# RegresionNoLineal

Cleber Perez

2024-09-17

## Cars

```
summary(cars)
```

```
##      speed           dist       
##  Min.   : 4.0   Min.   :  2.00  
##  1st Qu.:12.0   1st Qu.: 26.00  
##  Median :15.0   Median : 36.00  
##  Mean   :15.4   Mean   : 42.98  
##  3rd Qu.:19.0   3rd Qu.: 56.00  
##  Max.   :25.0   Max.   :120.00
```

```
data = cars
```

## Parte 1: Análisis de normalidad

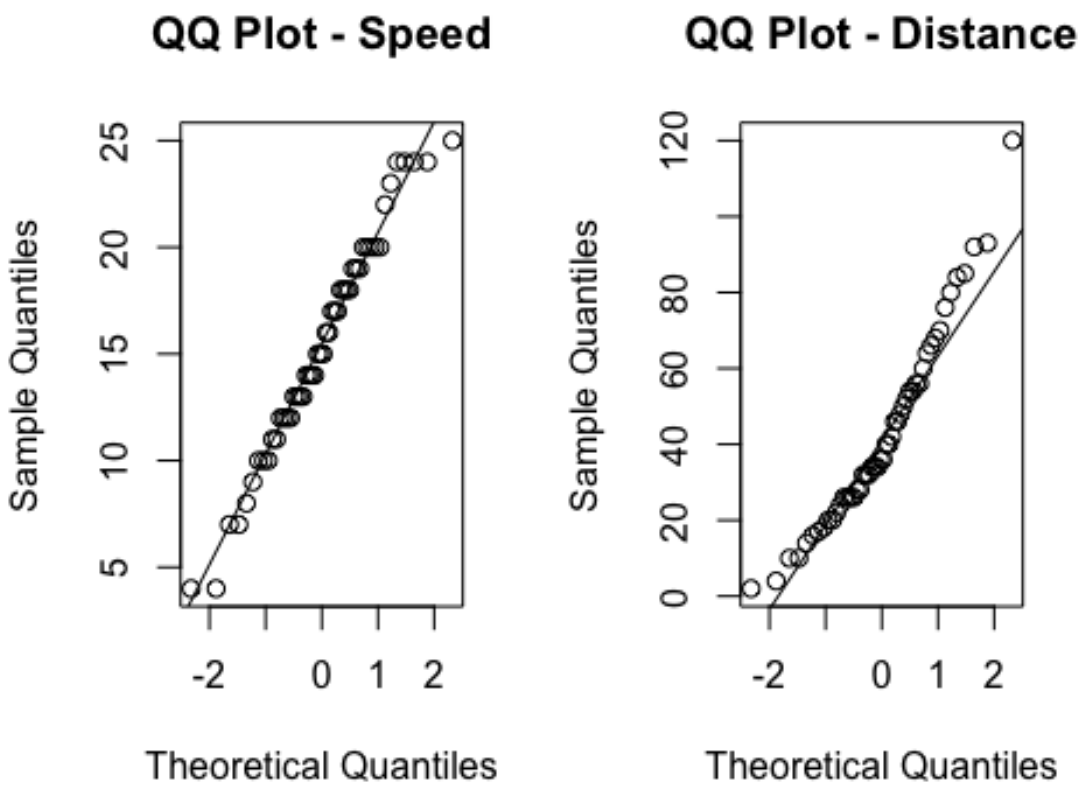
```
library(nortest)
ad.test(cars$speed)
```

```
##
##  Anderson-Darling normality test
##
## data:  cars$speed
## A = 0.26143, p-value = 0.6927
```

