

Trabajo Practico Regresión Avanzada

Modelo Lineal Multivariado

Jose Valdes

2023-07-08

Table of Contents

2.1. Modelo Aditivo.....	2
Ejercicio 2.1.....	2
a).....	2
b).....	4
c).....	6
d).....	6
e).....	6
f).....	6
g).....	6
2.2. Modelo con Interacción	7
Ejercicio 2.2.....	7
2.3. Regresoras Categóricas.....	7
Ejercicio 2.3.....	7
2.4. Regresión Polinómica	7
Ejercicio 2.4.....	7
2.5. Modelo Robusto.....	7
Ejercicio 2.6.....	7
2.6. Regresión Cuantiles.....	7
Ejercicio 2.7.....	7

```
##          used (Mb) gc trigger (Mb) max used (Mb)
## Ncells 455123 24.4   970168 51.9   644245 34.5
## Vcells 814599  6.3   8388608 64.0  1635138 12.5
```

Se realiza validación de la instalación de los paquetes necesarios para ejecutar el script

```
## [1] TRUE
```

2.1. Modelo Aditivo

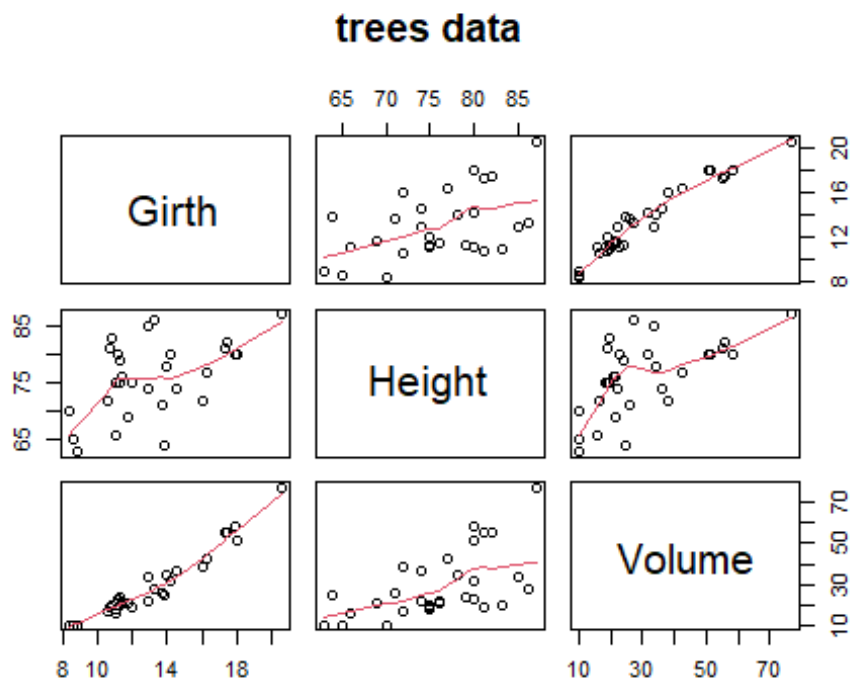
Ejercicio 2.1.

Con el set de datos `trees`, disponible en la biblioteca `dplyr` de R, pretendemos ajustar un modelo que estime el volumen (en pies cúbicos) de los árboles de cerezo en función de la longitud de su circunferencia (en pulgadas) y de su altura (en pies).

a)

Visualizar la asociación entre las variables de a pares.

```
##      Girth Height Volume
## 1      8.3      70  10.3
## 2      8.6      65  10.3
## 3      8.8      63  10.2
## 4     10.5      72  16.4
## 5     10.7      81  18.8
## 6     10.8      83  19.7
## 7     11.0      66  15.6
## 8     11.0      75  18.2
## 9     11.1      80  22.6
## 10    11.2      75  19.9
## 11    11.3      79  24.2
## 12    11.4      76  21.0
## 13    11.4      76  21.4
## 14    11.7      69  21.3
## 15    12.0      75  19.1
## 16    12.9      74  22.2
## 17    12.9      85  33.8
## 18    13.3      86  27.4
## 19    13.7      71  25.7
## 20    13.8      64  24.9
## 21    14.0      78  34.5
## 22    14.2      80  31.7
## 23    14.5      74  36.3
## 24    16.0      72  38.3
## 25    16.3      77  42.6
## 26    17.3      81  55.4
## 27    17.5      82  55.7
## 28    17.9      80  58.3
## 29    18.0      80  51.5
## 30    18.0      80  51.0
## 31    20.6      87  77.0
```



