Construccion de modelo estadistico

Héctor Francisco Marin Garrido

2022-09-08

Emepezamos cargando los datos y empezando con un analisis de correlacion para ver que factores son los que mas afectar a la concentracion de mercurio en los lagos. Para esto se retiraron variables categoricas y no numericas, de igual manera se hizo el analisis de la correlacion con la concentracion minima,maxima y la promedio de mercurio en grupos de peces.

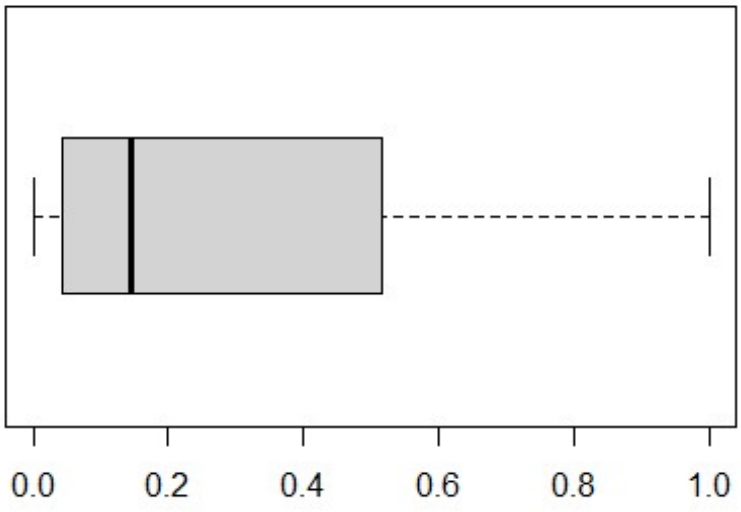
```
M=read.csv("D:\\AI_ITESM\\mercurio.csv")
M=M
Mtest=M[c("X3","X4","X5","X6","X8","X7","X9","X10","X11")]
min_max_norm<-function(x){(x-min(x))/(max(x)-min(x))}
M1=as.data.frame(lapply(Mtest,min_max_norm))
cor(Mtest)
```

```
##              X3          X4          X5          X6          X8
X7
## X3   1.00000000  0.71916568  0.83260419  0.47753085  0.01029074 -
0.59389671
## X4   0.71916568  1.00000000  0.57713272  0.60848276 -0.01860607 -
0.57540012
## X5   0.83260419  0.57713272  1.00000000  0.40991385 -0.08937901 -
0.40067958
## X6   0.47753085  0.60848276  0.40991385  1.00000000 -0.01182027 -
0.49137481
## X8   0.01029074 -0.01860607 -0.08937901 -0.01182027  1.00000000
0.07903426
## X7  -0.59389671 -0.57540012 -0.40067958 -0.49137481  0.07903426
1.00000000
## X9  -0.52535654 -0.54196524 -0.33247623 -0.40045856 -0.08165278
0.92720506
## X10 -0.60479558 -0.55181523 -0.40791663 -0.48497215  0.16109174
0.91586397
## X11 -0.62795845 -0.61284905 -0.46440947 -0.50644193  0.02580046
0.95921481
##              X9        X10         X11
## X3  -0.52535654 -0.6047956 -0.62795845
## X4  -0.54196524 -0.5518152 -0.61284905
## X5  -0.33247623 -0.4079166 -0.46440947
## X6  -0.40045856 -0.4849721 -0.50644193
## X8  -0.08165278  0.1610917  0.02580046
## X7   0.92720506  0.9158640  0.95921481
## X9   1.00000000  0.7653532  0.91908939
```

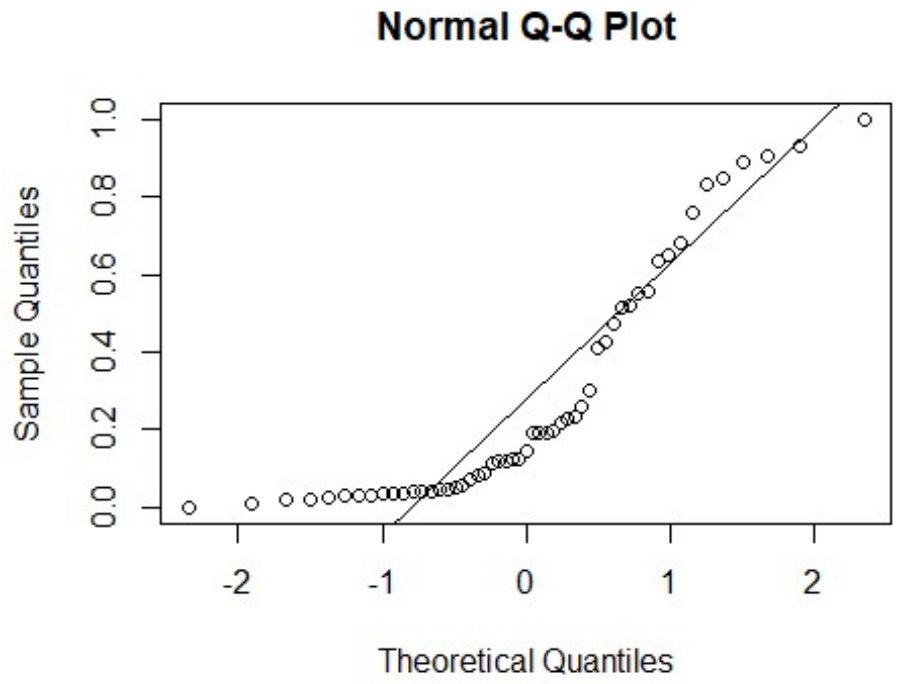
```
## X10  0.76535319  1.0000000  0.85975810
## X11  0.91908939  0.8597581  1.00000000
```

Ahora procedemos a analizar la distribucion de los datos y los intervalos de confianza de todas las variables cuantitativas en nuestra base de datos.

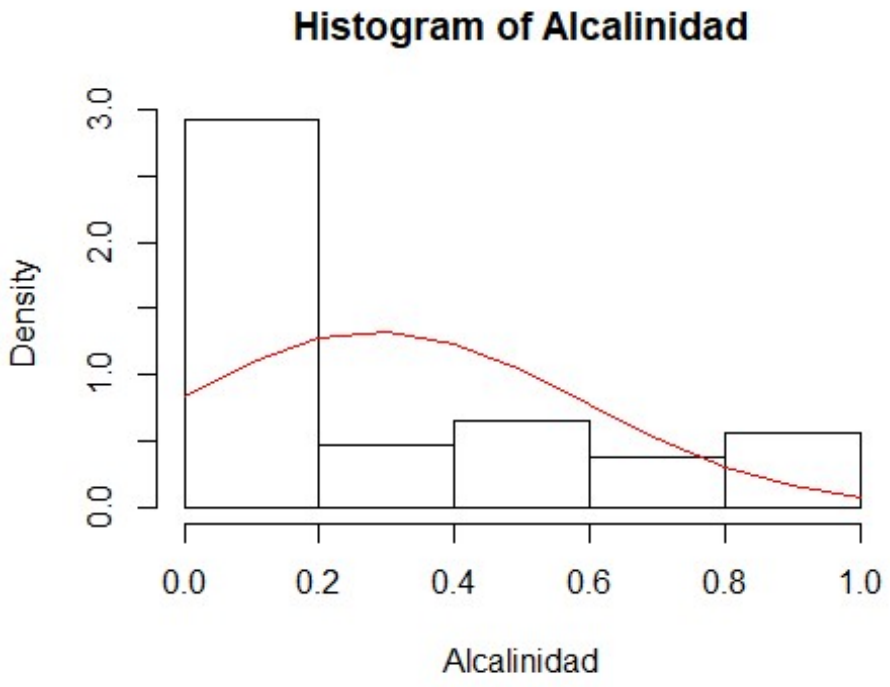
```
q1=quantile(M1$X3,0.25)  #Cuantil 1 de la variable X
q2=quantile(M1$X3,0.5)
q3=quantile(M1$X3,0.75)
y1=min(M1$X3)
y2=max(M1$X3)
ri=IQR(M1$X3)
boxplot(M1$X3,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```




```
Alcalinidad= M1[M1$X3<q3+1.5*ri,c("X3")]
qqnorm(Alcalinidad)
qqline(Alcalinidad)
```

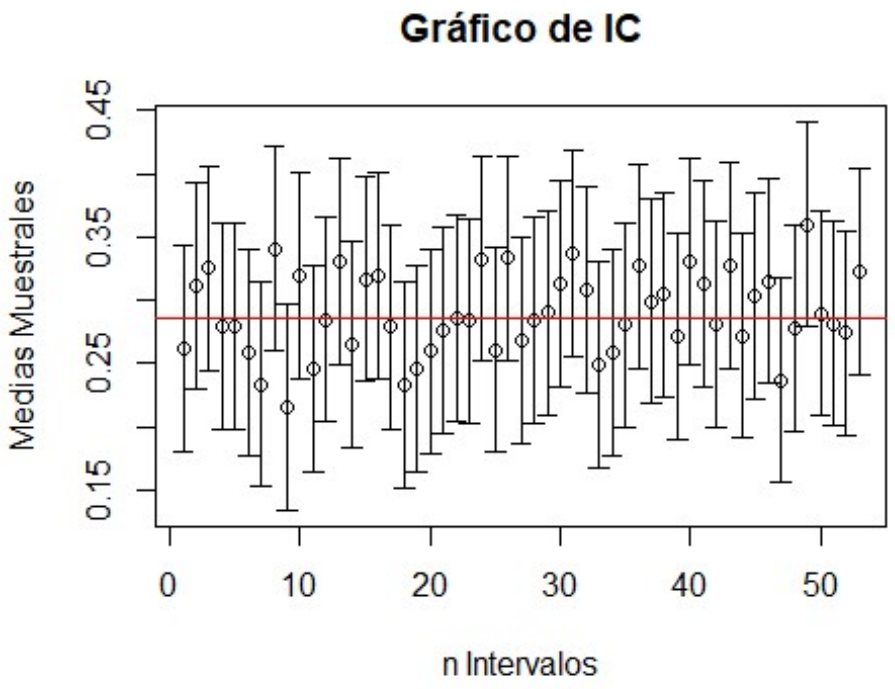



```
hist(Alcalinidad,prob=TRUE,col=0)
x=seq(min(Alcalinidad),max(Alcalinidad),0.1)
y=dnorm(x,mean(Alcalinidad),sd(Alcalinidad))
lines(x,y,col="red")
```

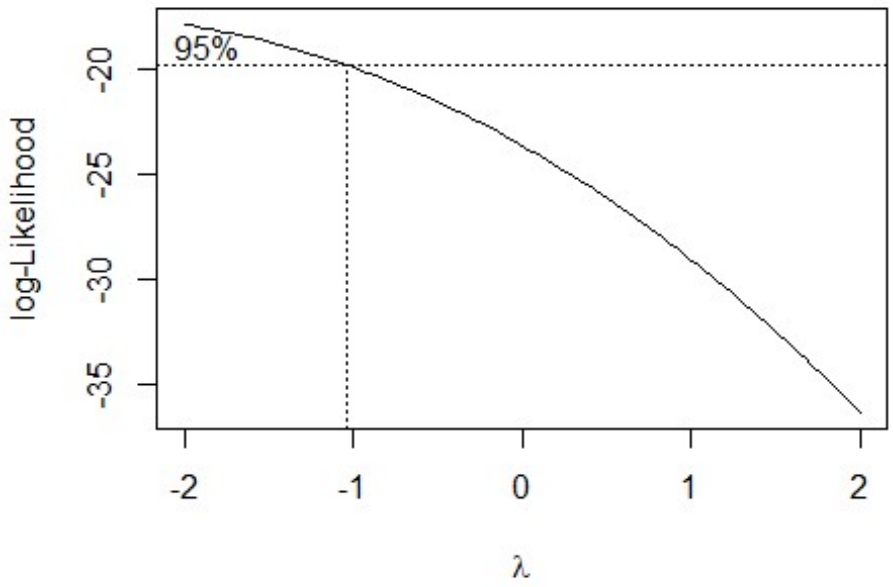



```
library(moments)

## Warning: package 'moments' was built under R version 4.1.3

skewness1=skewness(Alcalinidad)
cat("Sesgo: ",skewness1)

## Sesgo:  0.9959715

cat("\n")

kurtosis1=kurtosis(Alcalinidad)
cat("Curtosis: ",kurtosis1)

## Curtosis:  2.627688

cat("\n")

library(plotrix)
n=length(Alcalinidad)
media=mean(Alcalinidad)
DE=sd(Alcalinidad)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
```

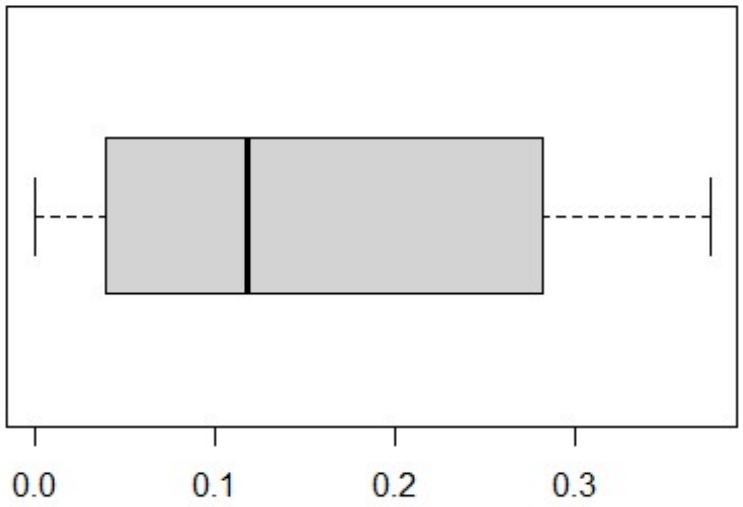
```
Muestrales")
abline(h=media,col="red")
```



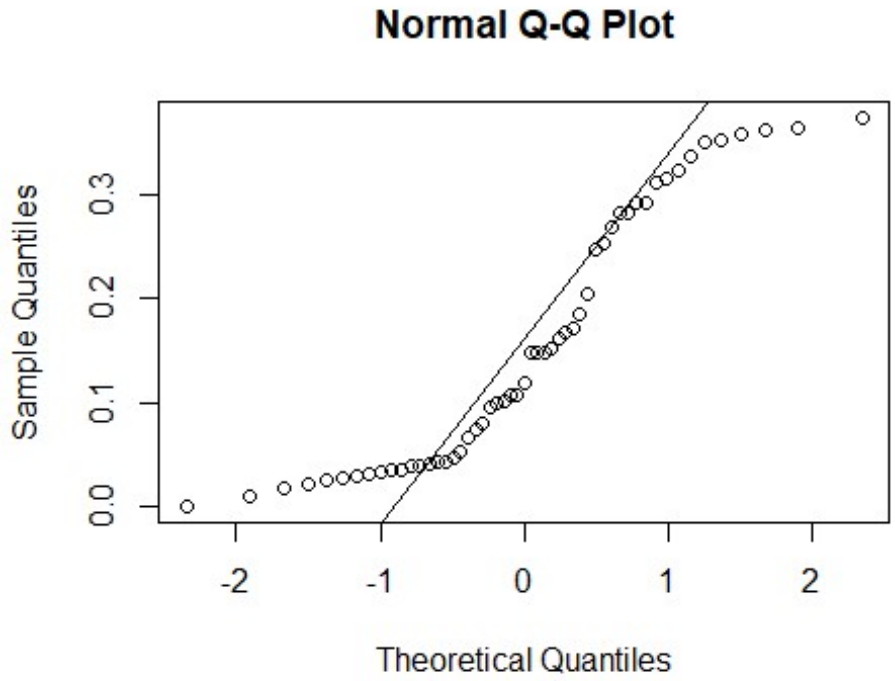

```
cat("El promedio es de: ",media," +- ",E)

## El promedio es de:  0.2865157  +-  0.08111373

library(MASS)
minim=min(M1$X3)
b <- boxcox(lm((M1$X3+1) ~ 1))
```



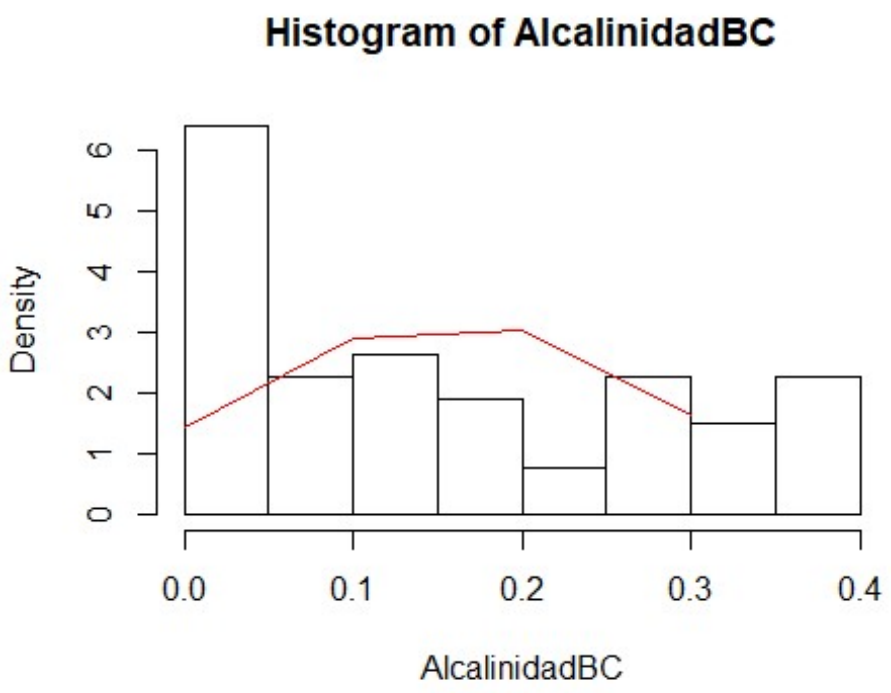

```
lambda <- b$x[which.max(b$y)]
X3<-((M1$X3+1)^lambda-1)/lambda
q1=quantile(X3,0.25)  #Cuantil 1 de la variable X
q2=quantile(X3,0.5)
q3=quantile(X3,0.75)
y1=min(X3)
y2=max(X3)
ri=IQR(X3)
boxplot(X3,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```




```
AlcalinidadBC= X3
qqnorm(AlcalinidadBC)
qqline(AlcalinidadBC)
```

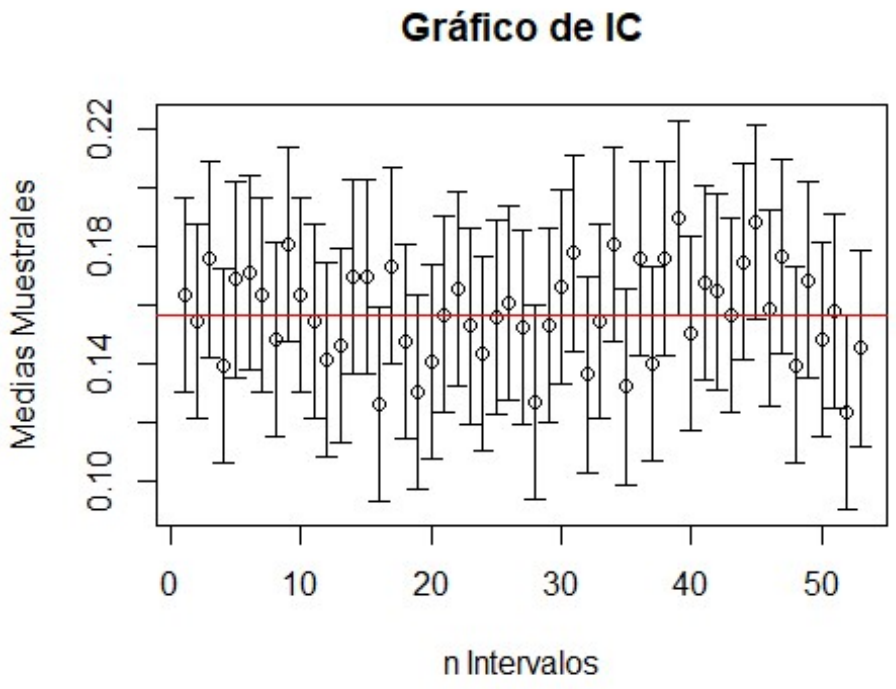


```
hist(AlcalinidadBC,prob=TRUE,col=0)
x=seq(min(AlcalinidadBC),max(AlcalinidadBC),0.1)
y=dnorm(x,mean(AlcalinidadBC),sd(AlcalinidadBC))
lines(x,y,col="red")
```




```
library(moments)
skewness1=skewness(AlcalinidadBC)
cat("Sesgo: ",skewness1)

## Sesgo:  0.4590799

cat("\n")

kurtosis1=kurtosis(AlcalinidadBC)
cat("Curtosis: ",kurtosis1)

## Curtosis:  1.723831

cat("\n")

library(plotrix)
n=length(AlcalinidadBC)
media=mean(AlcalinidadBC)
DE=sd(AlcalinidadBC)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
```

```
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```




```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.1562331  +-  0.03316211
```

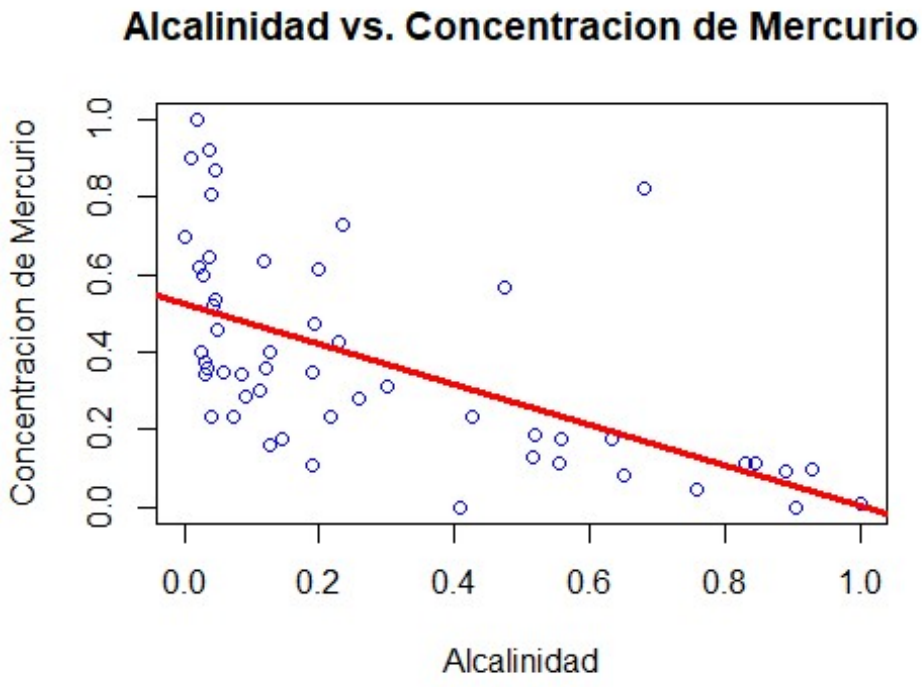
```
cor.test(M1$X7,M1$X3)
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and M1$X3
## t = -5.2717, df = 51, p-value = 2.763e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.7446538 -0.3854832
## sample estimates:
##        cor
## -0.5938967
```

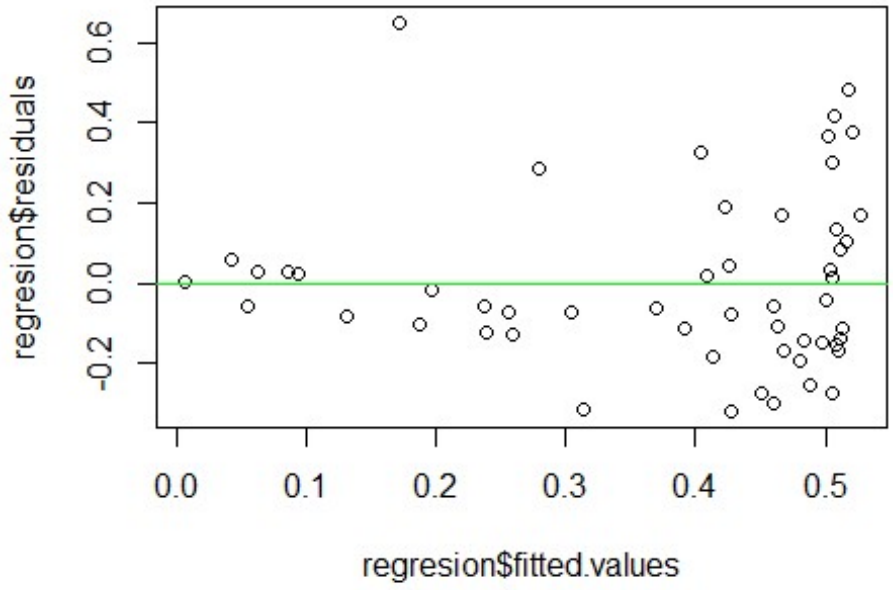
```
regresion=lm(M1$X7 ~ M1$X3)
regresion
```

```
##
## Call:
## lm(formula = M1$X7 ~ M1$X3)
```

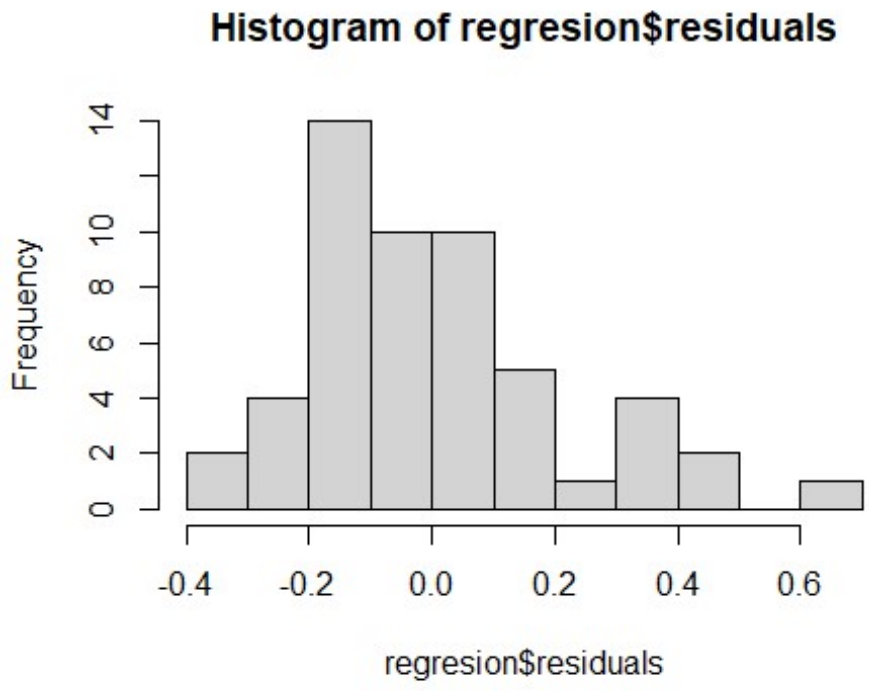
```
##
## Coefficients:
## (Intercept)        M1$X3
##      0.5270      -0.5211

PH=summary(regresion)
PH

##
## Call:
## lm(formula = M1$X7 ~ M1$X3)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.31898 -0.13925 -0.05769  0.08473  0.64983
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.52696    0.04090  12.885  < 2e-16 ***
## M1$X3       -0.52112    0.09885  -5.272 2.76e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2148 on 51 degrees of freedom
## Multiple R-squared:  0.3527, Adjusted R-squared:   0.34
## F-statistic: 27.79 on 1 and 51 DF,  p-value: 2.763e-06

plot(M1$X3,M1$X7,col="blue",xlab="Alcalinidad",ylab="Concentracion de
Mercurio",main="Alcalinidad vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```



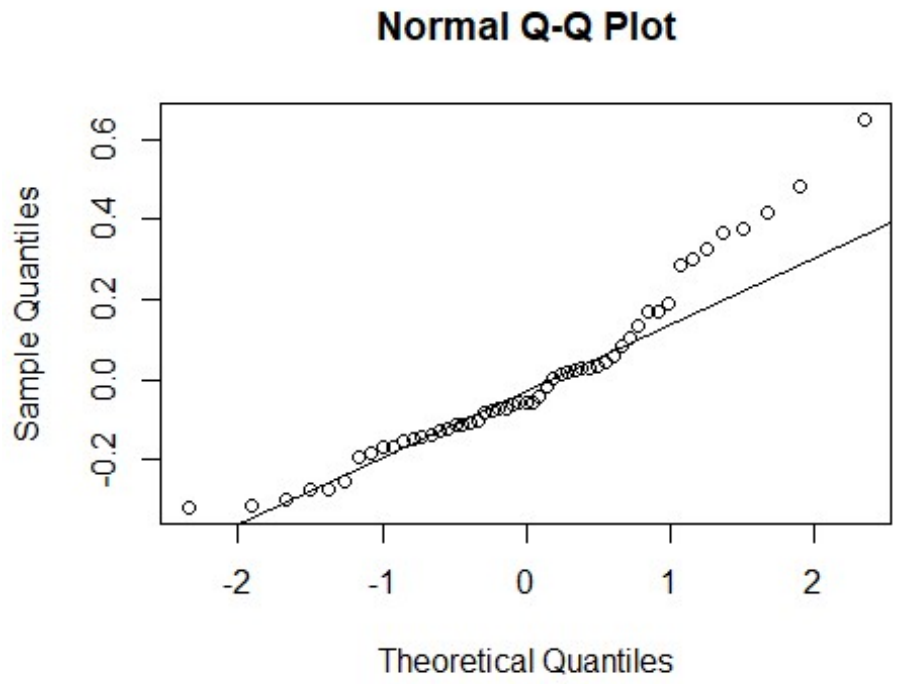

```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.9295, p-value = 0.003866

hist(regresion$residuals)
```




```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```




```
t.test(regresion$residuals)

##
##  One Sample t-test
##
## data:  regresion$residuals
## t = 1.2507e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.05862615  0.05862615
## sample estimates:
##   mean of x
## 3.65394e-18

cor.test(M1$X7,AlcalinidadBC)

##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and AlcalinidadBC
## t = -5.6648, df = 51, p-value = 6.838e-07
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.7635021 -0.4220611
## sample estimates:
##       cor
## -0.6214586
```

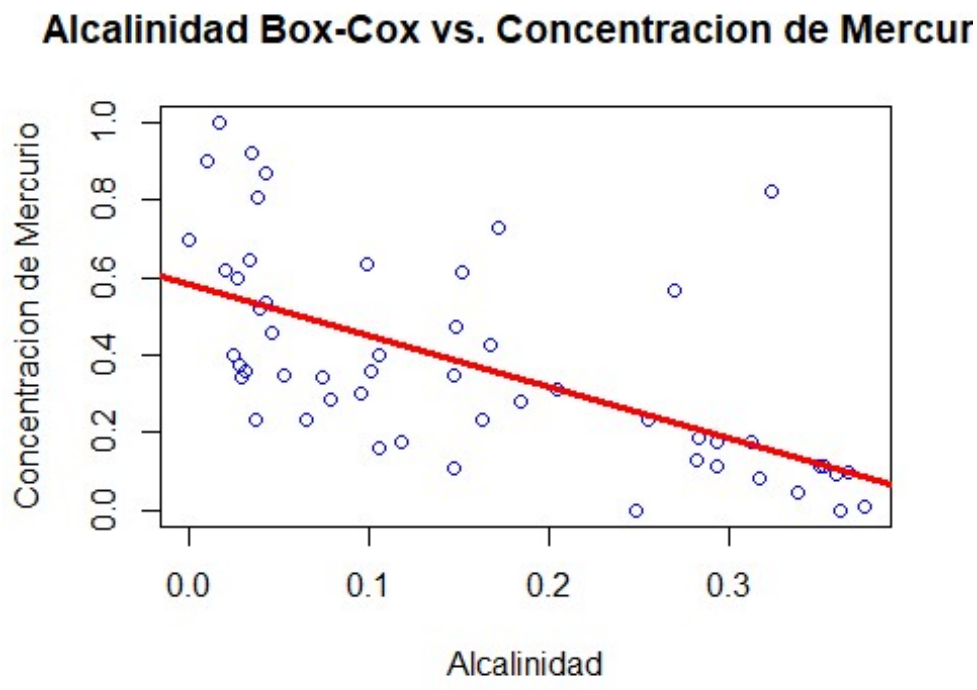
```r
regresion=lm(M1$X7 ~ AlcalinidadBC)
regresion
```

```
##
## Call:
## lm(formula = M1$X7 ~ AlcalinidadBC)
##
## Coefficients:
##   (Intercept)  AlcalinidadBC
##         0.586         -1.334
```

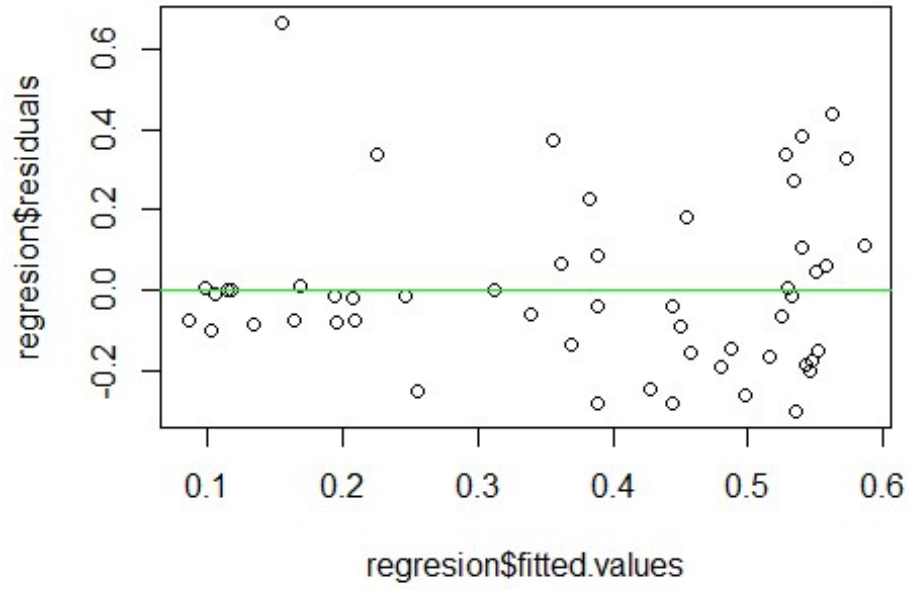
```r
PH=summary(regresion)
PH
```

```
##
## Call:
## lm(formula = M1$X7 ~ AlcalinidadBC)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.30290 -0.14602 -0.02215  0.06435  0.66667
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.58603    0.04667  12.556  < 2e-16 ***
## AlcalinidadBC -1.33380    0.23545  -5.665 6.84e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2091 on 51 degrees of freedom
## Multiple R-squared:  0.3862, Adjusted R-squared:  0.3742
## F-statistic: 32.09 on 1 and 51 DF,  p-value: 6.838e-07
```

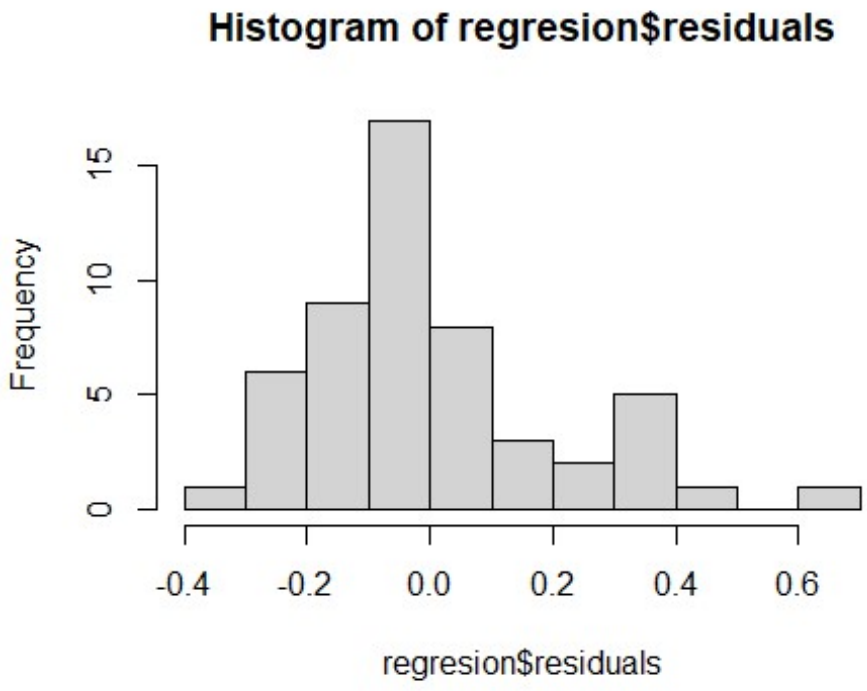
```r
plot(AlcalinidadBC,M1$X7,col="blue",xlab="Alcalinidad",ylab="Concentracion de
Mercurio",main="Alcalinidad Box-Cox vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```



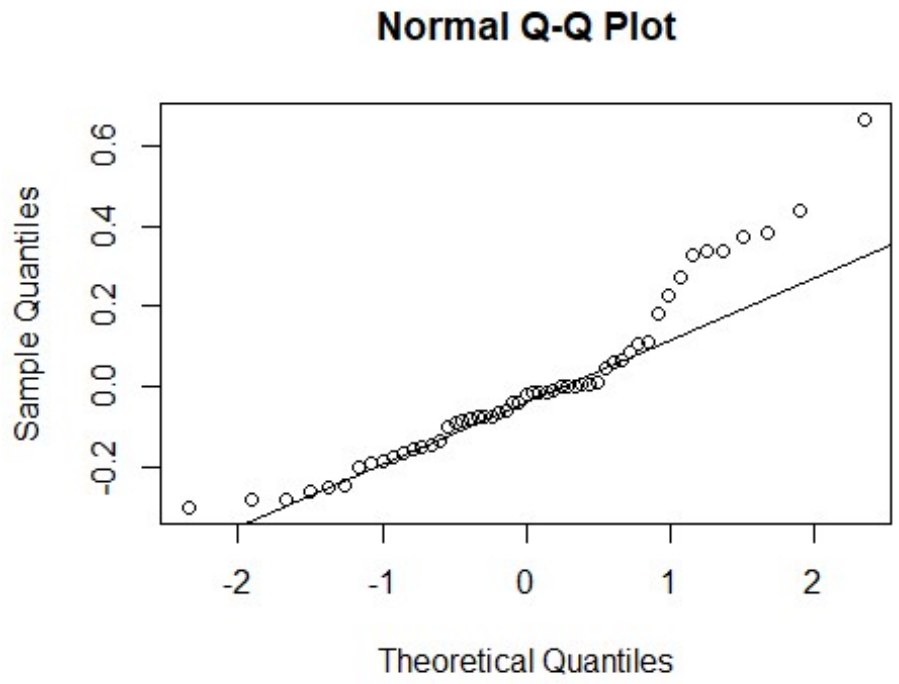

```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.91826, p-value = 0.001441

hist(regresion$residuals)
```




```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```



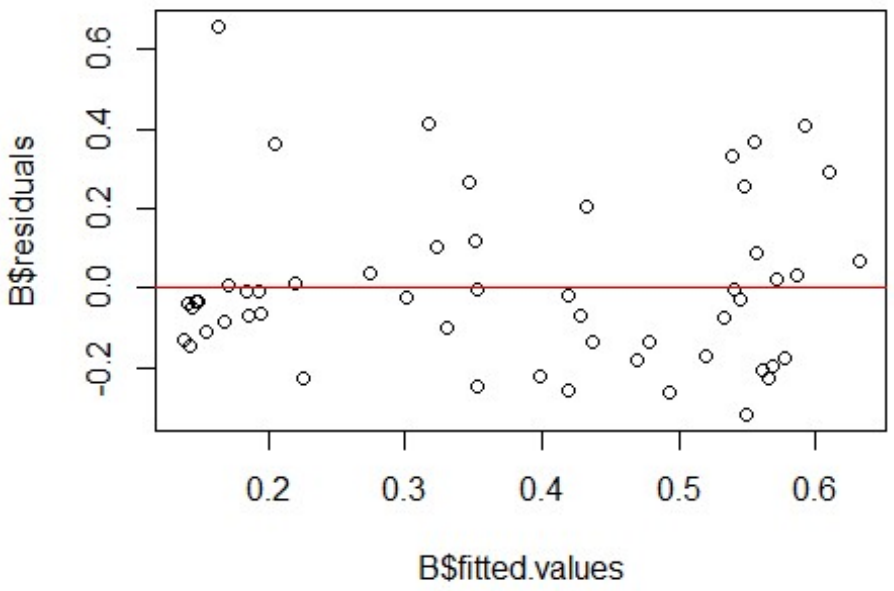

```
t.test(regresion$residuals)
```

```
##
##  One Sample t-test
##
## data:  regresion$residuals
## t = 3.6714e-17, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.05708903  0.05708903
## sample estimates:
##    mean of x
## 1.044503e-18
```

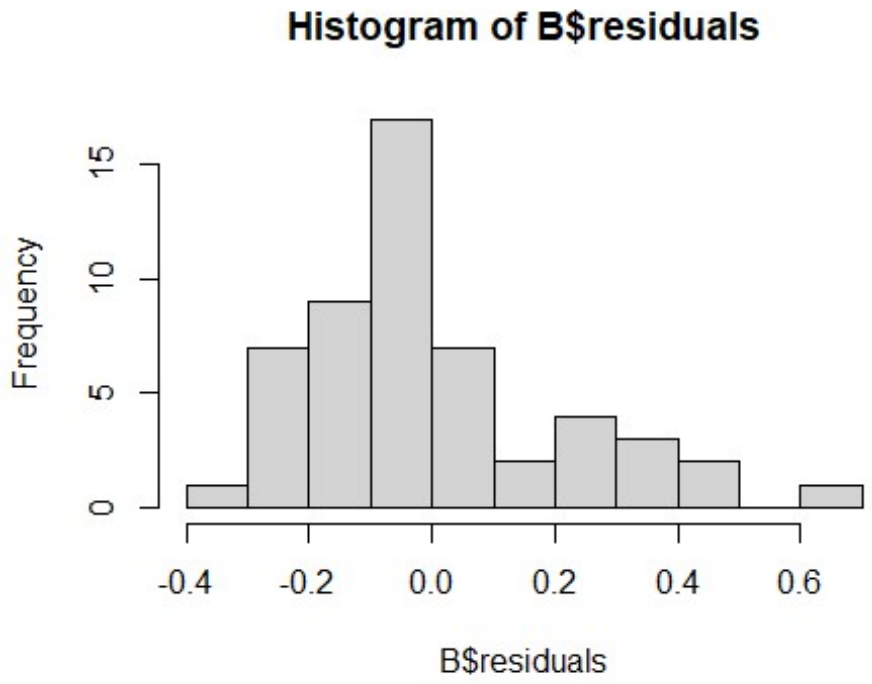
```
z1=AlcalinidadBC
z2=AlcalinidadBC^2
cor.test(M1$X7 ,z1+ z2)
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and z1 + z2
## t = -5.5505, df = 51, p-value = 1.028e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.7582009 -0.4116597
## sample estimates:
```

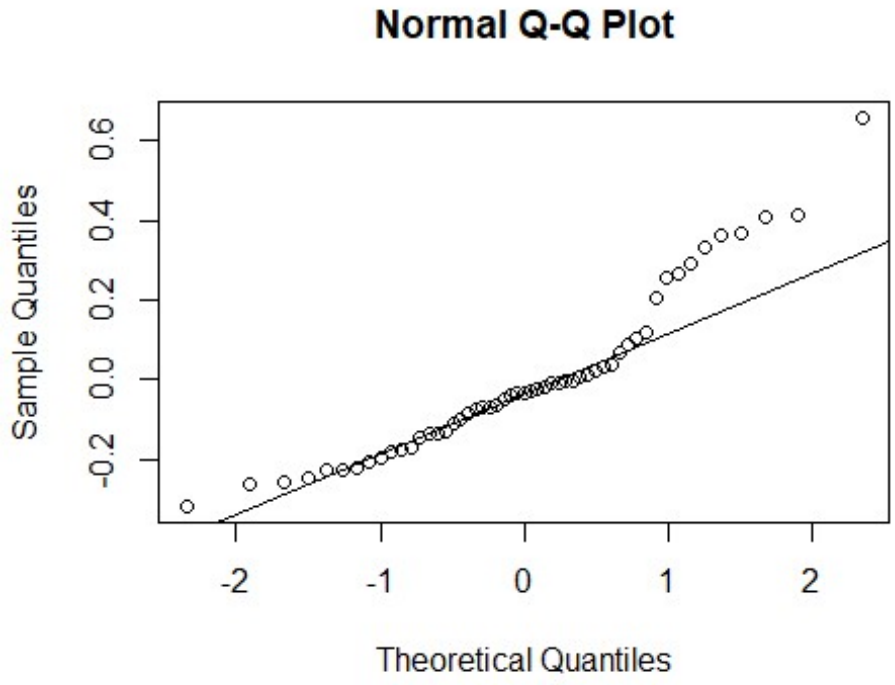
```
##       cor
## -0.613671