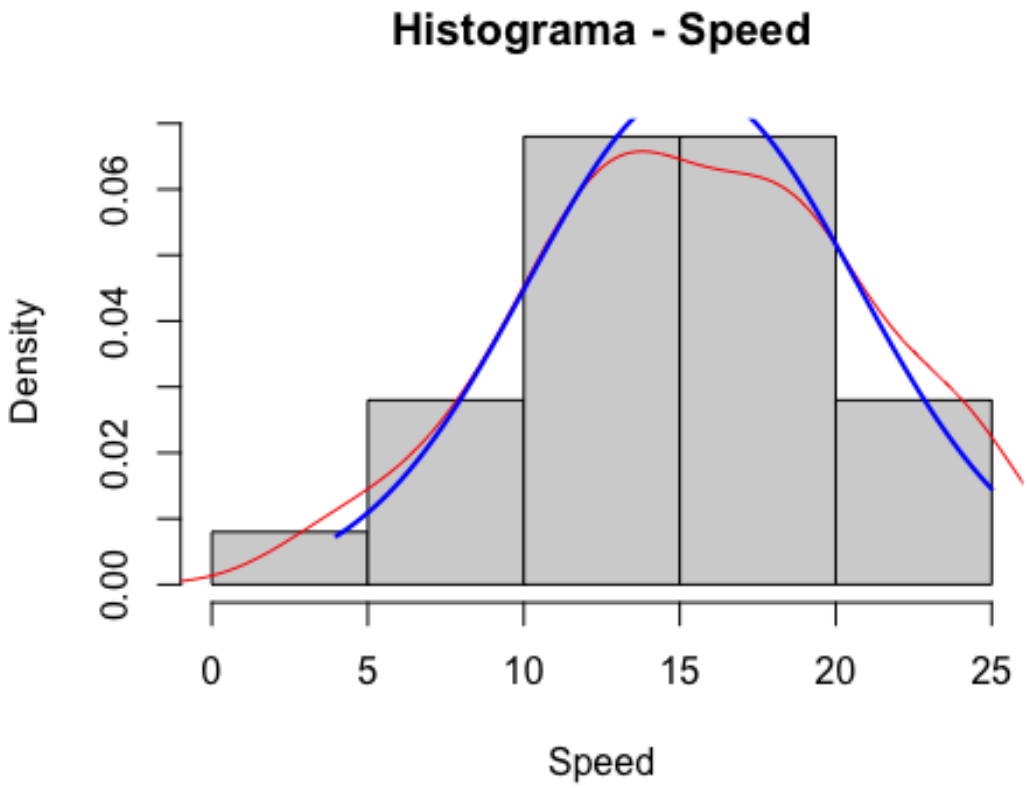
```
ad.test(cars$dist)
```

```
##
##  Anderson-Darling normality test
##
## data:  cars$dist
## A = 0.74067, p-value = 0.05021
```

