

# Clase 13 Supuestos de la regresión lineal

OCE 386 - Introducción al análisis de datos con R.

Dr. José A. Gallardo | Pontificia Universidad Católica de  
Valparaíso

09 November 2021

# PLAN DE LA CLASE

## 1.- Introducción

- ▶ Supuestos de la regresión lineal.
- ▶ Independencia.
- ▶ Linealidad.
- ▶ Homogeneidad de varianzas.
- ▶ Normalidad.
- ▶ Efecto de los valores atípicos.

## 2.- Práctica con R y Rstudio cloud

- ▶ Realizar análisis de regresión lineal.
- ▶ Realizar gráficas avanzadas con ggplot2.
- ▶ Elaborar un reporte dinámico en formato pdf.

# SUPUESTOS DE LA REGRESIÓN LINEAL SIMPLE

- ▶ ¿Cuales son los supuestos?

- Independencia.

- Linealidad entre variable independiente y dependiente.

- Homocedasticidad.

- Normalidad.

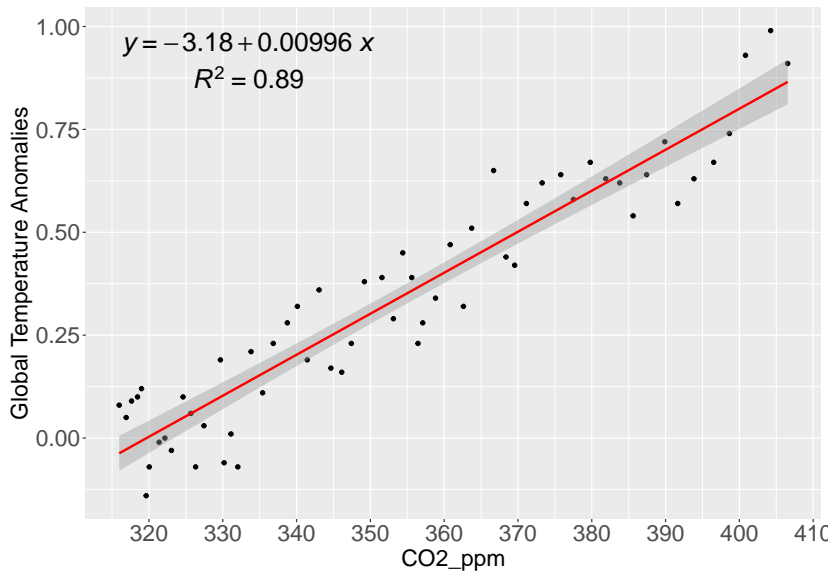
- ▶ ¿Por qué son importantes?

- Para validar el resultado obtenido.

- En caso de incumplimiento se pueden transformar datos o elaborar otros modelos (Regresión logística).

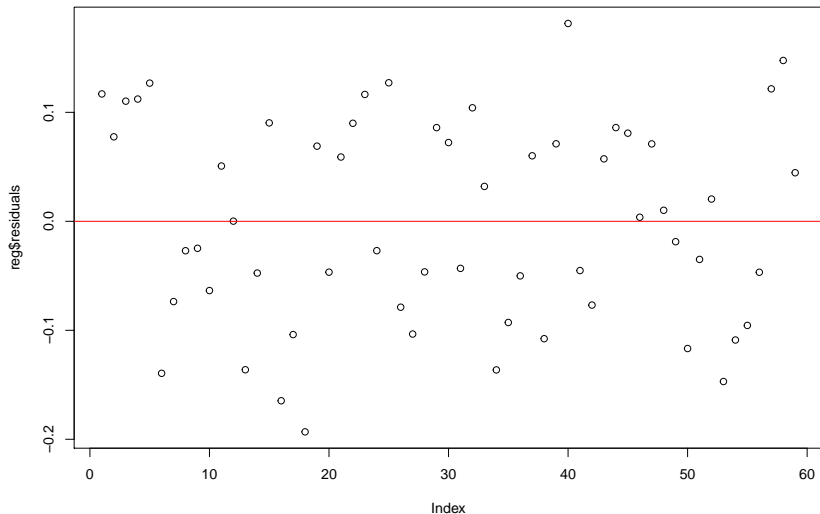
# CASO DE ESTUDIO: CALENTAMIENTO GLOBAL.

