Taller De Regresion Lineal Multiple

Econometria I 2019-02

Andres Felipe Sabogal Ramirez

30/5/2020

Table of Contents

# Algunas Formulas Importantes Para Comenzar

# Primer Punto

1. Se supone que la inversión es una protección contra la inflación si su precio y/o tasa de retorno, al menos, se mantiene al ritmo de la inflación. Para probar esta hipótesis, se decide ajustar los modelos

Modelo 1

Modelo 2

## Lectura de Datos

rm(list=ls()) # rm remueve objetos, ls = lista  
setwd('C://Users/f-pis/Desktop/Semestres/Archivos R') #Establece el workDirectory  
#install.packages("readxl") # instala el paquete para leer Excel   
library(readxl) #..library() Carga el paquete

## Warning: package 'readxl' was built under R version 3.6.3

library(normtest)  
Datos\_1P <- read\_excel("Datos\_E1\_Precio\_Oro\_BVNY.xls") #read\_excel lee el excel  
Datos\_1P=data.frame(Datos\_1P) # Aqui se estable que Tiempo es una hoja de Excel  
attach(Datos\_1P) # attach() carga los datos   
#install.packages("knitr")  
library(knitr)

## Warning: package 'knitr' was built under R version 3.6.3

knitr::kable(  
 head(Datos\_1P), caption = "Head Datos Primer Punto",align = "c")

Head Datos Primer Punto

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Poro | BVNY | IPC |
| 1974 | 59.26 | 511.354 | 49.3 |
| 1975 | 61.02 | 513.355 | 53.8 |
| 1976 | 24.84 | 522.585 | 56.9 |
| 1977 | 57.71 | 521.766 | 60.6 |
| 1978 | 93.22 | 521.781 | 65.2 |
| 1979 | 206.68 | 526.668 | 72.6 |

## Definimos los dos modelos de inversion

# el modelo 1 es Poro\_t= beta1 + beta2\*IPC\_t +U\_t  
  
Mod\_1<-lm(Datos\_1P$Poro~Datos\_1P$IPC)  
"Esta es la regresion del primer modelo"

## [1] "Esta es la regresion del primer modelo"

(summary(Mod\_1))

##   
## Call:  
## lm(formula = Datos\_1P$Poro ~ Datos\_1P$IPC)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -150.557 -90.711 5.178 37.052 310.944   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 116.8937 54.2597 2.154 0.0391 \*  
## Datos\_1P$IPC 1.0282 0.4022 2.556 0.0157 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 102.7 on 31 degrees of freedom  
## Multiple R-squared: 0.1741, Adjusted R-squared: 0.1474   
## F-statistic: 6.534 on 1 and 31 DF, p-value: 0.0157

plot(Mod\_1)  
e1<-residuals(Mod\_1)  
hist(e1)  
boxplot(e1, main="Boxplot residuals") # grafica de caja de los errores de estimacion (et)  
"Prueba de jarque-bera"

## [1] "Prueba de jarque-bera"

jb.norm.test(e1) # Prueba de normalidad Jarque-Bera

##   
## Jarque-Bera test for normality  
##   
## data: e1  
## JB = 5.5814, p-value = 0.03

"Intervalos de Confianza"

## [1] "Intervalos de Confianza"

(Intervalos\_1<-confint(Mod\_1))

## 2.5 % 97.5 %  
## (Intercept) 6.2302421 227.557098  
## Datos\_1P$IPC 0.2078211 1.848546

# El modelo 2 es BVNY\_t=beta1 + beta2\*IPC\_t + U\_t  
  
Mod\_2<-lm(Datos\_1P$BVNY~Datos\_1P$IPC)  
"Esta es la regresion del segundo modelo"

## [1] "Esta es la regresion del segundo modelo"

summary(Mod\_2)