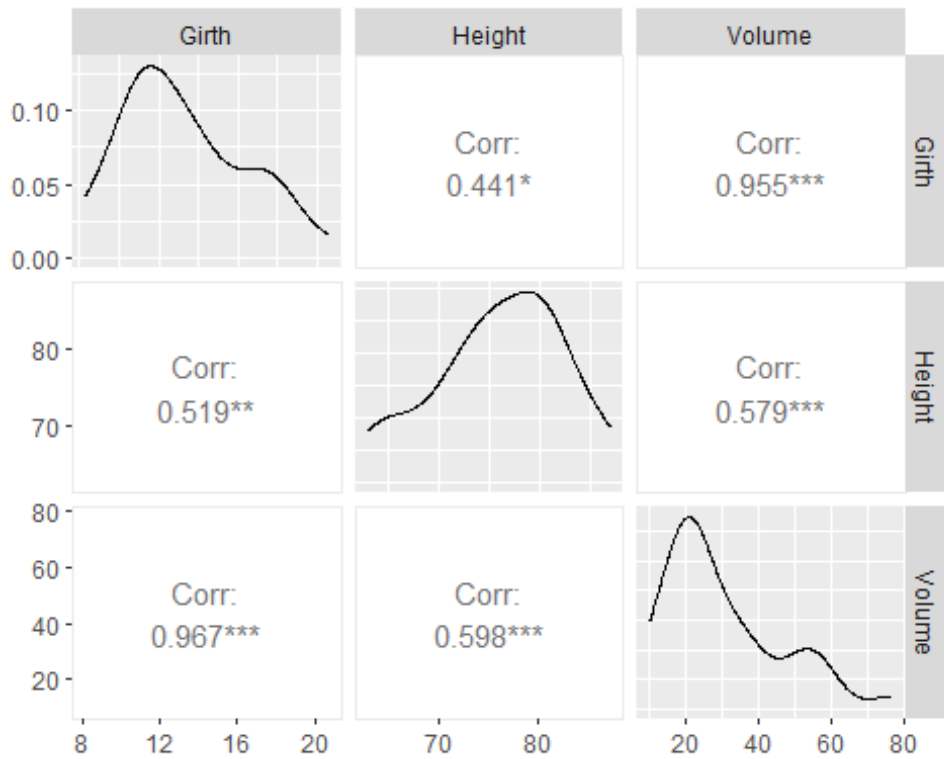
Se observa una relación positiva de las variables.

```
##           Test           HZ      p value MVN
## 1 Henze-Zirkler 0.6636803 0.127490598 YES
## 2 Henze-Zirkler 1.0912171 0.007796525 NO
## 3 Henze-Zirkler 0.9019672 0.026926833 NO

##           cor
## 0.5192801

##           rho
## 0.9547151

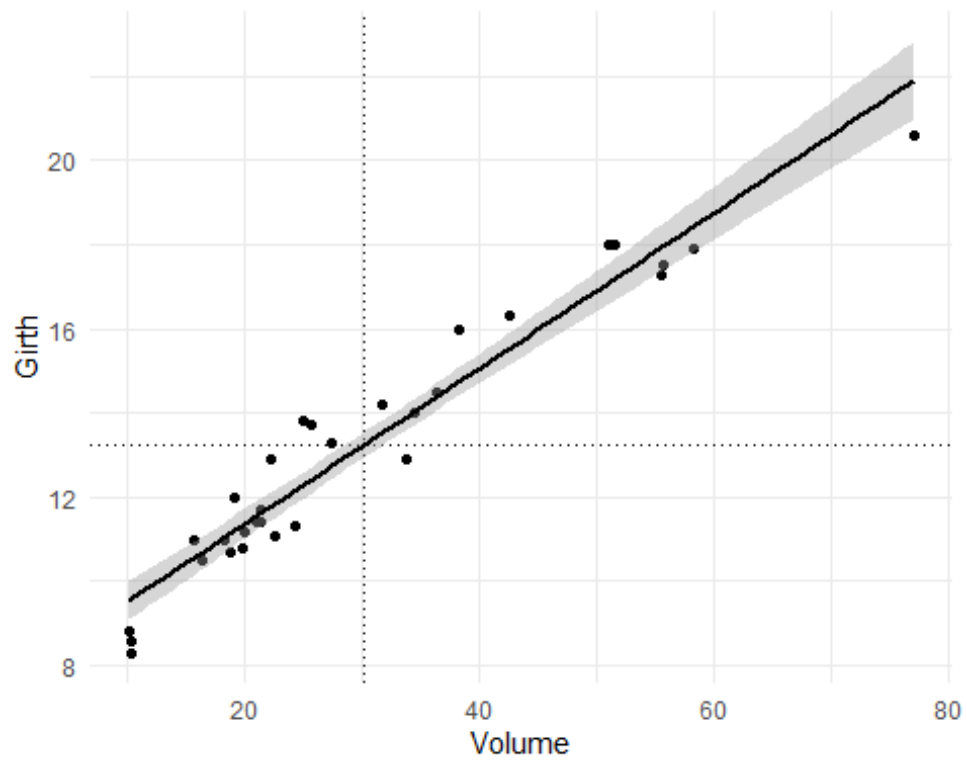
##           rho
## 0.5787101
```