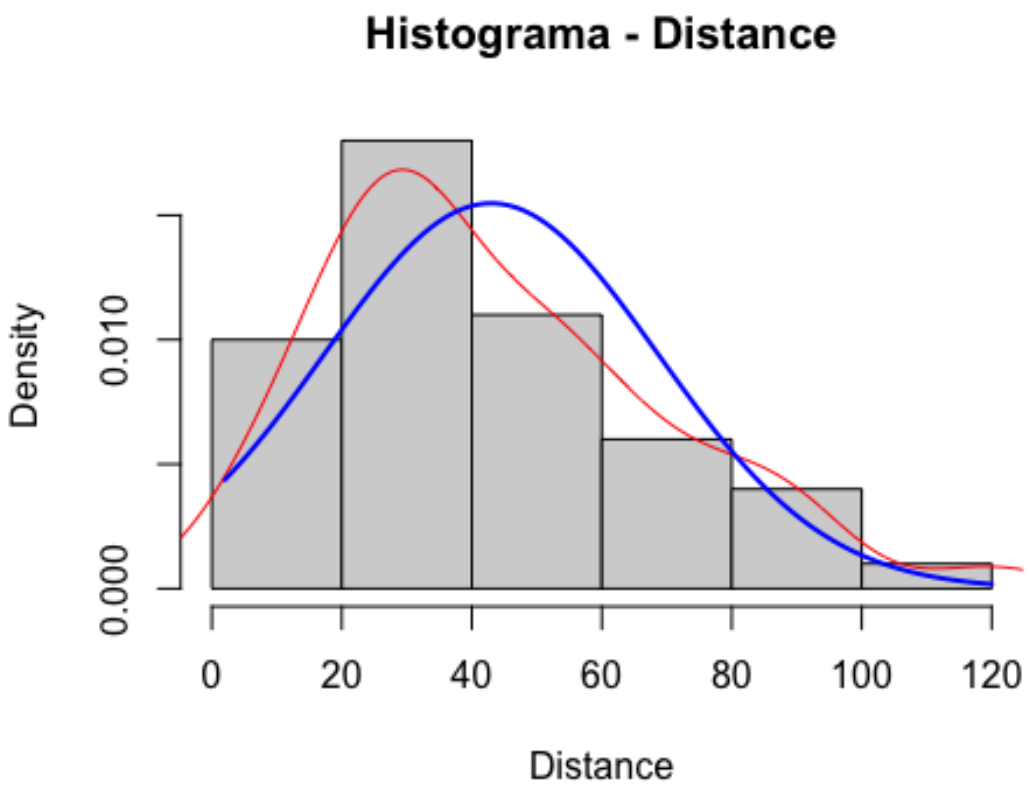
```
library(e1071)
par(mfrow=c(1,2))
qqnorm(cars$speed, main = "QQ Plot - Speed")
qqline(cars$speed)
qqnorm(cars$dist, main = "QQ Plot - Distance")
qqline(cars$dist)
```




```r
hist(cars$speed, freq = FALSE, main = "Histograma - Speed", xlab = "Speed")
lines(density(cars$speed), col = "red")  # Densidad observada
curve(dnorm(x, mean=mean(cars$speed), sd=sd(cars$speed)),
      from=min(cars$speed), to=max(cars$speed), add=TRUE, col="blue", lwd=2)
```




```r
hist(cars$dist, freq = FALSE, main = "Histograma - Distance", xlab =
"Distance")
lines(density(cars$dist), col = "red")  # Densidad observada
curve(dnorm(x, mean=mean(cars$dist), sd=sd(cars$dist)),
      from=min(cars$dist), to=max(cars$dist), add=TRUE, col="blue", lwd=2)
```




```
speedske <- skewness(cars$speed)
speedske
```

```
## [1] -0.1105533
```

```
speedkur <- kurtosis(cars$speed)
speedkur
```

```
## [1] -0.6730924
```

```
distske <- skewness(cars$dist)
distske
```

```
## [1] 0.7591268
```

```
distkur <- kurtosis(cars$dist)
distkur
```

