

Actividad 2.5 Intervalos de confianza de una regresión

Franco Mendoza Muraira A01383399

2023-11-13

```
## Warning: package 'threejs' was built under R version 4.2.3

## Loading required package: igraph

## Warning: package 'igraph' was built under R version 4.2.3

##
## Attaching package: 'igraph'

## The following objects are masked from 'package:stats':
##
##      decompose, spectrum

## The following object is masked from 'package:base':
##
##      union

## Warning: package 'car' was built under R version 4.2.3

## Loading required package: carData

## Warning: package 'carData' was built under R version 4.2.2
```

1. Modelo de regresión

```
## Resistencia Longitud Altura.matriz Altura.poste Altura.amarre
## 1      9.95      2      50      1      1
## 2     24.45      8     110      1      1
## 3     31.75     11     120      2      1
## 4     35.00     10     550      2      2
## 5     25.02      8     295      1      1
```

```
A = lm(y~x1)
B = lm(y~x1+x2)
alfa = 0.05
```

Modelo lineal simple

A

```
##  
## Call:  
## lm(formula = y ~ x1)  
##  
## Coefficients:  
## (Intercept)          x1  
##      5.115      2.903
```

Se hizo un modelo de regresion lineal simple con la variable predictora de longitud, debido a la correlacion que tiene con la variable a predecir de resistencia.

Modelo lineal multiple

B

```
##  
## Call:  
## lm(formula = y ~ x1 + x2)  
##  
## Coefficients:  
## (Intercept)          x1          x2  
##      2.26379      2.74427      0.01253
```

En este modelo de regresion lineal multiple se usaron las variables predictoras de longitud y altura.matriz para predecir resistencia tambien.

2. Intervalos de regresion

Modelo lineal simple

```
confint(A, level=1-alfa)
```

```
##              2.5 %   97.5 %  
## (Intercept) 2.744239 7.484792  
## x1          2.660587 3.144822
```

El intercepto en este intervalo de confianza va desde 2.74 a 7.48, y la x1 puede ir desde 2.66 hasta 3.14. Esto significa que, con un 95% de confianza, se espera que el valor real del intercepto y la pendiente de la longitud caigan dentro de estos intervalos respectivos en la relación de regresión entre la longitud y la resistencia.

```
predA1=predict(A, interval = "confidence", level=1-alfa)  
head(predA1,6)
```

```
##           fit           lwr           upr
## 1 10.91992   8.939885 12.89996
## 2 28.33615  27.054987 29.61731
## 3 37.04426  35.600466 38.48806
## 4 34.14156  32.792639 35.49048
## 5 28.33615  27.054987 29.61731
## 6 16.72533  15.084643 18.36602
```

En este resultado podemos ver los valores predichos del modelo de regresión lineal en la columna fit, y su intervalo de confianza para la predicción yendo de lwr a upr. Este enfoque de intervalos de confianza proporciona una medida de la incertidumbre asociada con las predicciones del modelo. Cuanto más estrecho sea el intervalo, mayor será la precisión de la predicción.

```
predA2=predict(A, interval = "prediction", level=1-alfa)
```

```
## Warning in predict.lm(A, interval = "prediction", level = 1 - alfa): predictions on current data ref
```

```
head(predA2,6)
```

```
##           fit           lwr           upr
## 1 10.91992   4.221369 17.61848
## 2 28.33615  21.809937 34.86237
## 3 37.04426  30.484184 43.60434
## 4 34.14156  27.601706 40.68141
## 5 28.33615  21.809937 34.86237
## 6 16.72533  10.119127 23.33154
```

Esta función hace algo parecido al pasado, en comparación con los intervalos de confianza, los intervalos de predicción son generalmente más amplios, ya que tienen en cuenta tanto la variabilidad en la estimación del modelo como la variabilidad inherente en la respuesta.

Modelo lineal multiple

```
confint(B, level=1-alfa)
```

```
##           2.5 %      97.5 %
## (Intercept) 0.065348613 4.46223426
## x1          2.550313061 2.93822623
## x2          0.006724246 0.01833138
```

El intercepto en este intervalo de confianza va desde 0.065 a 4.462, la x1 puede ir desde 2.55 hasta 2.938, y la x2 puede ir desde 0.0067 hasta 0.018. Esto significa que, con un 95% de confianza, se espera que el valor real del intercepto, la pendiente de la longitud, y la pendiente de la altura.matriz caigan dentro de estos intervalos respectivos en la relación de regresión entre la longitud y altura.matriz, y la resistencia.

```
predB1=predict(B, interval = "confidence", level=1-alfa)
head(predB1,6)
```

```
##          fit          lwr          upr
## 1  8.378721  6.496821 10.26062
## 2 25.596008 24.010499 27.18152
## 3 33.954095 32.166507 35.74168
## 4 36.596784 35.082138 38.11143
## 5 27.913653 26.943699 28.88361
## 6 15.746432 14.448068 17.04480
```

