REGRESIÓN LINEAL SIMPLE

Etna Carolina Cortés Martínez

23/11/2021

REGRESION LINEAL SIMPLE

Técnica estadistica para crear modelos de variables continuas con ayuda de una función predictora. Nos ayuda a predecir y comprender el comportamiento de
sistemas complejos,mediante la generacion de modelos matematicos.

Se parte de la ecuación: Y = beta0 + beta1(X1)+e

donde: Y = es la variable dependiente o variable respuesta **beta0** = es la intersección o término constante **beta1** = son los parámetros del modelo **X1** = es la variable explicativa o independiente **e** = error estandar del residual # LIBRERIAS

```
library(graphics)
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

LECTURA DE LA MATRIZ DE DATOS

1.- Lectura de los datos

```
BD<-read.csv("BD3_penguins.csv")
```

```
BD$Especies
```

```
##   [1] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##   [7] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [13] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [19] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [25] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [31] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [37] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [43] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [49] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [55] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [61] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [67] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
```

```
##  [73] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [79] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [85] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [91] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
##  [97] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
## [103] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
## [109] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
## [115] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
## [121] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
## [127] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
## [133] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
## [139] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
## [145] "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"    "Adelie"
## [151] "Adelie"    "Adelie"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [157] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [163] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [169] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [175] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [181] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [187] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [193] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [199] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [205] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [211] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [217] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [223] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [229] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [235] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [241] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [247] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [253] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [259] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [265] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [271] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [277] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [283] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [289] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [295] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [301] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [307] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [313] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [319] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [325] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [331] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [337] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [343] "Chinstrap" "Chinstrap"
```

2.- seleccionar la seccion de la especie Adelie y de la columna 4 a la 7.

```
adelie<-BD[1:152,4:7]
```

CALCULO DE CORRELACION DE PEARSON

```
cor(adelie)
```

```
##                 Largo_pico_mm Ancho_pico_mm Largo_aleta_mm Masa_corporal_g
## Largo_pico_mm      1.0000000      0.3917580      0.3256178       0.5486177
## Ancho_pico_mm      0.3917580      1.0000000      0.3075689       0.5760619
## Largo_aleta_mm     0.3256178      0.3075689      1.0000000       0.4682015
## Masa_corporal_g    0.5486177      0.5760619      0.4682015       1.0000000
```

CALCULO Y REPRESENTACION DE LA RECTA DE MÍNIMOS CUADRADOS

1.- Se genera el cálculo de la recta cn las varaibles que tienen mayor correlación.

```
regresion<-lm(adelie$Masa_corporal_g~adelie$Ancho_pico_mm, data=adelie)
```

2.- Se visualiza el resultado de los cálculos.

```
summary(regresion)
```

```
##
## Call:
## lm(formula = adelie$Masa_corporal_g ~ adelie$Ancho_pico_mm, data = adelie)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -846.20 -258.05  -44.58  266.62 1161.02
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -281.86     462.40  -0.610    0.543
## adelie$Ancho_pico_mm   217.09      25.15   8.631 8.18e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 374.8 on 150 degrees of freedom
## Multiple R-squared:  0.3318, Adjusted R-squared:  0.3274
## F-statistic:  74.5 on 1 and 150 DF,  p-value: 8.181e-15
```

3.- Interpretacion del modelo

Los parametro de la ecuación de la recta de minimos cuadrados que relaciona la la masa corporal y el ancho del pico se localizan en la columna **"ESTIMATE"**. para esto el estimador *beta0* es de **-281.86**, *beta1* es de **217.09** y el *error standar del residual* es **374.8**, Por lo tanto, la ecuación del modelo quedaría del siguiente modo:

```
Y = -281.86 + 217.09x + 374.8

     Visto de otro modo:

Ancho del pico = -281.86 + 217.09(masa_corporal) + 374.8
```

Sin embargo, R^2 (multiple R-squared) es de **0.3318**, lo que significa que **NO** se trata de un buen modelo.

GENERACION DEL GAFICO CON LA LINEA DE REGRESION

```
ggplot(adelie, aes(x=Masa_corporal_g, y=Ancho_pico_mm))+
  geom_point()+
  geom_smooth(method = "lm", formula=y~x, se=FALSE, col="dodgerblue1")+
  theme_light()
```