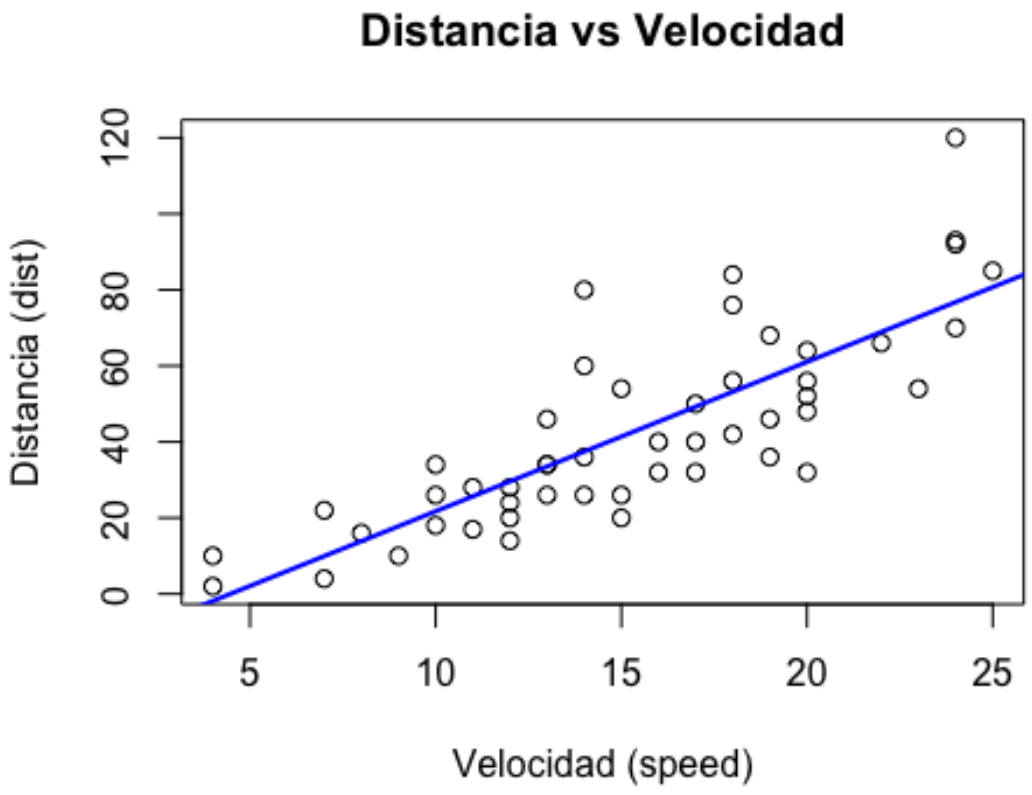
```
## [1] 0.1193971
```

## Parte 2: Regresión lineal

```
modelo1 =  lm(dist ~ speed, data = cars)
summary(modelo1)
```

```
##
## Call:
## lm(formula = dist ~ speed, data = cars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.069  -9.525  -2.272   9.215  43.201
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -17.5791     6.7584  -2.601   0.0123 *
## speed         3.9324     0.4155   9.464 1.49e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.38 on 48 degrees of freedom
## Multiple R-squared:  0.6511, Adjusted R-squared:  0.6438
## F-statistic: 89.57 on 1 and 48 DF,  p-value: 1.49e-12
```

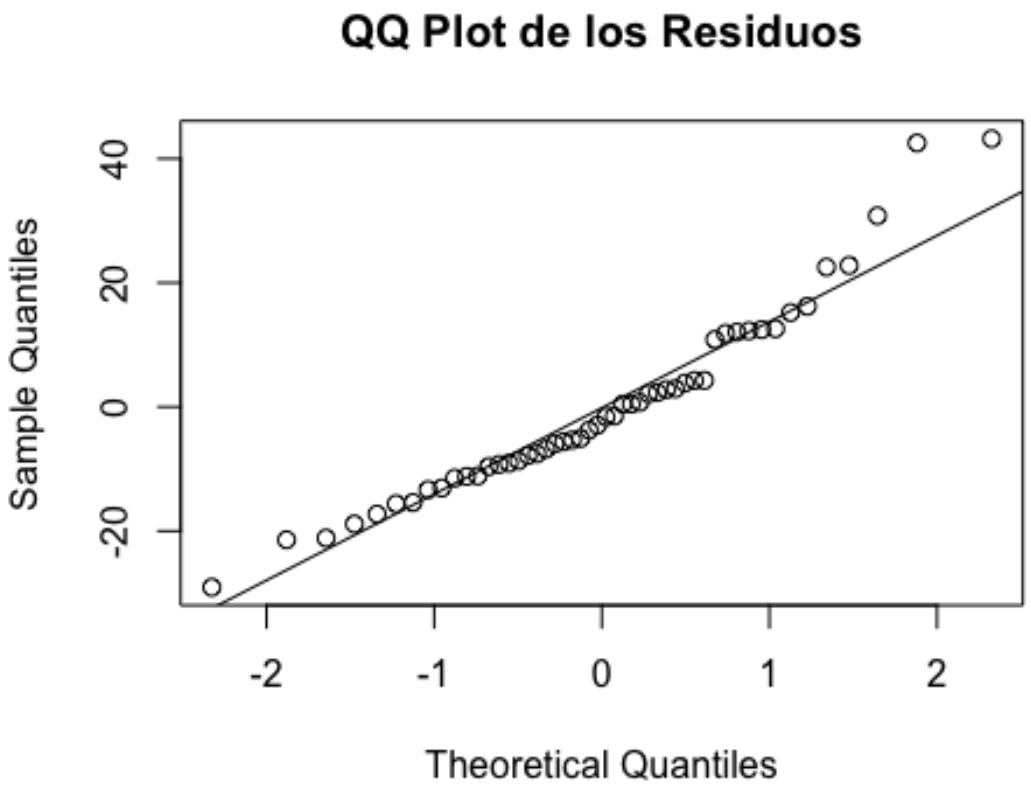
```
plot(cars$speed, cars$dist, main="Distancia vs Velocidad", xlab="Velocidad
(speed)", ylab="Distancia (dist)")
abline(modelo1, col="blue", lwd=2)
```



```r
t.test(modelo1$residuals)
```

```
##
##  One Sample t-test
##
## data:  modelo1$residuals
## t = -2.0629e-16, df = 49, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -4.326  4.326
## sample estimates:
##     mean of x
## -4.440892e-16
```

```r
qqnorm(residuals(modelo1), main="QQ Plot de los Residuos")
qqline(residuals(modelo1))
```




```r
library(lmtest)
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
bptest(modelo1)
```