B = lm(M1$X7 ~ z1+ z2)
summary(B)

##
## Call:
## lm(formula = M1$X7 ~ z1 + z2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.31733 -0.13708 -0.03098  0.06570  0.65757
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.63197    0.06793   9.304 1.72e-12 ***
## z1          -2.26013    1.02158  -2.212   0.0315 *
## z2           2.51394    2.69759   0.932   0.3559
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2094 on 50 degrees of freedom
## Multiple R-squared:  0.3967, Adjusted R-squared:  0.3726
## F-statistic: 16.44 on 2 and 50 DF,  p-value: 3.262e-06

plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```




```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.91627, p-value = 0.001216

hist(B$residuals)
```

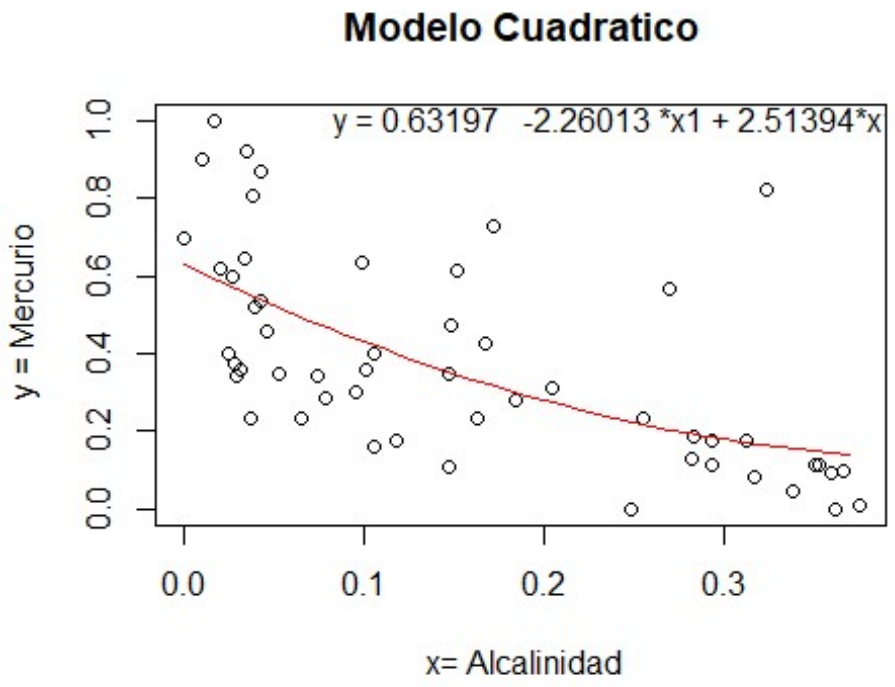



```
qqnorm(B$residuals)
qqline(B$residuals)
```



```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = -6.284e-17, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.05659959  0.05659959
## sample estimates:
##     mean of x
## -1.772464e-18
```

```
x=AlcalinidadBC
y=M1$X7
plot(x, y, main = "Modelo Cuadratico", xlab =" x= Alcalinidad", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = 0.63197   -2.26013 *x1 + 2.51394*x1^2
lines(x1, y1, col = "red")
text(.25, 1, "y = 0.63197   -2.26013 *x1 + 2.51394*x1^2")
```



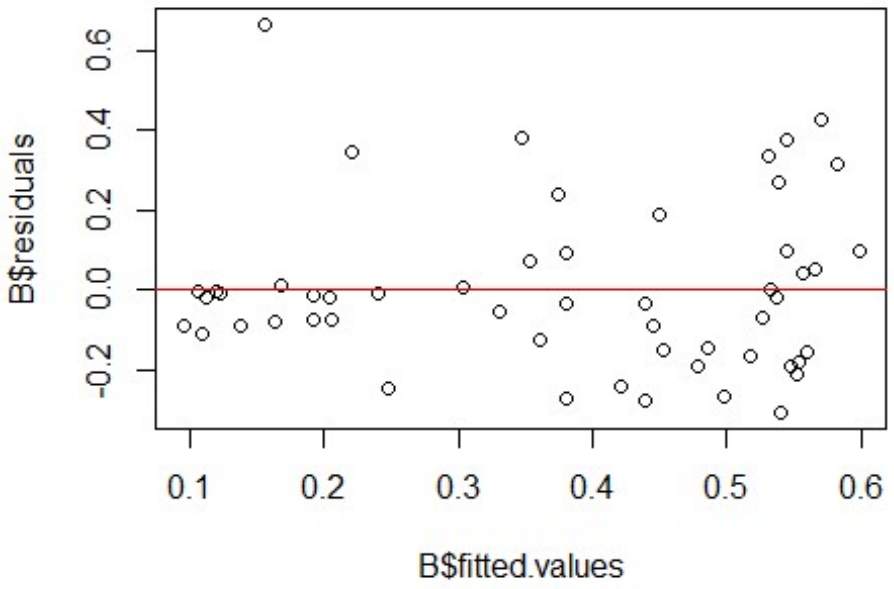

```
z1=AlcalinidadBC+1
z2=log(AlcalinidadBC+1)
cor.test(M1$X7, log(AlcalinidadBC+1))
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and log(AlcalinidadBC + 1)
## t = -5.7215, df = 51, p-value = 5.583e-07
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.7660753 -0.4271425
## sample estimates:
##       cor
## -0.6252489
```

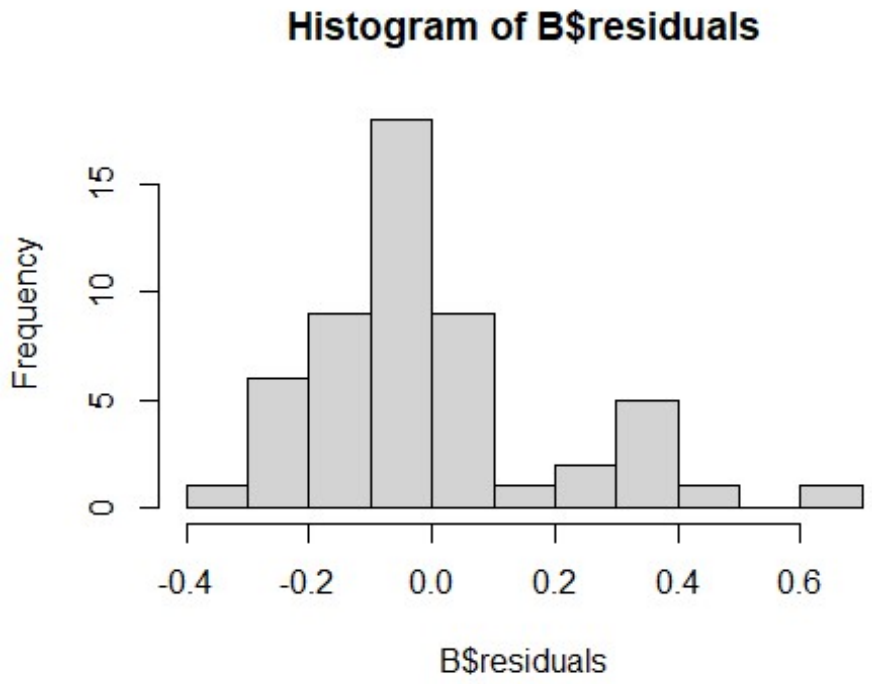
```
B = lm(M1$X7 ~ log(AlcalinidadBC+1))
summary(B)
```

```
##
## Call:
## lm(formula = M1$X7 ~ log(AlcalinidadBC + 1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.30699 -0.14424 -0.01974  0.07323  0.66559
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              0.59832    0.04803  12.458  < 2e-16 ***
## log(AlcalinidadBC + 1) -1.57928    0.27603  -5.721 5.58e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2083 on 51 degrees of freedom
## Multiple R-squared:  0.3909, Adjusted R-squared:  0.379
## F-statistic: 32.74 on 1 and 51 DF,  p-value: 5.583e-07
```

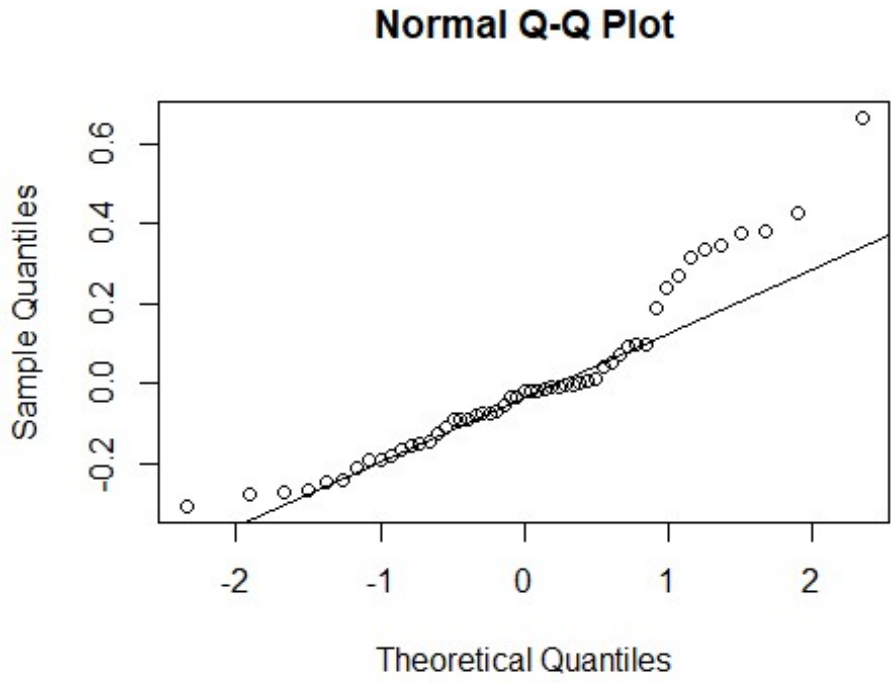
```
plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```



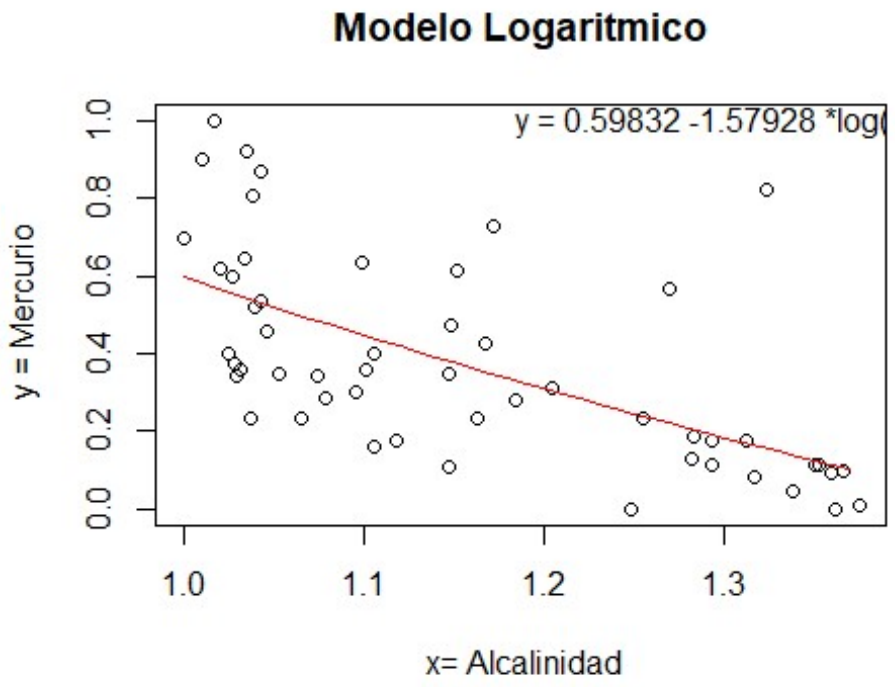

```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.91764, p-value = 0.001367

hist(B$residuals)
```




```
qqnorm(B$residuals)
qqline(B$residuals)
```



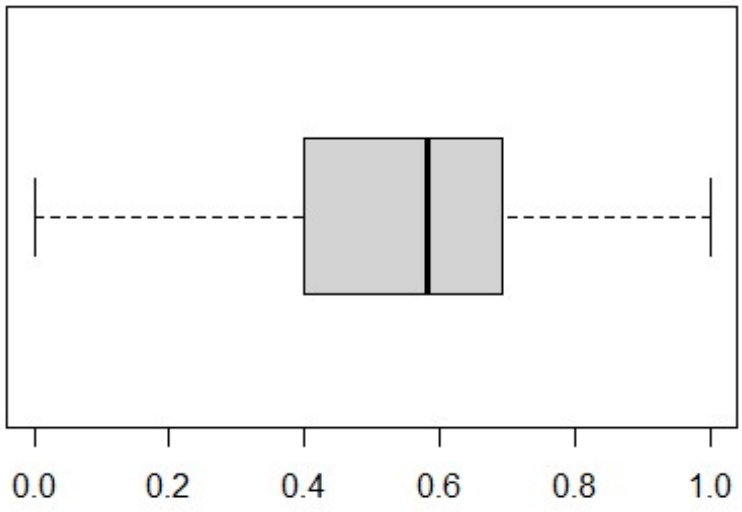
```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = 2.9567e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.05686885  0.05686885
## sample estimates:
##    mean of x
## 8.379489e-18

x=AlcalinidadBC+1
y=M1$X7
plot(x, y, main = "Modelo Logaritmico", xlab =" x= Alcalinidad", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = 0.59832 -1.57928 *log(x1)
lines(x1, y1, col = "red")
text(1.3, 1, "y = 0.59832 -1.57928 *log(x1)")
```



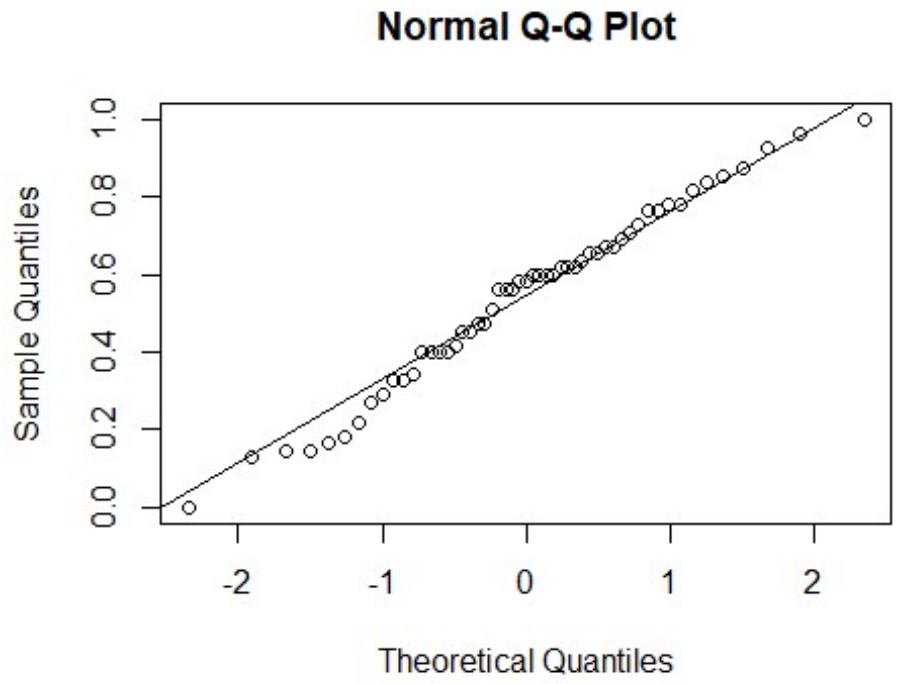
PH

```
q1=quantile(M1$X4,0.25)  #Cuantil 1 de la variable X
q2=quantile(M1$X4,0.5)
q3=quantile(M1$X4,0.75)
y1=min(M1$X4)
```

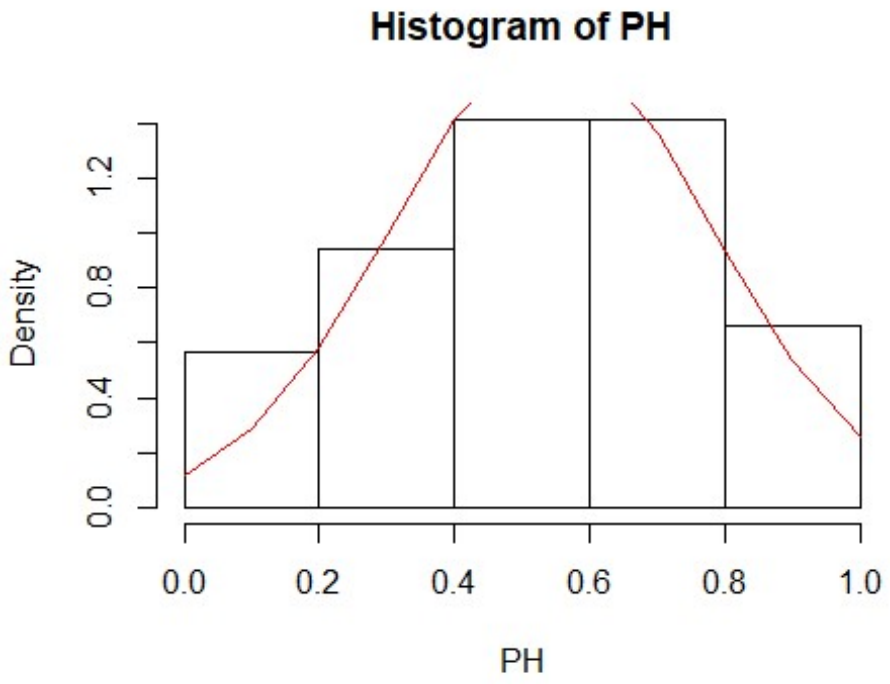
```
y2=max(M1$X4)
ri=IQR(M1$X4)
boxplot(M1$X4,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```



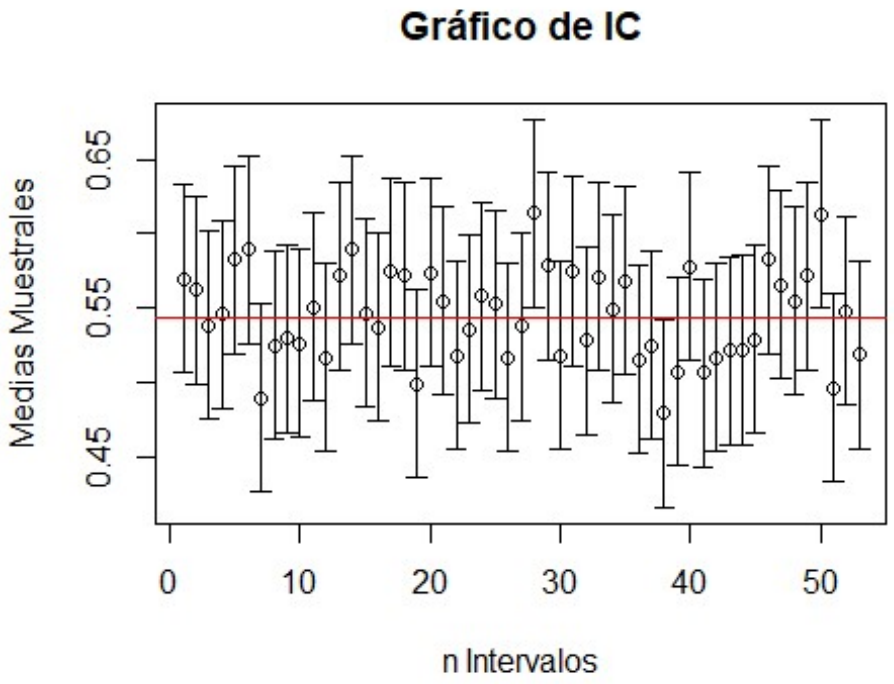

```
PH= M1[M1$X4<q3+1.5*ri,c("X4")]
qqnorm(PH)
qqline(PH)
```




```
hist(PH,prob=TRUE,col=0)
x=seq(min(PH),max(PH),0.1)
y=dnorm(x,mean(PH),sd(PH))
lines(x,y,col="red")
```



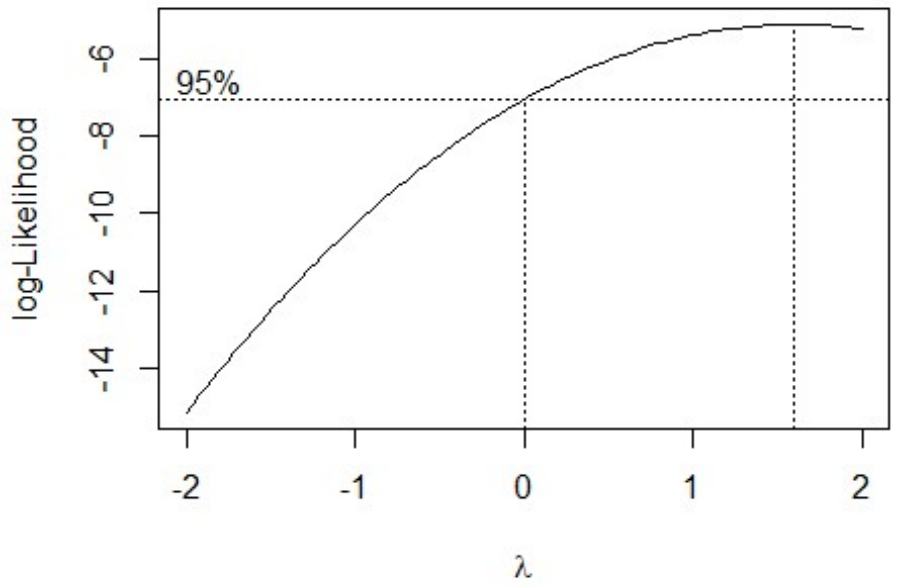

```
library(moments)
skewness1=skewness(PH)
cat("Sesgo: ",skewness1)

## Sesgo:  -0.2530037

cat("\n")

kurtosis1=kurtosis(PH)
cat("Curtosis: ",kurtosis1)

## Curtosis:  2.468301

cat("\n")

library(plotrix)
n=length(PH)
media=mean(PH)
DE=sd(PH)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```



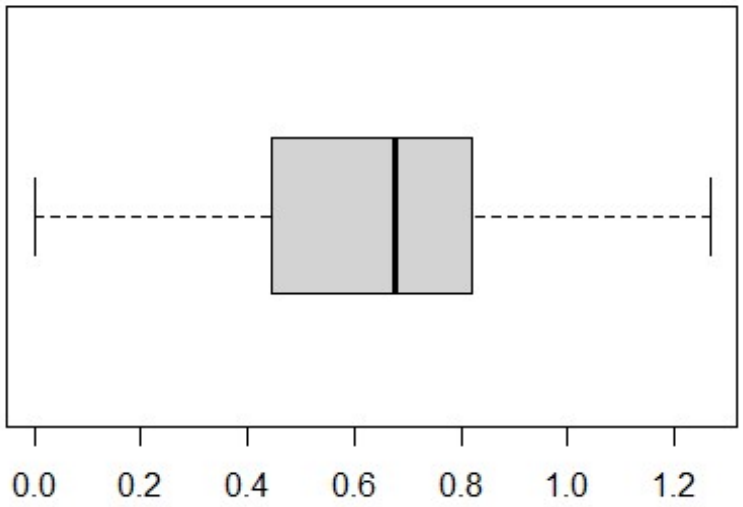

```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.5437393  +-  0.06306883
```

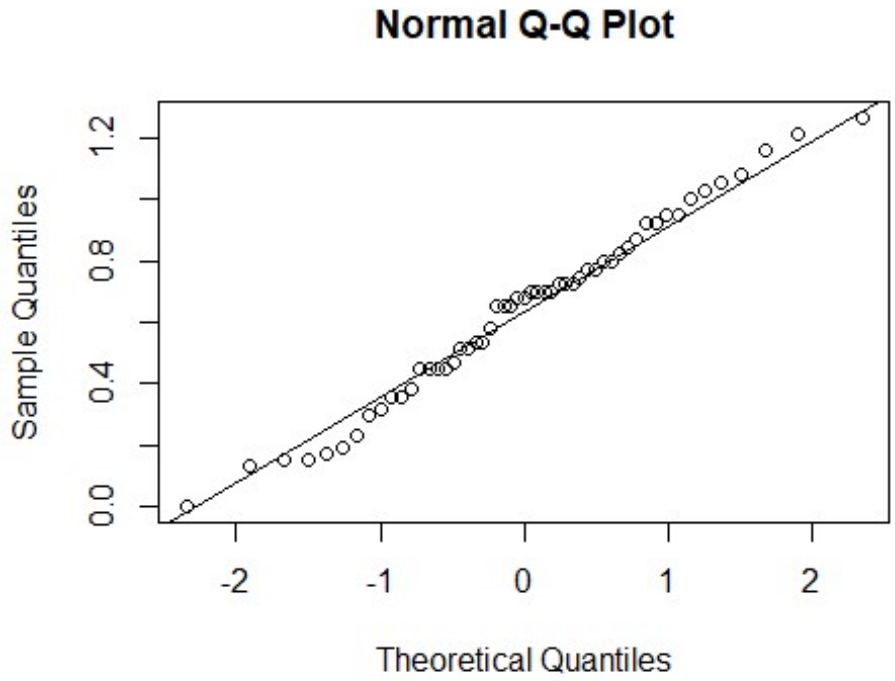
```
library(MASS)
minim=min(M1$X4)
b <- boxcox(lm((M1$X4+1) ~ 1))
```



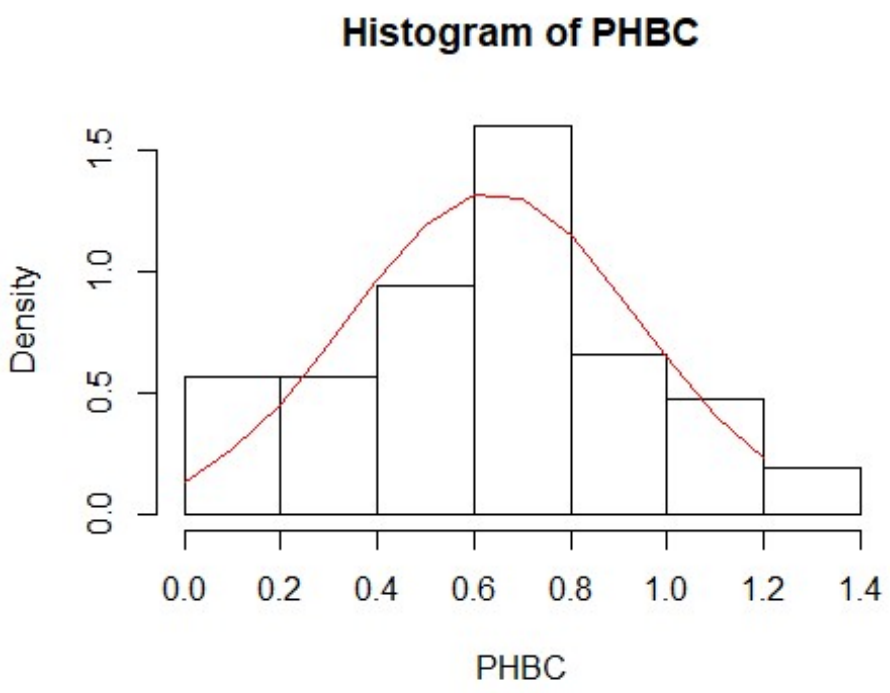

```
lambda <- b$x[which.max(b$y)]
X4<-((M1$X4+1)^lambda-1)/lambda
q1=quantile(X4,0.25)  #Cuantil 1 de la variable X
q2=quantile(X4,0.5)
q3=quantile(X4,0.75)
y1=min(X4)
y2=max(X4)
ri=IQR(X4)
boxplot(X4,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```




```
PHBC= X4
qqnorm(PHBC)
qqline(PHBC)
```

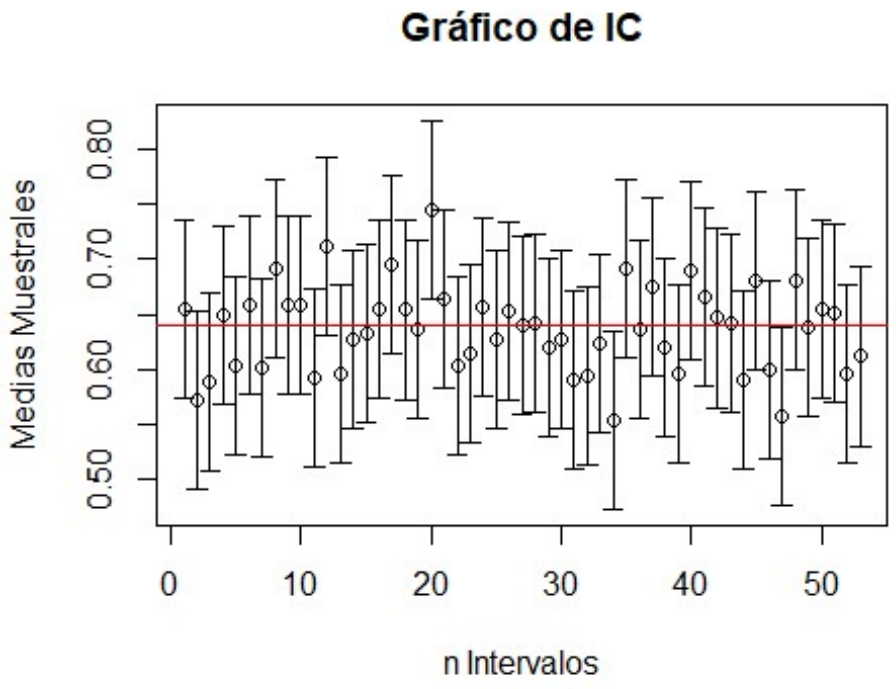


```
hist(PHBC,prob=TRUE,col=0)
x=seq(min(PHBC),max(PHBC),0.1)
y=dnorm(x,mean(PHBC),sd(PHBC))
lines(x,y,col="red")
```




```
library(moments)
skewness1=skewness(PHBC)
cat("Sesgo: ",skewness1)

## Sesgo:  -0.06020792

cat("\n")

kurtosis1=kurtosis(PHBC)
cat("Curtosis: ",kurtosis1)

## Curtosis:  2.413642

cat("\n")

library(plotrix)
n=length(PHBC)
media=mean(PHBC)
DE=sd(PHBC)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
```

```
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```




```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.6399393  +-  0.08069817
```

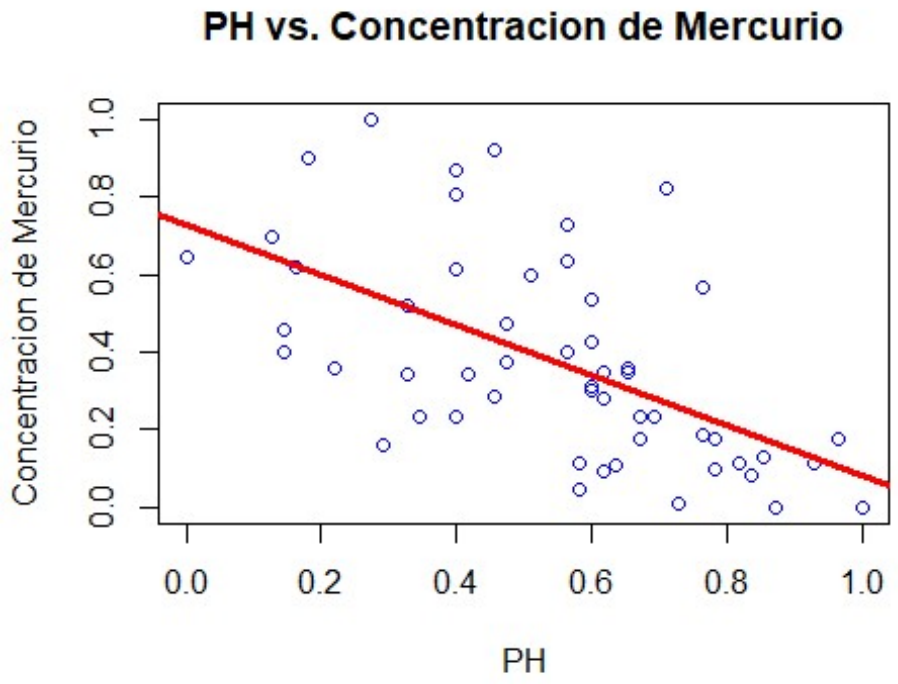
```
cor.test(M1$X7,M1$X4)
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and M1$X4
## t = -5.0242, df = 51, p-value = 6.573e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.7318685 -0.3612979
## sample estimates:
##        cor
## -0.5754001
```

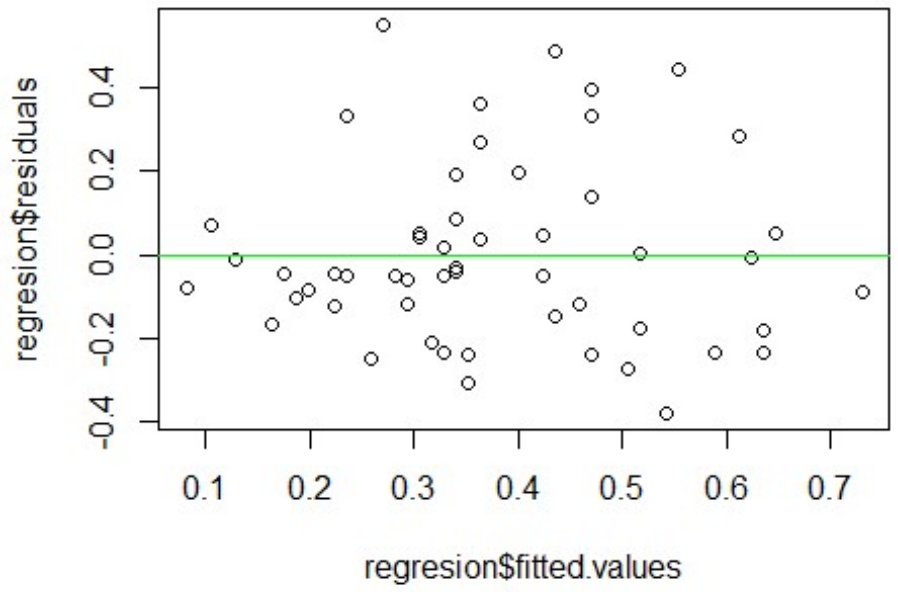
```
regresion=lm(M1$X7 ~ M1$X4)
regresion
```

```
##
## Call:
## lm(formula = M1$X7 ~ M1$X4)
```

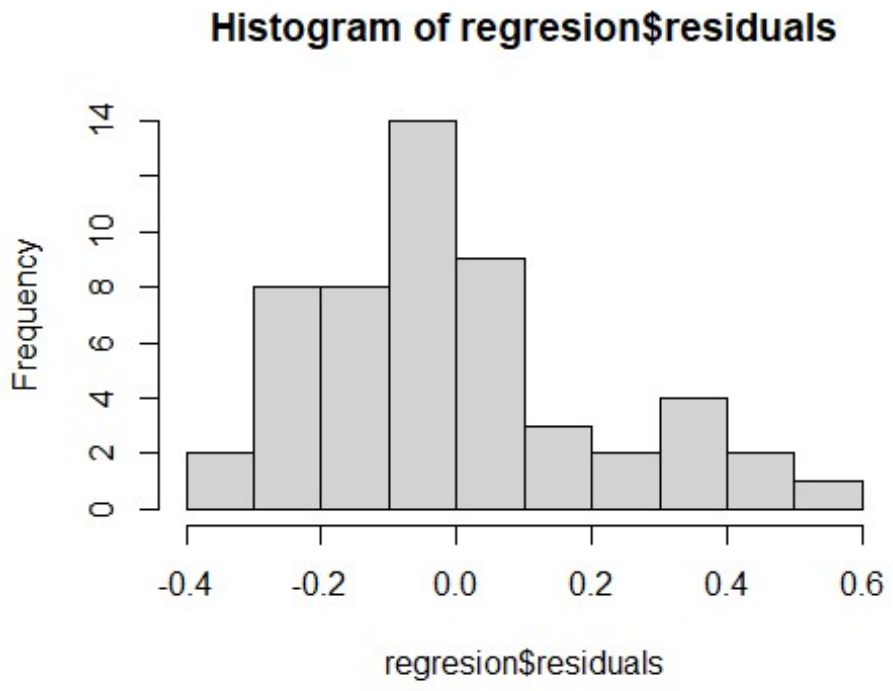
```
##
## Coefficients:
## (Intercept)        M1$X4
##      0.7307      -0.6493

PH=summary(regresion)
PH

##
## Call:
## lm(formula = M1$X7 ~ M1$X4)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.37903 -0.14875 -0.04476  0.07330  0.55142
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.73073    0.07641   9.564 5.77e-13 ***
## M1$X4       -0.64934    0.12924  -5.024 6.57e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2183 on 51 degrees of freedom
## Multiple R-squared:  0.3311, Adjusted R-squared:  0.318
## F-statistic: 25.24 on 1 and 51 DF,  p-value: 6.573e-06

plot(M1$X4,M1$X7,col="blue",xlab="PH",ylab="Concentracion de
Mercurio",main="PH vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```



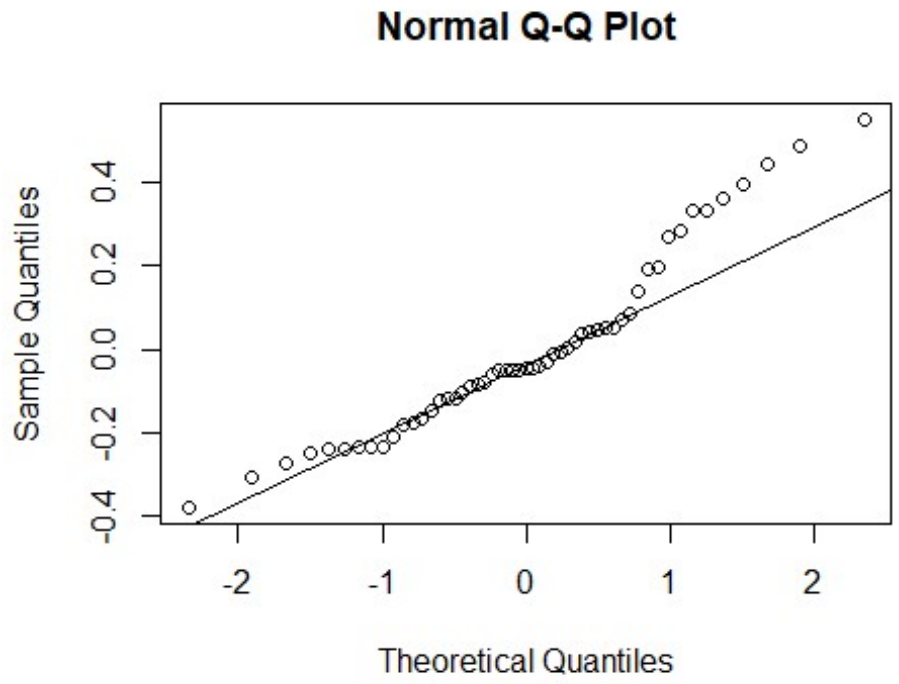

```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.93853, p-value = 0.008887

hist(regresion$residuals)
```




```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```




```
t.test(regresion$residuals)

##
##  One Sample t-test
##
## data:  regresion$residuals
## t = 8.3755e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.05959755  0.05959755
## sample estimates:
##    mean of x
## 2.487541e-17

cor.test(M1$X7,PHBC)

##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and PHBC
## t = -5.0981, df = 51, p-value = 5.081e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.7357608 -0.3686084
## sample estimates:
##       cor
## -0.5810142
```

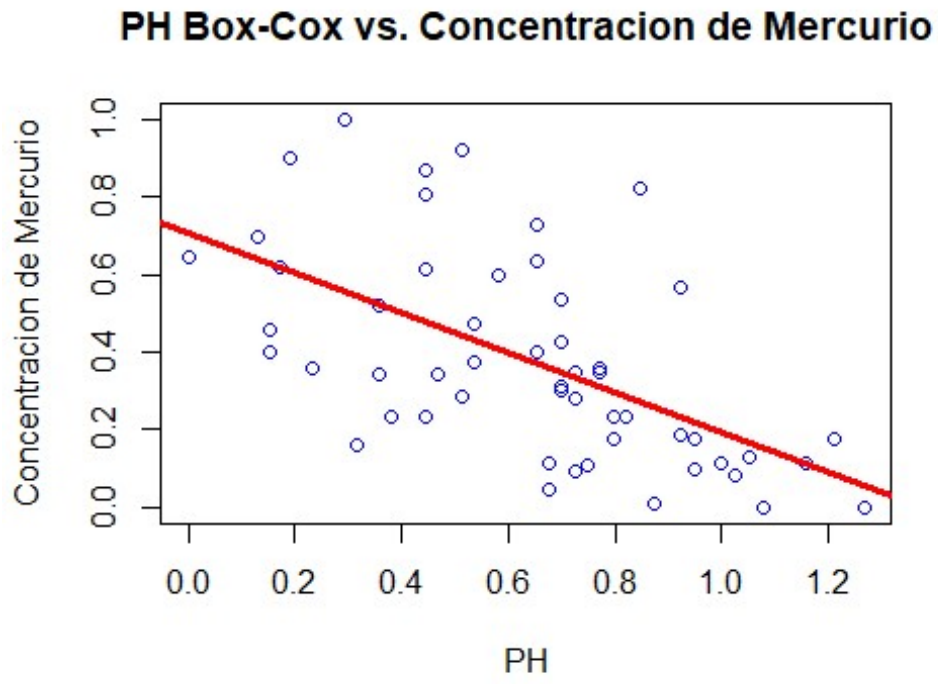
```r
regresion=lm(M1$X7 ~ PHBC)
regresion
```

```
##
## Call:
## lm(formula = M1$X7 ~ PHBC)
##
## Coefficients:
## (Intercept)         PHBC
##      0.7056      -0.5124
```

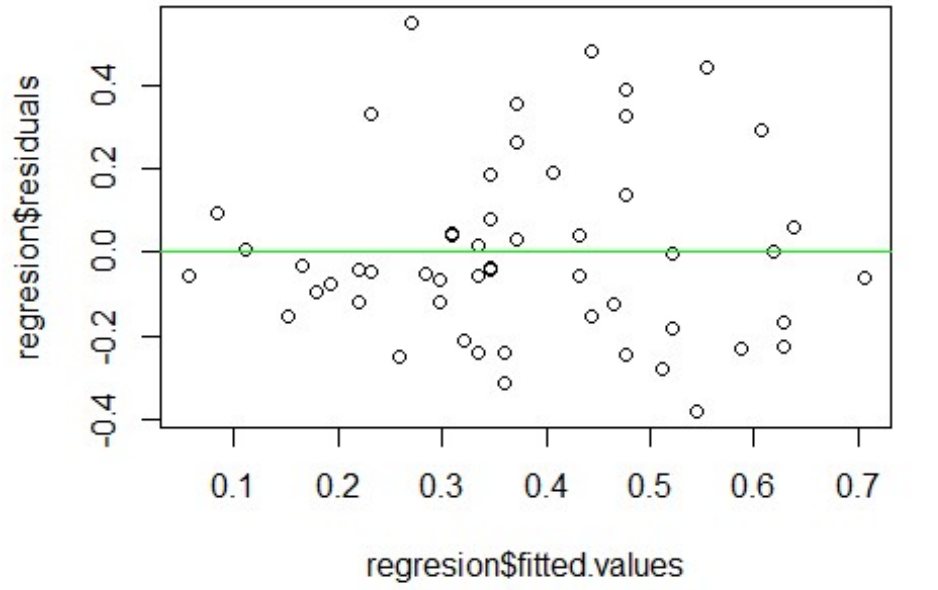
```r
PH=summary(regresion)
PH
```