En este resultado podemos ver los valores predichos del modelo de regresion multiple en la columna fit, y su intervalo de confianza para la prediccion yendo de lwr a upr. Este enfoque de intervalos de confianza proporciona una medida de la incertidumbre asociada con las predicciones del modelo. Cuanto más estrecho sea el intervalo, mayor será la precisión de la predicción.

```
predB2=predict(B, interval = "prediction", level=1-alfa)
```

```
## Warning in predict.lm(B, interval = "prediction", level = 1 - alfa): predictions on current data ref
```

```
head(predB2,6)
```

```
##          fit          lwr          upr
## 1  8.378721  3.274047 13.48340
## 2 25.596008 20.593010 30.59901
## 3 33.954095 28.883432 39.02476
## 4 36.596784 31.615790 41.57778
## 5 27.913653 23.070414 32.75689
## 6 15.746432 10.826890 20.66597
```

Esta funcion hace algo parecido al pasado, en comparación con los intervalos de confianza, los intervalos de predicción son generalmente más amplios, ya que tienen en cuenta tanto la variabilidad en la estimación del modelo como la variabilidad inherente en la respuesta.

3. Intervalos para una nueva observacion

Modelo de regresion simple

```
resistencia_real = 33.49
x = data.frame(x1=c(12),x2=c(535))
predict(object=A,newdata = x,interval="confidence",level = 1-alfa)
```

```
##          fit          lwr          upr
## 1 39.94697 38.37638 41.51756
```

Podemos ver que el resultado de resistencia real no queda dentro de el intervalo de confidence.

```
predict(object=A, newdata=x, interval="prediction", level=1-alfa)
```

```
##          fit          lwr          upr
## 1 39.94697 33.35782 46.53612
```

Podemos ver que el resultado de resistencia real si queda dentro de el intervalo de prediction debido a que se extiende mas.

Modelo de regresion multiple

```
predict(object=B,newdata = x,interval="confidence",level = 1-alfa)
```

```
##          fit          lwr          upr  
## 1 41.89741 40.42339 43.37143
```

Podemos ver que el resultado de resistencia real no queda dentro de el intervalo de confidence.

```
predict(object=B, newdata=x, interval="prediction", level=1-alfa)
```

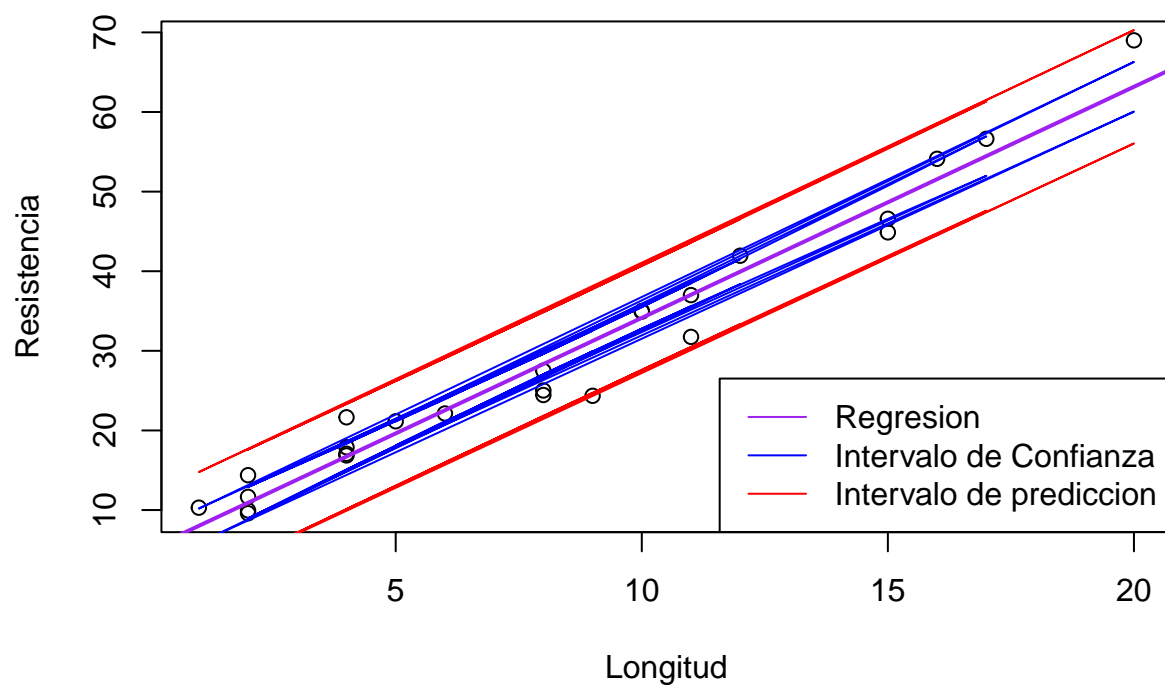
```
##          fit          lwr          upr  
## 1 41.89741 36.92862 46.8662
```

Podemos ver que el resultado de resistencia real no queda dentro de el intervalo de prediction, y queda lejos todavia en el modelo de regresion lineal multiple.