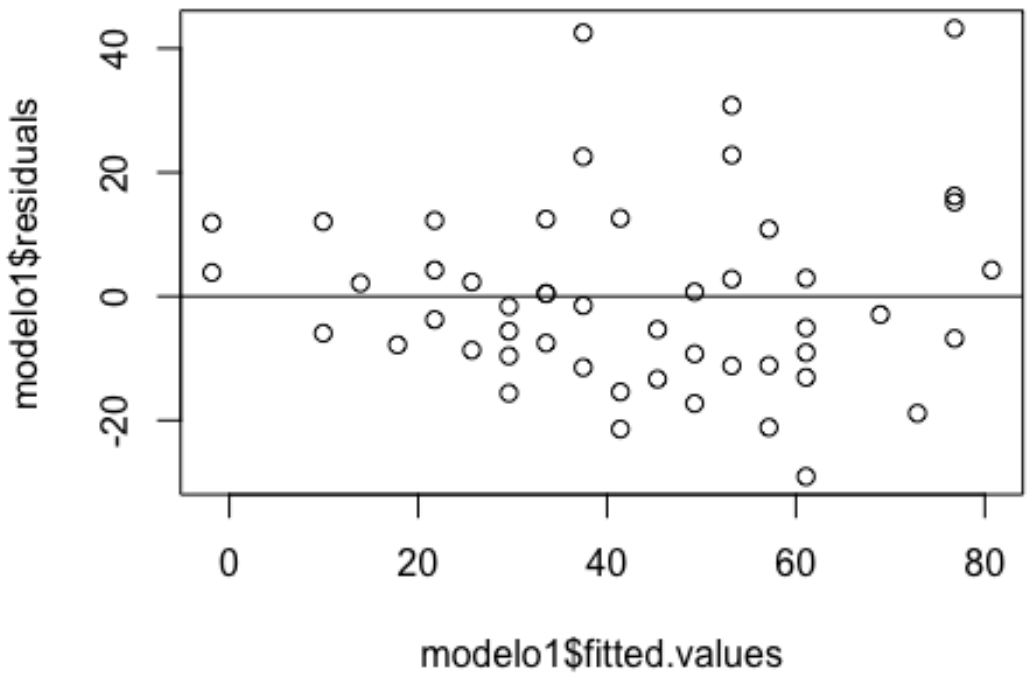
```
##
##  studentized Breusch-Pagan test
##
## data:  modelo1
## BP = 3.2149, df = 1, p-value = 0.07297
```

```
gqtest(modelo1)
```

```
##
##  Goldfeld-Quandt test
##
## data:  modelo1
## GQ = 1.5512, df1 = 23, df2 = 23, p-value = 0.1498
## alternative hypothesis: variance increases from segment 1 to 2
```

```
plot(modelo1$fitted.values,modelo1$residuals)
abline(h=0, color="blue")
```

```
## Warning in int_abline(a = a, b = b, h = h, v = v, untf = untf, ...):
"color" is
## not a graphical parameter
```