```
##
## Call:
## lm(formula = M1$X7 ~ PHBC)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.38126 -0.15273 -0.04453  0.07950  0.55031
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.70558    0.07091   9.950 1.53e-13 ***
## PHBC        -0.51244    0.10052  -5.098 5.08e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2173 on 51 degrees of freedom
## Multiple R-squared:  0.3376, Adjusted R-squared:  0.3246
## F-statistic: 25.99 on 1 and 51 DF,  p-value: 5.081e-06
```

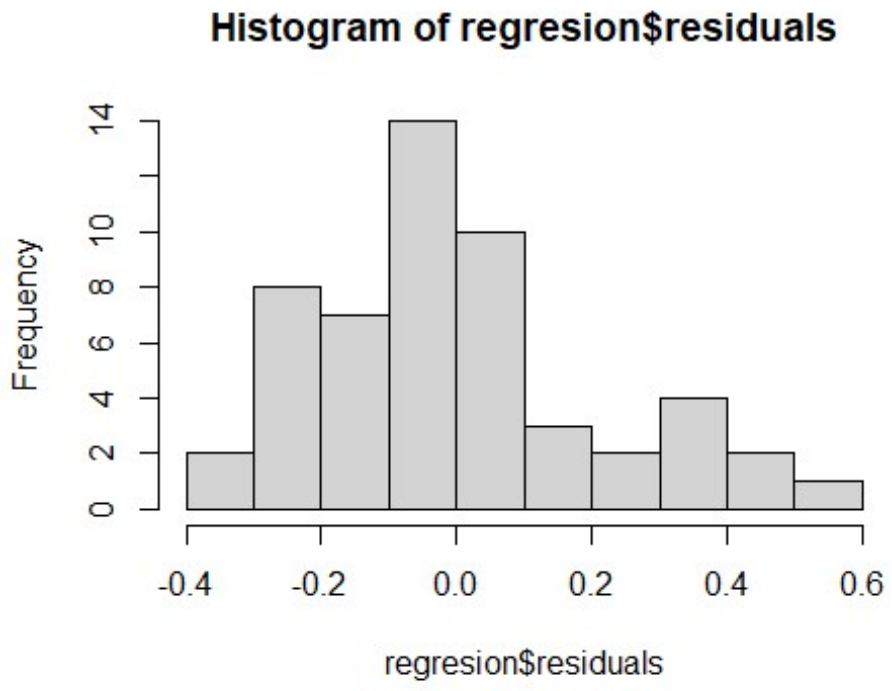
```r
plot(PHBC,M1$X7,col="blue",xlab="PH",ylab="Concentracion de
Mercurio",main="PH Box-Cox vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```



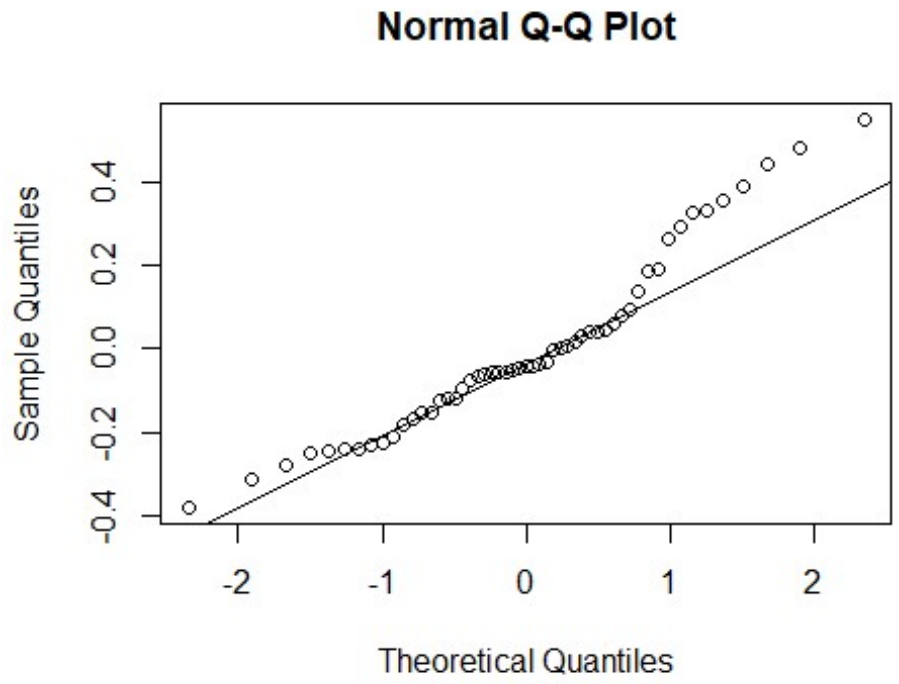

```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.94241, p-value = 0.01285

hist(regresion$residuals)
```




```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```




```
t.test(regresion$residuals)
```

```
##
##  One Sample t-test
##
## data:  regresion$residuals
## t = 9.3433e-17, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.05930763  0.05930763
## sample estimates:
##    mean of x
## 2.761455e-18
```

```
z1=PHBC
z2=PHBC^2
cor.test(M1$X7 , z1+ z2)
```

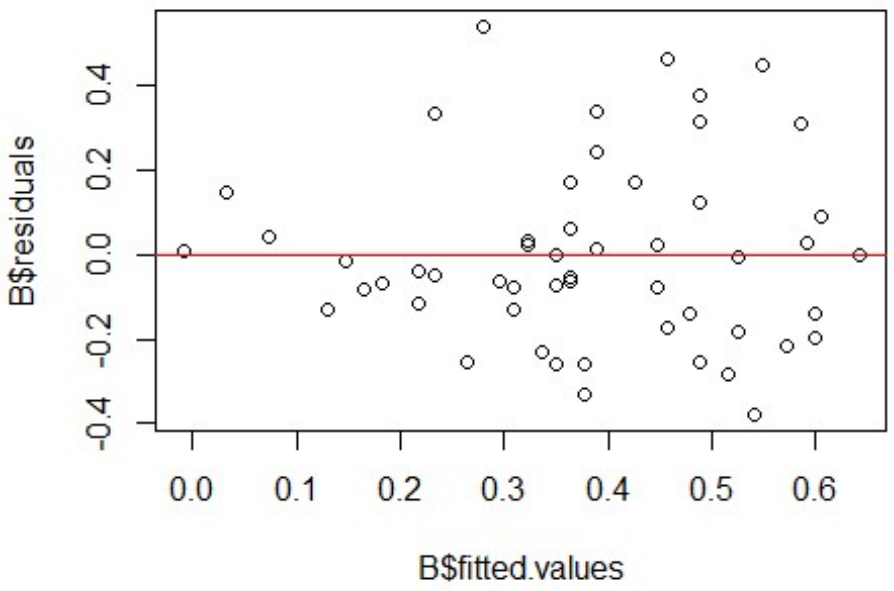
```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and z1 + z2
## t = -5.1745, df = 51, p-value = 3.888e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.7397192 -0.3760899
## sample estimates:
```

```
##        cor
## -0.5867389
```

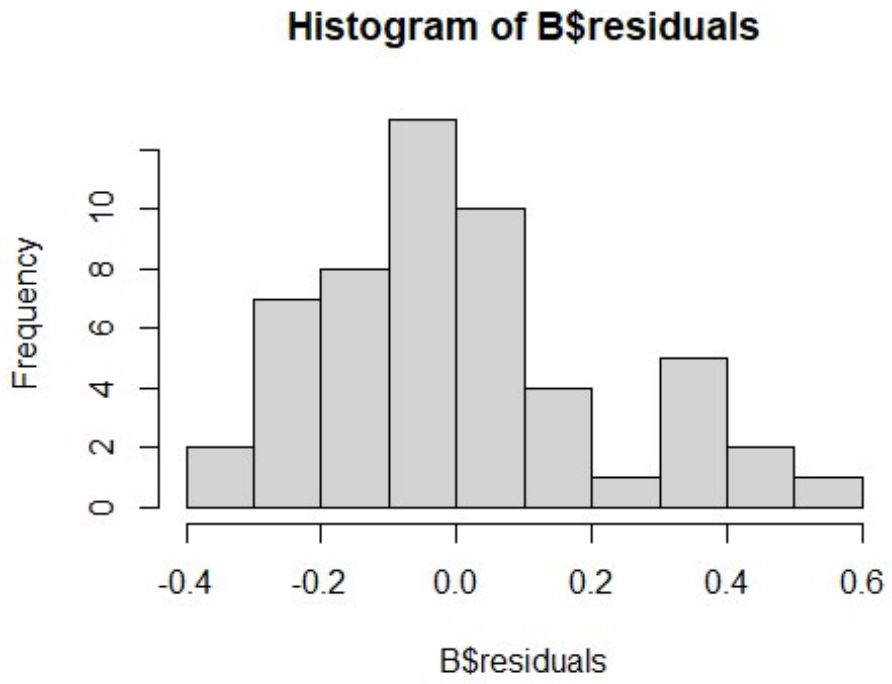
```
B = lm(M1$X7 ~ z1+ z2)
summary(B)
```

```
##
## Call:
## lm(formula = M1$X7 ~ z1 + z2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.37890 -0.13748 -0.03855  0.09279  0.54179
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.6419     0.1138   5.638 7.94e-07 ***
## z1           -0.2530     0.3754  -0.674    0.503
## z2           -0.2056     0.2865  -0.718    0.476
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2183 on 50 degrees of freedom
## Multiple R-squared:  0.3443, Adjusted R-squared:  0.3181
## F-statistic: 13.13 on 2 and 50 DF,  p-value: 2.613e-05
```

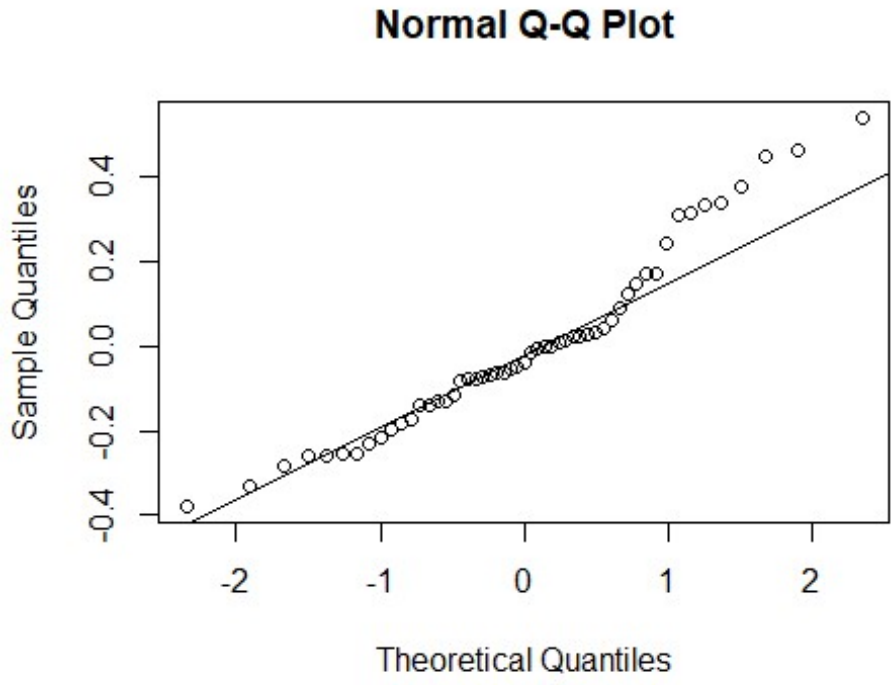
```
plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```



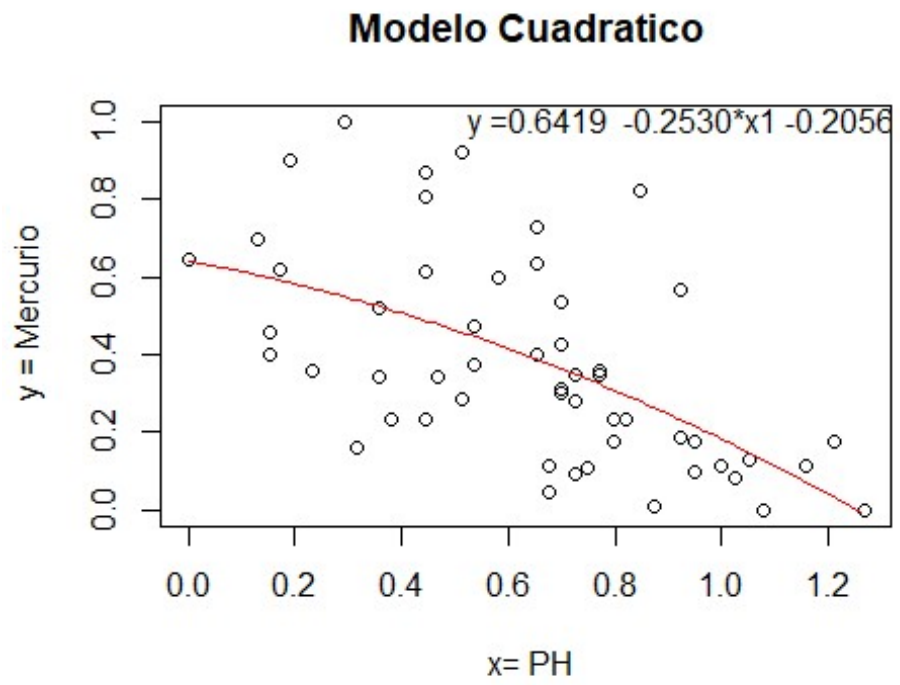

```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.951, p-value = 0.02971

hist(B$residuals)
```




```
qqnorm(B$residuals)
qqline(B$residuals)
```



```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = 2.2312e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.05900453  0.05900453
## sample estimates:
##    mean of x
## 6.560702e-18

x=PHBC
y=M1$X7
plot(x, y, main = "Modelo Cuadratico", xlab =" x= PH", ylab =" y = Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = 0.6419  -0.2530*x1 -0.2056*x1^2
lines(x1, y1, col = "red")
text(1, 1, "y =0.6419  -0.2530*x1 -0.2056*x1^2")
```



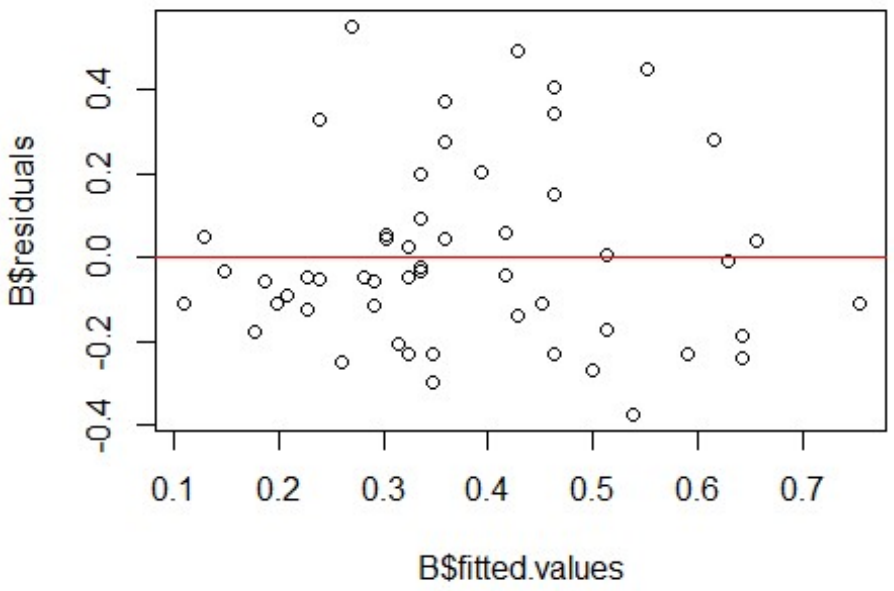

```
z1=PHBC+1
z2=log(PHBC+1)
cor.test(M1$X7 , log(PHBC+1))
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and log(PHBC + 1)
## t = -4.9253, df = 51, p-value = 9.268e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.7265433 -0.3513693
## sample estimates:
##       cor
## -0.5677429
```

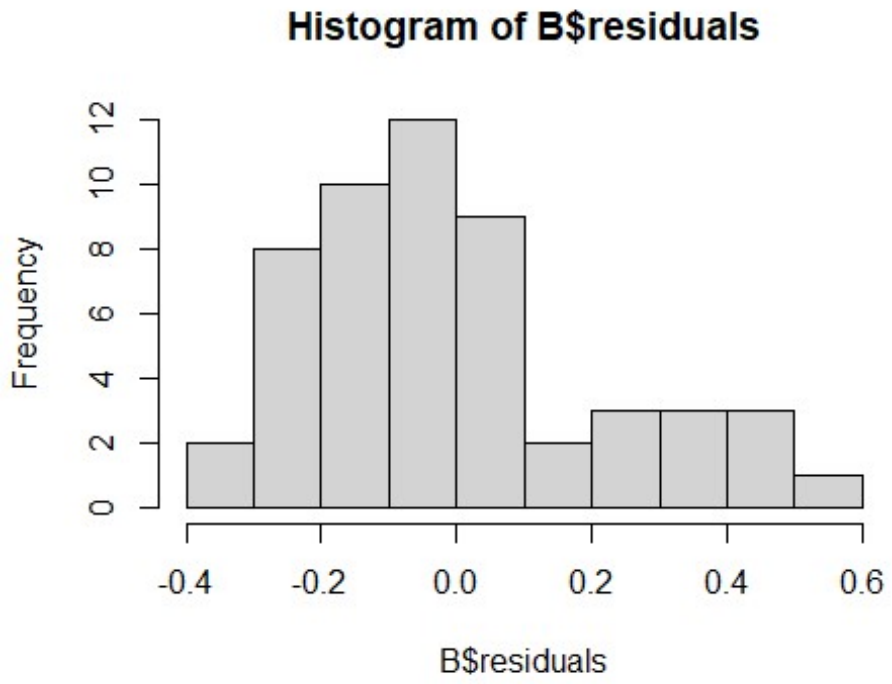
```
B = lm(M1$X7 ~ log(PHBC+1))
summary(B)
```

```
##
## Call:
## lm(formula = M1$X7 ~ log(PHBC + 1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.37547 -0.14093 -0.04545  0.05693  0.55164
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.75459    0.08227   9.172 2.25e-12 ***
## log(PHBC + 1) -0.78949    0.16029  -4.925 9.27e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2198 on 51 degrees of freedom
## Multiple R-squared:  0.3223, Adjusted R-squared:  0.309
## F-statistic: 24.26 on 1 and 51 DF,  p-value: 9.268e-06
```

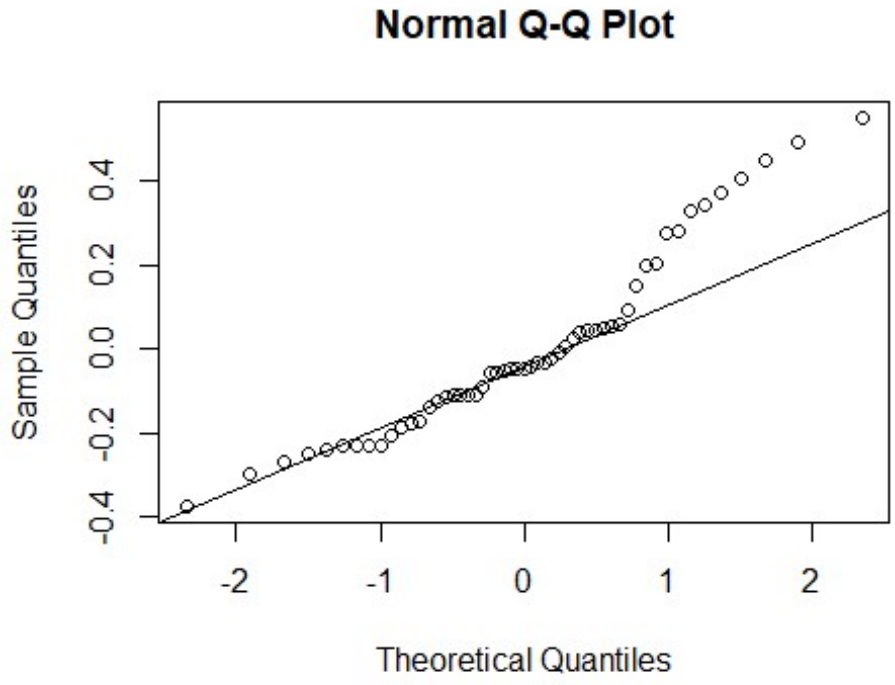
```
plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```



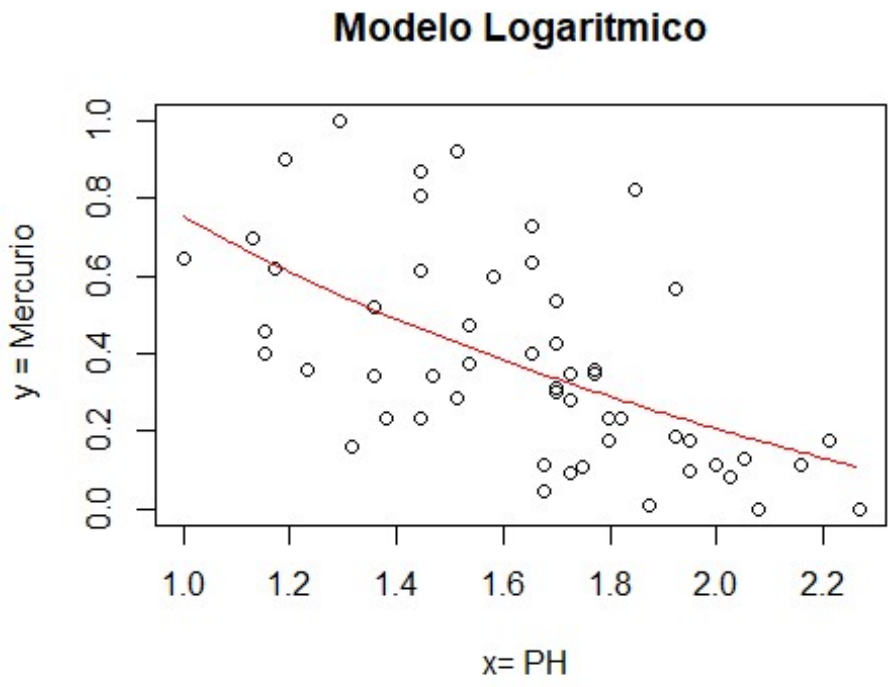

```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.93411, p-value = 0.005889

hist(B$residuals)
```




```
qqnorm(B$residuals)
qqline(B$residuals)
```



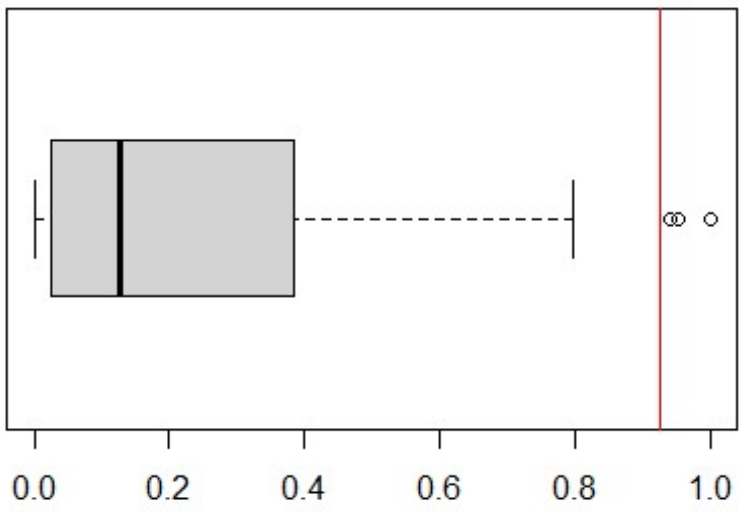
```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = -1.5739e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.05998623  0.05998623
## sample estimates:
##     mean of x
## -4.705092e-18

x=PHBC+1
y=M1$X7
plot(x, y, main = "Modelo Logaritmico", xlab =" x= PH", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = 0.75459 -0.78949*log(x1)
lines(x1, y1, col = "red")
text(1, 1.8, "y = 0.75459 -0.78949*log(x1)")
```



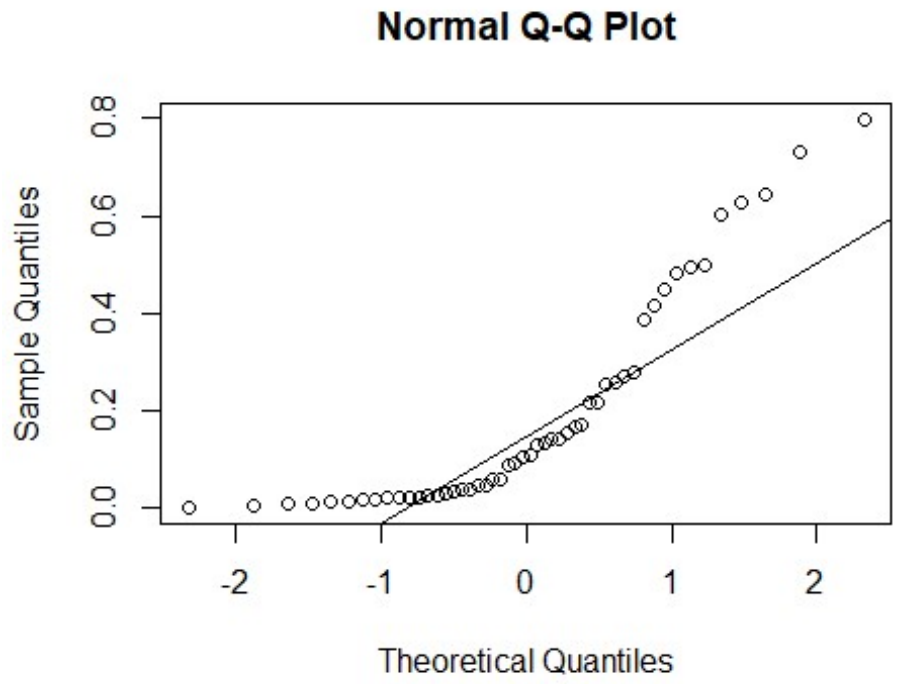
CALCIO

```
q1=quantile(M1$X5,0.25)  #Cuantil 1 de la variable X
q2=quantile(M1$X5,0.5)
```

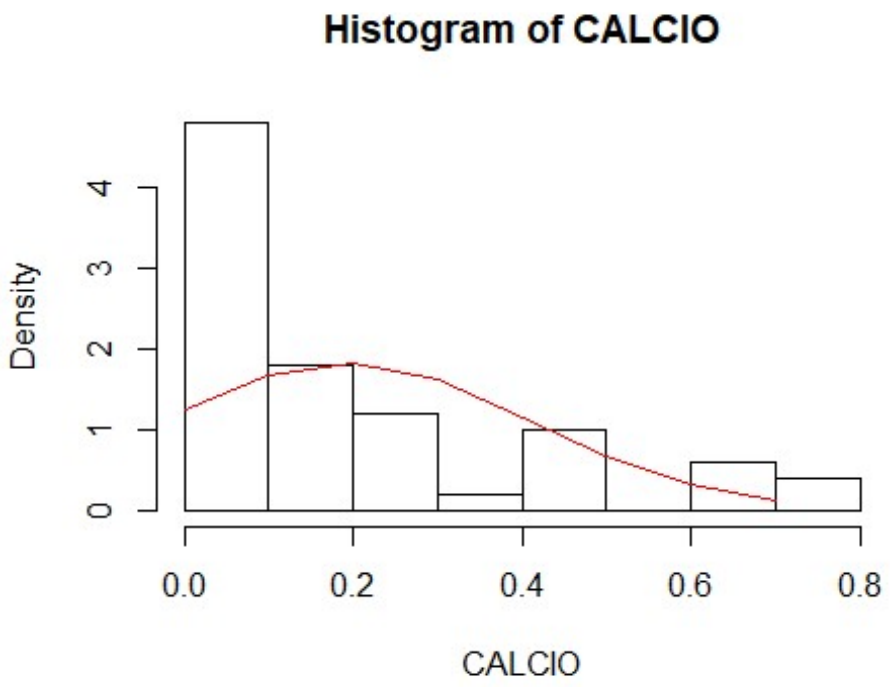
```
q3=quantile(M1$X5,0.75)
y1=min(M1$X5)
y2=max(M1$X5)
ri=IQR(M1$X5)
boxplot(M1$X5,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```



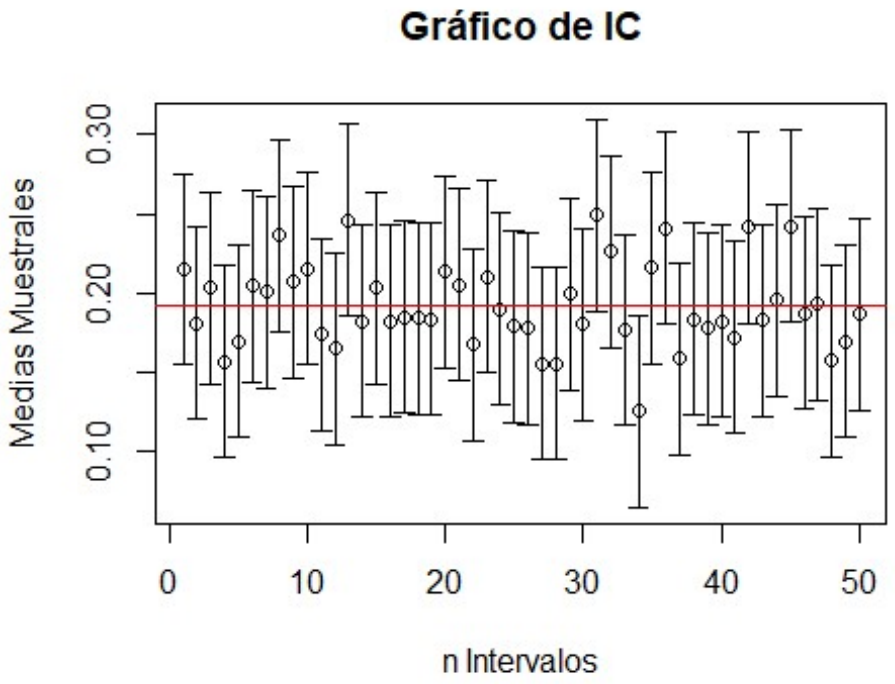

```
CALCIO= M1[M1$X5<q3+1.5*ri,c("X5")]
qqnorm(CALCIO)
qqline(CALCIO)
```




```
hist(CALCIO,prob=TRUE,col=0)
x=seq(min(CALCIO),max(CALCIO),0.1)
y=dnorm(x,mean(CALCIO),sd(CALCIO))
lines(x,y,col="red")
```



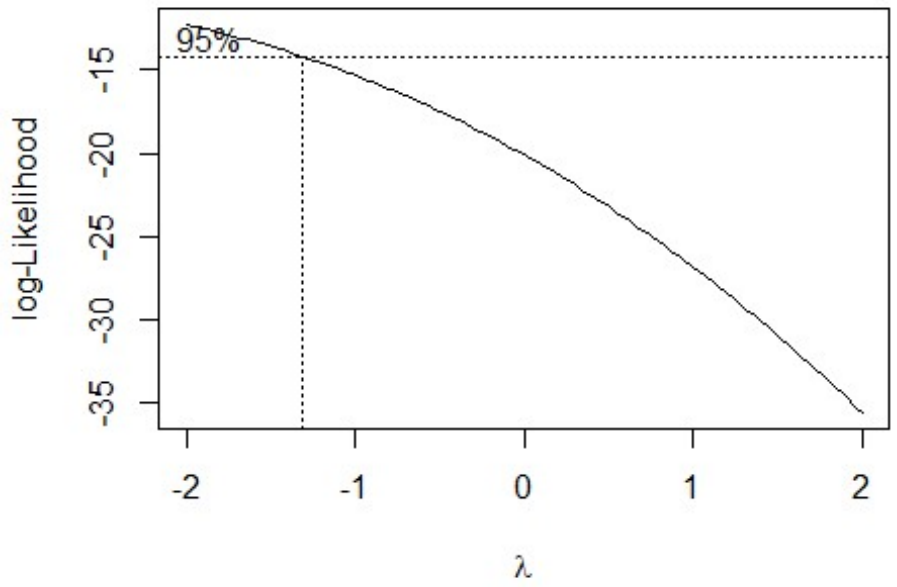

```
library(moments)
skewness1=skewness(CALCIO)
cat("Sesgo: ",skewness1)

## Sesgo:  1.253402

cat("\n")

kurtosis1=kurtosis(CALCIO)
cat("Curtosis: ",kurtosis1)

## Curtosis:  3.454808

cat("\n")

library(plotrix)
n=length(CALCIO)
media=mean(CALCIO)
DE=sd(CALCIO)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```



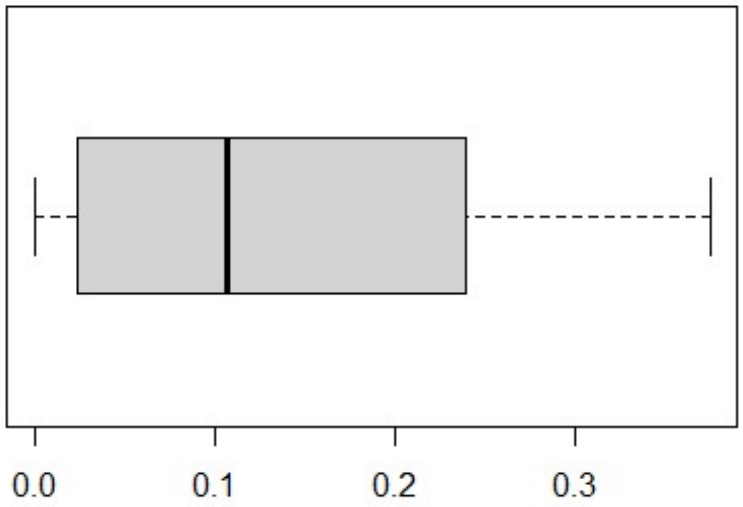

```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.1917411  +-  0.06047063
```

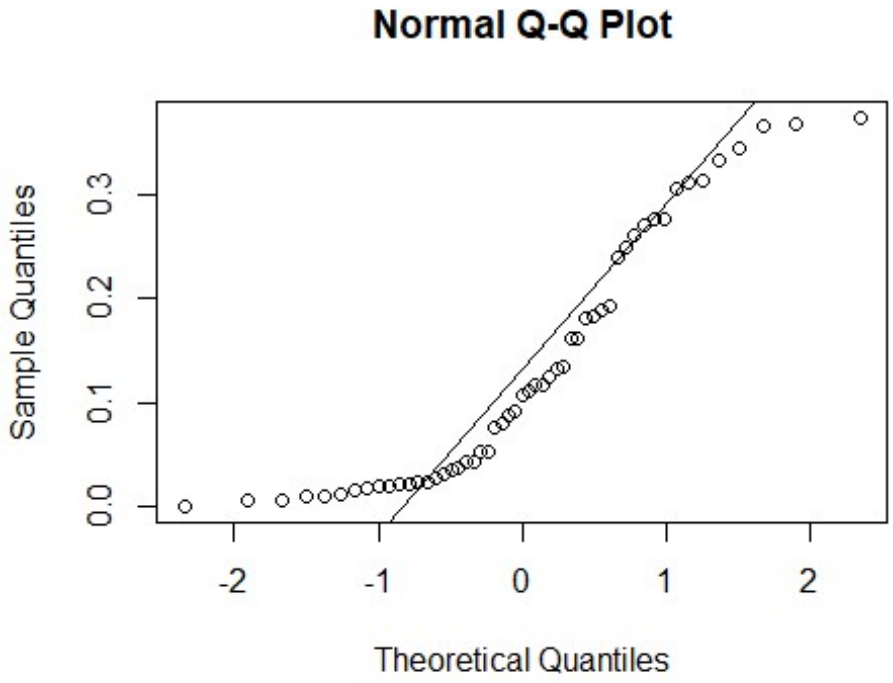
```
library(MASS)
minim=min(M1$X5)
b <- boxcox(lm((M1$X5+1) ~ 1))
```



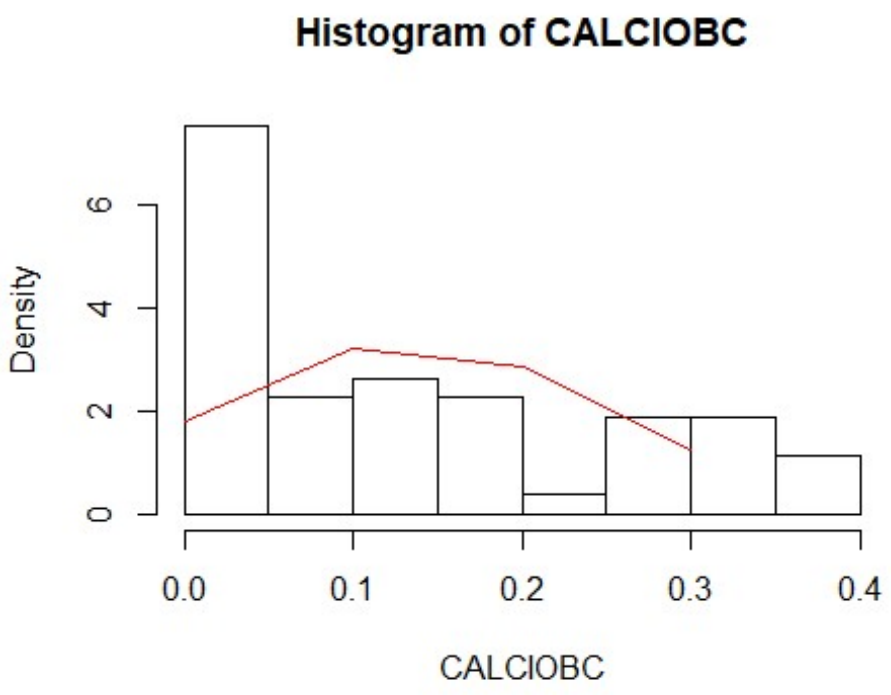

```
lambda <- b$x[which.max(b$y)]
X5<-((M1$X5+1)^lambda-1)/lambda
q1=quantile(X5,0.25)  #Cuantil 1 de la variable X
q2=quantile(X5,0.5)
q3=quantile(X5,0.75)
y1=min(X5)
y2=max(X5)
ri=IQR(X5)
boxplot(X5,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```




```
CALCIOBC= X5
qqnorm(CALCIOBC)
qqline(CALCIOBC)
```

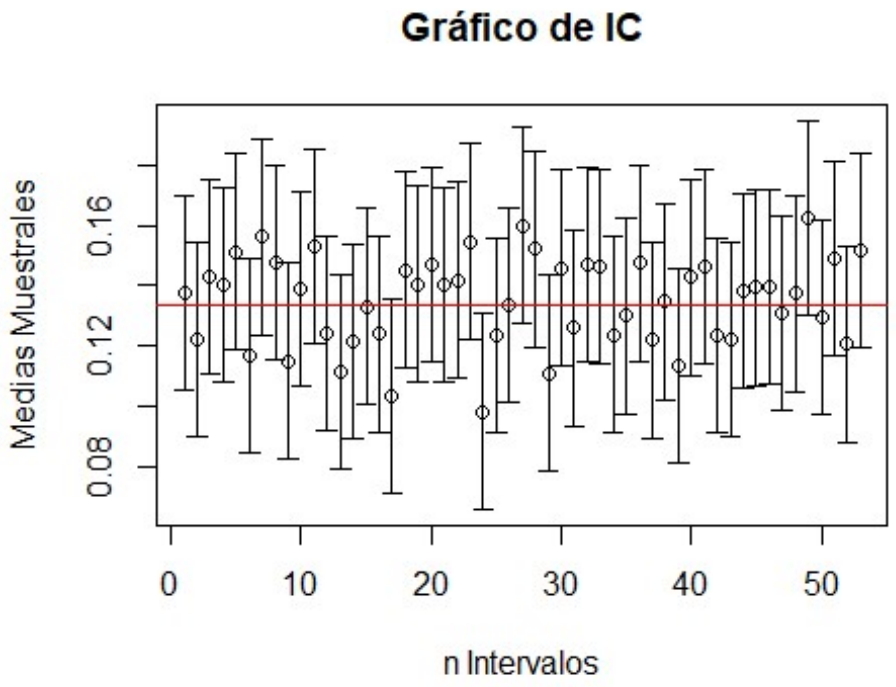


```
hist(CALCIOBC,prob=TRUE,col=0)
x=seq(min(CALCIOBC),max(CALCIOBC),0.1)
y=dnorm(x,mean(CALCIOBC),sd(CALCIOBC))
lines(x,y,col="red")
```




```
library(moments)
skewness1=skewness(CALCIOBC)
cat("Sesgo: ",skewness1)

## Sesgo:  0.6486829

cat("\n")

kurtosis1=kurtosis(CALCIOBC)
cat("Curtosis: ",kurtosis1)

## Curtosis:  2.05392

cat("\n")

library(plotrix)
n=length(CALCIOBC)
media=mean(CALCIOBC)
DE=sd(CALCIOBC)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
```

```
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```




```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.1334733  +-  0.03217819
```

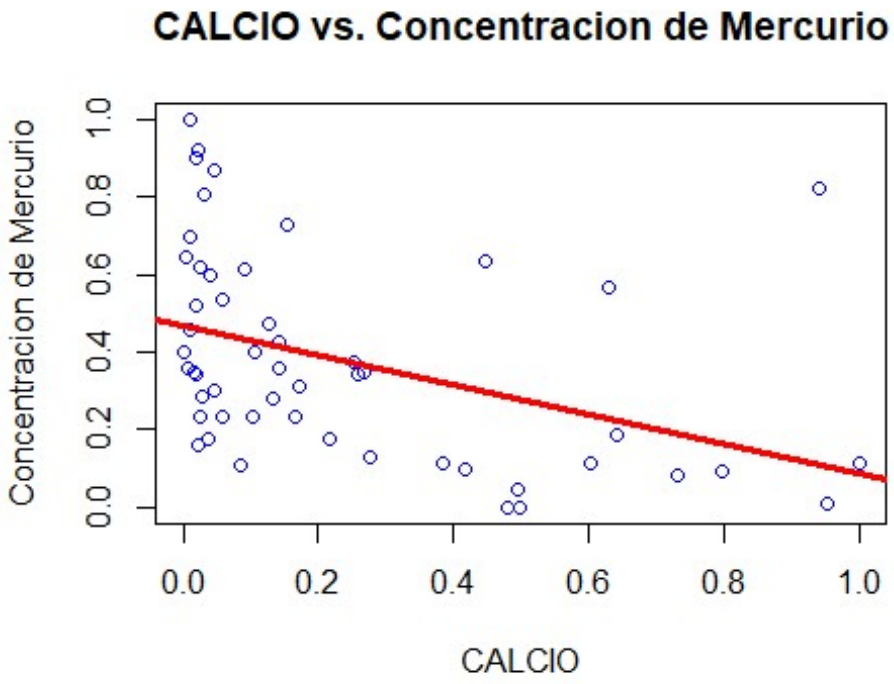
```
cor.test(M1$X7,M1$X5)
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and M1$X5
## t = -3.1231, df = 51, p-value = 0.002948
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.6054071 -0.1462218
## sample estimates:
##        cor
## -0.4006796
```

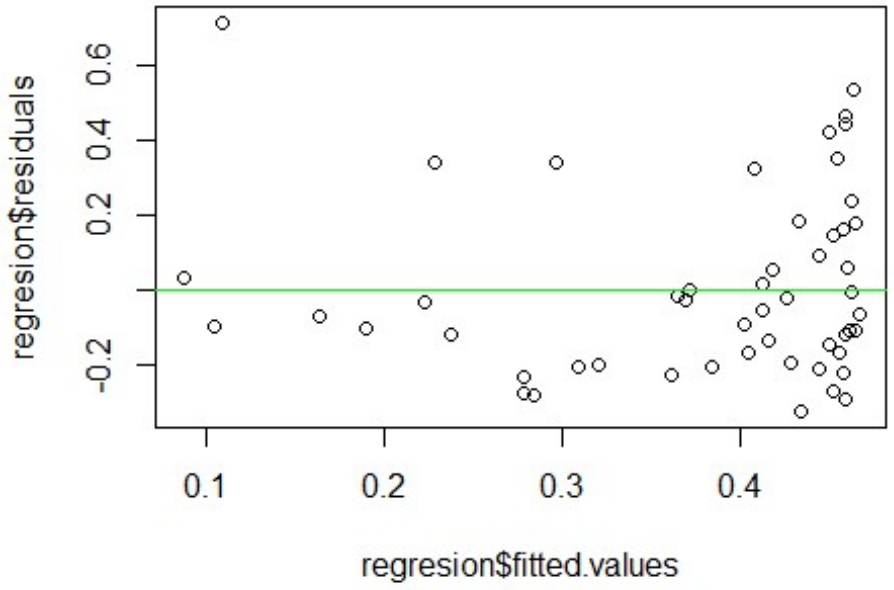
```
regresion=lm(M1$X7 ~ M1$X5)
regresion
```