Calentamiento global Fuente: climate.gov



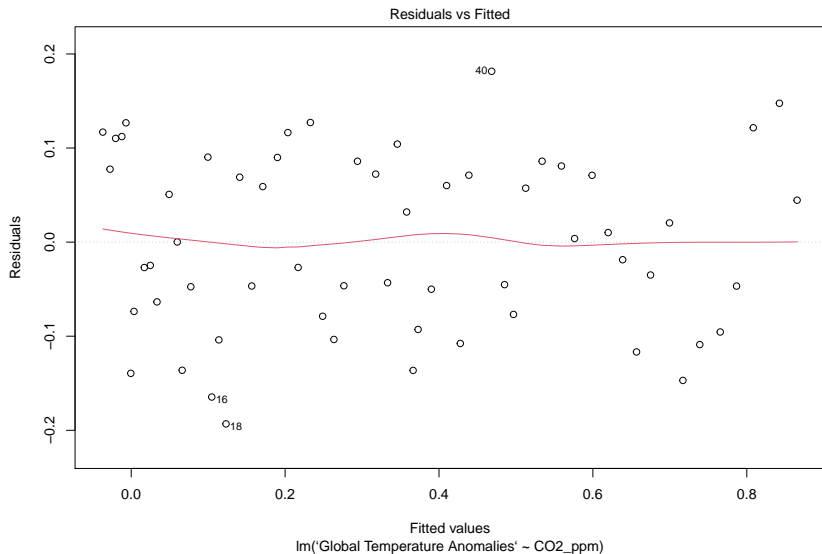
# INDEPENDENCIA: MÉTODO GRÁFICO

```
plot(reg$residuals)  
abline(h=0, col="red")
```



# LINEALIDAD: MÉTODO GRÁFICO

```
plot(reg, which=1)
```



# INDEPENDENCIA: Durbin Watson

$H_0$  : No existe autocorrelación entre los datos (lo que deseamos).

$H_0$  : Existe autocorrelación entre los datos.

Dado que  $p < 0,05$  se rechaza independencia.

```
# durbin watson test
```

```
durbinWatsonTest(reg) # library(car)
```

```
## lag Autocorrelation D-W Statistic p-value
```

```
## 1 0.2338696 1.501215 0.042
```

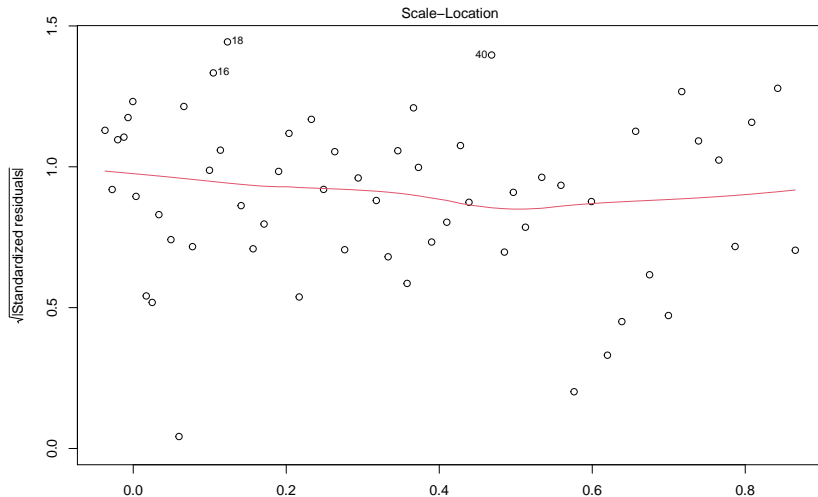
```
## Alternative hypothesis: rho != 0
```

# HOMOGENEIDAD DE VARIANZAS: MÉTODO GRÁFICO

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_A: \sigma_1^2 \neq \sigma_2^2$$

```
plot(reg, which=3)
```





# HOMOGENEIDAD DE VARIANZAS: PRUEBAS DE HIPÓTESIS

$p > .05$ , No tenemos evidencias para rechazar que nuestros datos son homocedasticos.

```
ncvTest(reg) # library(car)
```