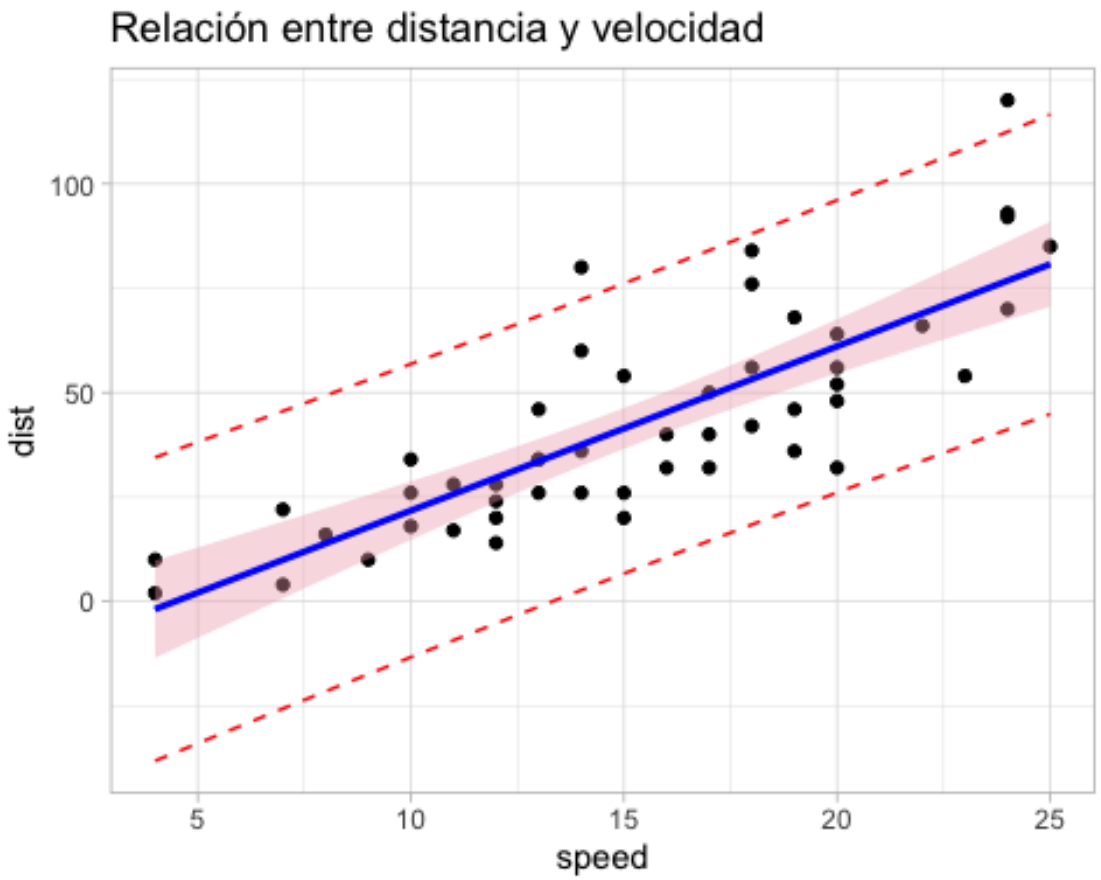

```
library(lmtest)
bptest(modelo1)

##
##  studentized Breusch-Pagan test
##
## data:  modelo1
## BP = 3.2149, df = 1, p-value = 0.07297

gqtest(modelo1)

##
##  Goldfeld-Quandt test
##
## data:  modelo1
## GQ = 1.5512, df1 = 23, df2 = 23, p-value = 0.1498
## alternative hypothesis: variance increases from segment 1 to 2

library(lmtest)
resettest(modelo1)

##
##  RESET test
##
```

```
## data:  modelo1
## RESET = 1.5554, df1 = 2, df2 = 46, p-value = 0.222
```

```
modelo1 =  lm(dist ~ speed, data = cars)
Ip = predict(object = modelo1, interval = "prediction", level = 0.97)
```