```
##
## Call:
## lm(formula = M1$X7 ~ M1$X5)
```

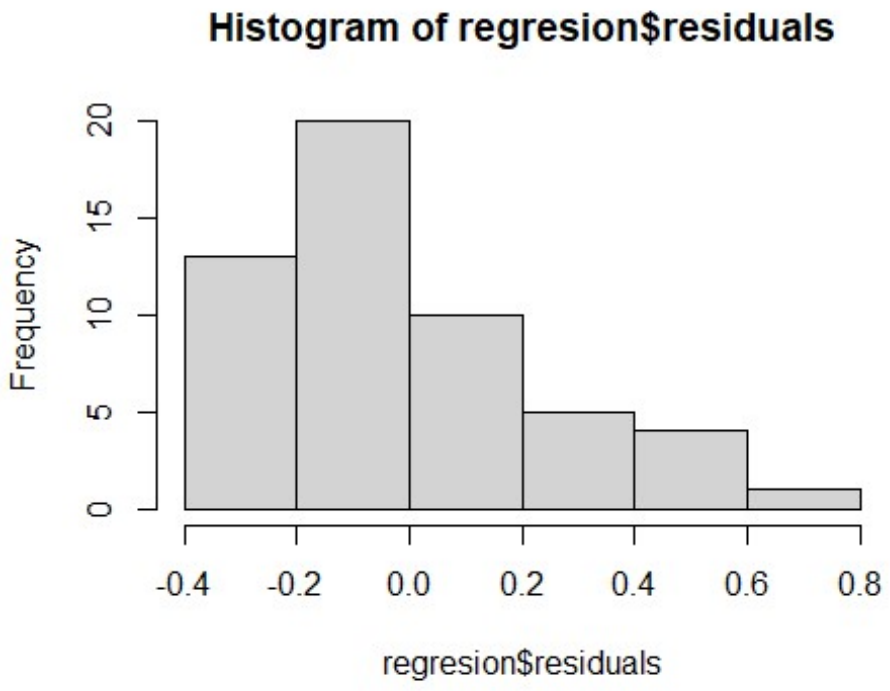
```
##
## Coefficients:
## (Intercept)        M1$X5
##      0.4673      -0.3807

CALCIO=summary(regresion)
CALCIO

##
## Call:
## lm(formula = M1$X7 ~ M1$X5)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3261 -0.1961 -0.0642  0.1445  0.7130
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.46730    0.04419  10.575 1.85e-14 ***
## M1$X5       -0.38067    0.12189  -3.123  0.00295 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2446 on 51 degrees of freedom
## Multiple R-squared:  0.1605, Adjusted R-squared:  0.1441
## F-statistic: 9.754 on 1 and 51 DF,  p-value: 0.002948

plot(M1$X5,M1$X7,col="blue",xlab="CALCIO",ylab="Concentracion de
Mercurio",main="CALCIO vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```



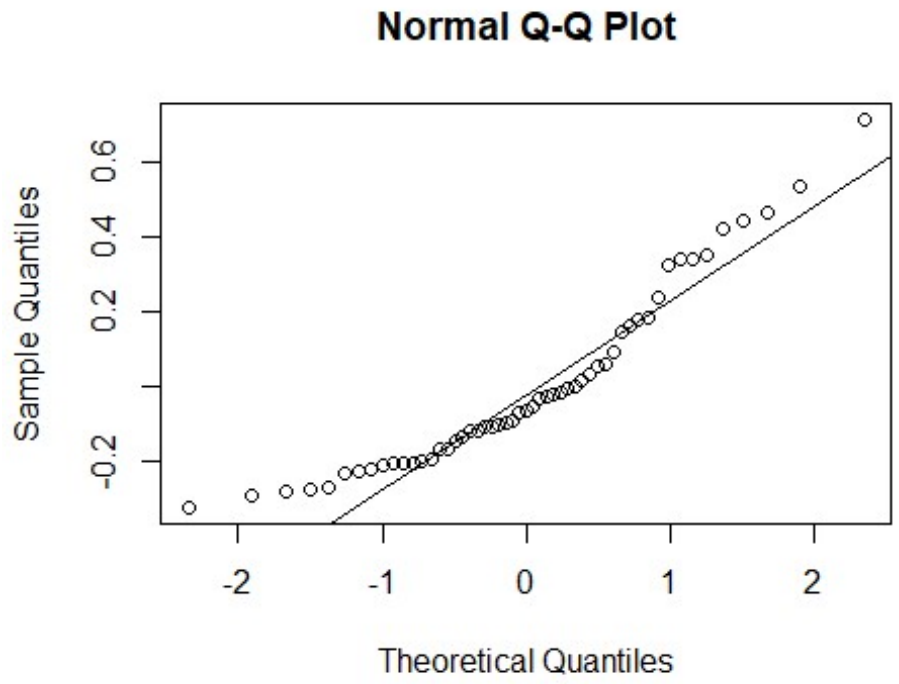

```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.90943, p-value = 0.0006884

hist(regresion$residuals)
```




```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```




```
t.test(regresion$residuals)

##
##  One Sample t-test
##
## data:  regresion$residuals
## t = 9.756e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.06676394  0.06676394
## sample estimates:
##    mean of x
## 3.245958e-17

cor.test(M1$X7,CALCIOBC)

##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and CALCIOBC
## t = -3.7845, df = 51, p-value = 0.0004064
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.6555727 -0.2266433
## sample estimates:
##       cor
## -0.4682519
```

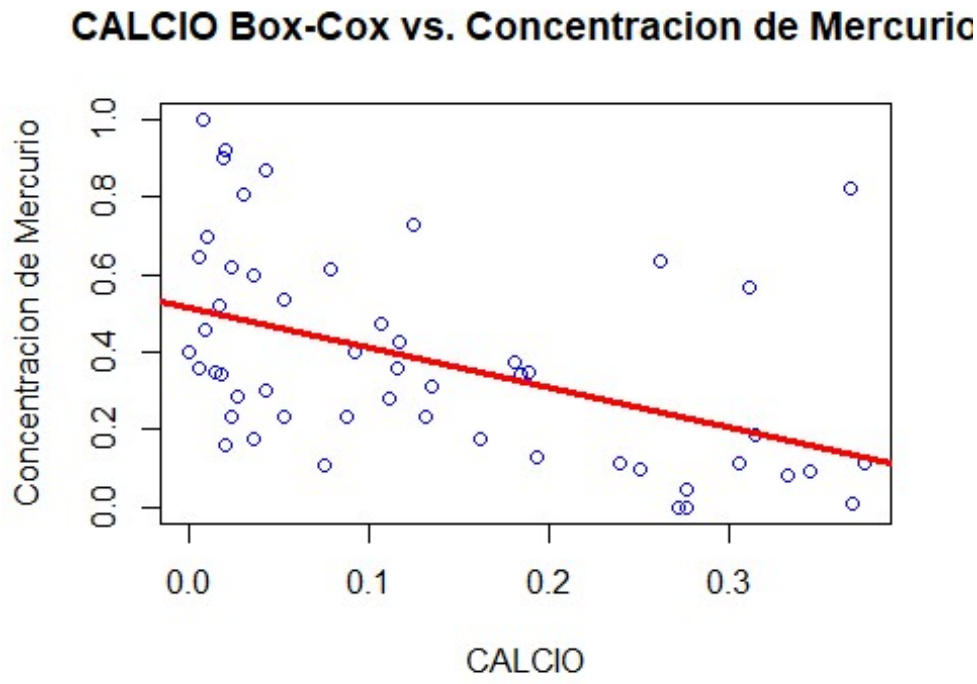
```r
regresion=lm(M1$X7 ~ CALCIOBC)
regresion
```

```
##
## Call:
## lm(formula = M1$X7 ~ CALCIOBC)
##
## Coefficients:
## (Intercept)     CALCIOBC
##      0.5159      -1.0357
```

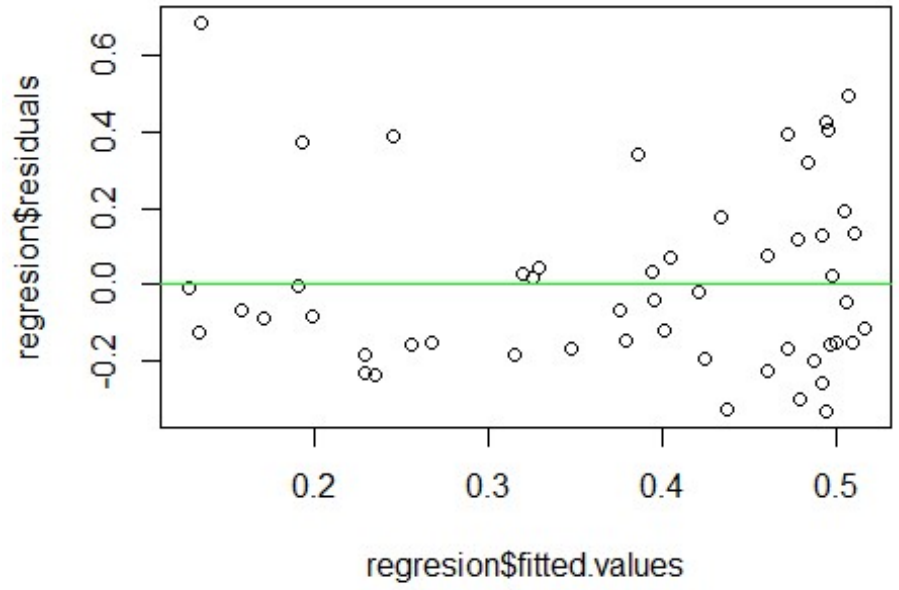
```r
CALCIO=summary(regresion)
CALCIO
```

```
##
## Call:
## lm(formula = M1$X7 ~ CALCIOBC)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3318 -0.1692 -0.0652  0.1192  0.6864
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.51589    0.04883  10.566 1.91e-14 ***
## CALCIOBC    -1.03571    0.27367  -3.785 0.000406 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2359 on 51 degrees of freedom
## Multiple R-squared:  0.2193, Adjusted R-squared:  0.204
## F-statistic: 14.32 on 1 and 51 DF,  p-value: 0.0004064
```

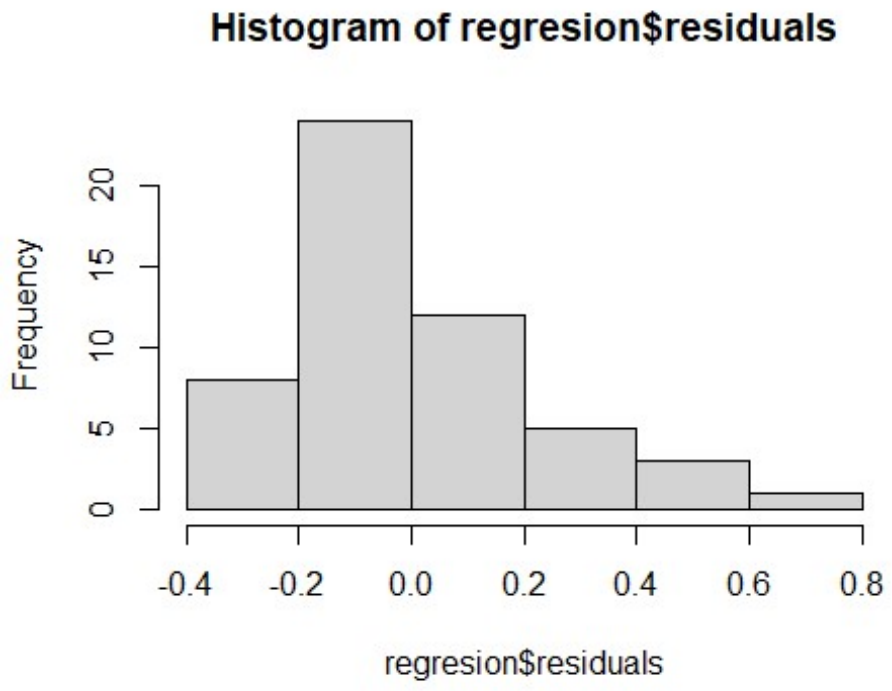
```r
plot(CALCIOBC,M1$X7,col="blue",xlab="CALCIO",ylab="Concentracion de
Mercurio",main="CALCIO Box-Cox vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```



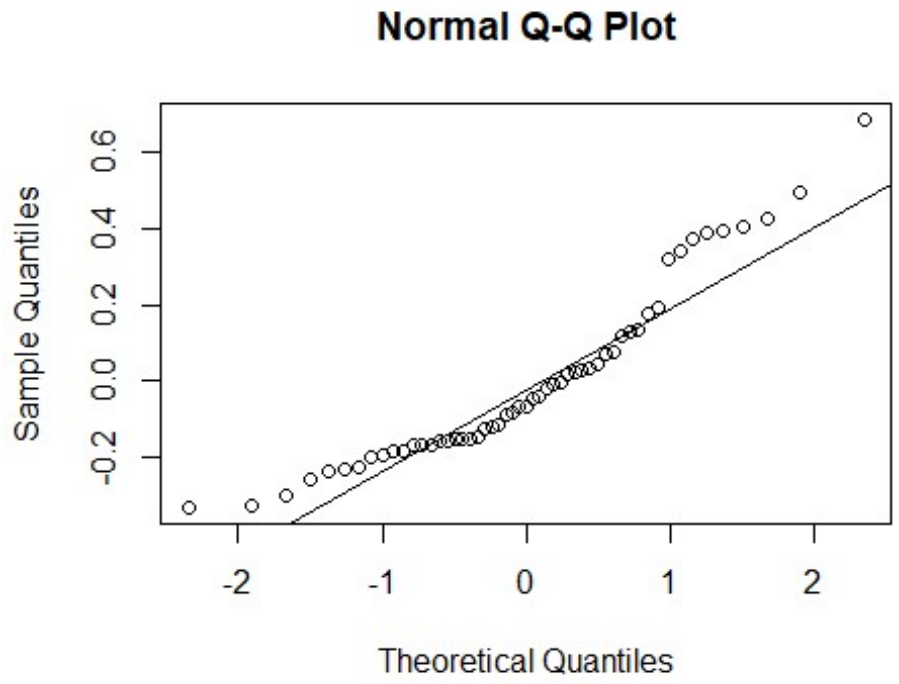

```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.91093, p-value = 0.0007785

hist(regresion$residuals)
```




```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```




```
t.test(regresion$residuals)
```

```
##
##  One Sample t-test
##
## data:  regresion$residuals
## t = 4.7755e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.06438672  0.06438672
## sample estimates:
##    mean of x
## 1.532305e-17
```

```
z1=CALCIOBC
z2=CALCIOBC^2
cor.test(M1$X7 , z1+ z2)
```

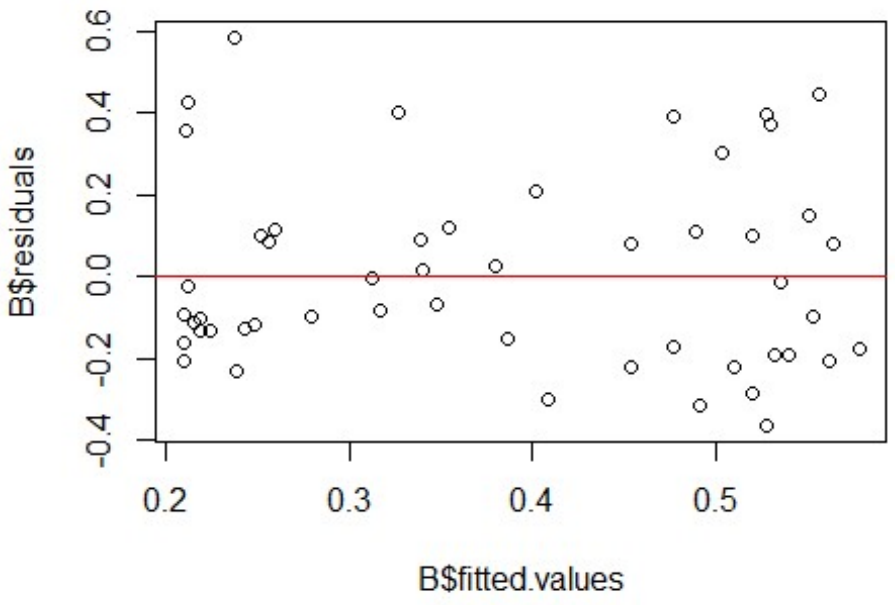
```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and z1 + z2
## t = -3.6364, df = 51, p-value = 0.0006438
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.6449477 -0.2091075
## sample estimates:
```

```
##        cor
## -0.4537549
```

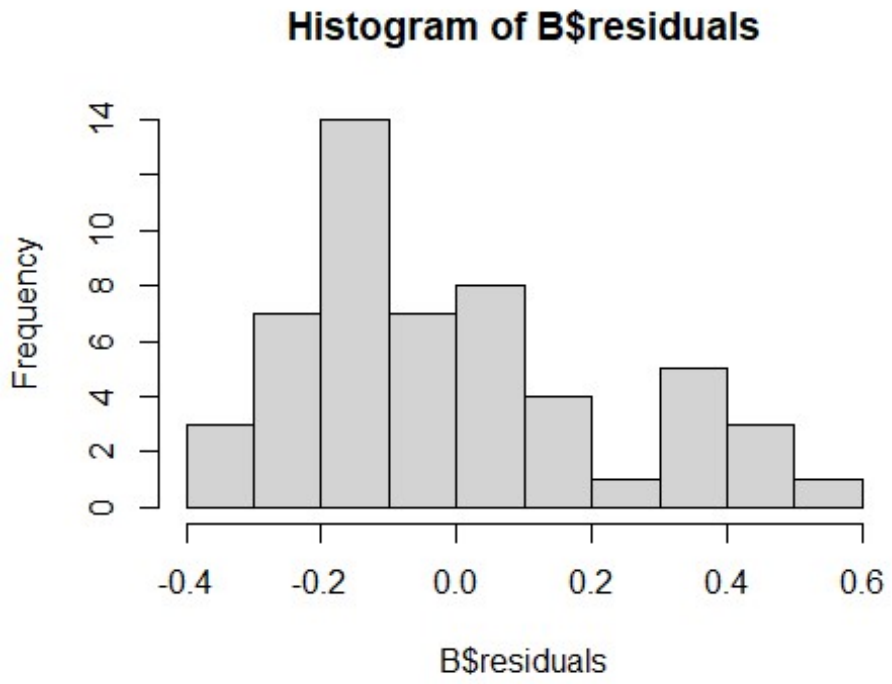
```
B = lm(M1$X7 ~ z1+ z2)
summary(B)
```

```
##
## Call:
## lm(formula = M1$X7 ~ z1 + z2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36502 -0.17448 -0.08471  0.10698  0.58439
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.57887    0.06275   9.224 2.26e-12 ***
## z1          -2.57653    1.02104  -2.523   0.0148 *
## z2           4.48246    2.86473   1.565   0.1240
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2326 on 50 degrees of freedom
## Multiple R-squared:  0.2557, Adjusted R-squared:  0.2259
## F-statistic: 8.589 on 2 and 50 DF,  p-value: 0.0006218
```

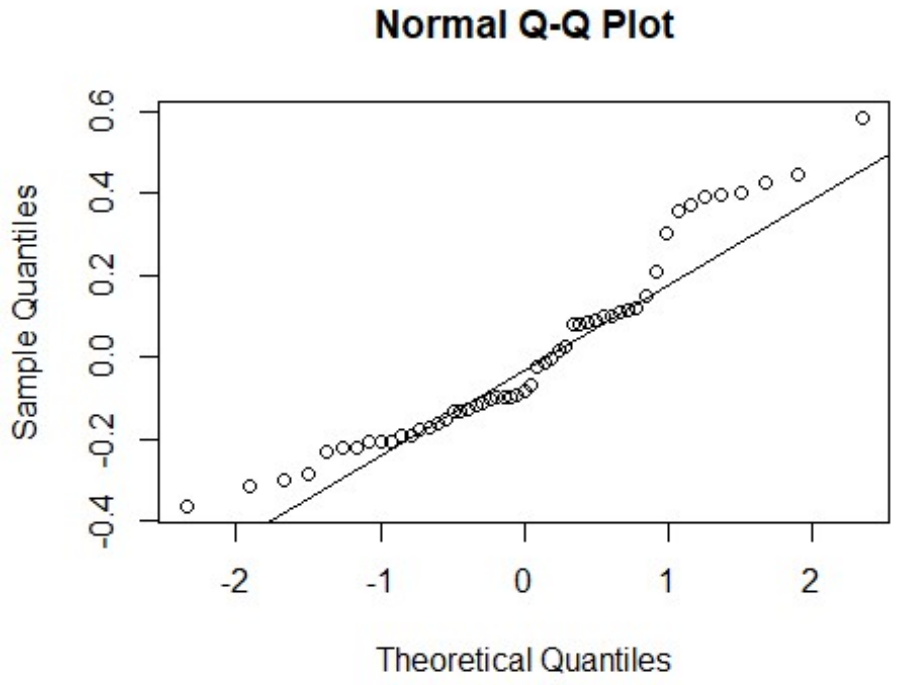
```
plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```



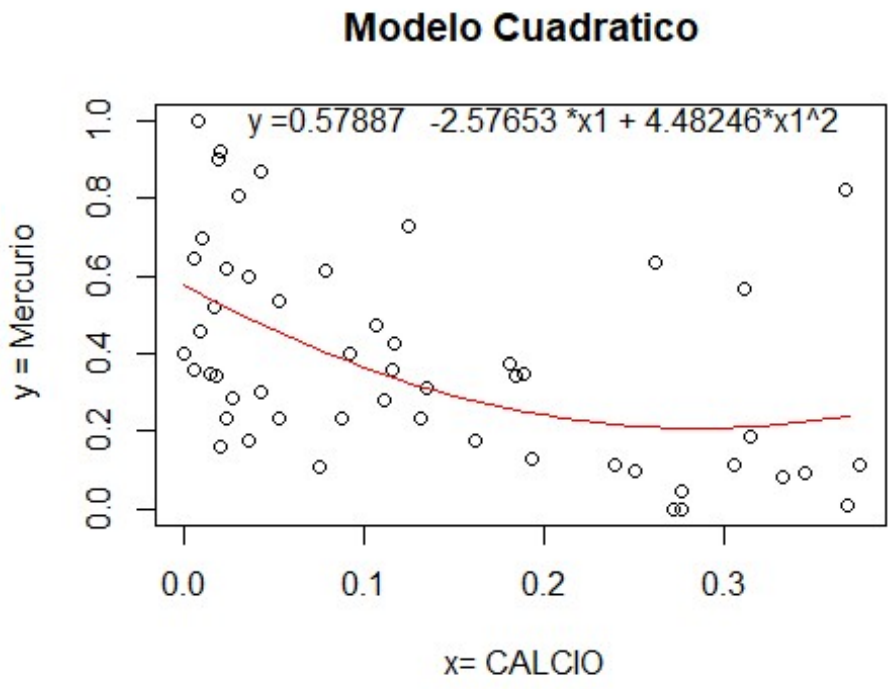

```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.92778, p-value = 0.003314

hist(B$residuals)
```




```
qqnorm(B$residuals)
qqline(B$residuals)
```



```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = 5.2817e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.06286597  0.06286597
## sample estimates:
##    mean of x
## 1.654687e-17

x=CALCIOBC
y=M1$X7
plot(x, y, main = "Modelo Cuadratico", xlab =" x= CALCIO", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = 0.57887   -2.57653 *x1 + 4.48246*x1^2
lines(x1, y1, col = "red")
text(0.2, 1, "y =0.57887   -2.57653 *x1 + 4.48246*x1^2")
```



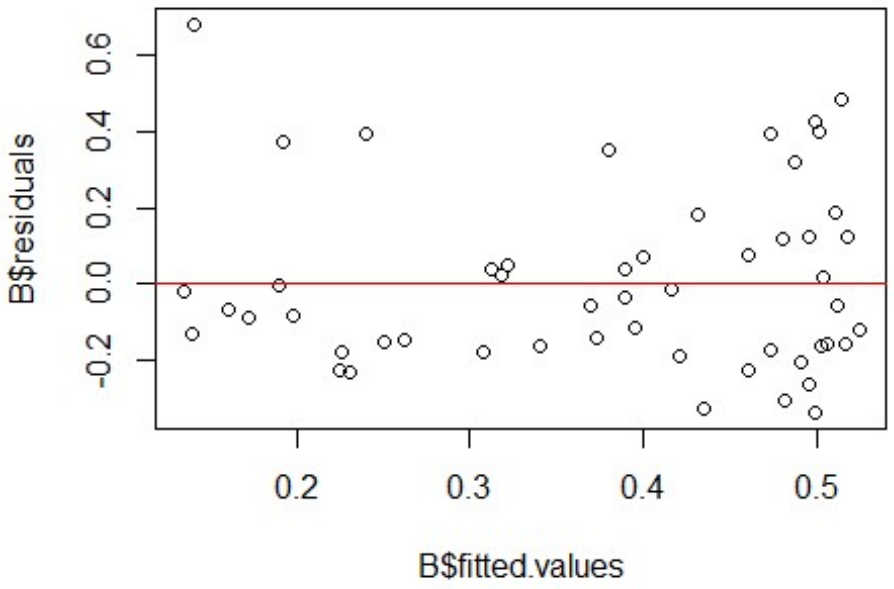

```
z1=CALCIOBC+1
z2=log(CALCIOBC+1)
cor.test(M1$X7 , log(CALCIOBC+1))
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and log(CALCIOBC + 1)
## t = -3.8589, df = 51, p-value = 0.0003216
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.6607747 -0.2353314
## sample estimates:
##       cor
## -0.4753865
```

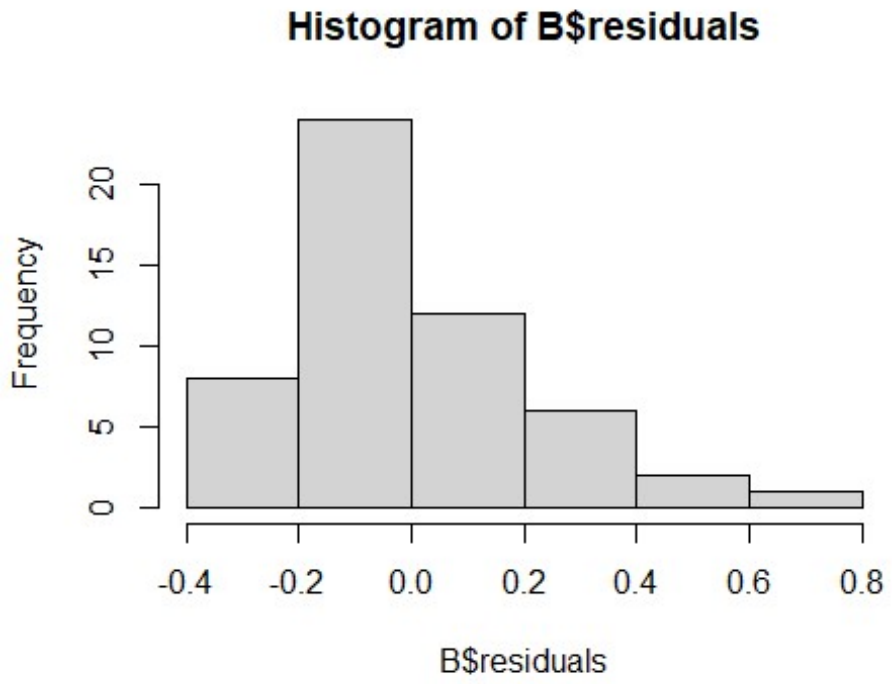
```
B = lm(M1$X7 ~ log(CALCIOBC+1))
summary(B)
```

```
##
## Call:
## lm(formula = M1$X7 ~ log(CALCIOBC + 1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.33687 -0.16234 -0.05925  0.11668  0.68017
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.5246     0.0499  10.512 2.29e-14 ***
## log(CALCIOBC + 1)  -1.2240     0.3172  -3.859 0.000322 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2349 on 51 degrees of freedom
## Multiple R-squared:  0.226,  Adjusted R-squared:  0.2108
## F-statistic: 14.89 on 1 and 51 DF,  p-value: 0.0003216
```

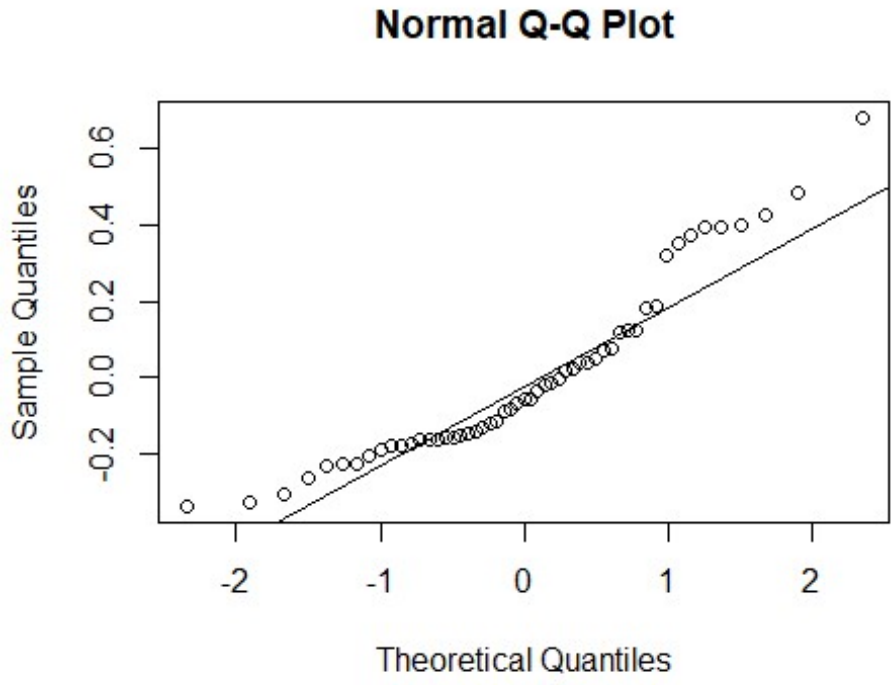
```
plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```




```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.91214, p-value = 0.0008609

hist(B$residuals)
```

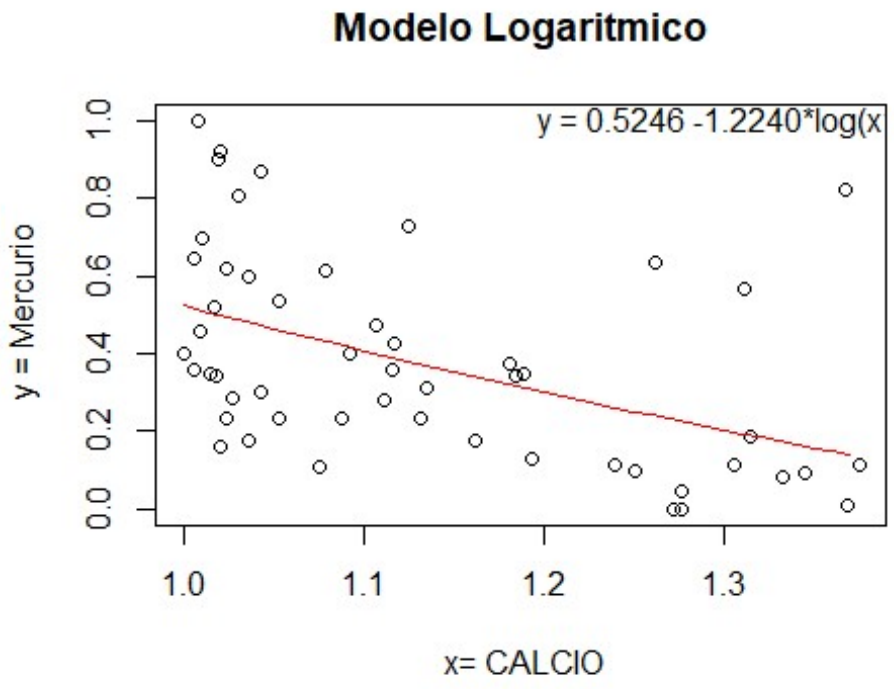



```
qqnorm(B$residuals)
qqline(B$residuals)
```



```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = 2.8711e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.06410851  0.06410851
## sample estimates:
##   mean of x
## 9.17276e-18
```

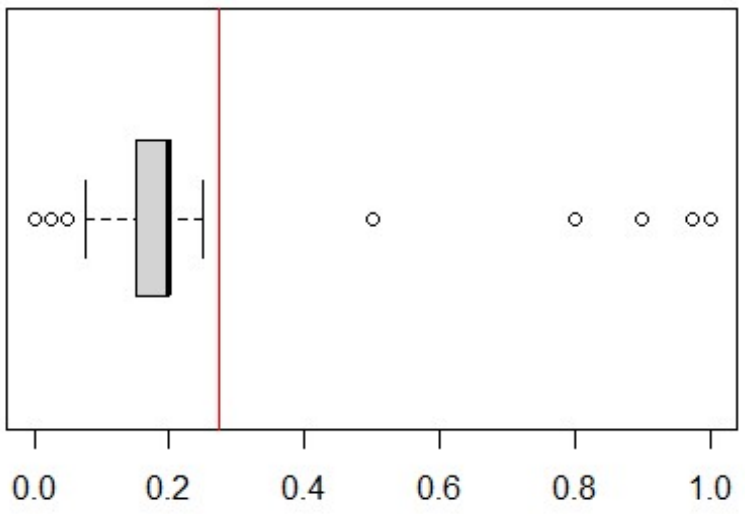
```
x=CALCIOBC+1
y=M1$X7
plot(x, y, main = "Modelo Logaritmico", xlab =" x= CALCIO", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = 0.5246 -1.2240 *log(x1)
lines(x1, y1, col = "red")
text(1.3, 1, "y = 0.5246 -1.2240*log(x1)")
```



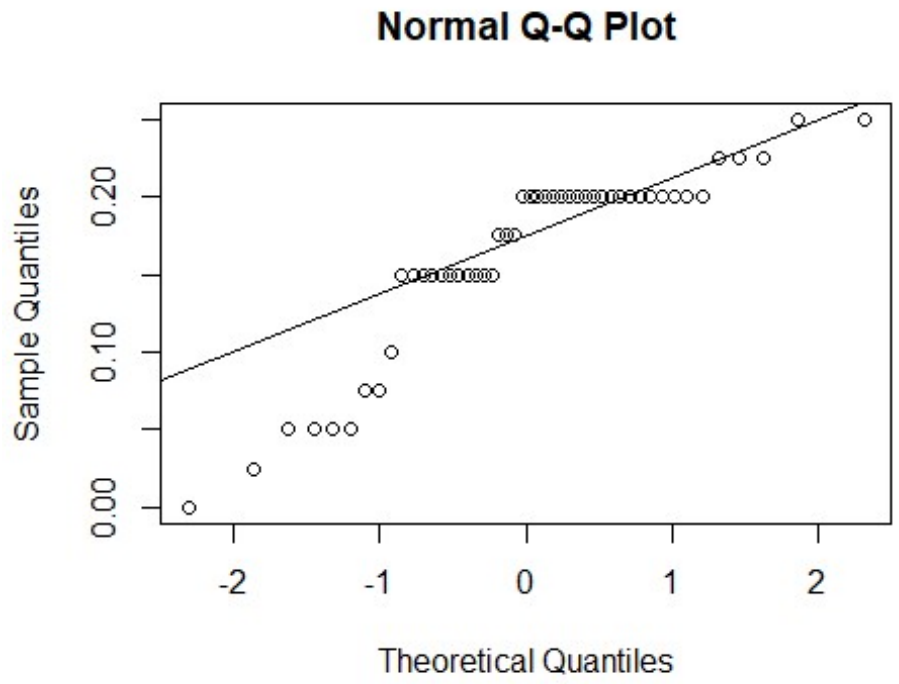
Clorofila

```
q1=quantile(M1$X8,0.25)  #Cuantil 1 de la variable X
q2=quantile(M1$X8,0.5)
```

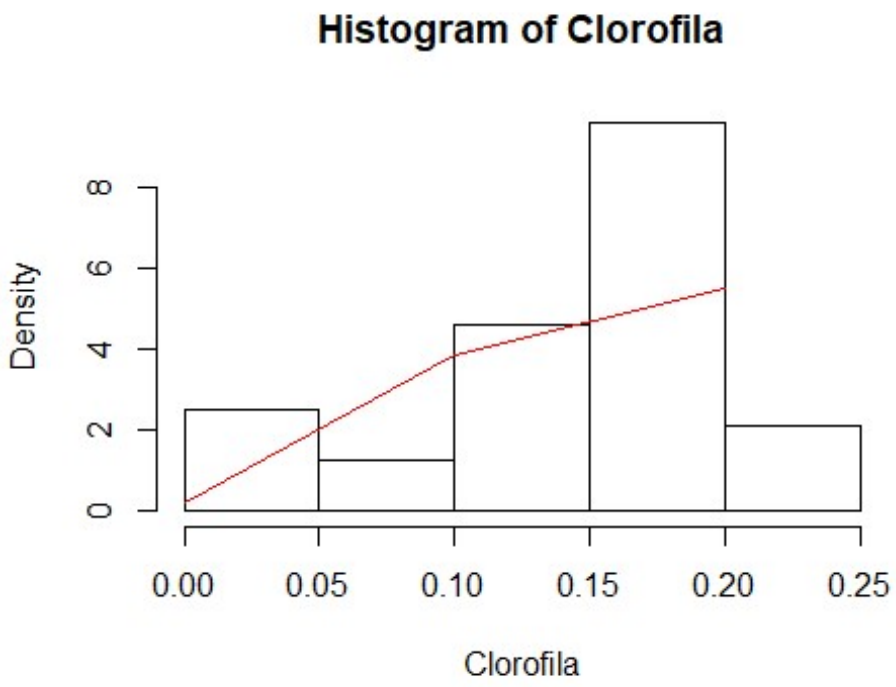
```
q3=quantile(M1$X8,0.75)
y1=min(M1$X8)
y2=max(M1$X8)
ri=IQR(M1$X8)
boxplot(M1$X8,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```



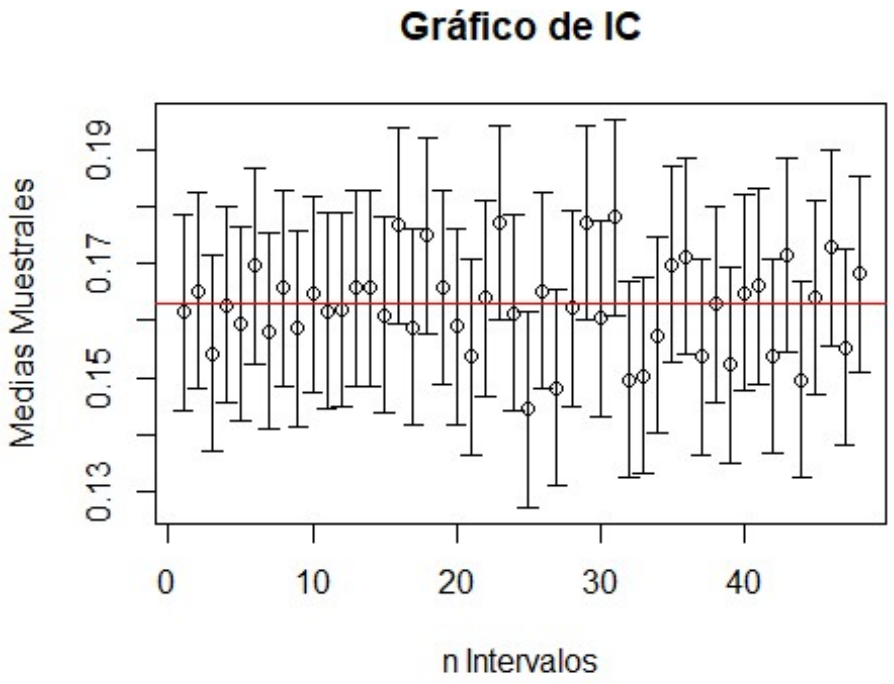

```
Clorofila= M1[M1$X8<q3+1.5*ri,c("X8")]
qqnorm(Clorofila)
qqline(Clorofila)
```




```
hist(Clorofila,prob=TRUE,col=0)
x=seq(min(Clorofila),max(Clorofila),0.1)
y=dnorm(x,mean(Clorofila),sd(Clorofila))
lines(x,y,col="red")
```



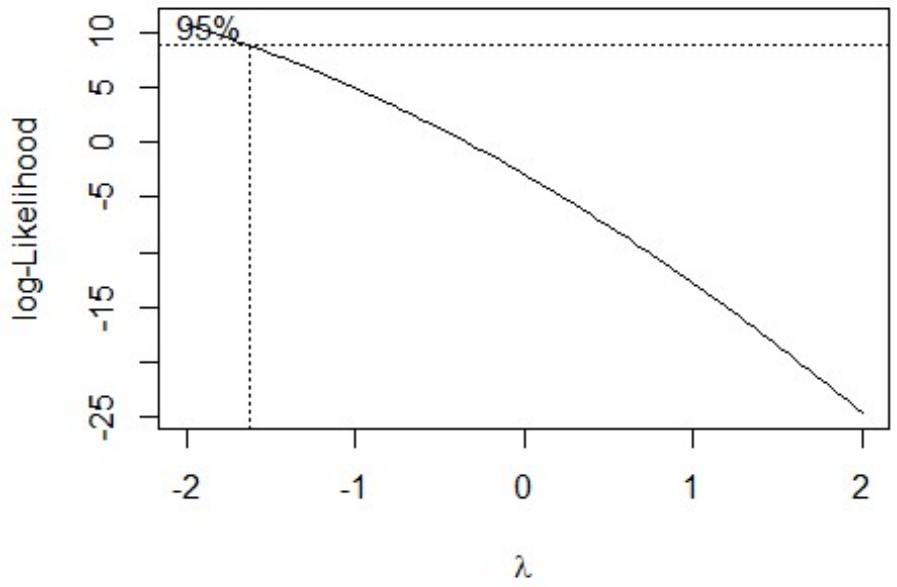

```
library(moments)
skewness1=skewness(Clorofila)
cat("Sesgo: ",skewness1)

## Sesgo:  -1.095579

cat("\n")

kurtosis1=kurtosis(Clorofila)
cat("Curtosis: ",kurtosis1)

## Curtosis:  3.31411

cat("\n")

library(plotrix)
n=length(Clorofila)
media=mean(Clorofila)
DE=sd(Clorofila)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```



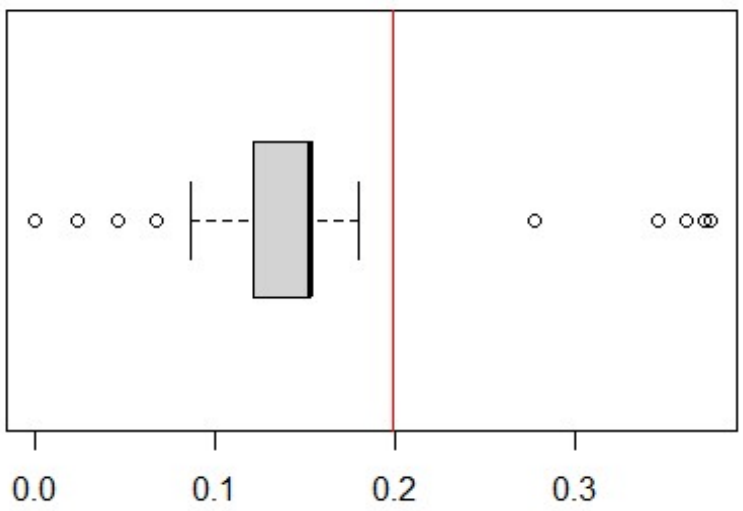