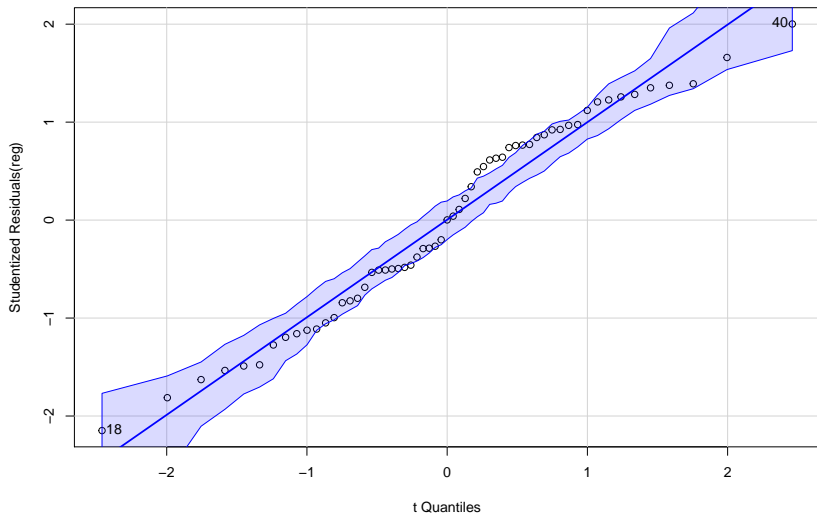
```
## Non-constant Variance Score Test  
## Variance formula: ~ fitted.values  
## Chisquare = 0.1563339, Df = 1, p = 0.69255
```

```
bptest(reg) # library(lmtest)
```

```
##  
## studentized Breusch-Pagan test  
##  
## data: reg  
## BP = 0.34329, df = 1, p-value = 0.5579
```

# NORMALIDAD: GRÁFICO DE CUANTILES

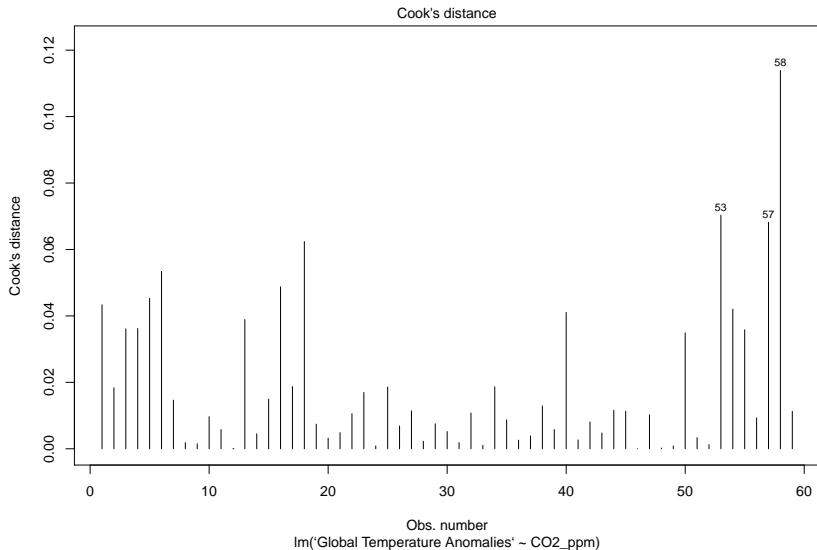
```
qqPlot(reg) # library(car)
```



```
## [1] 18 40
```

# VALORES ATÍPICOS

```
plot(reg, which=4)
```



# PRÁCTICA ANÁLISIS DE DATOS

- ▶ Guía de trabajo práctico disponible en drive y Rstudio.cloud.

## **Clase\_13**

- ▶ El trabajo práctico se realiza en Rstudio.cloud.

## **Guía 13 Regresión lineal**

# RESUMEN DE LA CLASE

- ▶ **Evaluar supuestos de análisis de regresión lineal**
- ▶ **Independencia**
- ▶ **Normalidad**
- ▶ **Linealidad**
- ▶ **Homogeneidad de varianzas**