4. Representación gráfica

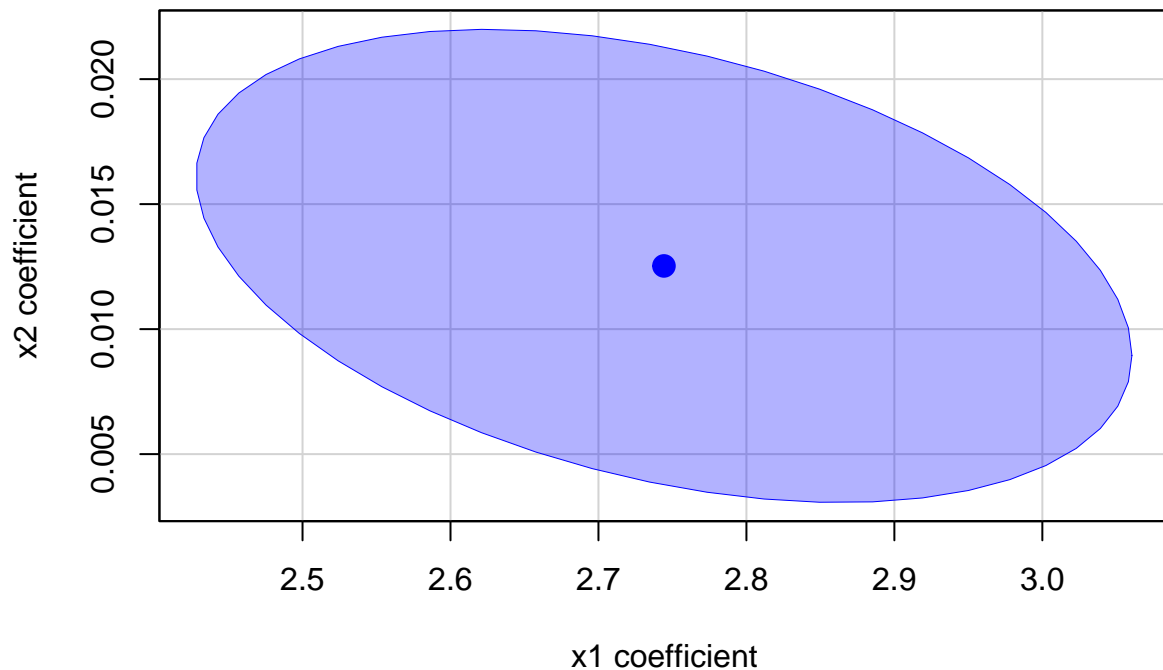
```
plot(x1,y,xlab="Longitud",ylab="Resistencia",main="Modelo regresion lineal simple")  
abline(A,col = "purple",lwd=2)  
  
lines(x1,predA1[, "lwr"],col = "blue")  
lines(x1,predA1[, "upr"],col = "blue")  
  
lines(x1,predA2[, "lwr"],col = "red")  
lines(x1,predA2[, "upr"],col = "red")  
  
legend("bottomright",legend=c("Regresion","Intervalo de Confianza","Intervalo de prediccion"),col = c("purple","blue","red"))
```

Modelo regresion lineal simple



```
confidenceEllipse(B,fill=TRUE,lwd=0,levels=0.99,which.coef=c("x1","x2"),main="Elipse de 99% de confianza")
```

Elipse de 99% de confianza regresion multiple



Podemos observar que las variables tienen una correlación debido a la forma elíptica que mantienen. También se nota una correlación negativa, ya que al aumentar la variable x_1 , la x_2 disminuye, como se puede apreciar en la pendiente inclinándose hacia la izquierda.

5. Conclusiones

Pudimos crear modelos de regresión lineal y múltiple para predecir la variable resistencia, eligiendo variables que mostraron una correlación aceptable con ella para establecer una conexión.

También exploramos cómo funcionan los intervalos de confianza, en nuestro caso con un 95% de confianza, y cómo son de gran utilidad para proporcionar información sobre los datos, teniendo en cuenta la variabilidad de las variables independientes del modelo.

Al realizar pruebas de predicciones en el modelo múltiple, pudimos observar que el modelo no funcionó de la mejor manera, a pesar de que se observaban buenos ajustes en actividades anteriores. Esto puede indicar que hay una mayor variabilidad o incertidumbre en los datos de lo que se espera.