```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.1630208  +-  0.01713786
```

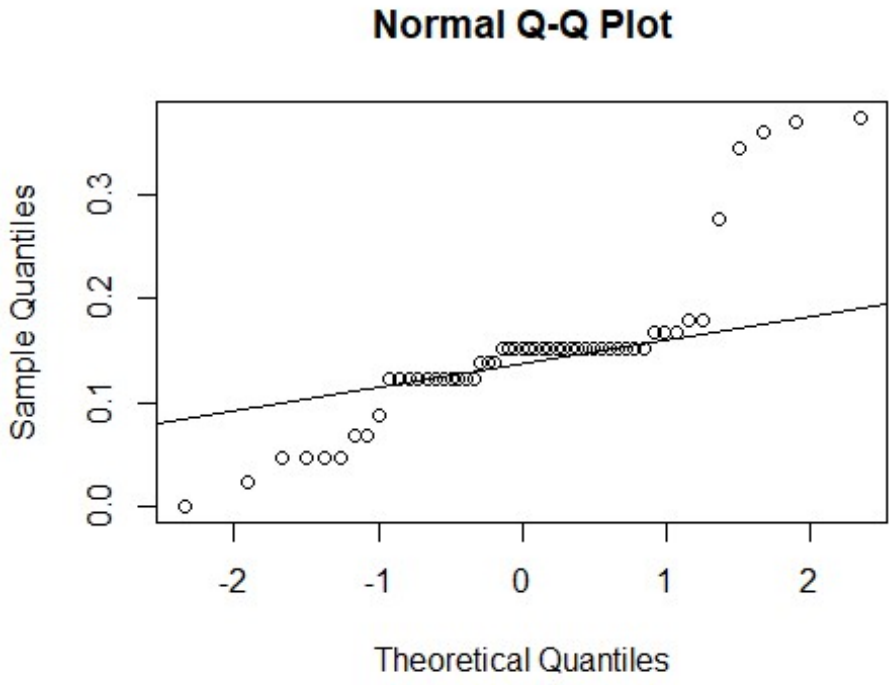
```
library(MASS)
minim=min(M1$X8)
b <- boxcox(lm((M1$X8+1) ~ 1))
```



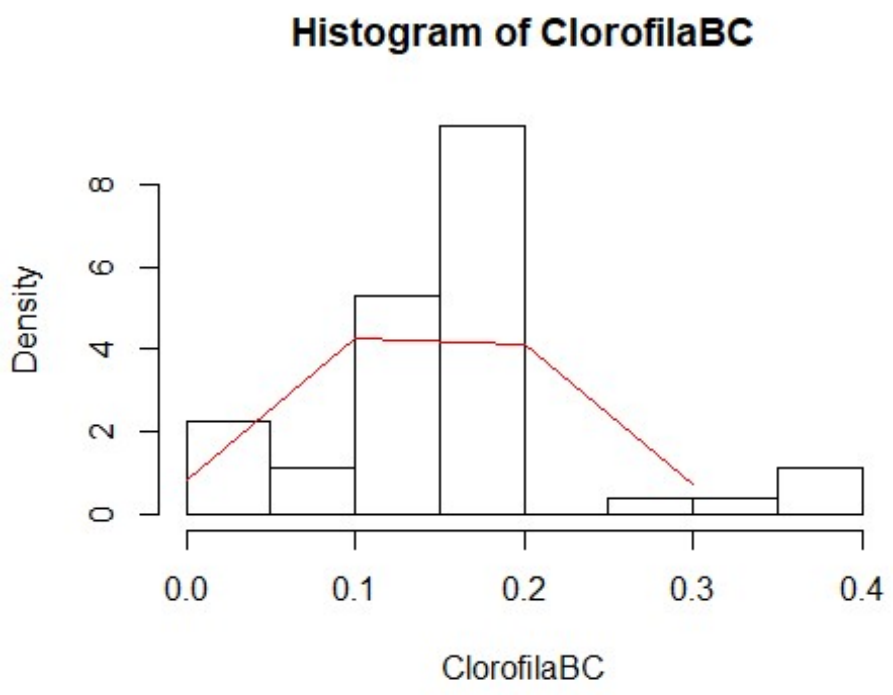

```
lambda <- b$x[which.max(b$y)]
X8<-((M1$X8+1)^lambda-1)/lambda
q1=quantile(X8,0.25)  #Cuantil 1 de la variable X
q2=quantile(X8,0.5)
q3=quantile(X8,0.75)
y1=min(X8)
y2=max(X8)
ri=IQR(X8)
boxplot(X8,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```




```
ClorofilaBC= X8
qqnorm(ClorofilaBC)
qqline(ClorofilaBC)
```

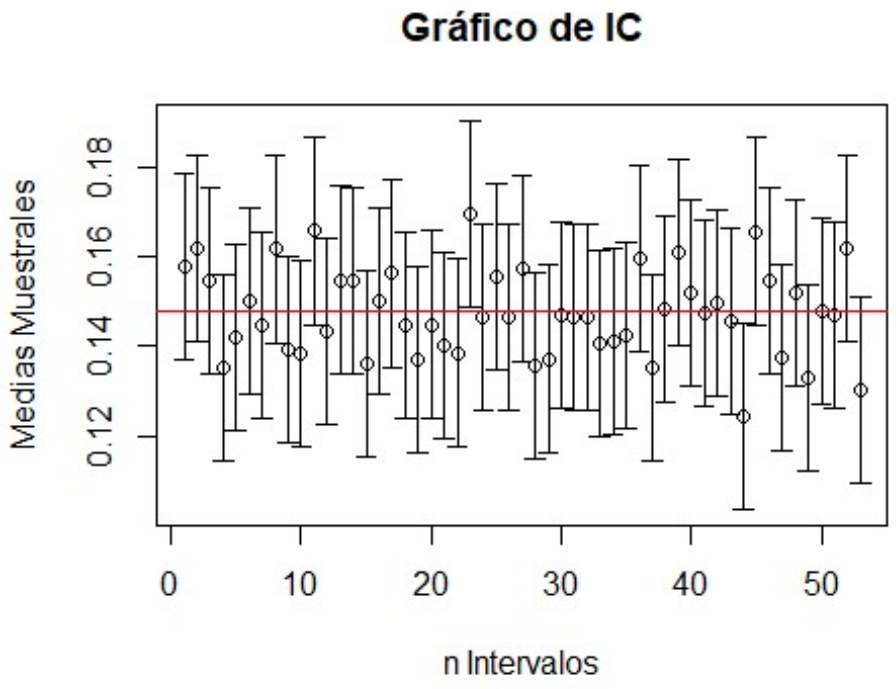


```
hist(ClorofilaBC,prob=TRUE,col=0)
x=seq(min(ClorofilaBC),max(ClorofilaBC),0.1)
y=dnorm(x,mean(ClorofilaBC),sd(ClorofilaBC))
lines(x,y,col="red")
```




```
library(moments)
skewness1=skewness(ClorofilaBC)
cat("Sesgo: ",skewness1)

## Sesgo:  1.351944

cat("\n")

kurtosis1=kurtosis(ClorofilaBC)
cat("Curtosis: ",kurtosis1)

## Curtosis:  5.70415

cat("\n")

library(plotrix)
n=length(ClorofilaBC)
media=mean(ClorofilaBC)
DE=sd(ClorofilaBC)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
```

```
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```




```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.1478106  +-  0.02077724
```

```
cor.test(M1$X7,M1$X8)
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and M1$X8
## t = 0.56619, df = 51, p-value = 0.5737
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.1954345  0.3420216
## sample estimates:
##        cor
## 0.07903426
```

```
regresion=lm(M1$X7 ~ M1$X8)
regresion
```

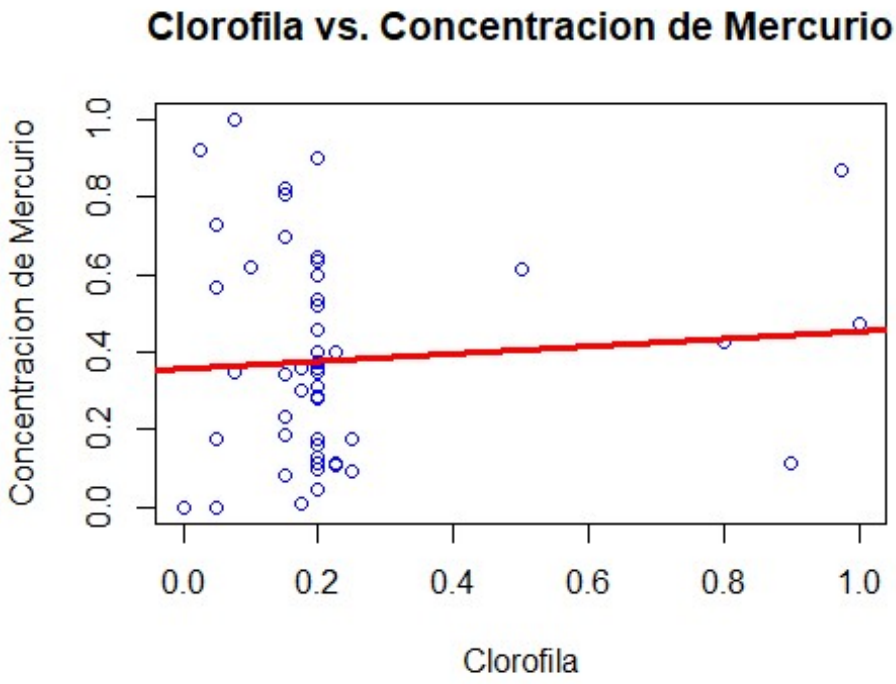
```
##
## Call:
## lm(formula = M1$X7 ~ M1$X8)
```

```
##
## Coefficients:
## (Intercept)        M1$X8
##     0.35555      0.09763
```

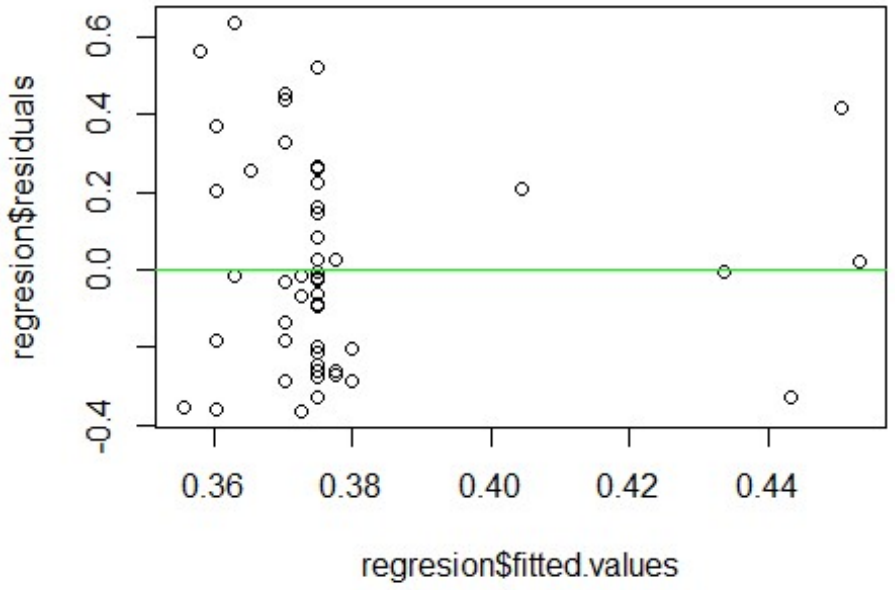
```
Clorofila=summary(regresion)
Clorofila
```

```
##
## Call:
## lm(formula = M1$X7 ~ M1$X8)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36488 -0.20166 -0.02911  0.20546  0.63713
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.35555    0.05348   6.648 1.96e-08 ***
## M1$X8        0.09763    0.17243   0.566    0.574
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2661 on 51 degrees of freedom
## Multiple R-squared:  0.006246,   Adjusted R-squared:  -0.01324
## F-statistic: 0.3206 on 1 and 51 DF,  p-value: 0.5737
```

```
plot(M1$X8,M1$X7,col="blue",xlab="Clorofila",ylab="Concentracion de
Mercurio",main="Clorofila vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```

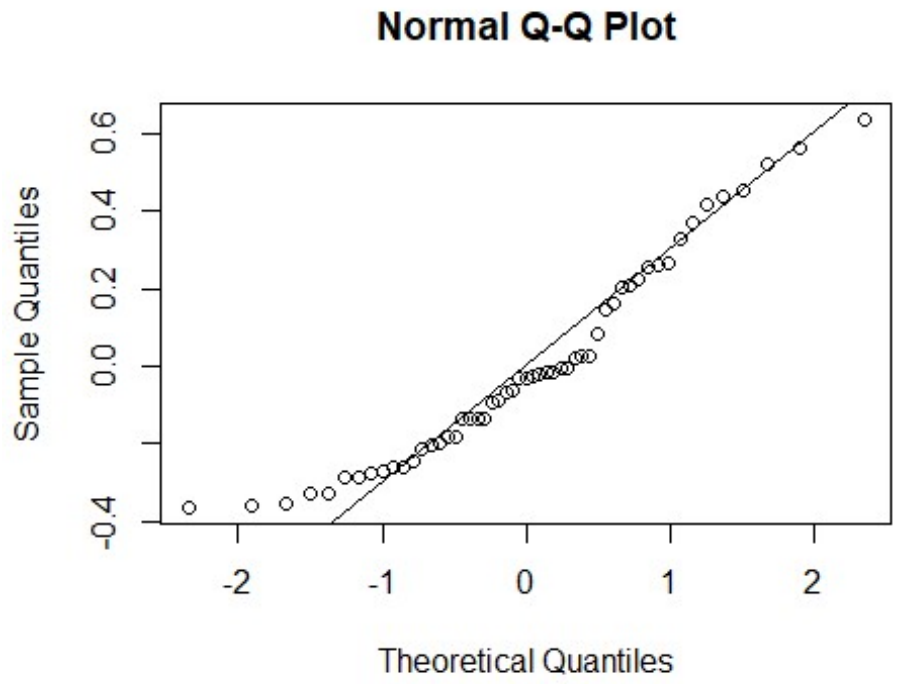



```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.93925, p-value = 0.009517

hist(regresion$residuals)
```


```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```




```
t.test(regresion$residuals)

##
##  One Sample t-test
##
## data:  regresion$residuals
## t = -5.8417e-17, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.0726411  0.0726411
## sample estimates:
##     mean of x
## -2.114706e-18

cor.test(M1$X7,ClorofilaBC)

##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and ClorofilaBC
## t = 0.15268, df = 51, p-value = 0.8793
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.2503659  0.2899928
## sample estimates:
##        cor
## 0.02137436
```

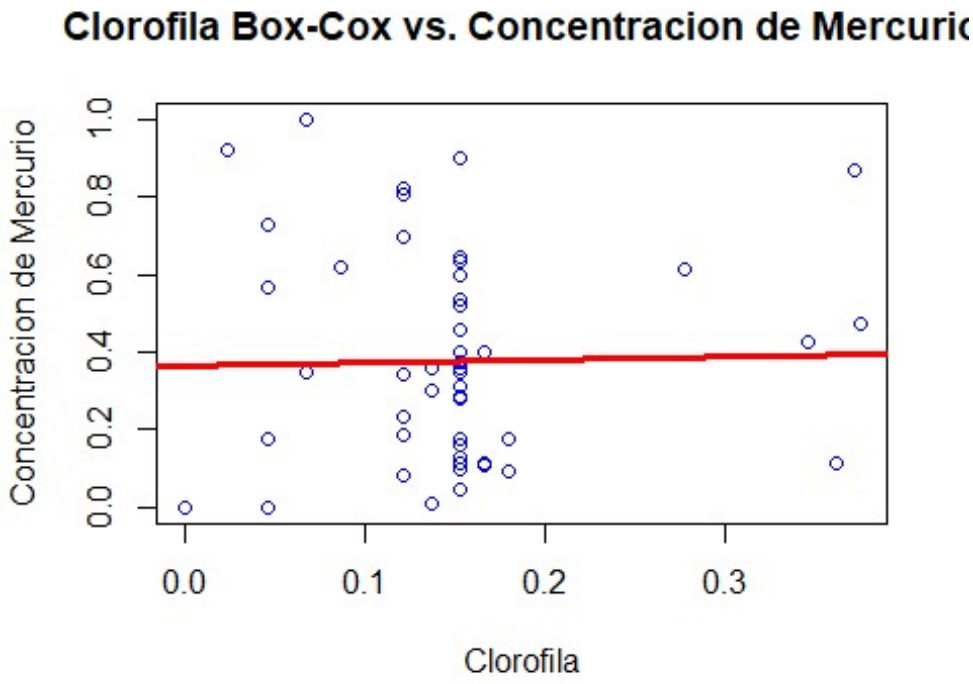
```r
regresion=lm(M1$X7 ~ ClorofilaBC)
regresion
```

```
##
## Call:
## lm(formula = M1$X7 ~ ClorofilaBC)
##
## Coefficients:
## (Intercept)  ClorofilaBC
##     0.36683      0.07322
```

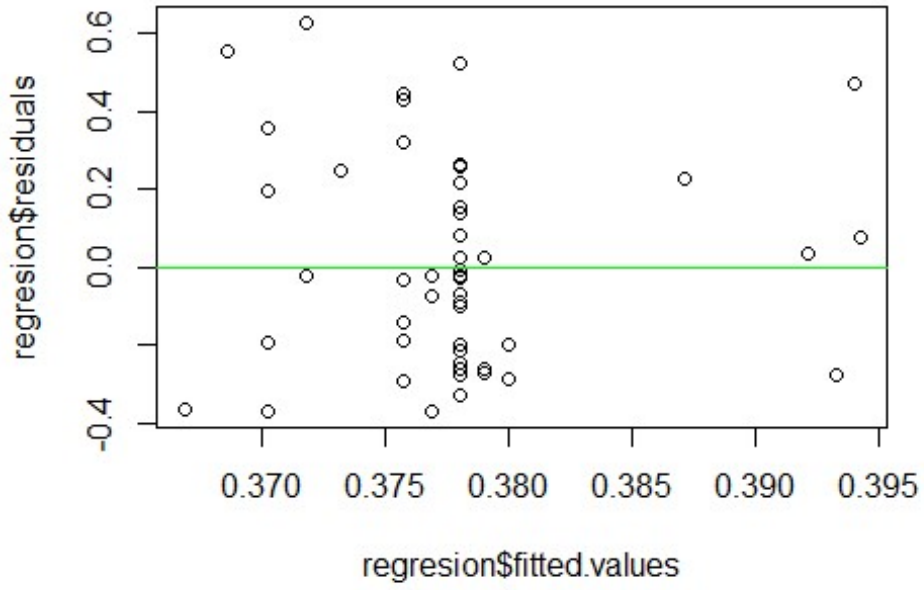
```r
Clorofila=summary(regresion)
Clorofila
```

```
##
## Call:
## lm(formula = M1$X7 ~ ClorofilaBC)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.37023 -0.20171 -0.03467  0.19566  0.62824
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.36683    0.07980   4.597 2.86e-05 ***
## ClorofilaBC  0.07322    0.47957   0.153    0.879
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2669 on 51 degrees of freedom
## Multiple R-squared:  0.0004569,  Adjusted R-squared:  -0.01914
## F-statistic: 0.02331 on 1 and 51 DF,  p-value: 0.8793
```

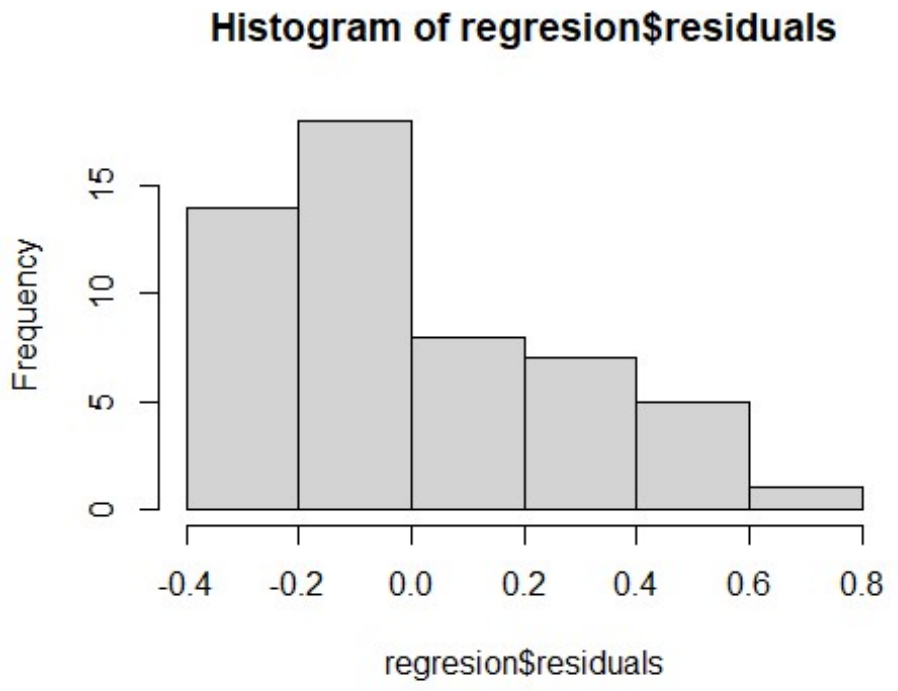
```r
plot(ClorofilaBC,M1$X7,col="blue",xlab="Clorofila",ylab="Concentracion de
Mercurio",main="Clorofila Box-Cox vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```



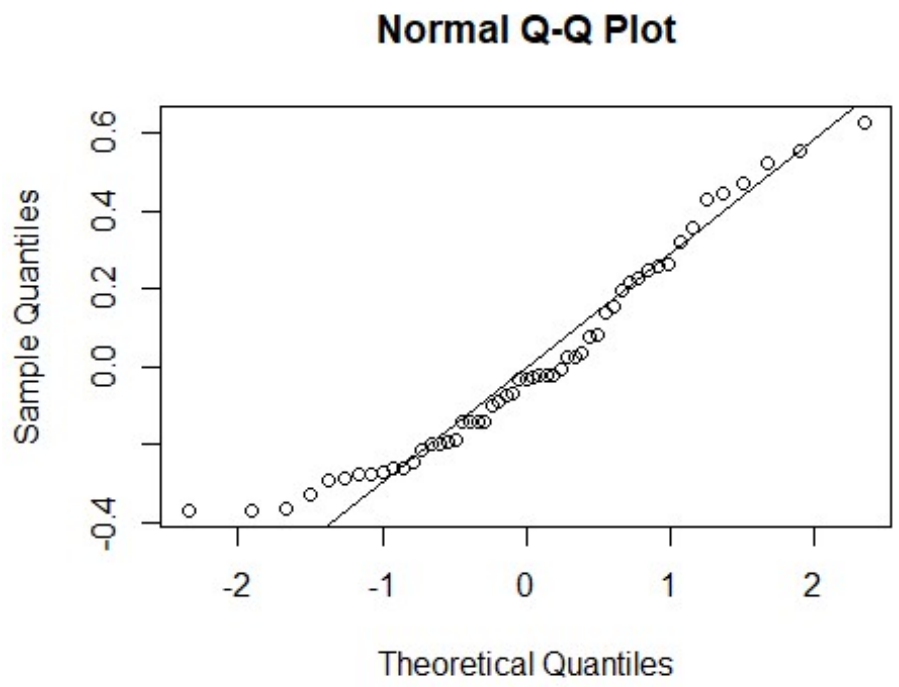

```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.94026, p-value = 0.01047

hist(regresion$residuals)
```




```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```




```
t.test(regresion$residuals)
```

```
##
##  One Sample t-test
##
## data:  regresion$residuals
## t = 1.3671e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.07285239  0.07285239
## sample estimates:
##    mean of x
## 4.963293e-18
```

```
z1=ClorofilaBC
z2=ClorofilaBC^2
cor.test(M1$X7 , z1+ z2)
```

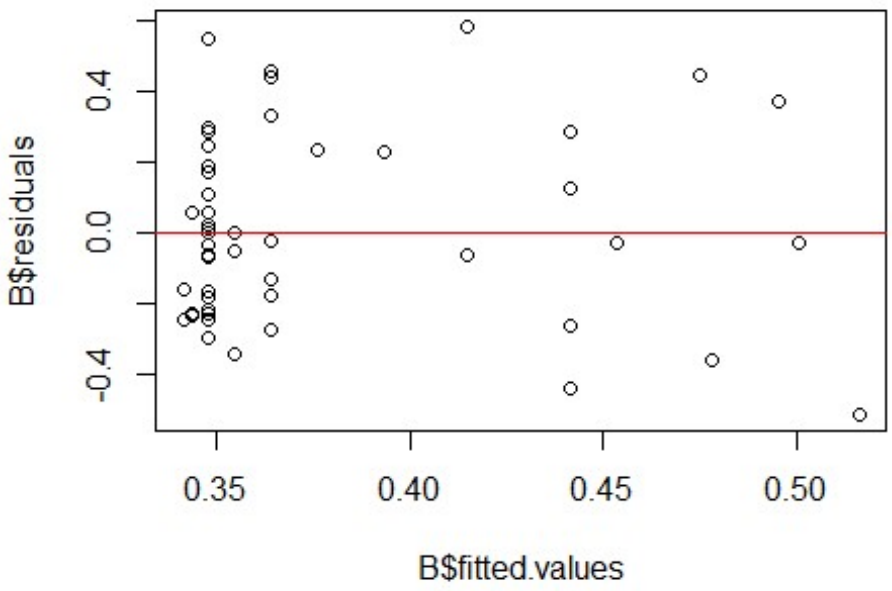
```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and z1 + z2
## t = 0.27137, df = 51, p-value = 0.7872
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.2347315  0.3051335
## sample estimates:
```

```
##        cor
## 0.03797145
```

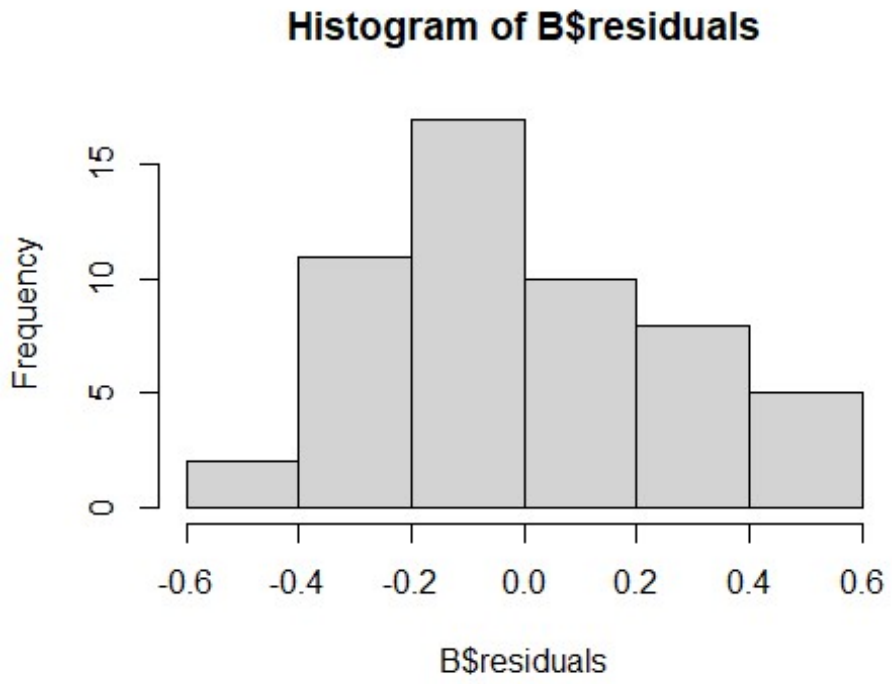
```
B = lm(M1$X7 ~ z1+ z2)
summary(B)
```

```
##
## Call:
## lm(formula = M1$X7 ~ z1 + z2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.51630 -0.18497 -0.02794  0.18713  0.58548
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.5163     0.1396   3.698  0.00054 ***
## z1           -1.8332     1.5413  -1.189  0.23992
## z2            4.7783     3.6742   1.301  0.19938
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2651 on 50 degrees of freedom
## Multiple R-squared:  0.03316,    Adjusted R-squared:  -0.005511
## F-statistic: 0.8575 on 2 and 50 DF,  p-value: 0.4304
```

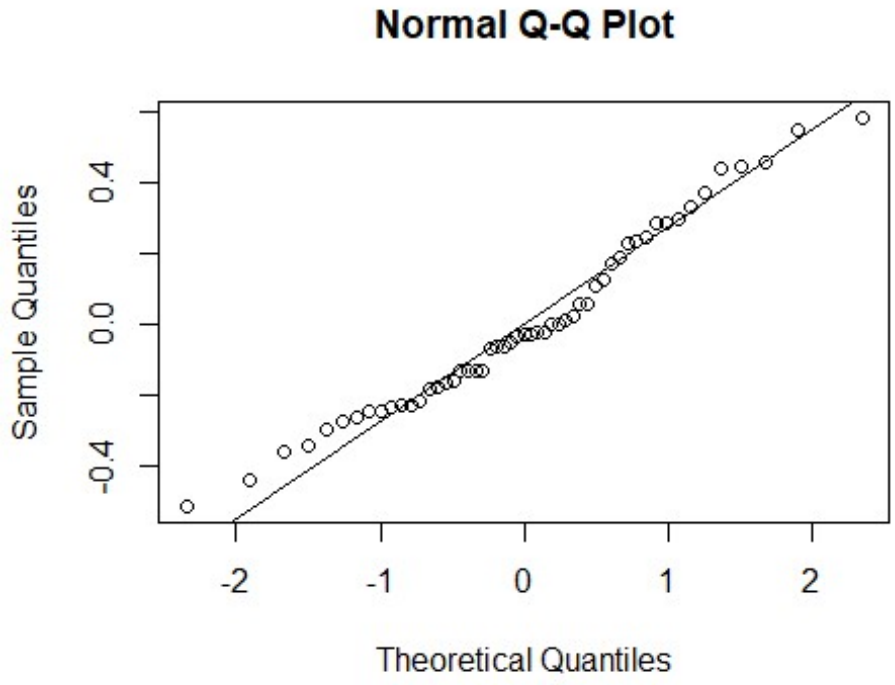
```
plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```



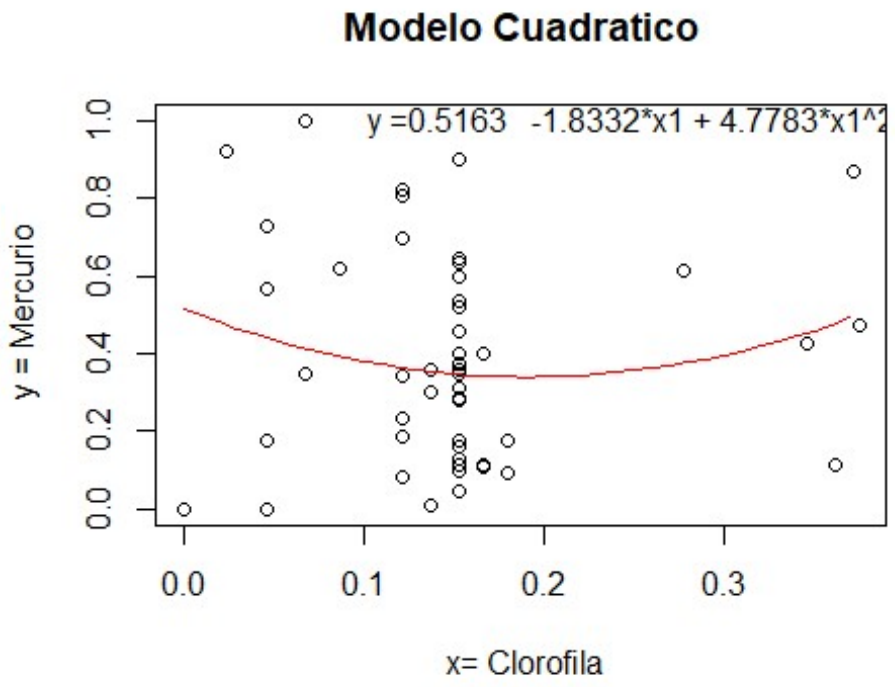

```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.97045, p-value = 0.2111

hist(B$residuals)
```




```
qqnorm(B$residuals)
qqline(B$residuals)
```



```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = 4.4509e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.07165059  0.07165059
## sample estimates:
##    mean of x
## 1.589277e-17

x=ClorofilaBC
y=M1$X7
plot(x, y, main = "Modelo Cuadratico", xlab =" x= Clorofila", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = 0.5163   -1.8332*x1 + 4.7783*x1^2
lines(x1, y1, col = "red")
text(0.25, 1, "y =0.5163   -1.8332*x1 + 4.7783*x1^2")
```



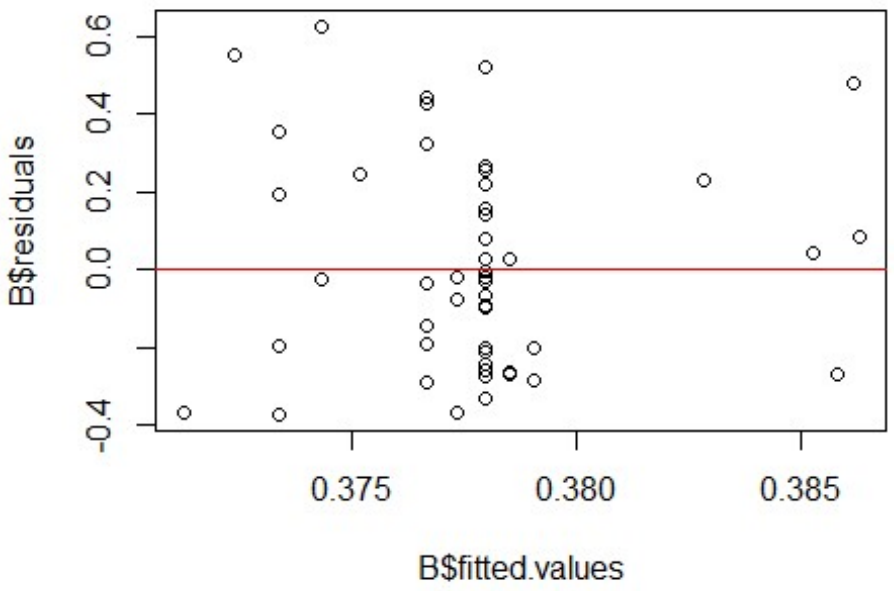

```
z1=ClorofilaBC+1
z2=log(ClorofilaBC+1)
cor.test(M1$X7,log(ClorofilaBC+1))
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and log(ClorofilaBC + 1)
## t = 0.083013, df = 51, p-value = 0.9342
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.2594857  0.2810342
## sample estimates:
##       cor
## 0.01162333
```

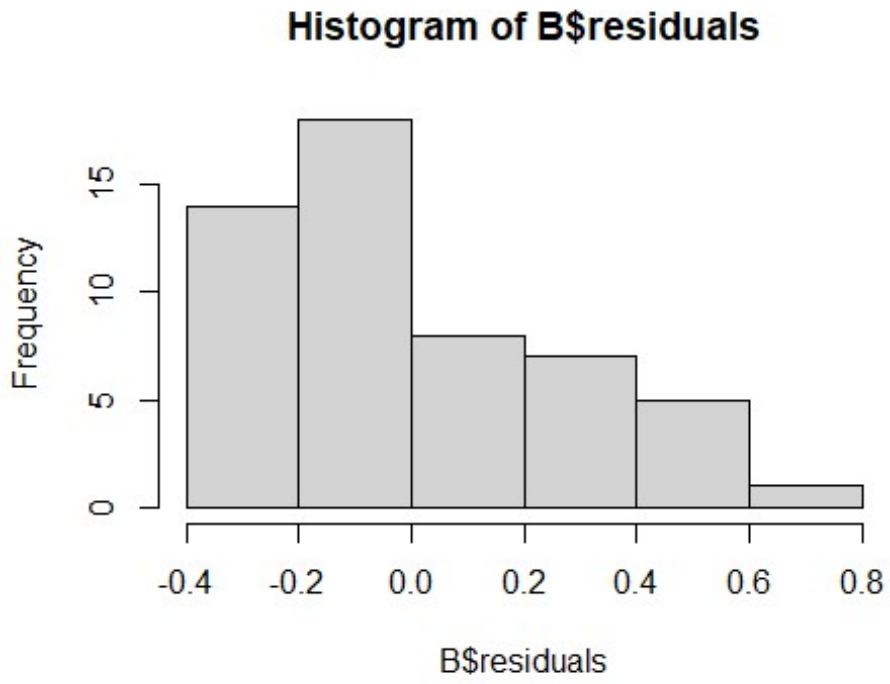
```
B = lm(M1$X7 ~ log(ClorofilaBC+1))
summary(B)
```

```
##
## Call:
## lm(formula = M1$X7 ~ log(ClorofilaBC + 1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.37337 -0.20077 -0.03558  0.19252  0.62569
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.37122    0.08575   4.329 7.01e-05 ***
## log(ClorofilaBC + 1)  0.04740    0.57104   0.083    0.934
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2669 on 51 degrees of freedom
## Multiple R-squared:  0.0001351,  Adjusted R-squared:  -0.01947
## F-statistic: 0.006891 on 1 and 51 DF,  p-value: 0.9342
```

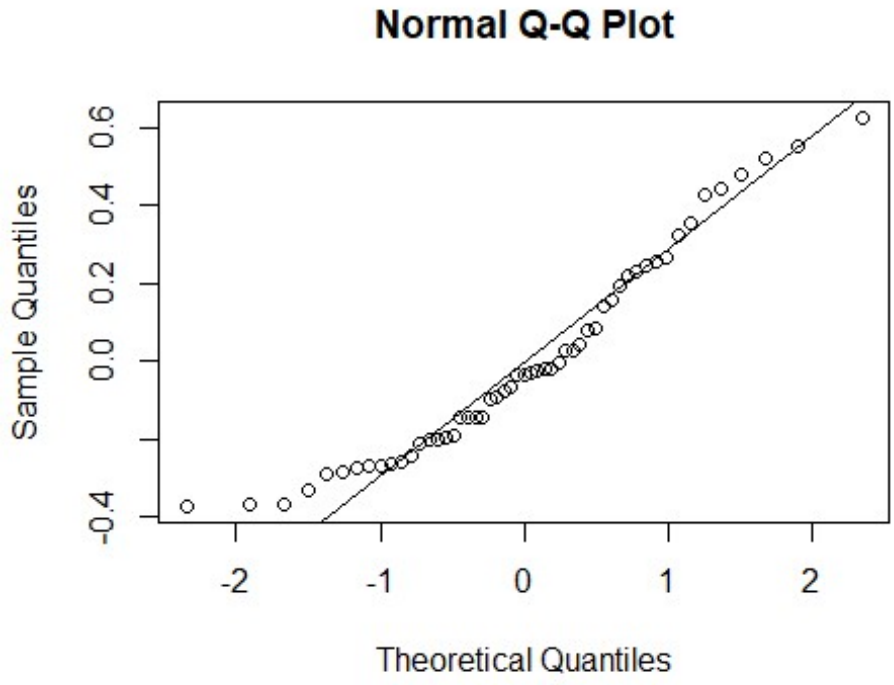
```
plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```




```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.94104, p-value = 0.01127

hist(B$residuals)
```

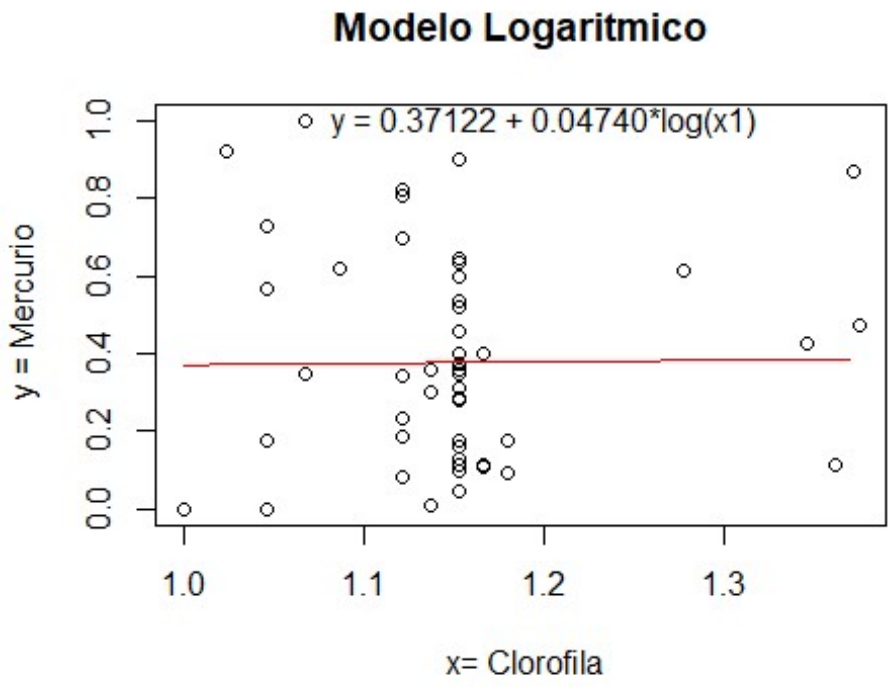



```
qqnorm(B$residuals)
qqline(B$residuals)
```



```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = 5.4407e-17, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.07286412  0.07286412
## sample estimates:
##  mean of x
## 1.9756e-18
```

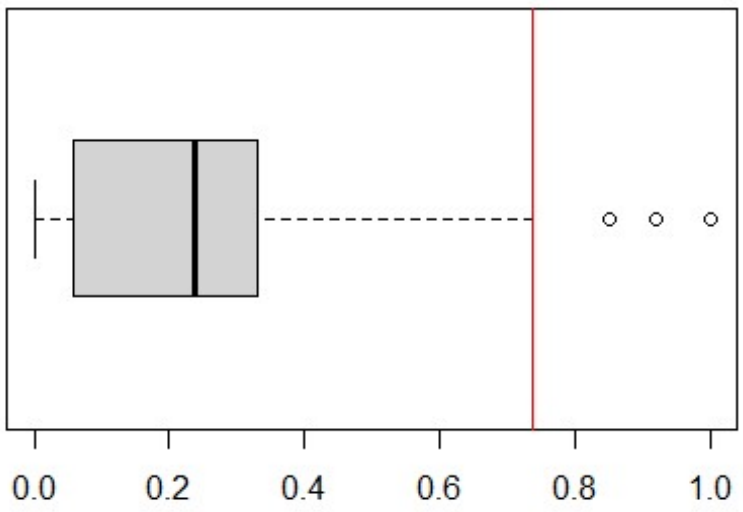
```
x=ClorofilaBC+1
y=M1$X7
plot(x, y, main = "Modelo Logaritmico", xlab =" x= Clorofila", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = 0.37122 + 0.04740*log(x1)
lines(x1, y1, col = "red")
text(1.2, 1, "y = 0.37122 + 0.04740*log(x1)")
```



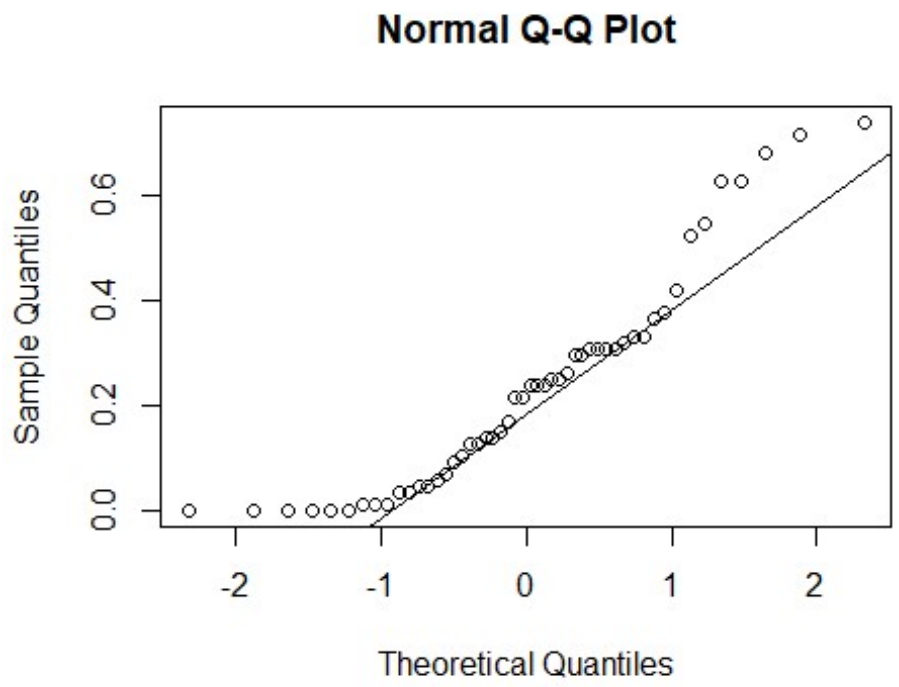
N_peces