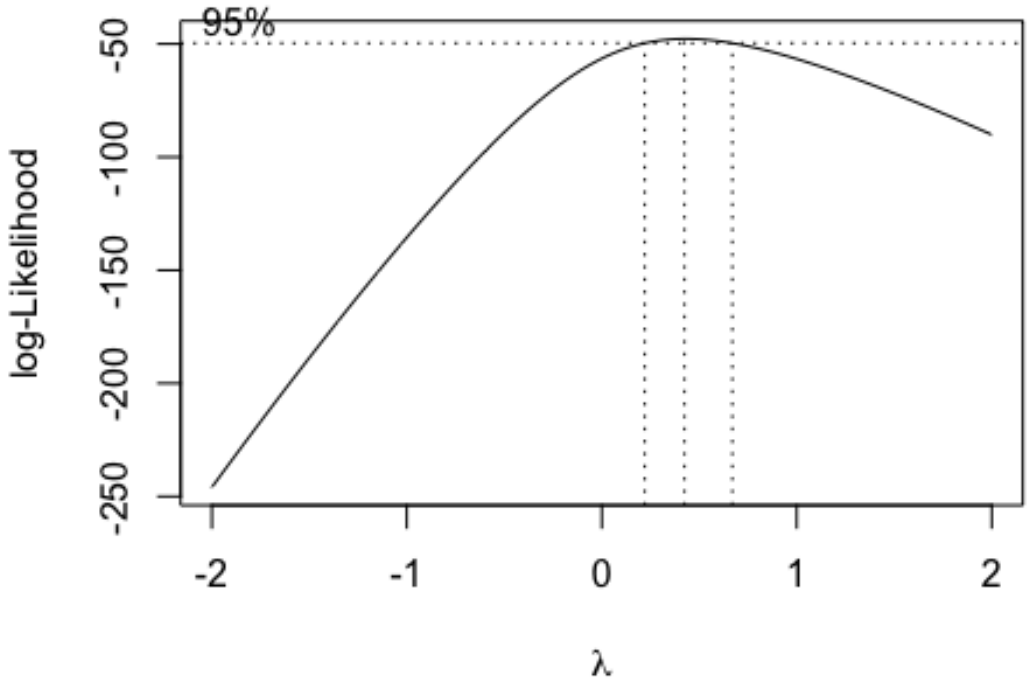
```
## Warning in predict.lm(object = modelo1, interval = "prediction", level =
0.97): predictions on current data refer to _future_ responses
```

```
datos1 = cbind(cars, Ip)
library(ggplot2)
ggplot(datos1, aes(x = speed, y = dist)) +
  geom_point() +
  geom_line(aes(y = fit), color = "blue") +  # Línea de ajuste del modelo
  geom_line(aes(y = lwr), color = "red", linetype = "dashed") +  # Línea
inferior del intervalo de predicción
  geom_line(aes(y = upr), color = "red", linetype = "dashed") +  # Línea
superior del intervalo de predicción
  geom_smooth(method = "lm", formula = y ~ x, se = TRUE, level = 0.97, col =
"blue", fill = "pink2") +
  ggtitle("Relación entre distancia y velocidad") +
  theme_light()
```



## Parte 3: Regresión no lineal

```
library(MASS)
modelo2 = lm(dist ~ speed, data = cars)
bc = boxcox(modelo2)
```




```
lambda_opt = bc$x[which.max(bc$y)]
lambda_opt
```

