.3)

## [1] 66.47

sum(Cuadro1$residual2)

## [1] 0.09

An_lm <- anova(lm.erup)
An_lm

## Analysis of Variance Table
##
## Response: erupciones$eruptions
##              Df Sum Sq Mean Sq F value    Pr(>F)
## erupciones$waiting  1 286.478  286.478  1162.1 < 2.2e-16 ***
## Residuals          270  66.562    0.247
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

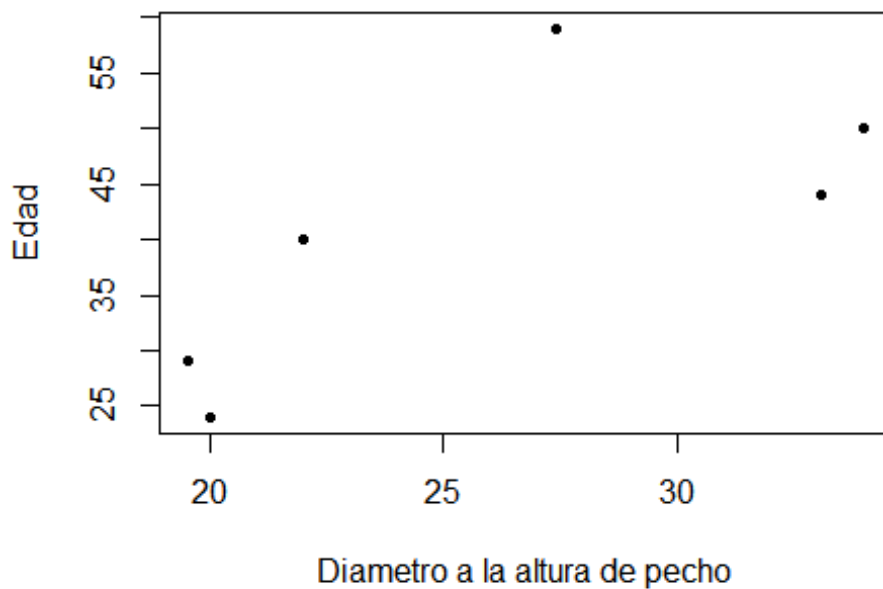
#Regresión significativa

Ejercicio 2. -----
--

```

p_ariz <- read.csv("p_arizonica.csv")
plot(p_ariz$DAP, p_ariz$edad, xlab = "Diametro a la altura de pecho",
ylab= "Edad", pch= 20)

```



```
lm.diam <- lm(p_ariz$edad ~ p_ariz$DAP)
lm.diam

##
## Call:
## lm(formula = p_ariz$edad ~ p_ariz$DAP)
##
## Coefficients:
## (Intercept)  p_ariz$DAP
##          3.658          1.436

summary(lm.diam)

##
## Call:
## lm(formula = p_ariz$edad ~ p_ariz$DAP)
##
## Residuals:
##      1      2      3      4      5      6
## 15.989 -2.664 -8.383  4.745 -2.490 -7.197
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.6579    18.6829   0.196   0.854
## p_ariz$DAP    1.4362     0.7006   2.050   0.110
##
## Residual standard error: 10.17 on 4 degrees of freedom
```

```
## Multiple R-squared:  0.5123, Adjusted R-squared:  0.3904
## F-statistic: 4.202 on 1 and 4 DF,  p-value: 0.1097

#Valor intercepto= 3.65
#valor de la pendiente= 1.43
#p-valor= 0.10
#Por lo tanto alfa y beta no son significativos.

resid_1 <- resid(lm.diam)
pred <- fitted(lm.diam)
resid_2 <- resid_1^2

cuad_2 <- round(data.frame(p_ariz$edad, p_ariz$DAP, resid_1,
pred,resid_2), 2)
cuad_2

##   p_ariz.edad p_ariz.DAP resid_1  pred resid_2
## 1          59        27.4   15.99 43.01  255.66
## 2          29        19.5   -2.66 31.66    7.10
## 3          24        20.0   -8.38 32.38   70.27
## 4          40        22.0    4.74 35.26   22.51
## 5          50        34.0   -2.49 52.49    6.20
## 6          44        33.1   -7.20 51.20   51.80

sum(cuad_2$resid_2)

## [1] 413.54

sum(cuad_2$resid_1)

## [1] -8.881784e-16

Av_lm <- anova(lm.diam)
Av_lm

## Analysis of Variance Table
##
## Response: p_ariz$edad
##           Df Sum Sq Mean Sq F value Pr(>F)
## p_ariz$DAP  1 434.46  434.46   4.2024 0.1097
## Residuals   4 413.54  103.38
```