```
q1=quantile(M1$X9,0.25)  #Cuantil 1 de la variable X
q2=quantile(M1$X9,0.5)
```

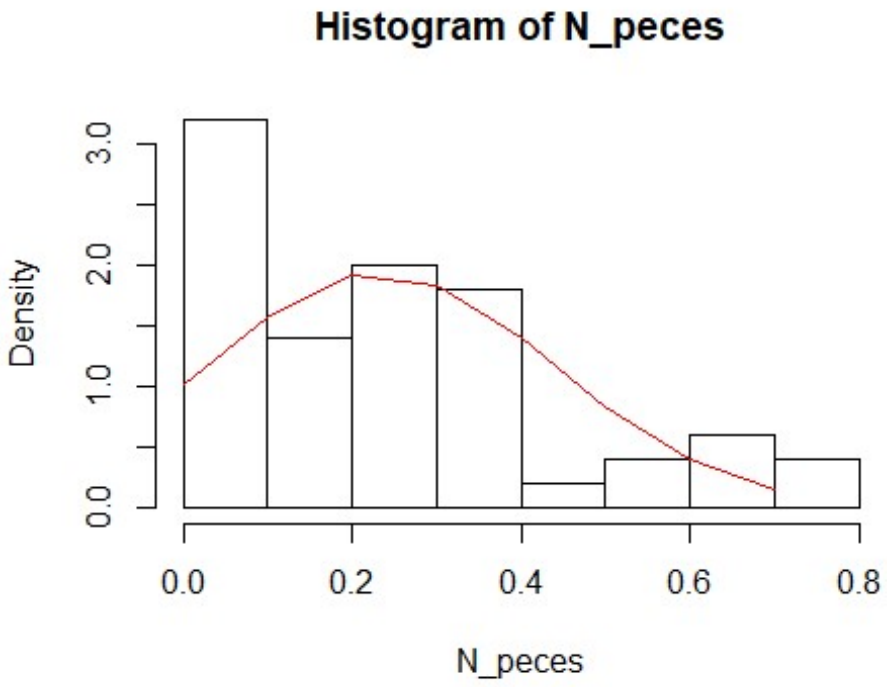
```r
q3=quantile(M1$X9,0.75)
y1=min(M1$X9)
y2=max(M1$X9)
ri=IQR(M1$X9)
boxplot(M1$X9,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```



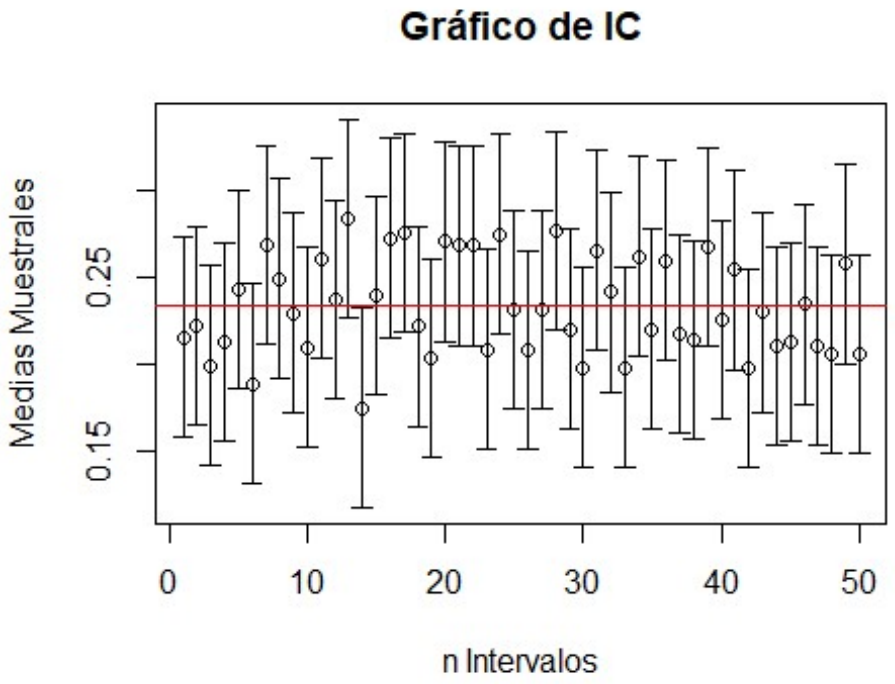

```r
N_peces= M1[M1$X9<q3+1.5*ri,c("X9")]
qqnorm(N_peces)
qqline(N_peces)
```




```
hist(N_peces,prob=TRUE,col=0)
x=seq(min(N_peces),max(N_peces),0.1)
y=dnorm(x,mean(N_peces),sd(N_peces))
lines(x,y,col="red")
```



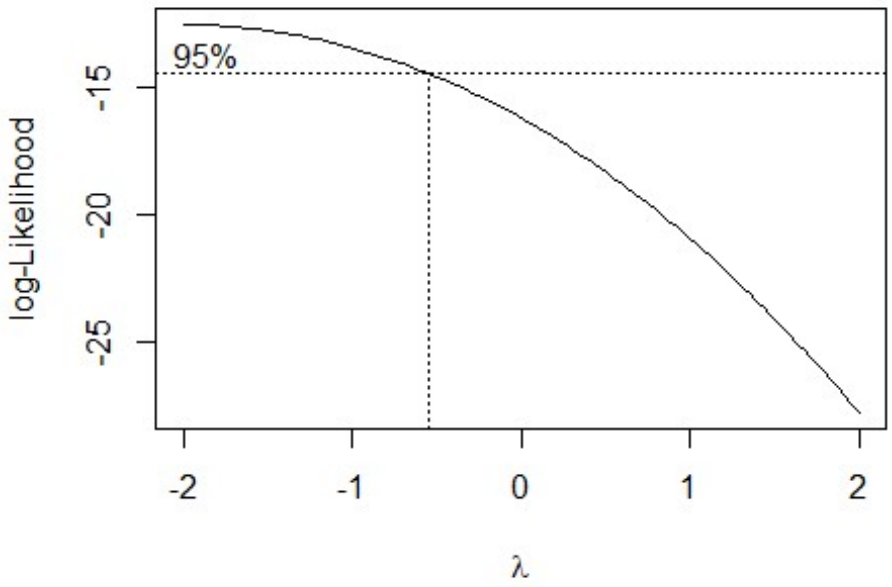

```
library(moments)
skewness1=skewness(N_peces)
cat("Sesgo: ",skewness1)

## Sesgo:  0.8566714

cat("\n")

kurtosis1=kurtosis(N_peces)
cat("Curtosis: ",kurtosis1)

## Curtosis:  2.995464

cat("\n")

library(plotrix)
n=length(N_peces)
media=mean(N_peces)
DE=sd(N_peces)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```



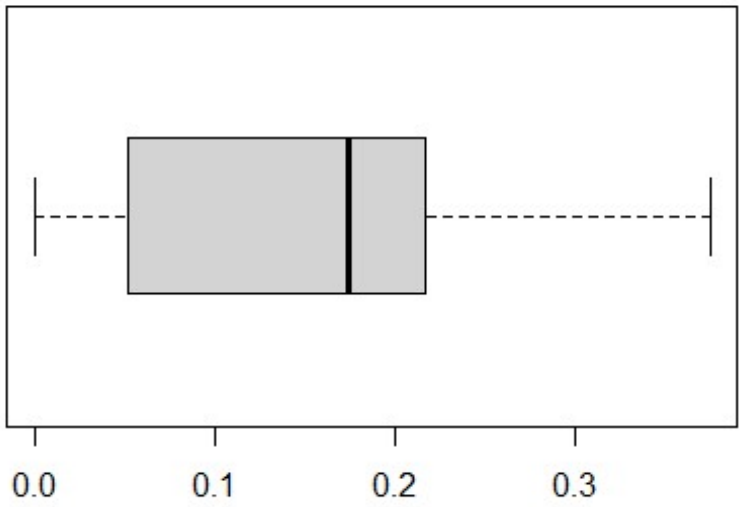

```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.2334091  +-  0.05711311
```

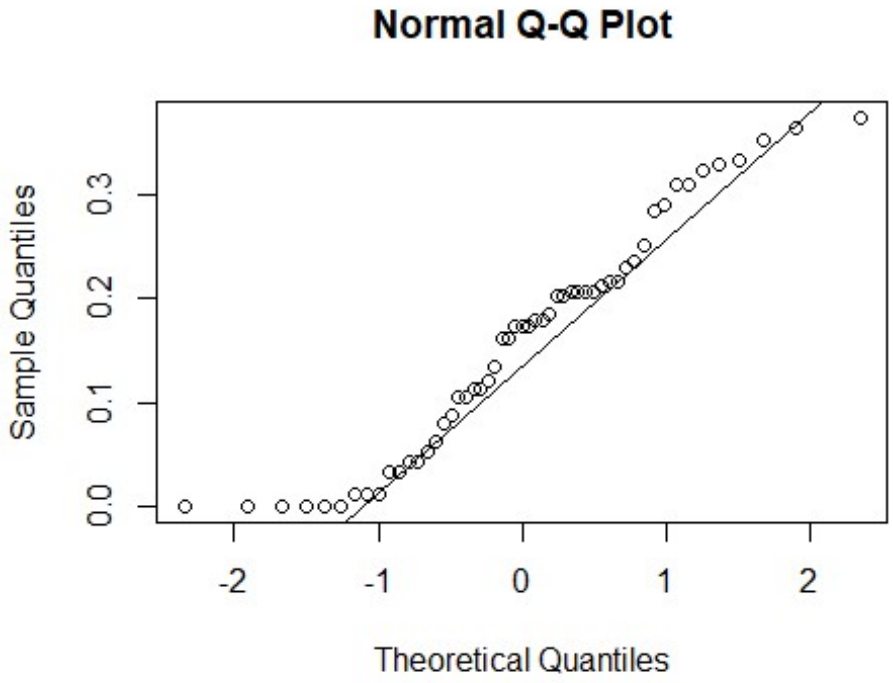
```
library(MASS)
minim=min(M1$X9)
b <- boxcox(lm((M1$X9+1) ~ 1))
```



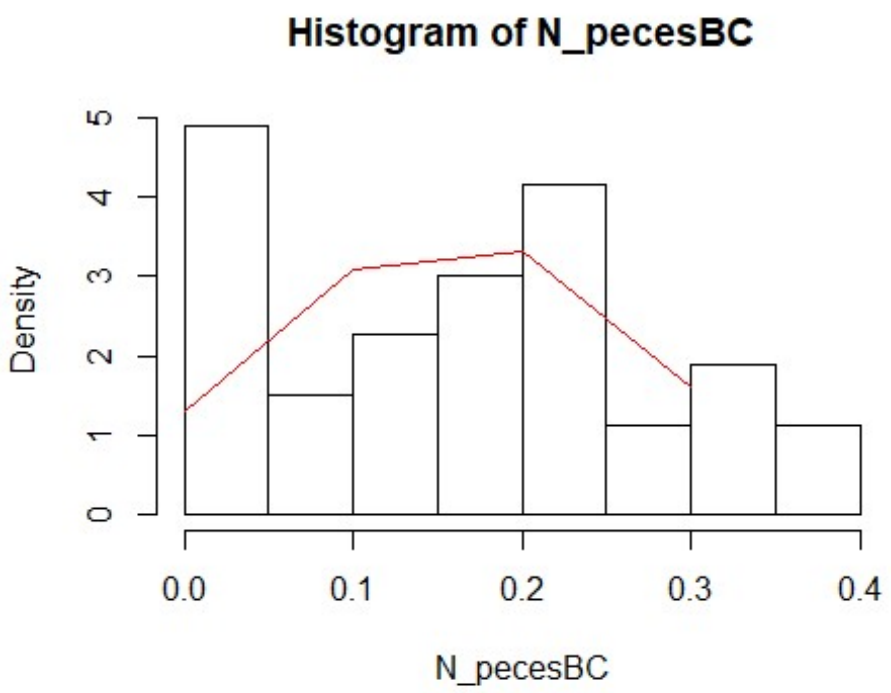

```
lambda <- b$x[which.max(b$y)]
X9<-((M1$X9+1)^lambda-1)/lambda
q1=quantile(X9,0.25)  #Cuantil 1 de la variable X
q2=quantile(X9,0.5)
q3=quantile(X9,0.75)
y1=min(X9)
y2=max(X9)
ri=IQR(X9)
boxplot(X9,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```




```
N_pecesBC= X9
qqnorm(N_pecesBC)
qqline(N_pecesBC)
```

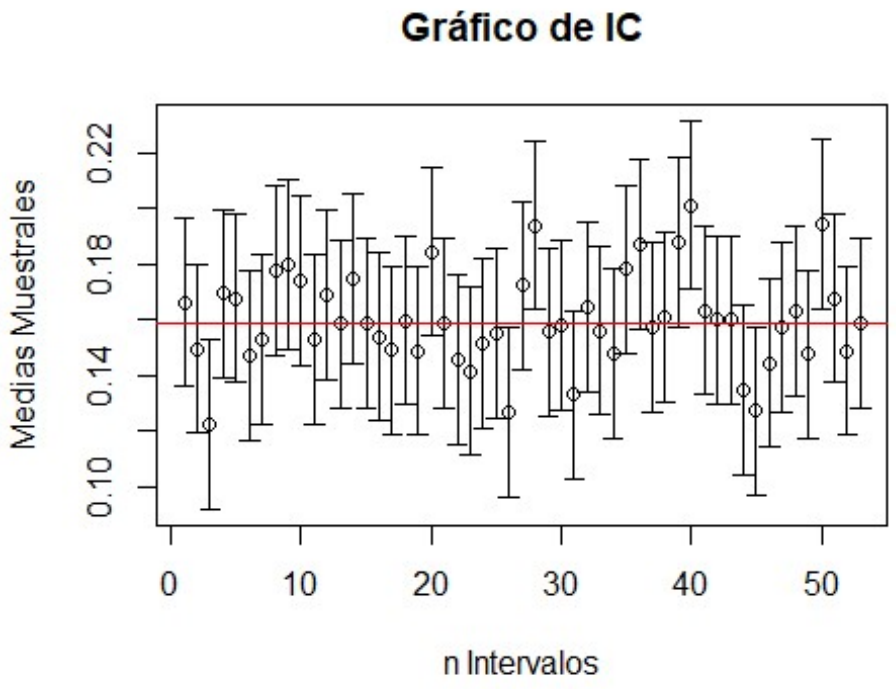


```r
hist(N_pecesBC,prob=TRUE,col=0)
x=seq(min(N_pecesBC),max(N_pecesBC),0.1)
y=dnorm(x,mean(N_pecesBC),sd(N_pecesBC))
lines(x,y,col="red")
```




```r
library(moments)
skewness1=skewness(N_pecesBC)
cat("Sesgo: ",skewness1)

## Sesgo:  0.1444805

cat("\n")

kurtosis1=kurtosis(N_pecesBC)
cat("Curtosis: ",kurtosis1)

## Curtosis:  1.981965

cat("\n")

library(plotrix)
n=length(N_pecesBC)
media=mean(N_pecesBC)
DE=sd(N_pecesBC)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
```

```
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```




```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.15896  +-  0.03029483
```

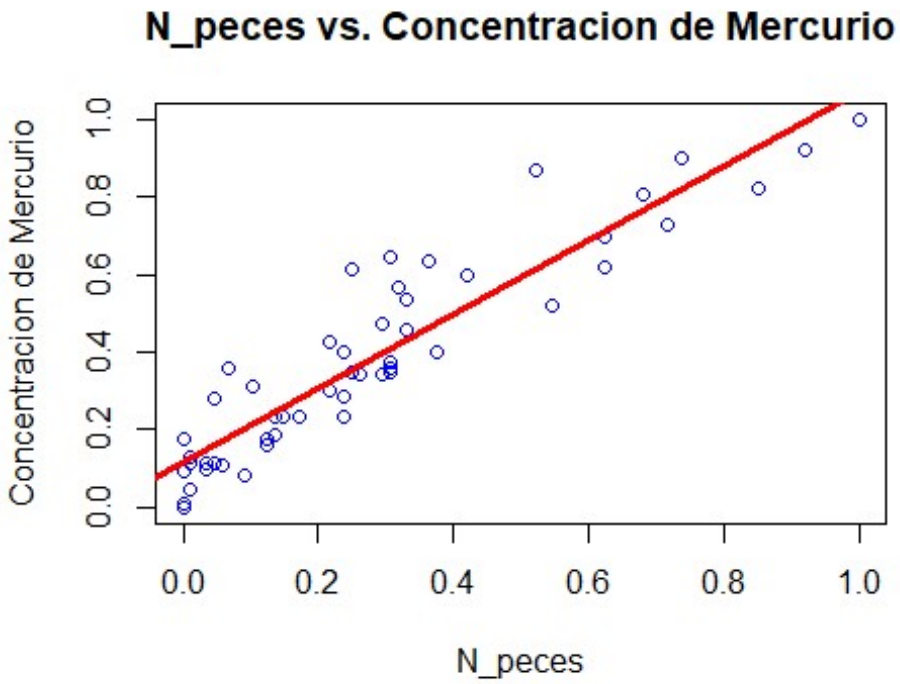
```
cor.test(M1$X7,M1$X9)
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and M1$X9
## t = 17.679, df = 51, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.8766037 0.9575259
## sample estimates:
##       cor
## 0.9272051
```

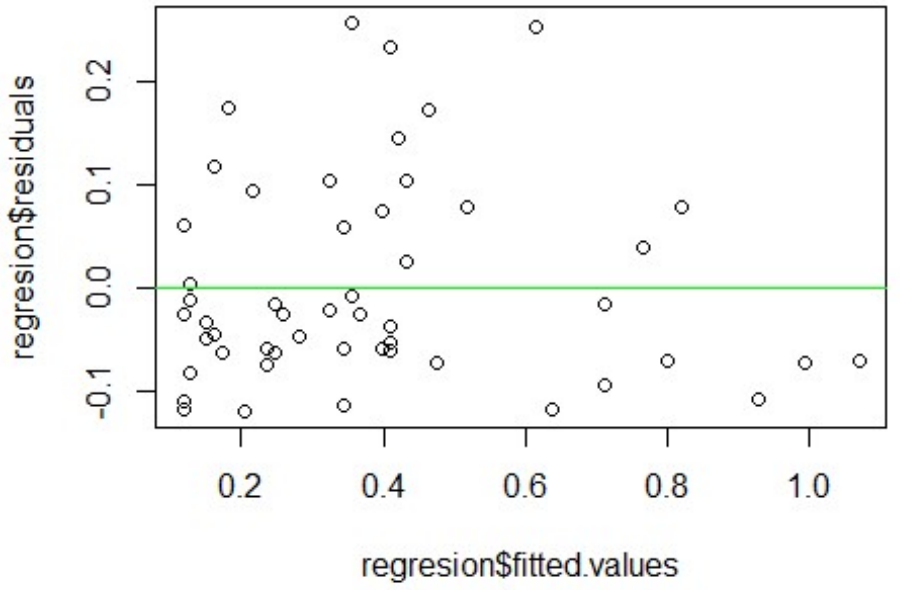
```
regresion=lm(M1$X7 ~ M1$X9)
regresion
```

```
##
## Call:
## lm(formula = M1$X7 ~ M1$X9)
```

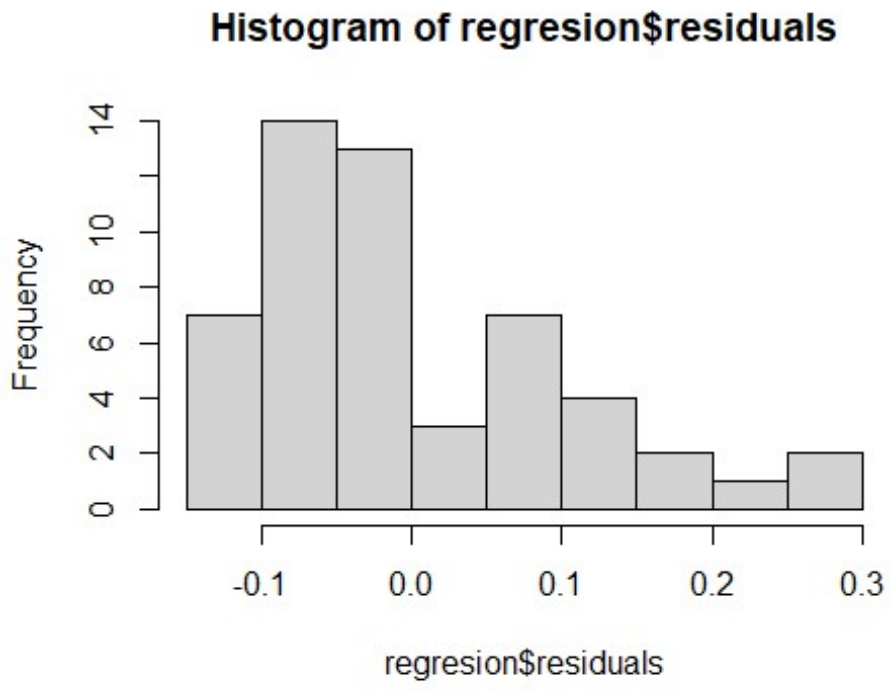
```
##
## Coefficients:
## (Intercept)        M1$X9
##      0.1180       0.9528

N_peces=summary(regresion)
N_peces

##
## Call:
## lm(formula = M1$X7 ~ M1$X9)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.11936 -0.07077 -0.02620  0.06028  0.25620
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.11801    0.02011   5.869 3.29e-07 ***
## M1$X9        0.95275    0.05389  17.679  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09999 on 51 degrees of freedom
## Multiple R-squared:  0.8597, Adjusted R-squared:  0.857
## F-statistic: 312.5 on 1 and 51 DF,  p-value: < 2.2e-16

plot(M1$X9,M1$X7,col="blue",xlab="N_peces",ylab="Concentracion de
Mercurio",main="N_peces vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```



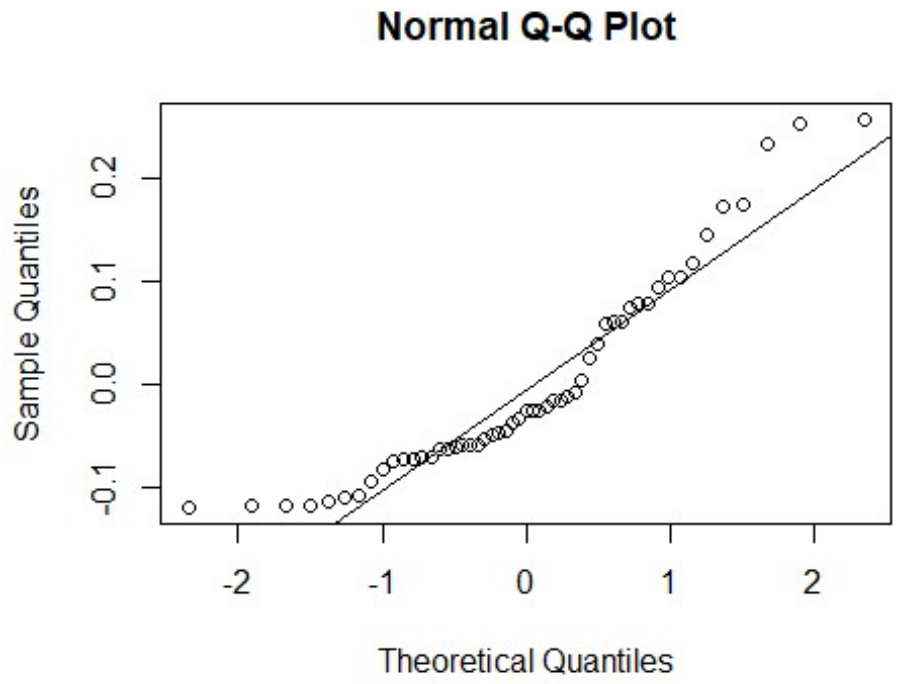

```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.89509, p-value = 0.0002206

hist(regresion$residuals)
```




```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```




```
t.test(regresion$residuals)

##
##  One Sample t-test
##
## data:  regresion$residuals
## t = -2.5991e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.0272934  0.0272934
## sample estimates:
##     mean of x
## -3.535228e-18

cor.test(M1$X7,N_pecesBC)

##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and N_pecesBC
## t = 16.933, df = 51, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.8670586 0.9540872
## sample estimates:
##       cor
## 0.9214102
```

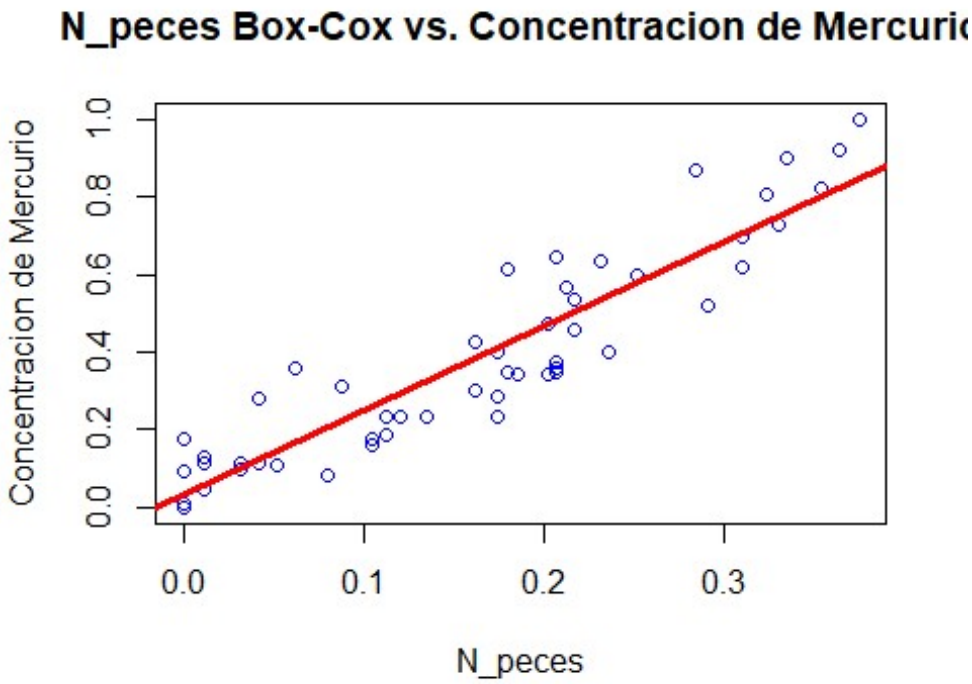
```r
regresion=lm(M1$X7 ~ N_pecesBC)
regresion
```

```
##
## Call:
## lm(formula = M1$X7 ~ N_pecesBC)
##
## Coefficients:
## (Intercept)    N_pecesBC
##     0.03354      2.16473
```

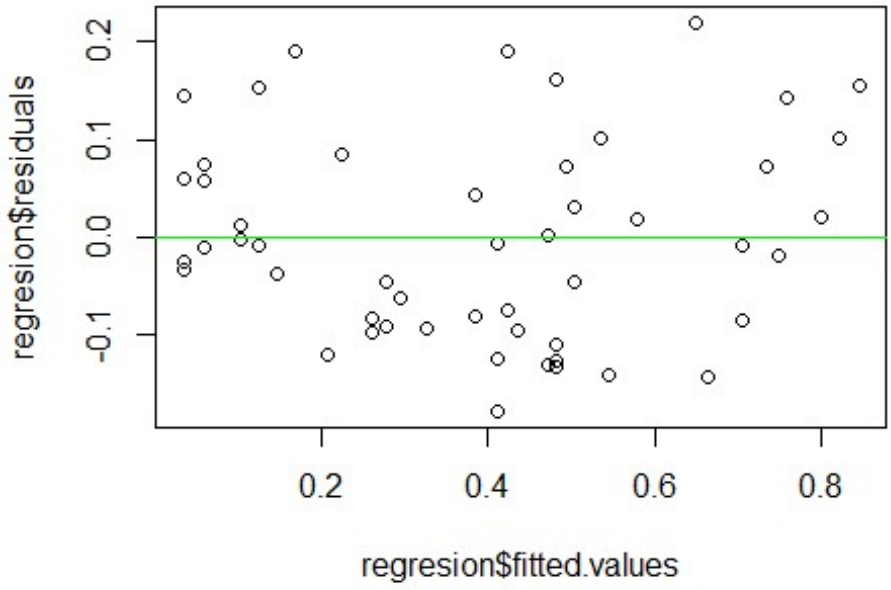
```r
N_peces=summary(regresion)
N_peces
```

```
##
## Call:
## lm(formula = M1$X7 ~ N_pecesBC)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.177870 -0.085867 -0.009339  0.072952  0.219104
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.03354    0.02482   1.352    0.182
## N_pecesBC    2.16473    0.12784  16.933   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1037 on 51 degrees of freedom
## Multiple R-squared:  0.849,  Adjusted R-squared:  0.846
## F-statistic: 286.7 on 1 and 51 DF,  p-value: < 2.2e-16
```

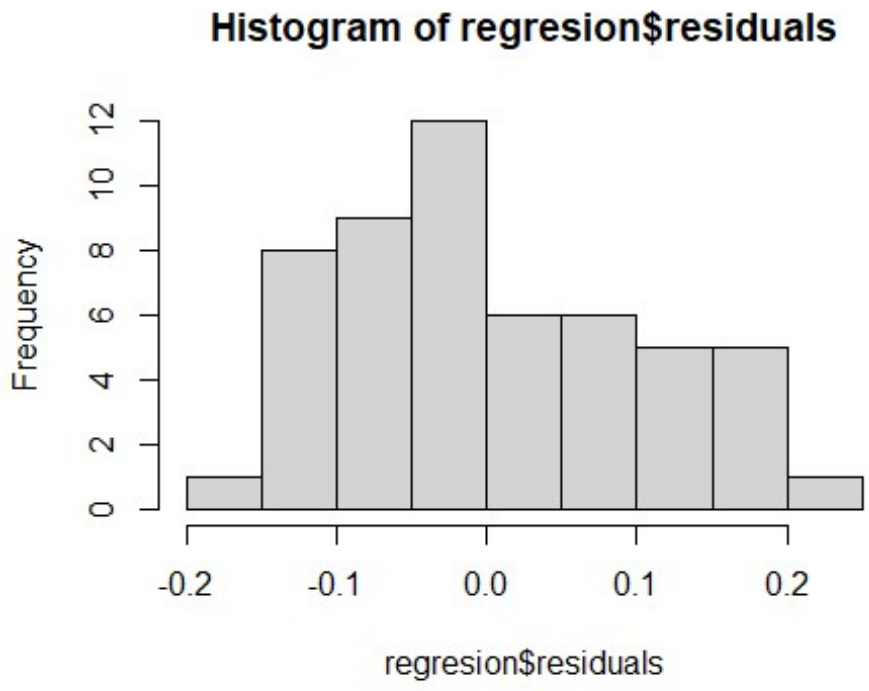
```r
plot(N_pecesBC,M1$X7,col="blue",xlab="N_peces",ylab="Concentracion de
Mercurio",main="N_peces Box-Cox vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```



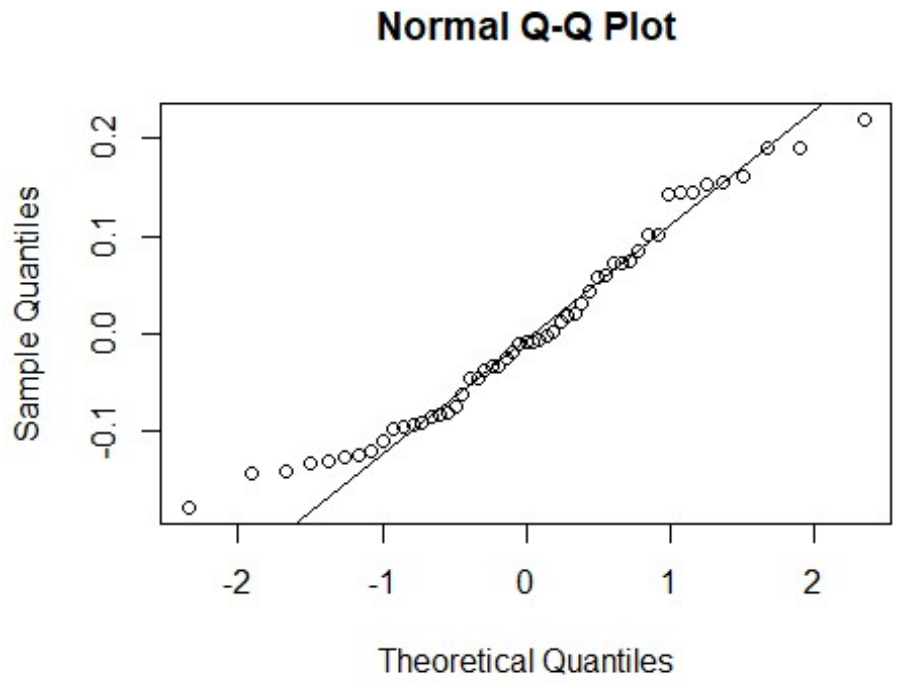

```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.95979, p-value = 0.07185

hist(regresion$residuals)
```




```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```



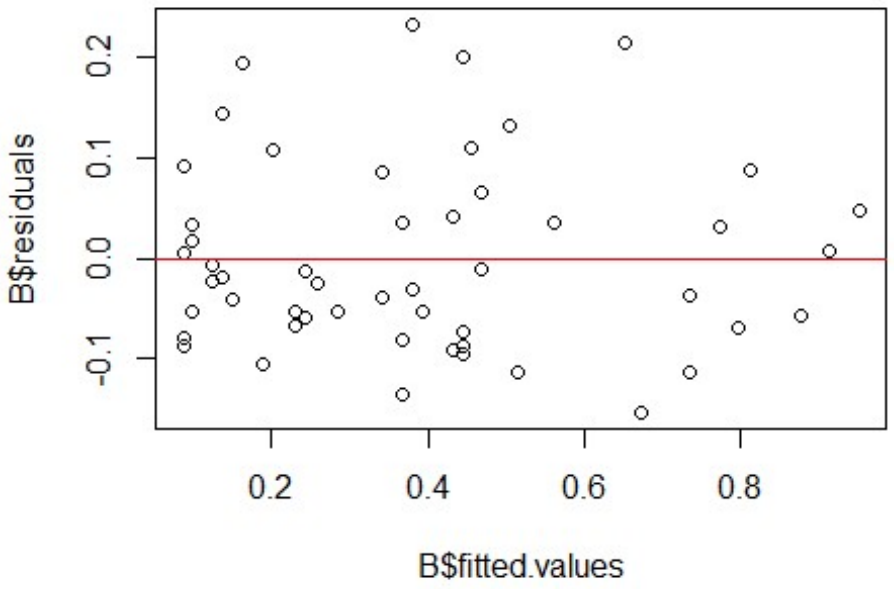

```
t.test(regresion$residuals)
```

```
##
##  One Sample t-test
##
## data:  regresion$residuals
## t = 2.5081e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.02831628  0.02831628
## sample estimates:
##    mean of x
## 3.539255e-18
```

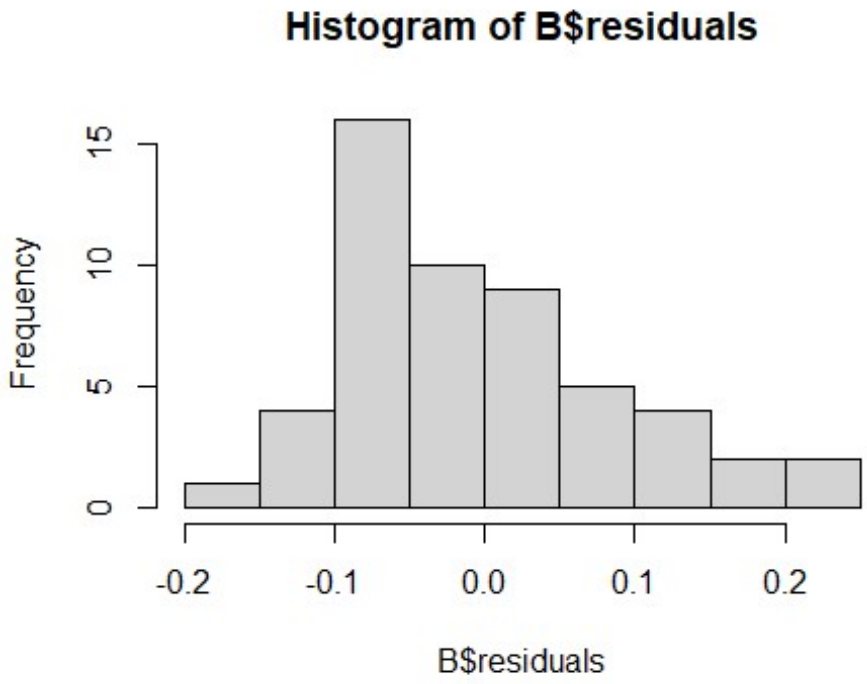
```
z1=N_pecesBC
z2=N_pecesBC^2
cor.test(M1$X7 , z1+ z2)
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and z1 + z2
## t = 18.29, df = 51, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.8837265 0.9600769
## sample estimates:
```

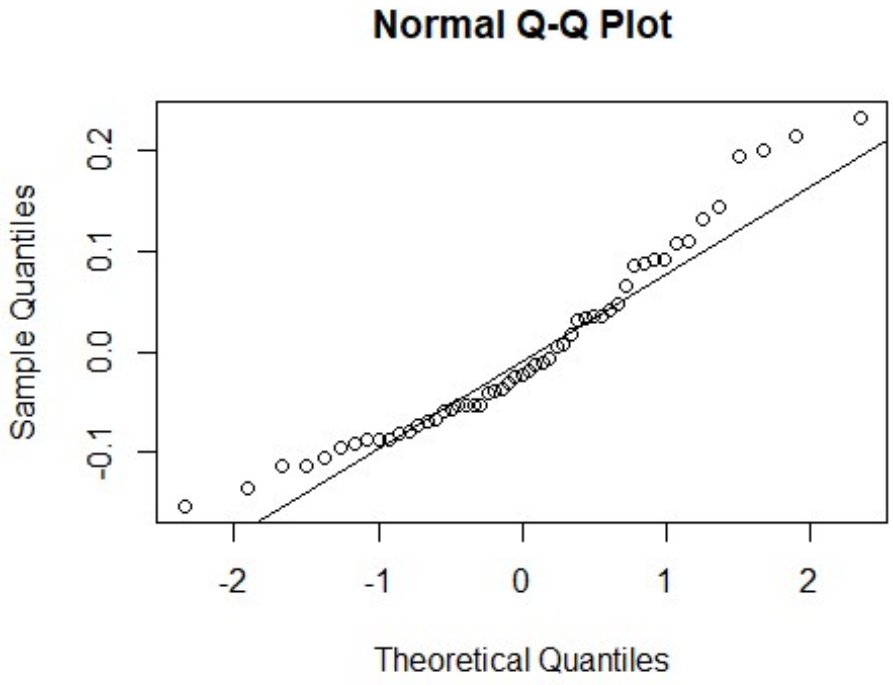
```
##       cor
## 0.9315136

B = lm(M1$X7 ~ z1+ z2)
summary(B)

##
## Call:
## lm(formula = M1$X7 ~ z1 + z2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.15308 -0.06869 -0.02248  0.04788  0.23197
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.08701    0.02818   3.088  0.00328 **
## z1           1.00538    0.37772   2.662  0.01043 *
## z2           3.47088    1.07478   3.229  0.00219 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0953 on 50 degrees of freedom
## Multiple R-squared:  0.8751, Adjusted R-squared:  0.8701
## F-statistic: 175.1 on 2 and 50 DF,  p-value: < 2.2e-16

plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```



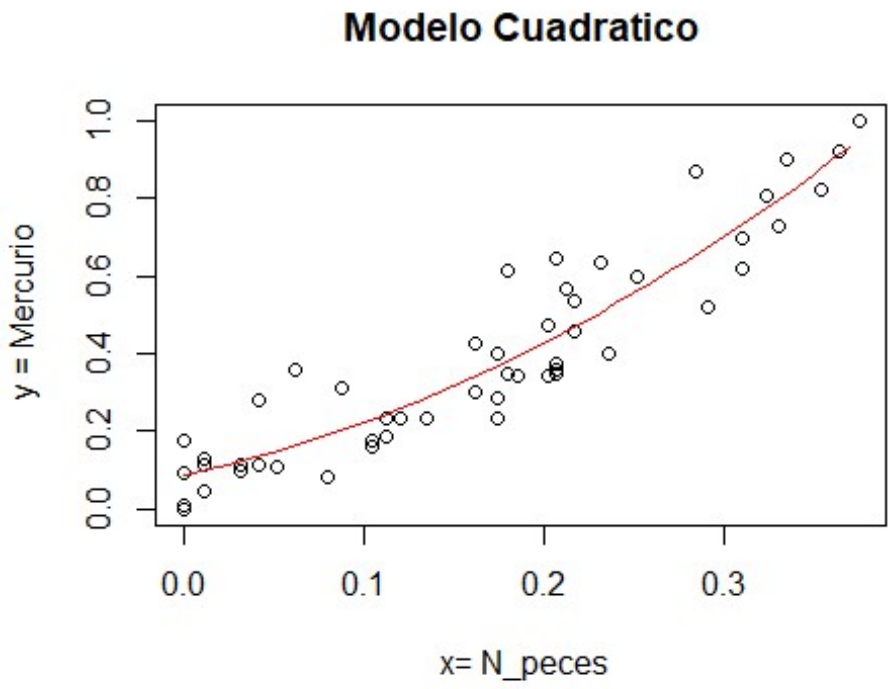

```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.94025, p-value = 0.01046

hist(B$residuals)
```




```
qqnorm(B$residuals)
qqline(B$residuals)
```



```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = 8.1642e-17, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.02575719  0.02575719
## sample estimates:
##    mean of x
## 1.047956e-18

x=N_pecesBC
y=M1$X7
plot(x, y, main = "Modelo Cuadratico", xlab =" x= N_peces", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = 0.08701  +1.00538*x1 +3.47088*x1^2
lines(x1, y1, col = "red")
text(1, .1, "0.08701  +1.00538*x1 +3.47088*x1^2")
```