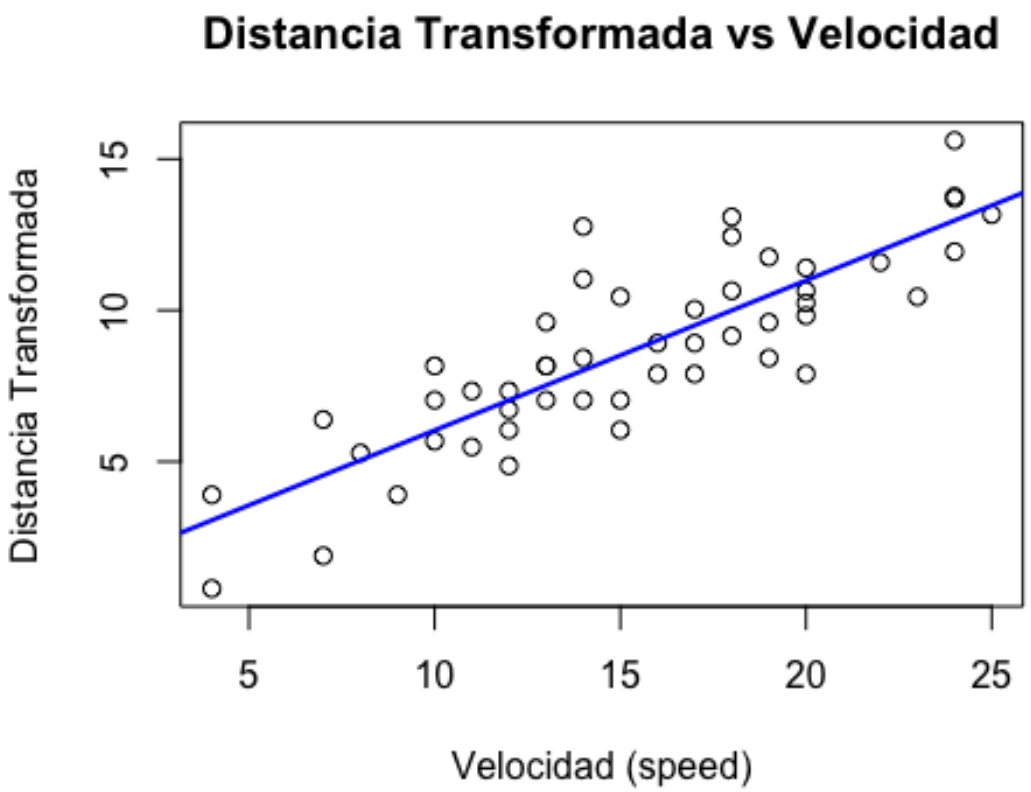
```
## [1] 0.4242424
```

```
if (lambda_opt == 0) {
  cars$dist_transformed <- log(cars$dist)
} else {
  cars$dist_transformed <- (cars$dist^lambda_opt - 1) / lambda_opt
}

dist2skew <- skewness(cars$dist_transformed)
dist2kurt <- kurtosis(cars$dist_transformed)
cat("Sesgo de la transformación:", dist2skew, "\n")
```

```
## Sesgo de la transformación: -0.1701619
```

```
cat("Curtosis de la transformación:", dist2kurt, "\n")
```

```
## Curtosis de la transformación: -0.186884

modelo_transformed <- lm(dist_transformed ~ speed, data = cars)
summary(modelo_transformed)

##
## Call:
## lm(formula = dist_transformed ~ speed, data = cars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0926 -1.0444 -0.3055  0.7999  4.7520
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.08227    0.73856   1.465    0.149
## speed        0.49541    0.04541  10.910 1.35e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.681 on 48 degrees of freedom
## Multiple R-squared:  0.7126, Adjusted R-squared:  0.7066
## F-statistic:   119 on 1 and 48 DF,  p-value: 1.354e-14

plot(cars$speed, cars$dist_transformed, main="Distancia Transformada vs
Velocidad", xlab="Velocidad (speed)", ylab="Distancia Transformada")
abline(modelo_transformed, col="blue", lwd=2)
```


Distancia Transformada vs Velocidad
Distancia Transformada
5
10
15
5
10
15
20
25
Velocidad (speed)