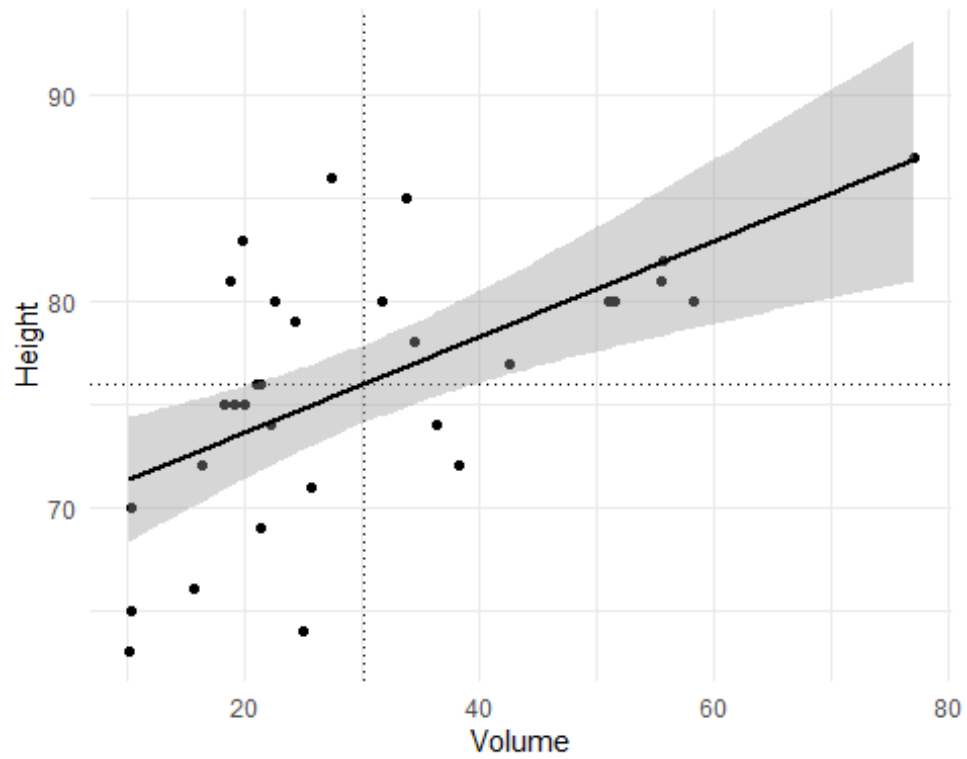
b)

Ajuste un modelo lineal simple para cada una de las dos predictoras disponibles.

****Volume ~ Girth**



****Volume ~ Height**



c)

Realice un análisis diagnóstico en cada caso y señale en caso de haberlos puntos influyentes y outliers.

d)

Estime un intervalo de confianza para los coeficientes del modelo lineal estimado en cada caso.

e)

Ajuste un nuevo modelo sin la/s observaciones influyentes.

f)

Construya el intervalo de confianza y el de predicción del 95% para un árbol cuyo diámetro es 16.1 pulgadas.

g)

Ajuste un modelo utilizando conjuntamente las dos variables predictoras y compare este ajuste con el mejor de los modelos anteriores mediante un test de modelos anidados. Concluya.

2.2. Modelo con Interacción

Ejercicio 2.2.

2.3. Regresoras Categóricas

Ejercicio 2.3.

2.4. Regresión Polinómica

Ejercicio 2.4.

2.5. Modelo Robusto

Ejercicio 2.6.

2.6. Regresión Cuantiles

Ejercicio 2.7.