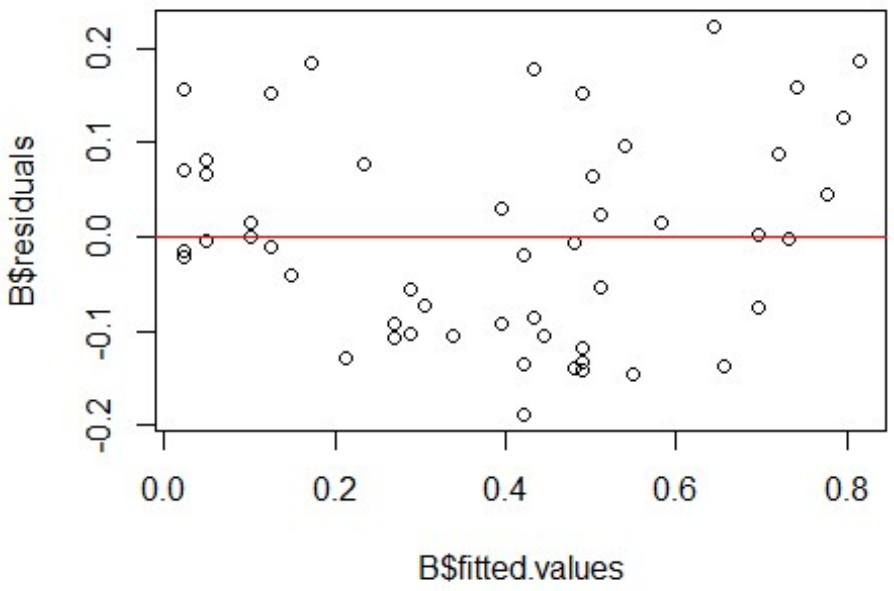

```
z1=N_pecesBC+1
z2=log(N_pecesBC+1)
cor.test(M1$X7 , log(N_pecesBC+1))
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and log(N_pecesBC + 1)
## t = 15.98, df = 51, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.8532455 0.9490700
## sample estimates:
##       cor
## 0.9129812
```

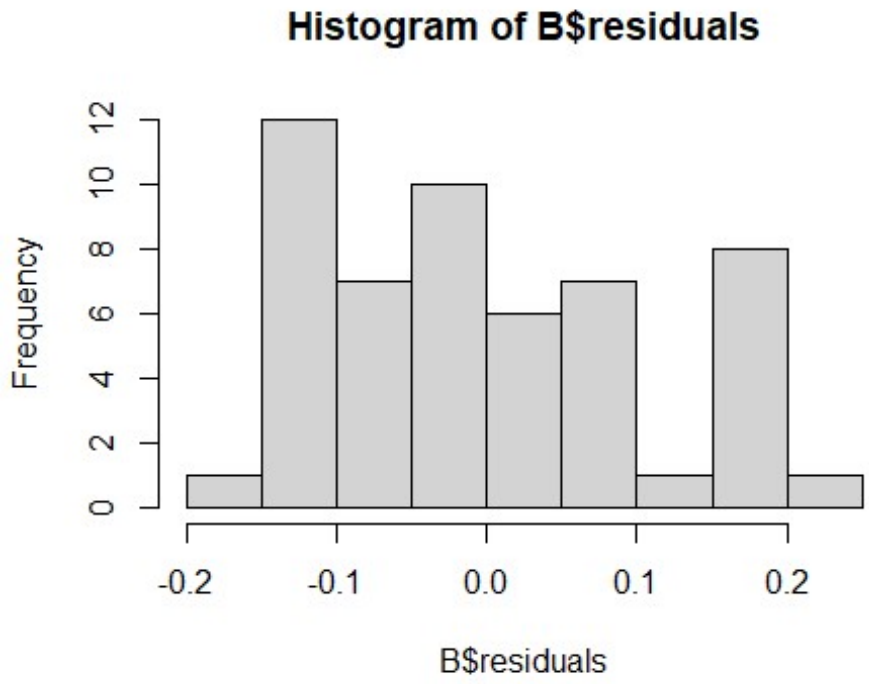
```
B = lm(M1$X7 ~ log(N_pecesBC+1))
summary(B)
```

```
##
## Call:
## lm(formula = M1$X7 ~ log(N_pecesBC + 1))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.188848 -0.092910 -0.007039  0.076887  0.223703
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.02245    0.02679   0.838    0.406
## log(N_pecesBC + 1)  2.48563    0.15554  15.980   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1089 on 51 degrees of freedom
## Multiple R-squared:  0.8335, Adjusted R-squared:  0.8303
## F-statistic: 255.4 on 1 and 51 DF,  p-value: < 2.2e-16
```

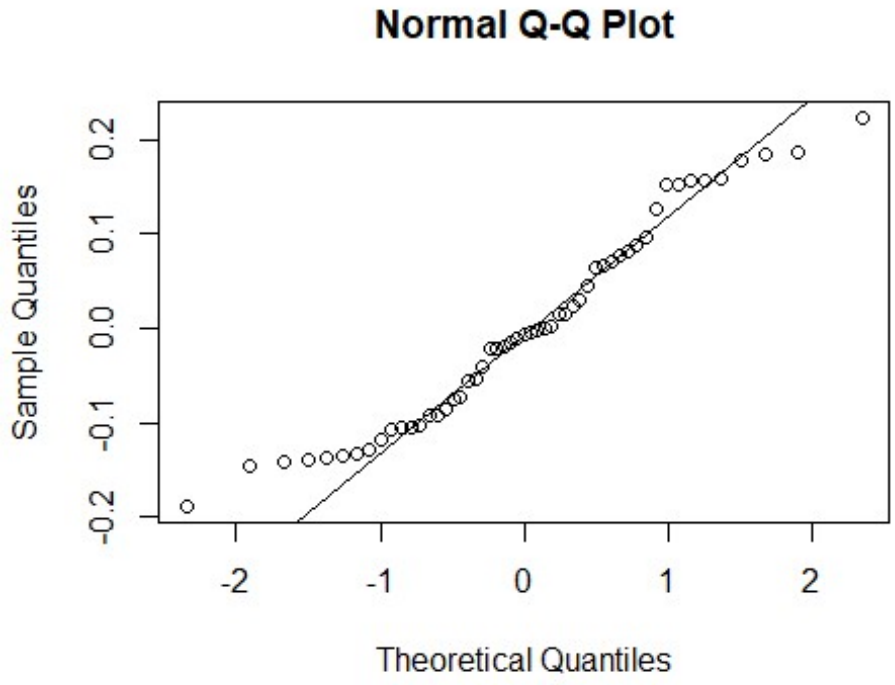
```
plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```



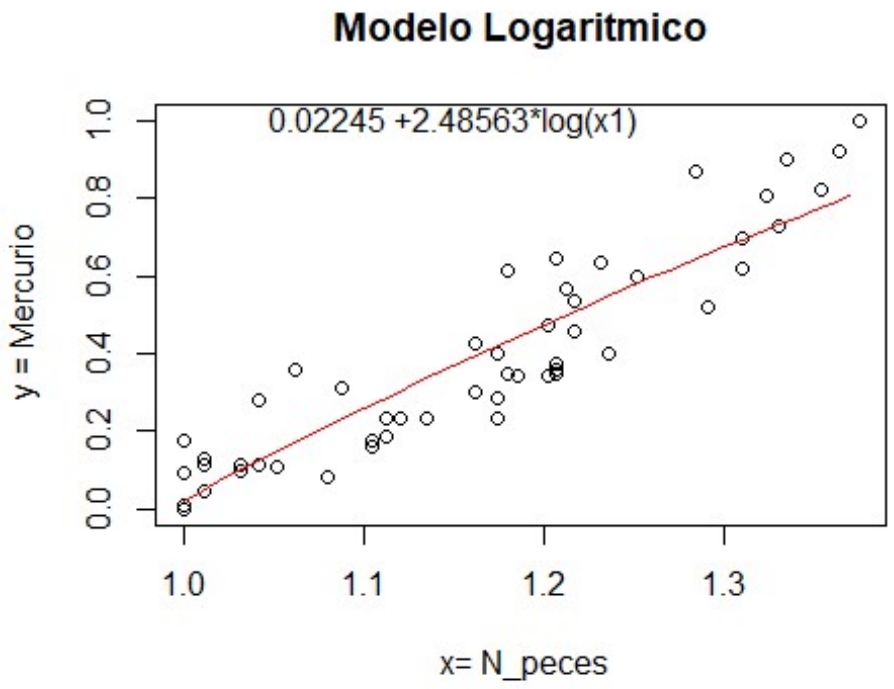

```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.95466, p-value = 0.04281

hist(B$residuals)
```




```
qqnorm(B$residuals)
qqline(B$residuals)
```



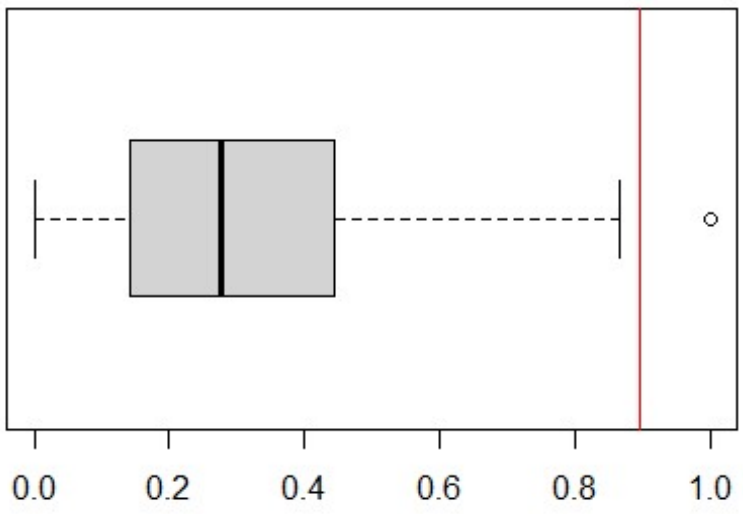
```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = 1.0775e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.02973068  0.02973068
## sample estimates:
##    mean of x
## 1.596385e-18

x=N_pecesBC+1
y=M1$X7
plot(x, y, main = "Modelo Logaritmico", xlab =" x= N_peces", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = 0.02245 +2.48563*log(x1)
lines(x1, y1, col = "red")
text(1.15, 1, "0.02245 +2.48563*log(x1)")
```



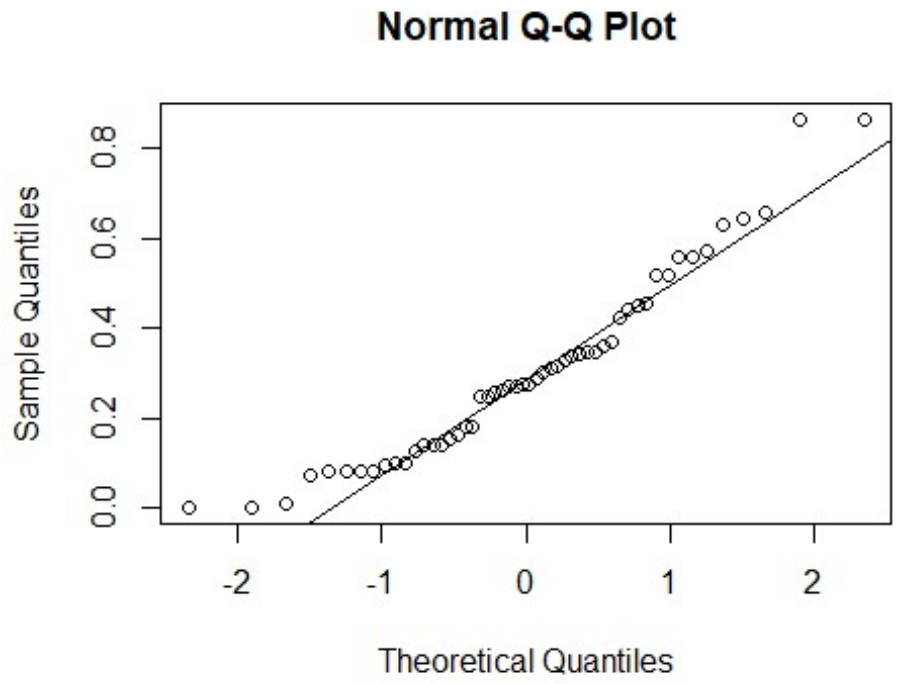
Estimacion

```
q1=quantile(M1$X11,0.25)  #Cuantil 1 de la variable X
q2=quantile(M1$X11,0.5)
```

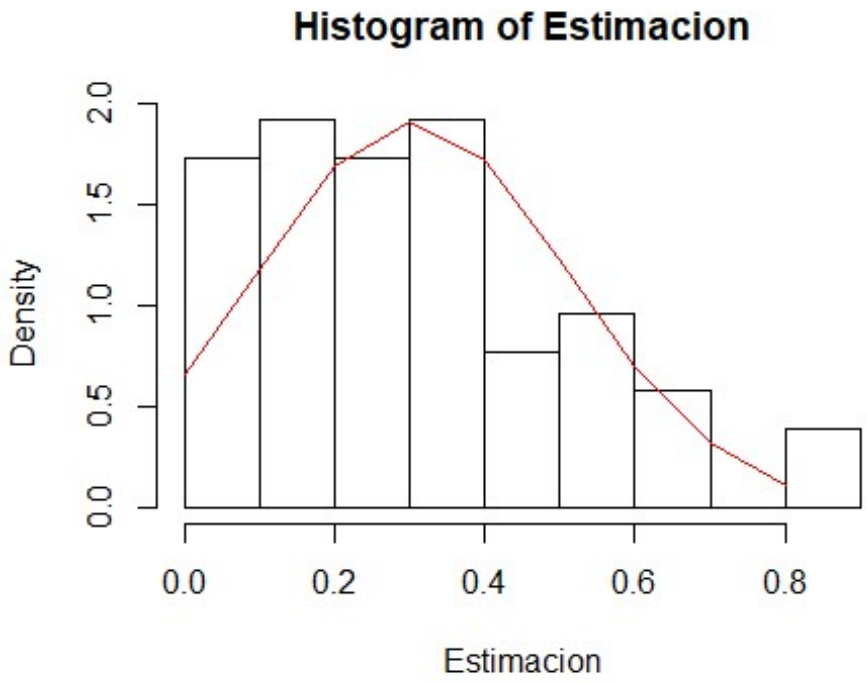
```
q3=quantile(M1$X11,0.75)
y1=min(M1$X11)
y2=max(M1$X11)
ri=IQR(M1$X11)
boxplot(M1$X11,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```



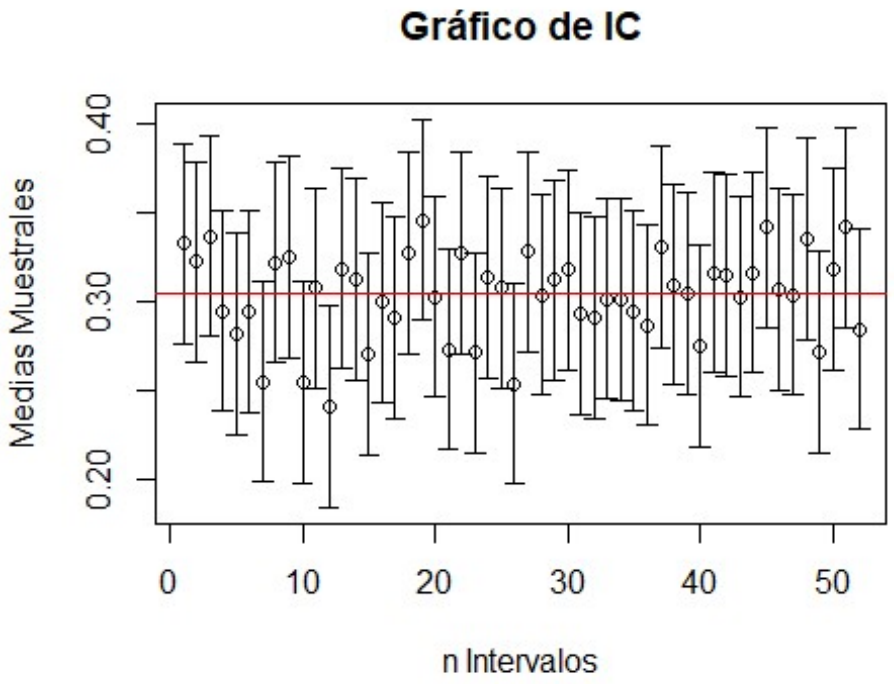

```
Estimacion= M1[M1$X11<q3+1.5*ri,c("X11")]
qqnorm(Estimacion)
qqline(Estimacion)
```




```
hist(Estimacion,prob=TRUE,col=0)
x=seq(min(Estimacion),max(Estimacion),0.1)
y=dnorm(x,mean(Estimacion),sd(Estimacion))
lines(x,y,col="red")
```



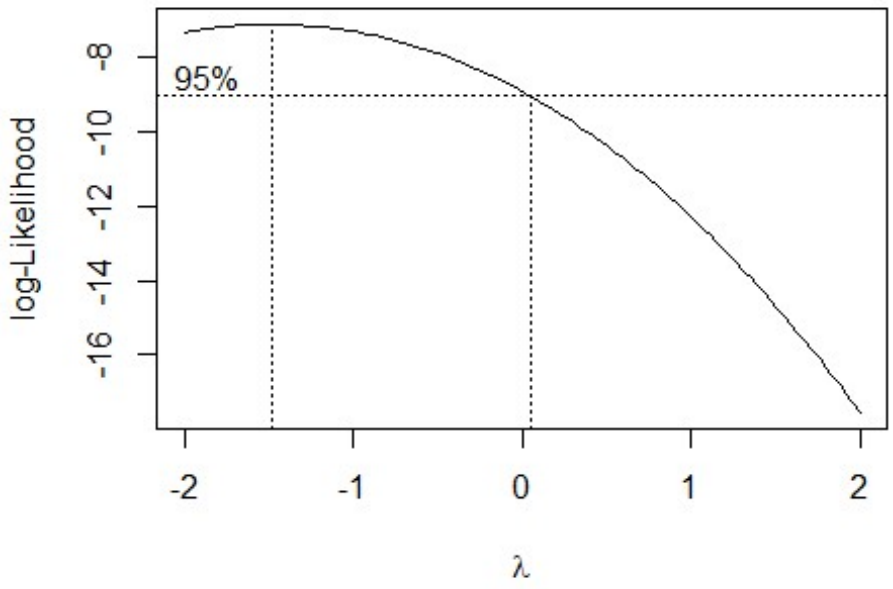

```
library(moments)
skewness1=skewness(Estimacion)
cat("Sesgo: ",skewness1)

## Sesgo:  0.7840477

cat("\n")

kurtosis1=kurtosis(Estimacion)
cat("Curtosis: ",kurtosis1)

## Curtosis:  3.244217

cat("\n")

library(plotrix)
n=length(Estimacion)
media=mean(Estimacion)
DE=sd(Estimacion)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```



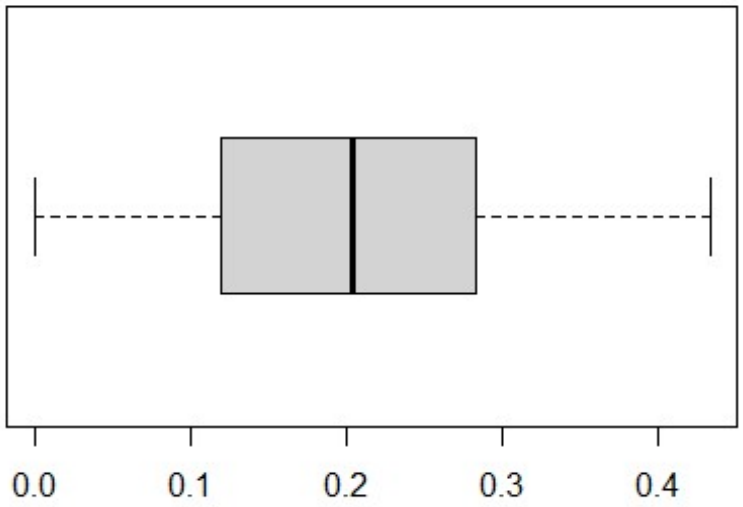

```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.3044657  +-  0.05661494
```

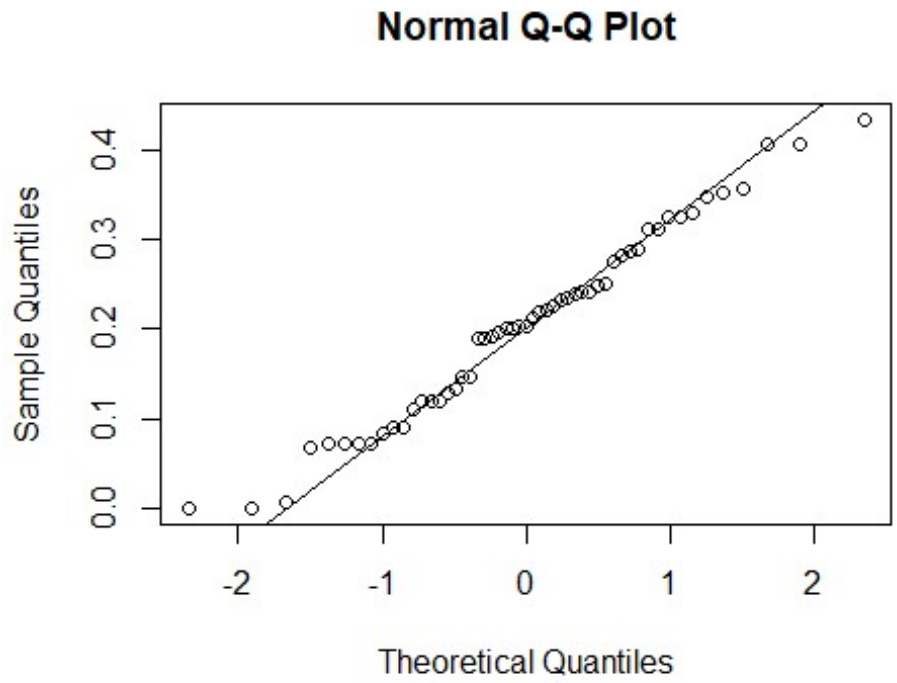
```
library(MASS)
minim=min(M1$X11)
b <- boxcox(lm((M1$X11+1) ~ 1))
```



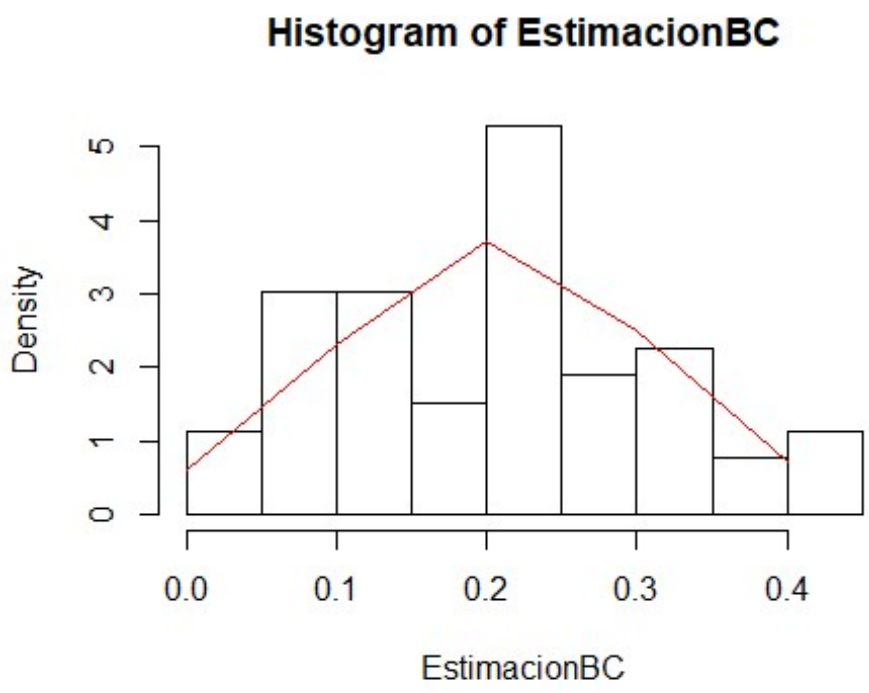

```
lambda <- b$x[which.max(b$y)]
X11<-((M1$X11+1)^lambda-1)/lambda
q1=quantile(X11,0.25)  #Cuantil 1 de la variable X
q2=quantile(X11,0.5)
q3=quantile(X11,0.75)
y1=min(X11)
y2=max(X11)
ri=IQR(X11)
boxplot(X11,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")
```




```
EstimacionBC= X11
qqnorm(EstimacionBC)
qqline(EstimacionBC)
```

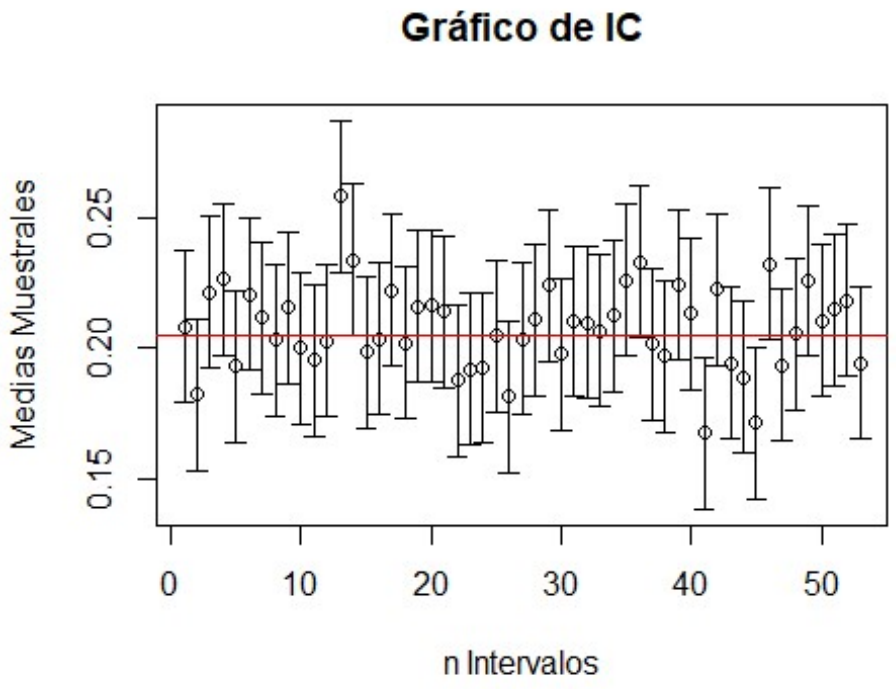


```
hist(EstimacionBC,prob=TRUE,col=0)
x=seq(min(EstimacionBC),max(EstimacionBC),0.1)
y=dnorm(x,mean(EstimacionBC),sd(EstimacionBC))
lines(x,y,col="red")
```




```
library(moments)
skewness1=skewness(EstimacionBC)
cat("Sesgo: ",skewness1)

## Sesgo:  0.05524268

cat("\n")

kurtosis1=kurtosis(EstimacionBC)
cat("Curtosis: ",kurtosis1)

## Curtosis:  2.3504

cat("\n")

library(plotrix)
n=length(EstimacionBC)
media=mean(EstimacionBC)
DE=sd(EstimacionBC)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
```

```
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```




```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.2046482  +-  0.02901399
```

```
cor.test(M1$X7,M1$X11)
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and M1$X11
## t = 24.233, df = 51, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.9300566 0.9763663
## sample estimates:
##       cor
## 0.9592148
```

```
regresion=lm(M1$X7 ~ M1$X11)
regresion
```

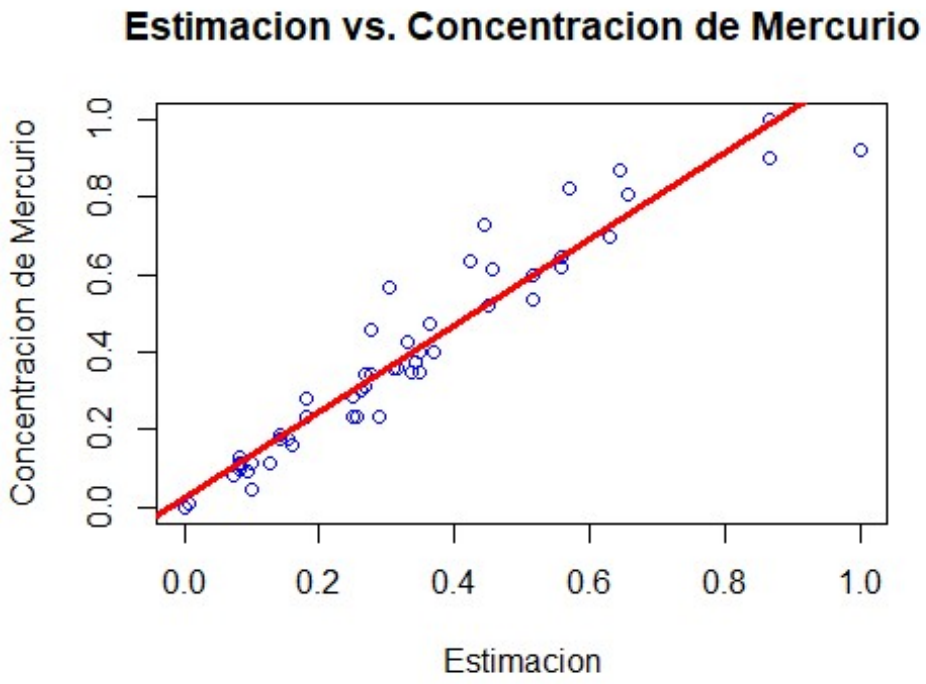
```
##
## Call:
## lm(formula = M1$X7 ~ M1$X11)
```

```
##
## Coefficients:
## (Intercept)       M1$X11
##     0.02339      1.11547
```

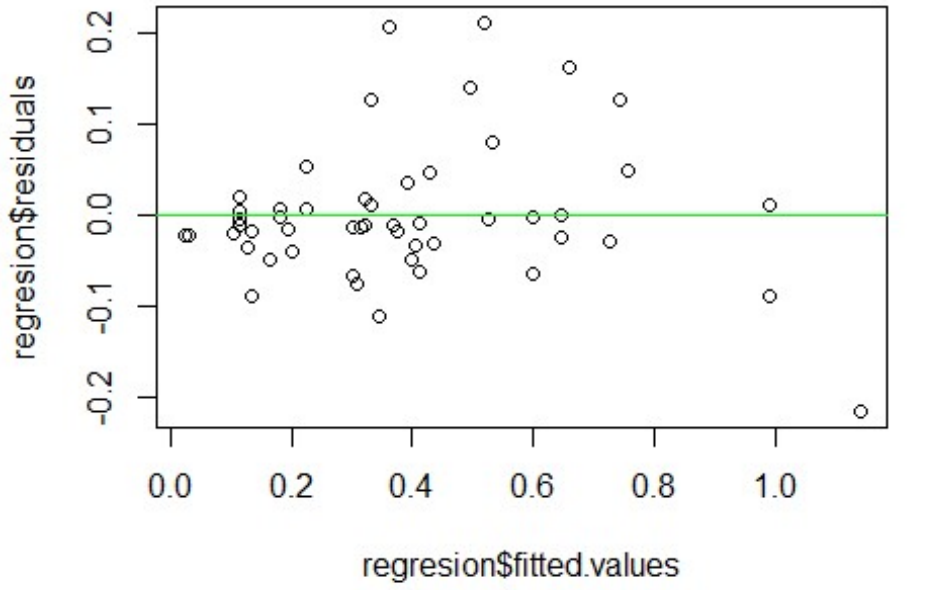
```
Estimacion=summary(regresion)
Estimacion
```

```
##
## Call:
## lm(formula = M1$X7 ~ M1$X11)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.21638 -0.03204 -0.01245  0.01086  0.21119
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.02339    0.01792   1.305    0.198
## M1$X11       1.11547    0.04603  24.233   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07546 on 51 degrees of freedom
## Multiple R-squared:  0.9201, Adjusted R-squared:  0.9185
## F-statistic: 587.2 on 1 and 51 DF,  p-value: < 2.2e-16
```

```
plot(M1$X11,M1$X7,col="blue",xlab="Estimacion",ylab="Concentracion de
Mercurio",main="Estimacion vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```

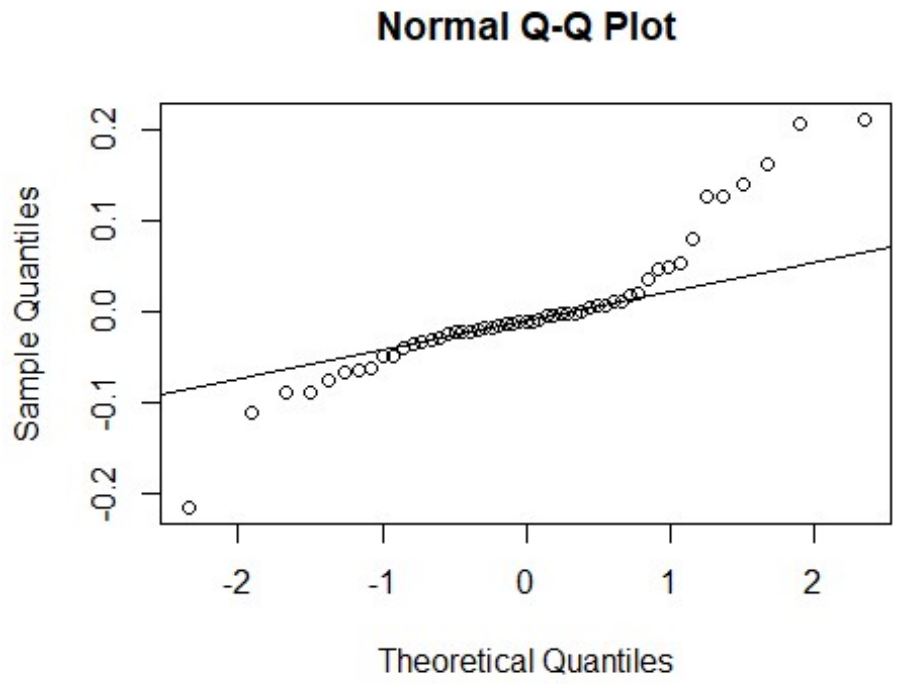



```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.89095, p-value = 0.0001611

hist(regresion$residuals)
```


```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```




```
t.test(regresion$residuals)

##
##  One Sample t-test
##
## data:  regresion$residuals
## t = -6.1197e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.02059849  0.02059849
## sample estimates:
##     mean of x
## -6.281992e-18

cor.test(M1$X7,EstimacionBC)

##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and EstimacionBC
## t = 23.044, df = 51, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.9232584 0.9740086
## sample estimates:
##       cor
## 0.9551849
```

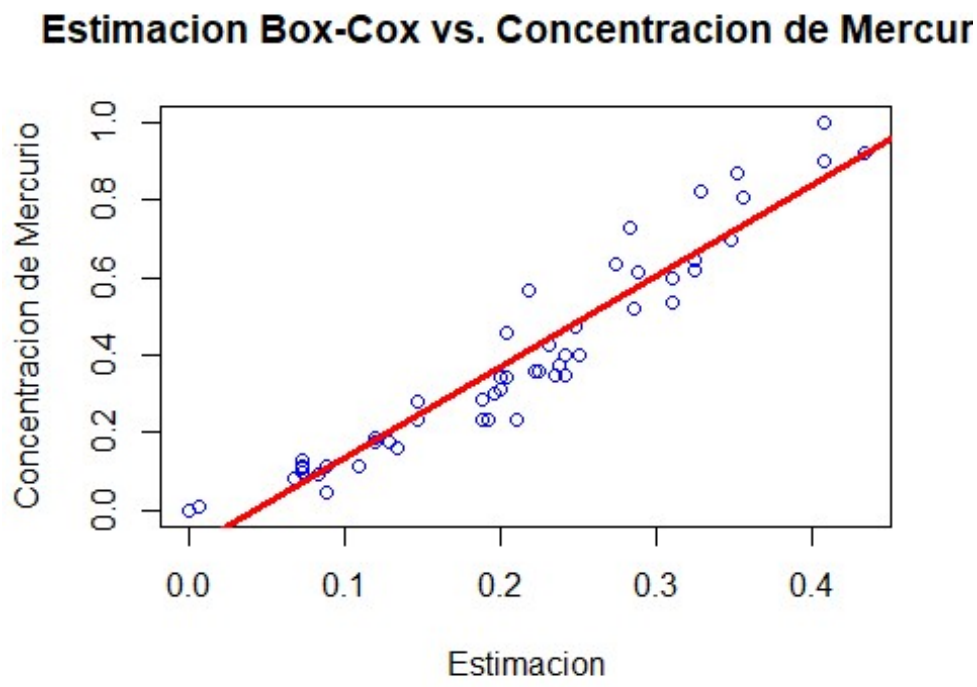
```r
regresion=lm(M1$X7 ~ EstimacionBC)
regresion
```

```
##
## Call:
## lm(formula = M1$X7 ~ EstimacionBC)
##
## Coefficients:
##  (Intercept)  EstimacionBC
##      -0.1019        2.3432
```

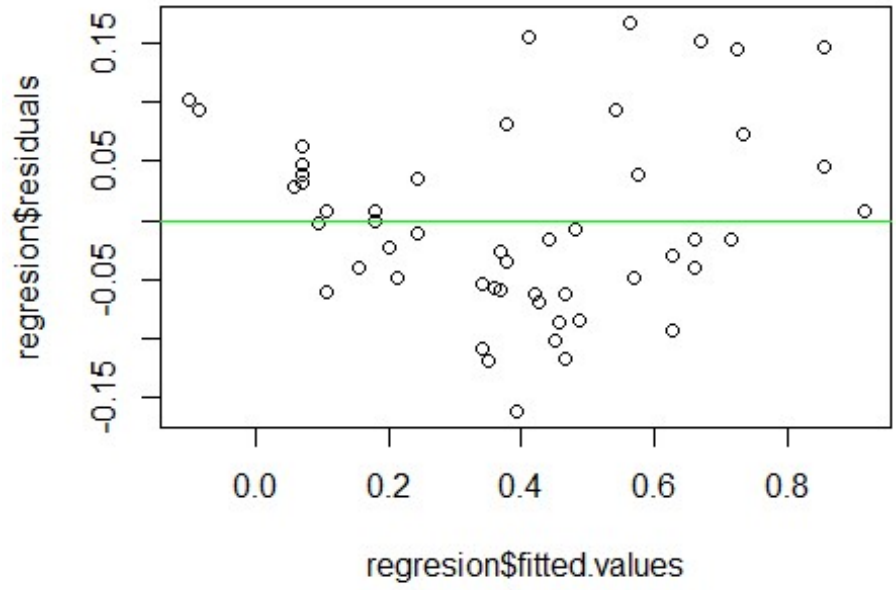
```r
Estimacion=summary(regresion)
Estimacion
```

```
##
## Call:
## lm(formula = M1$X7 ~ EstimacionBC)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.16124 -0.05700 -0.01157  0.04558  0.16688
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.10187    0.02347  -4.341 6.74e-05 ***
## EstimacionBC  2.34315    0.10168  23.044  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07902 on 51 degrees of freedom
## Multiple R-squared:  0.9124, Adjusted R-squared:  0.9107
## F-statistic:   531 on 1 and 51 DF,  p-value: < 2.2e-16
```

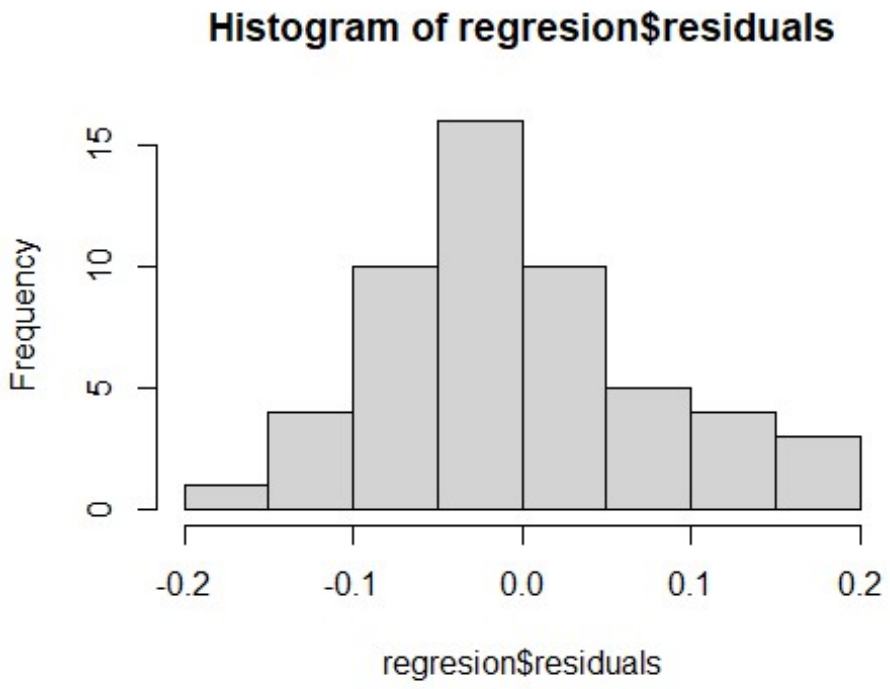
```r
plot(EstimacionBC,M1$X7,col="blue",xlab="Estimacion",ylab="Concentracion de
Mercurio",main="Estimacion Box-Cox vs. Concentracion de Mercurio")
abline(regresion,col="red",lwd=3)
text(80,1.2,"B1=-0.005302 B0=0.726140")
```



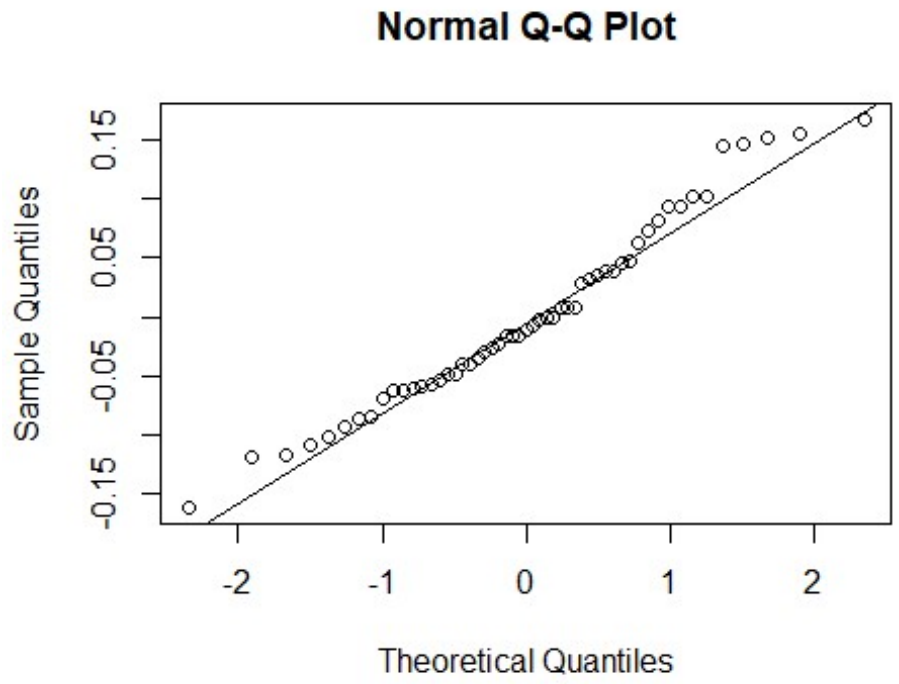

```
plot(regresion$fitted.values, regresion$residuals)
abline(h=0,col="green")
```



```
shapiro.test(regresion$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  regresion$residuals
## W = 0.97095, p-value = 0.2217

hist(regresion$residuals)
```



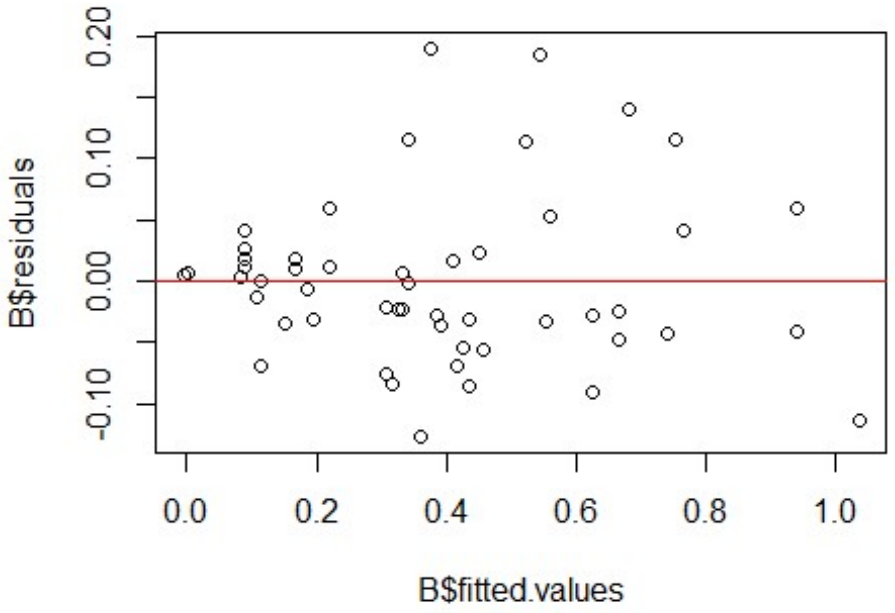

```
qqnorm(regresion$residuals)
qqline(regresion$residuals)
```



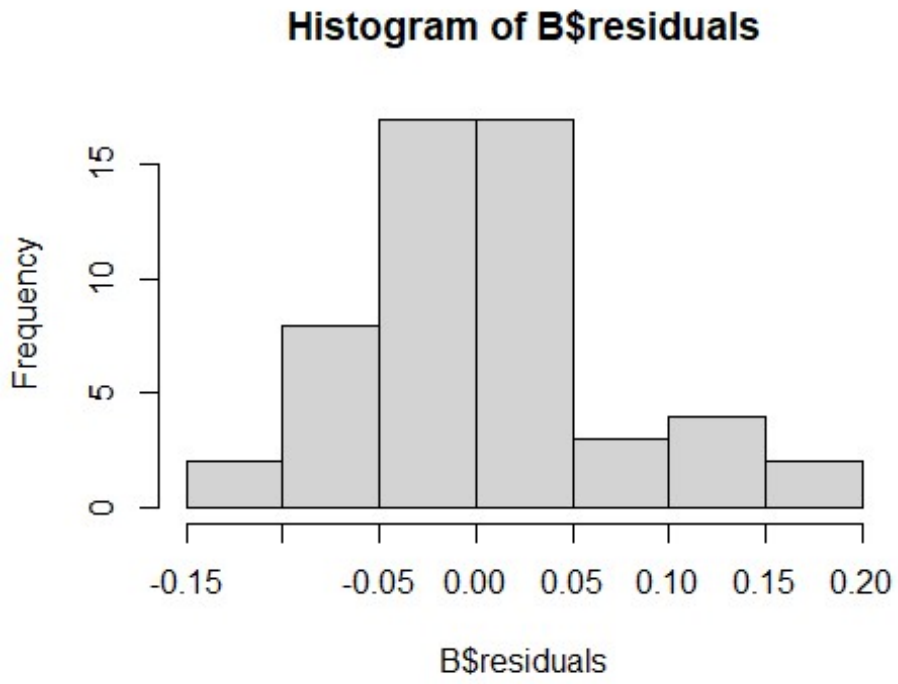

```
t.test(regresion$residuals)

##
##  One Sample t-test
##
## data:  regresion$residuals
## t = -3.2619e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.02156996  0.02156996
## sample estimates:
##     mean of x
## -3.506264e-18
```

```
z1=EstimacionBC
z2=EstimacionBC^2
cor.test(M1$X7 , z1+ z2)

##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and z1 + z2
## t = 26.242, df = 51, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.9396946 0.9796901
## sample estimates:
```

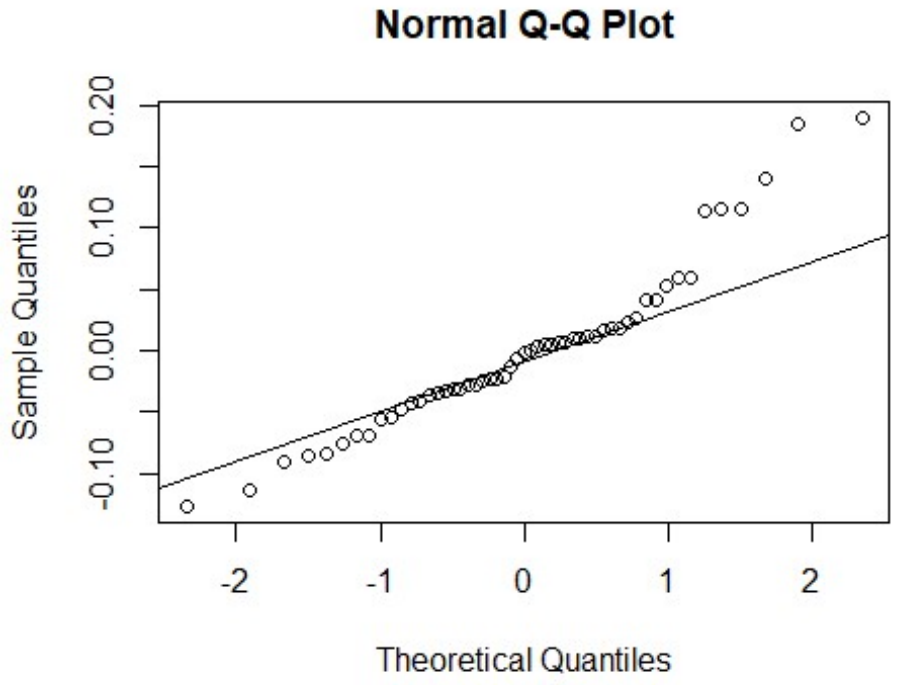
```
##       cor
## 0.9649079

B = lm(M1$X7 ~ z1+ z2)
summary(B)

##
## Call:
## lm(formula = M1$X7 ~ z1 + z2)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.126699 -0.036336 -0.001146  0.018686  0.189734
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.005741   0.030272  -0.190  0.85036
## z1           1.083784   0.307253   3.527  0.00091 ***
## z2           3.033206   0.709119   4.277 8.51e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06828 on 50 degrees of freedom
## Multiple R-squared:  0.9359, Adjusted R-squared:  0.9333
## F-statistic: 364.7 on 2 and 50 DF,  p-value: < 2.2e-16

plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```



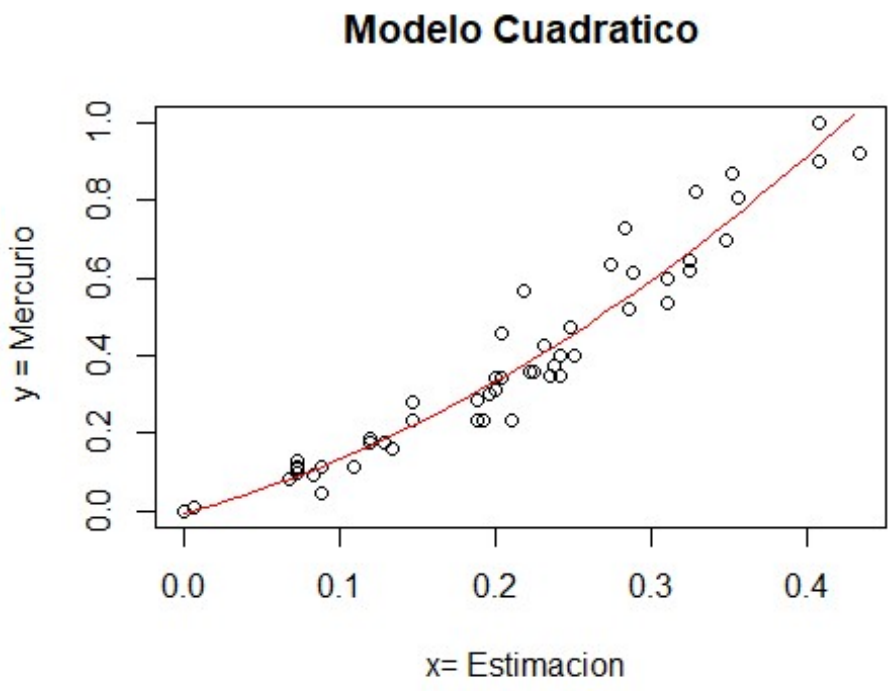

```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.93517, p-value = 0.006496

hist(B$residuals)
```




```
qqnorm(B$residuals)
qqline(B$residuals)
```



```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = -1.7447e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.01845591  0.01845591
## sample estimates:
##    mean of x
## -1.60468e-18

x=EstimacionBC
y=M1$X7
plot(x, y, main = "Modelo Cuadratico", xlab =" x= Estimacion", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = -0.005741  +1.083784*x1 +3.033206*x1^2
lines(x1, y1, col = "red")
text(1, .1, "y =-0.005741  +1.083784*x1 +3.033206*x1^2")
```



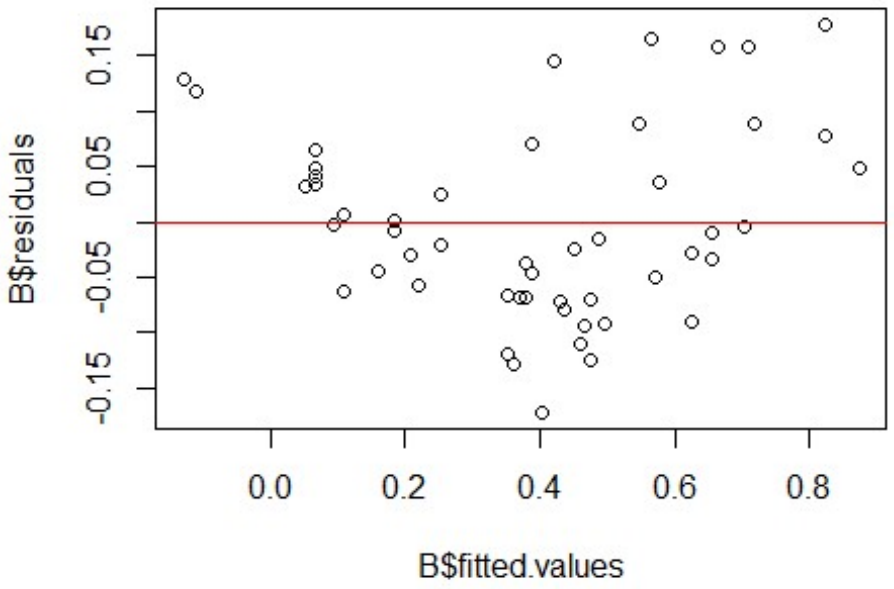

```
z1=EstimacionBC+1
z2=log(EstimacionBC+1)
cor.test(M1$X7 , log(EstimacionBC+1))
```

```
##
##  Pearson's product-moment correlation
##
## data:  M1$X7 and log(EstimacionBC + 1)
## t = 20.826, df = 51, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.9077255 0.9685800
## sample estimates:
##       cor
## 0.9459323
```

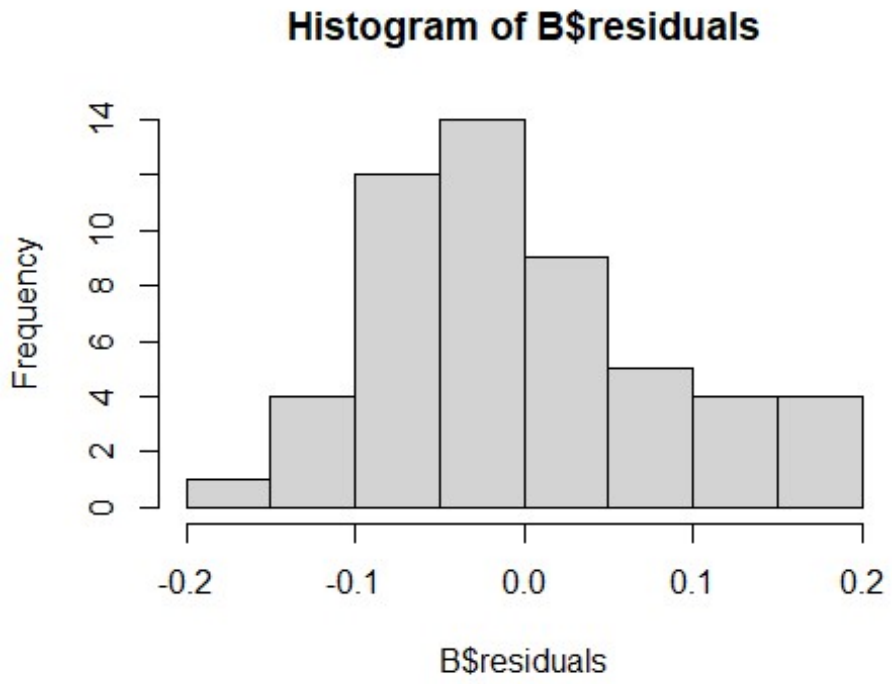
```
B = lm(M1$X7 ~ log(EstimacionBC+1))
summary(B)
```

```
##
## Call:
## lm(formula = M1$X7 ~ log(EstimacionBC + 1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.17195 -0.06586 -0.01040  0.04901  0.17794
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -0.12921    0.02709   -4.77 1.58e-05 ***
## log(EstimacionBC + 1)  2.78131    0.13355   20.83  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08659 on 51 degrees of freedom
## Multiple R-squared:  0.8948, Adjusted R-squared:  0.8927
## F-statistic: 433.7 on 1 and 51 DF,  p-value: < 2.2e-16
```

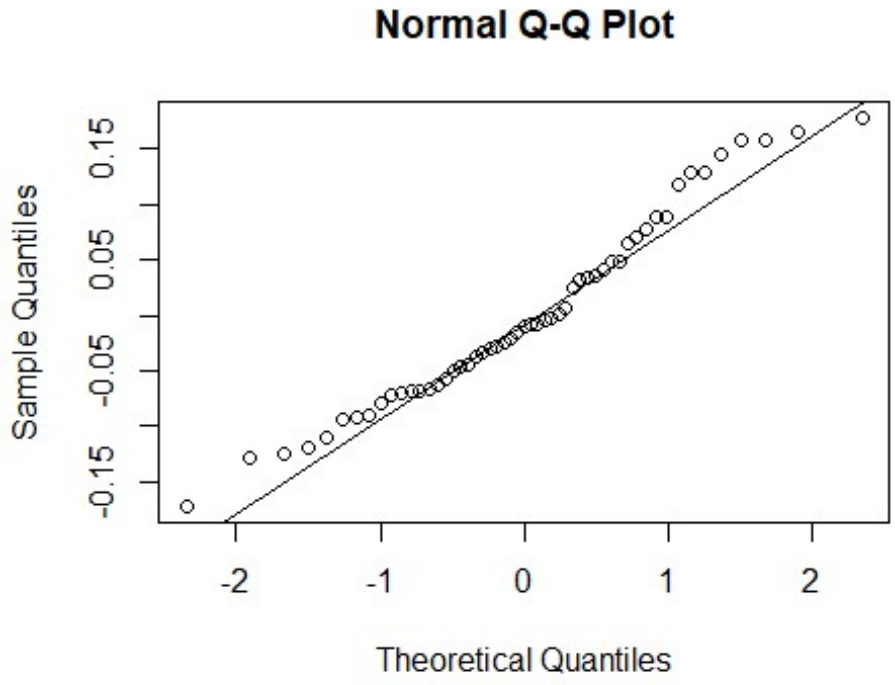
```
plot(B$fitted.values, B$residuals)
abline(h=0,col="red")
```



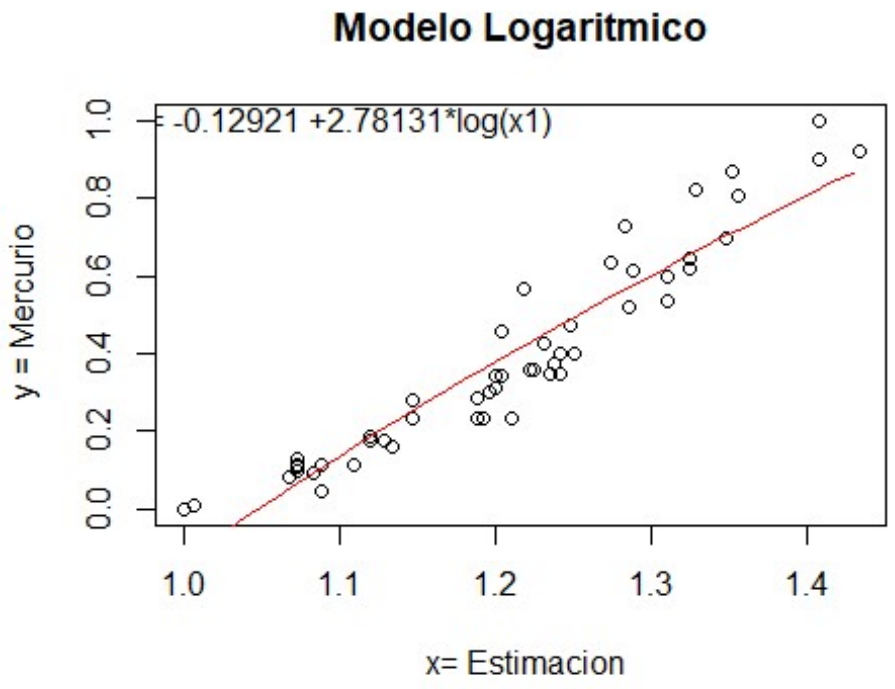

```
shapiro.test(B$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  B$residuals
## W = 0.9675, p-value = 0.1572

hist(B$residuals)
```




```
qqnorm(B$residuals)
qqline(B$residuals)
```

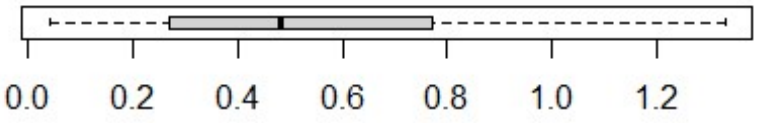


```
t.test(B$residuals)

##
##  One Sample t-test
##
## data:  B$residuals
## t = 2.8057e-16, df = 52, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.02363609  0.02363609
## sample estimates:
##    mean of x
## 3.304787e-18

x=EstimacionBC+1
y=M1$X7
plot(x, y, main = "Modelo Logaritmico", xlab =" x= Estimacion", ylab =" y =
Mercurio")
x1 = seq(min(x), max(x), 0.01)
y1 = -0.12921 +2.78131*log(x1)
lines(x1, y1, col = "red")
text(1.1, 1, "y = -0.12921 +2.78131*log(x1)")
```

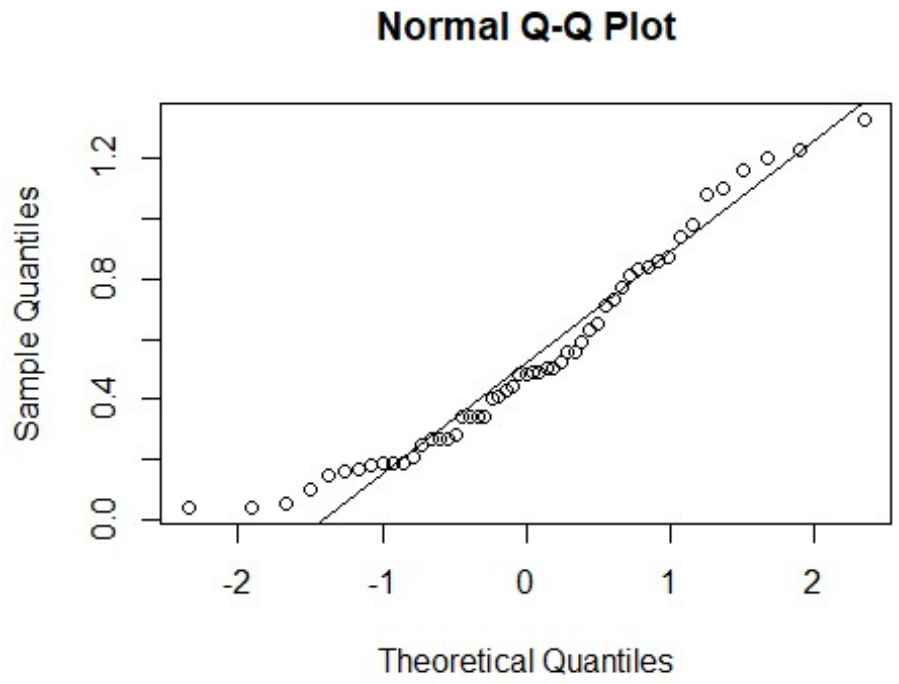



```
q1=quantile(M$X7,0.25)  #Cuantil 1 de la variable X
q2=quantile(M$X7,0.5)
q3=quantile(M$X7,0.75)
```

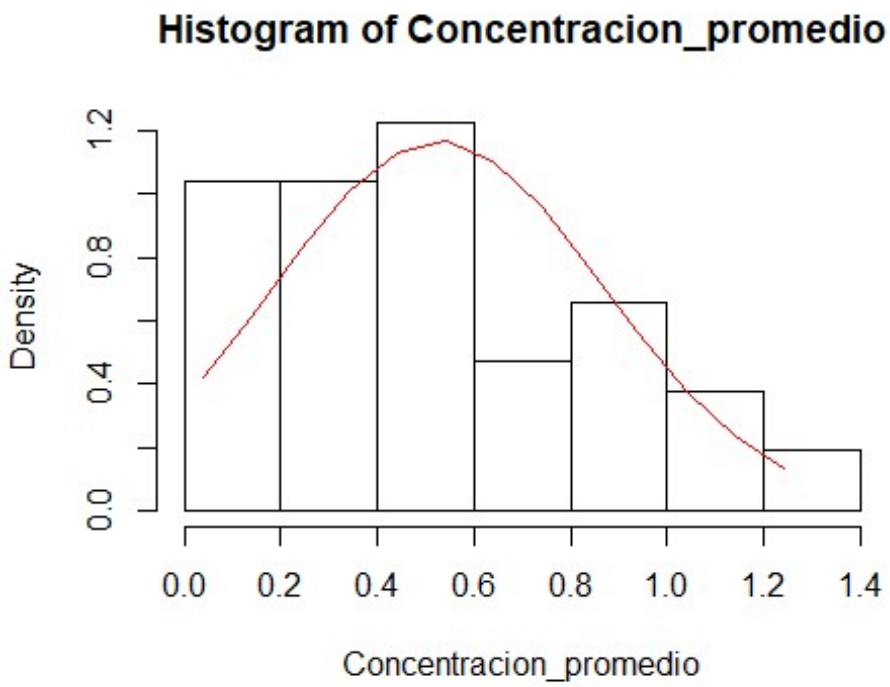
```r
y1=min(M$X7)
y2=max(M$X7)
ri=IQR(M$X7)    #Rango intercuartílico de X 2x1

par(mfrow=c(2,1))  #Matriz de gráficos de 2x1
boxplot(M$X7,horizontal=TRUE,ylim=c(y1,y2))
abline(v=q3+1.5*ri,col="red")  #linea vertical en el límite de los datos atípicos o extremos
```



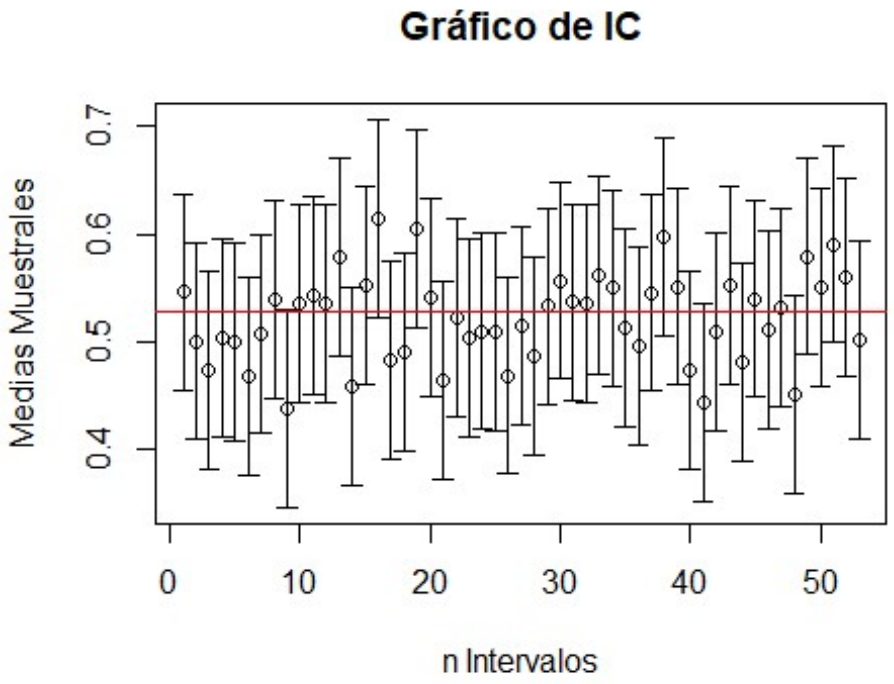

```r
Concentracion_promedio= M[M$X7<q3+1.5*ri,c("X7")]  #En la matriz M, quitar datos más allá de 3 rangos intercuartílicos arriba de q3 de la variable X

qqnorm(Concentracion_promedio)
qqline(Concentracion_promedio)
```




```
hist(Concentracion_promedio,prob=TRUE,col=0)
x=seq(min(Concentracion_promedio),max(Concentracion_promedio),0.1)
y=dnorm(x,mean(Concentracion_promedio),sd(Concentracion_promedio))
lines(x,y,col="red")
```




```
library(moments)
skewness(Concentracion_promedio)

## [1] 0.6159853

kurtosis(Concentracion_promedio)

## [1] 2.460721

library(plotrix)
n=length(Concentracion_promedio)
media=mean(Concentracion_promedio)
DE=sd(Concentracion_promedio)
alfa=0.05
ErrorEst=DE/sqrt(n)
X_=rnorm(n,media,ErrorEst)
E=abs(qnorm(alfa/2))*DE/sqrt(n)
plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=media,col="red")
```




```
cat("El promedio es de: ",media," +- ",E)
```

```
## El promedio es de:  0.5271698  +-  0.09181421
```

El valor promedio de concentracion media de mercurio en los peces es de 0.5271 ppm con un intervalo de confianza de 0.091 ppm haciendo uso de un 95% de confianza lo que nos determina que en los lagos con menor contentracion el valor no sobrepasa la medida propuesta como nociva para el consumo humano que es de 0.5 ppm pero en general los datos si llegan a sobrepasar este valor por lo que no recomendaria el consumir los alimentos recolectados en estos lagos en especifico.

```
age=M$X12
conc=M$X7
anova<-aov(conc ~ age)
summary(anova)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## age          1  0.072 0.07151    0.61  0.438
## Residuals   51  5.976 0.11718
```

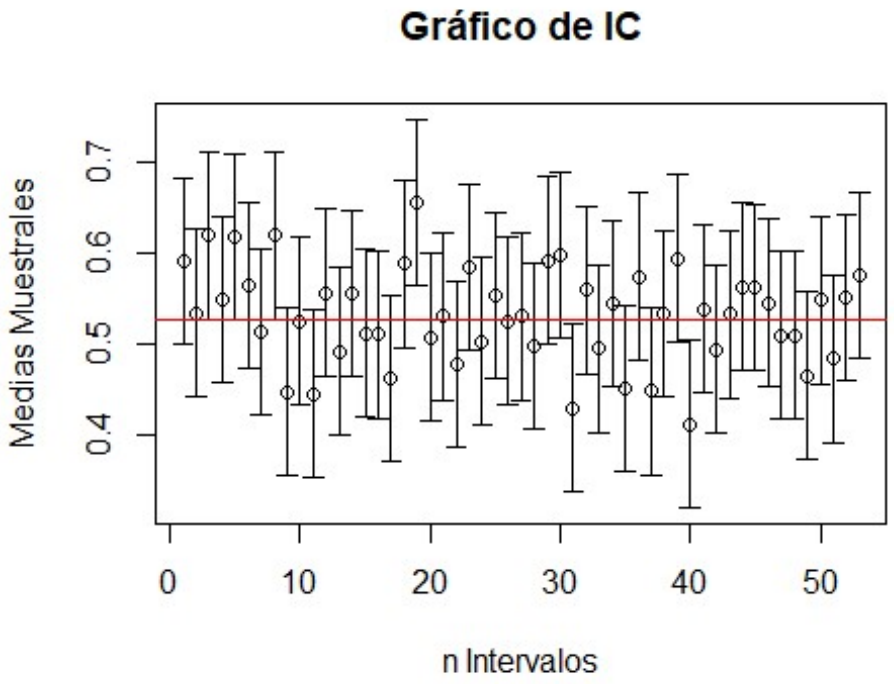
```
library(BSDA)
```

```
## Warning: package 'BSDA' was built under R version 4.1.3
```

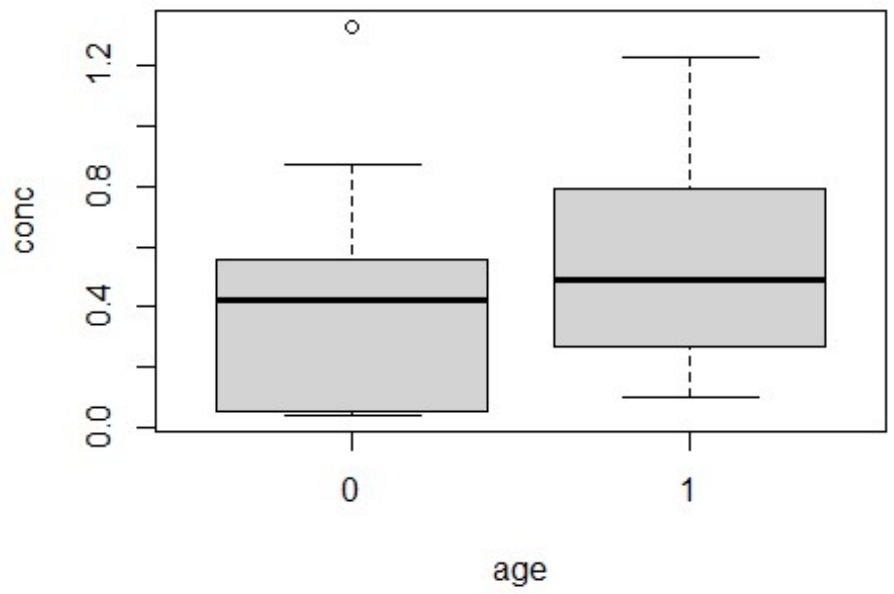
```
## Loading required package: lattice
```

```
##
## Attaching package: 'BSDA'
```

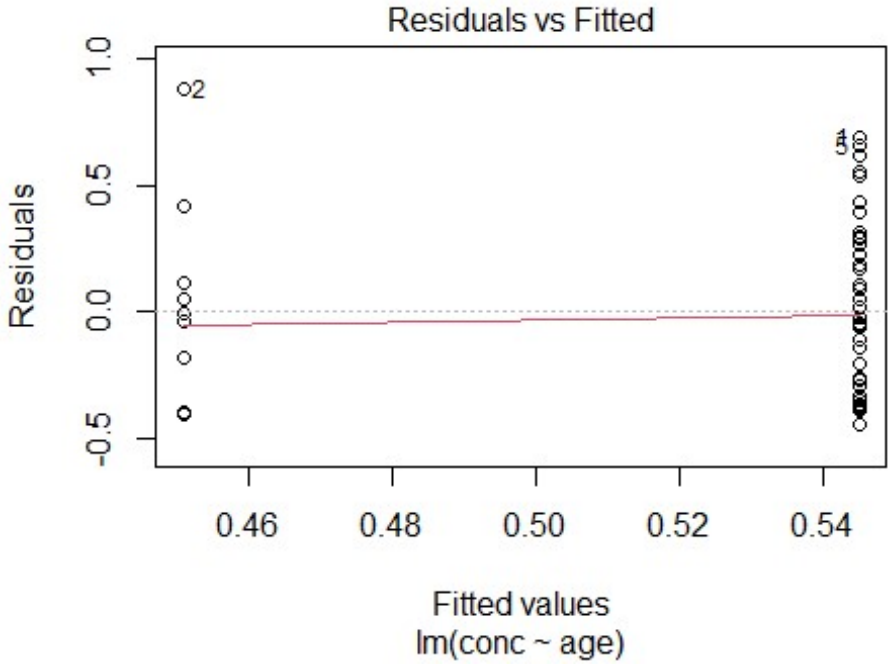
```
## The following object is masked from 'package:datasets':
##
##     Orange

library(plotrix)
tapply(conc,age)

##  [1] 2 1 1 1 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 1 2 2 2 2 2 1 2 2 2 1 1 2
2 2 1
## [39] 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2

M=mean(conc)
M

## [1] 0.5271698

SD=sd(conc)
alfa=0.05
n=length(conc)
ErrorEst=SD/sqrt(n)
X_=rnorm(n,M,ErrorEst)
E=abs(qnorm(alfa/2))*SD/sqrt(n)
z.test(conc,conf.level=0.95,sigma.x = SD)

##
##  One-sample z-Test
##
## data:  conc
## z = 11.254, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.4353556 0.6189840
## sample estimates:
## mean of x
## 0.5271698

plotCI(1:n,X_,E,main="Gráfico de IC",,xlab="n Intervalos",ylab="Medias
Muestrales")
abline(h=M,col="red")
```

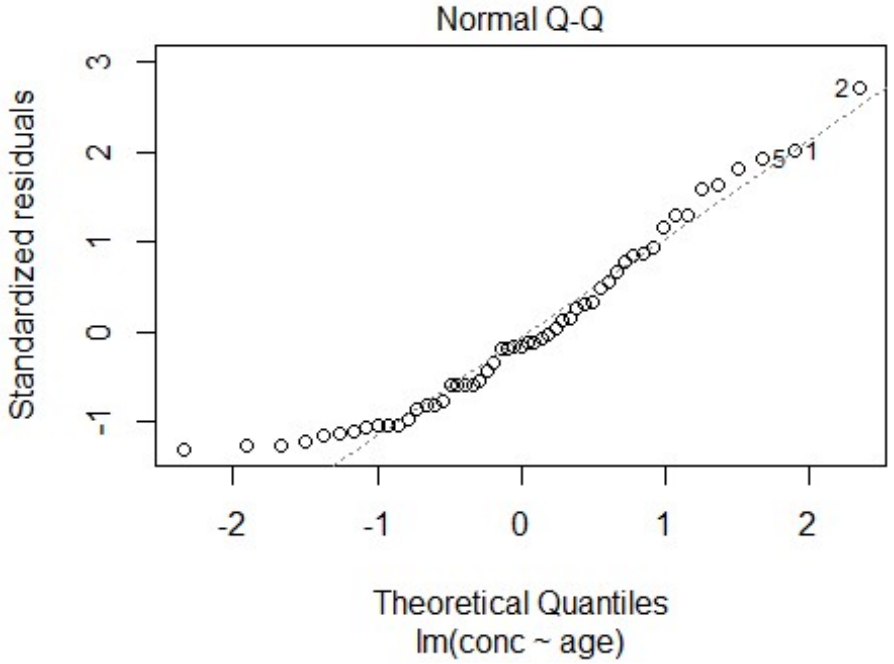


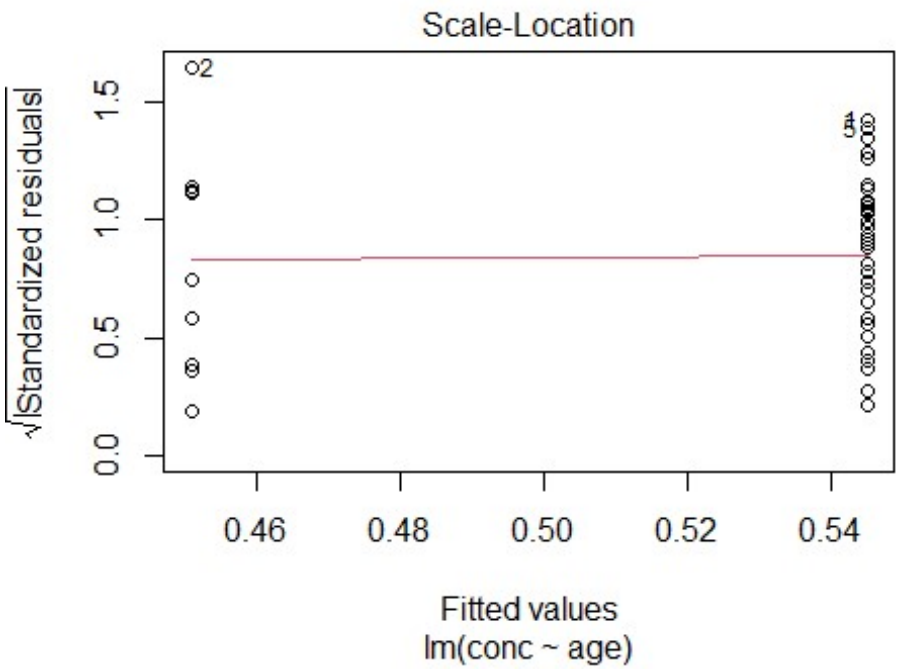

```
boxplot(conc ~ age)
```

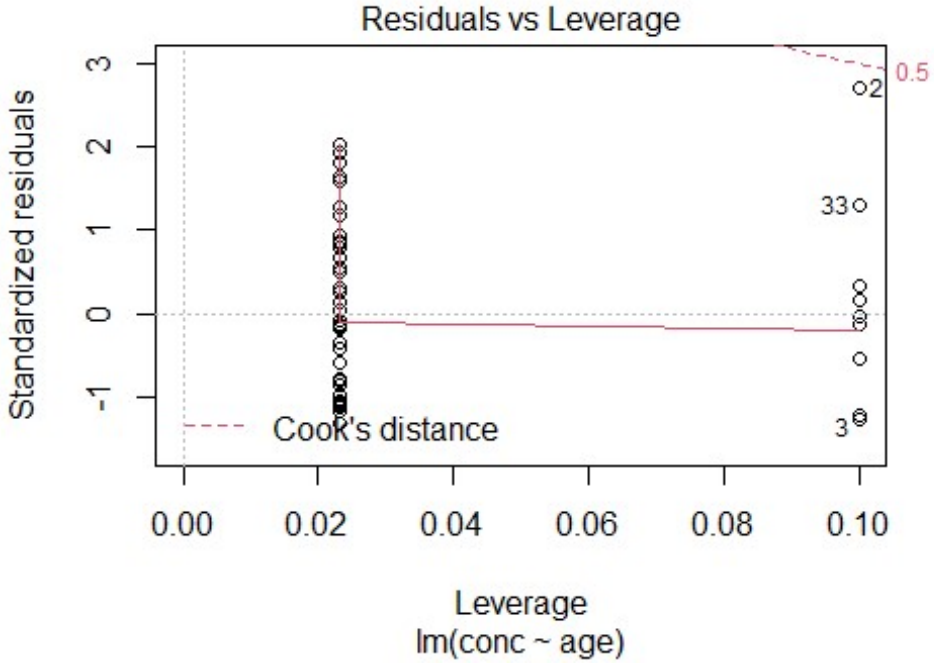



```
plot(lm(conc~age))
```


Residuals vs Fitted
Residuals
1.0
0.5
0.0
-0.5
2
0.46
0.48
0.50
0.52
0.54
Fitted values
lm(conc ~ age)


Normal Q-Q
Standardized residuals
3
2
1
0
-1
2
5
1
-2
-1
0
1
2
Theoretical Quantiles
lm(conc ~ age)

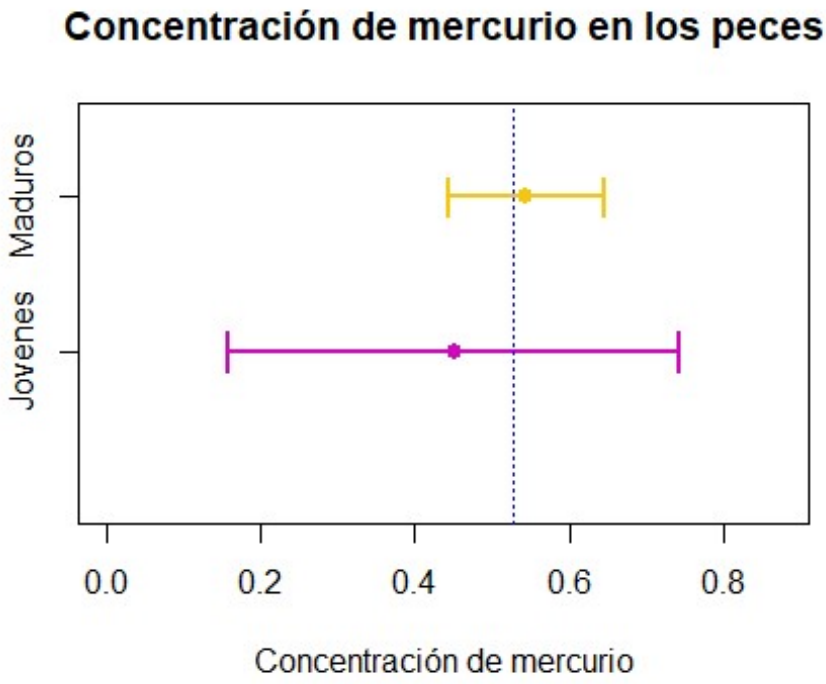





```
CD=  150/(150+76) #coeficiente de determinación para el modelo.
CD
```

```
## [1] 0.6637168
```

```r
## Análisis de cada tratamiento
m = tapply(conc, age, mean)
cat("Media de la concentración de mercurio por la edad de los peces:", m,
"\n")

## Media de la concentración de mercurio por la edad de los peces: 0.451
0.5448837

s = tapply(conc,age , sd)
cat("Desviación estandar de la concentración de mercurio por la edad de los
peces:", s, "\n")

## Desviación estandar de la concentración de mercurio por la edad de los
peces: 0.4076614 0.3266232

n = tapply(conc, age, length)
cat("Tamaño de la muestra de la concentración de mercurio por la edad de los
peces:", n, "\n")

## Tamaño de la muestra de la concentración de mercurio por la edad de los
peces: 10 43

## Intervalos de confianza
sm = s/sqrt(n)
E = abs(qt(0.025,n-1))*sm

In = m - E
cat("Intervalos de confianza inferiores:", In, "\n")

## Intervalos de confianza inferiores: 0.1593766 0.4443639

Sup = m + E
cat("Intervalos de confianza superiores:", Sup, "\n")

## Intervalos de confianza superiores: 0.7426234 0.6454035

plot(0, ylim=c(0,2.5), xlim=c(0,.87597), yaxt="n",
ylab="",xlab="Concentración de mercurio", main="Concentración de mercurio en
los peces")
axis(2,at=c(1:2),labels=c("Jovenes","Maduros"))

for(i in 1:2)
  {
arrows(In[i],i,Sup[i],i, angle=90, code=3, length = 0.1, lwd = 2,col=i+5)
points(m[i], i, pch=19, cex=1.1,col=i+5)
}

abline(v=mean(conc),lty=3,col="blue")
```



Basandonos en este modelo podemos llegar a la conclusion de que la edad del pez no afecta a la concentracion de mercurio en estos ya que todos los datos se encuentran dentro del mismo intervalo de confianza. Generando asi una distribucion similar.