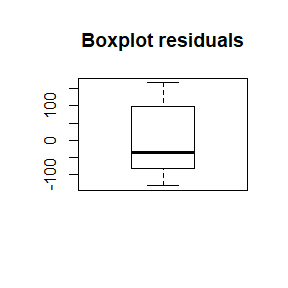
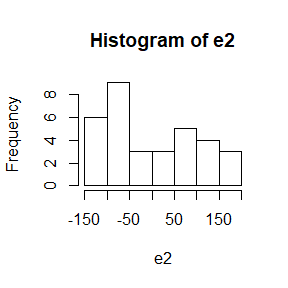
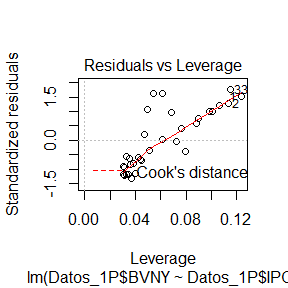
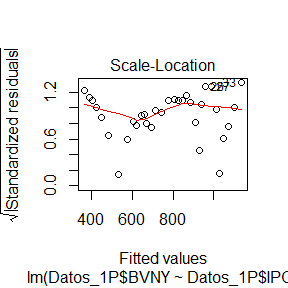
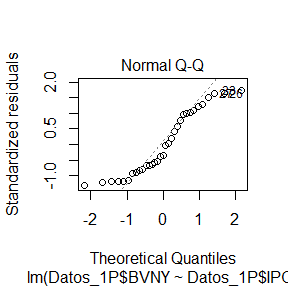
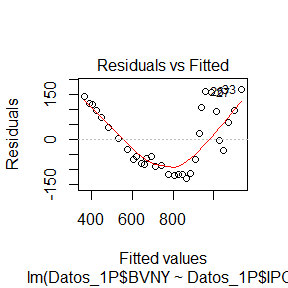
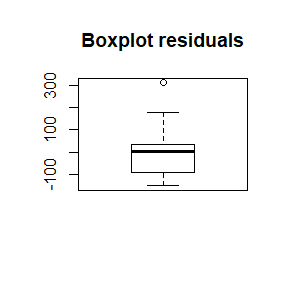
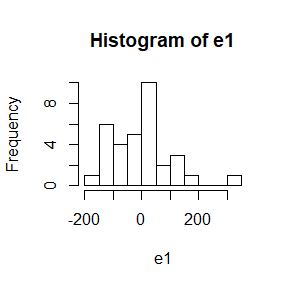
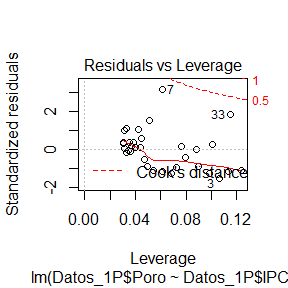
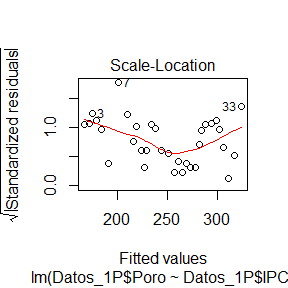
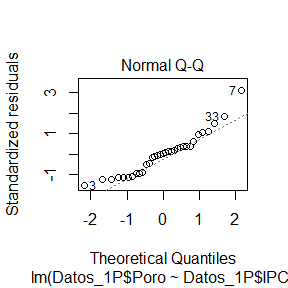
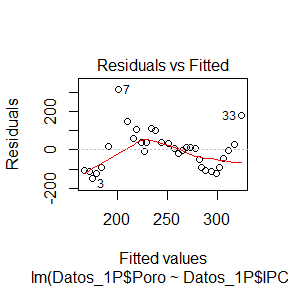
##   
## Call:  
## lm(formula = Datos\_1P$BVNY ~ Datos\_1P$IPC)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -132.20 -82.40 -34.44 96.46 166.31   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 120.5008 53.3967 2.257 0.0312 \*   
## Datos\_1P$IPC 5.0297 0.3958 12.707 7.9e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 101 on 31 degrees of freedom  
## Multiple R-squared: 0.8389, Adjusted R-squared: 0.8337   
## F-statistic: 161.5 on 1 and 31 DF, p-value: 7.895e-14

plot(Mod\_2)  
e2<-residuals(Mod\_2)  
hist(e2)  
boxplot(e2, main="Boxplot residuals") # grafica de caja de los errores de estimacion (et)  
jb.norm.test(e2) # Prueba de normalidad Jarque-Bera

##   
## Jarque-Bera test for normality  
##   
## data: e2  
## JB = 3.0839, p-value = 0.083

(Intervalos\_2<-confint(Mod\_2))

## 2.5 % 97.5 %  
## (Intercept) 11.597433 229.404209  
## Datos\_1P$IPC 4.222402 5.837033



Interpretacion: Se observa que enel modelo 1 tanto el intercepto Beta1, como la pendiente Beta2 son significaticas debido a p\_val<alpha, sin embargo se observa que el coeficiente de determinacion R\_cuadrado es muy bajo 0.1741 por lo que solo el 17,41% de la varianza del modelo esta siedo explicada, ademas de esto la prueba de normalidad de los errores Jarque -Bera arroja un p\_val<alpha 0.035 por lo que se puede pensar en la no normalidad en la distribucion de los errores Por otra partese observa que en el modelo dos Beta1 y Beta2 son significativos al 5% con un R\_cuadrado alto 0.8389 que explica el 83.39% de la varianza del modelo, y ademas la prueba J-B arroja un P\_val>alpha 0.0855 por lo que rechazamos Ho y inferimos normalidad en la distribucion

Como conclusion se obsrva que el segundo modelo es mejor ya que tiene un R2 cercano a uno y en su grafica de cuantiles se observa que se ajusta a la distribucion normal, sin embargo en la grafica Residuals Vs fitted del modelo 2 se observa que puede que el modelo no sea lineal.

# SEGUNDO PUNTO

Se formula el siguiente modelo para estudiar la relación entre X e Y: Se quiere estimar el modelo:

##REGRESION GASTO-INGRESOS POR HOGARES

setwd('C://Users/f-pis/Desktop/Semestres/Archivos R')  
Datos\_2P<-read\_excel("Datos\_E2\_Gastos\_Ingreso\_Hogares.xls")  
Datos\_2P=data.frame(Datos\_2P)  
  
library(knitr)  
knitr::kable(  
 head(Datos\_2P), caption = "Head Datos Segundo Punto",align = "c")

Head Datos Segundo Punto

|  |  |  |
| --- | --- | --- |
| Hogar | Y | X |
| 1 | 9.46 | 25.83 |
| 2 | 10.56 | 34.31 |
| 3 | 14.81 | 42.50 |
| 4 | 21.71 | 46.75 |
| 5 | 22.79 | 48.29 |
| 6 | 18.19 | 48.77 |

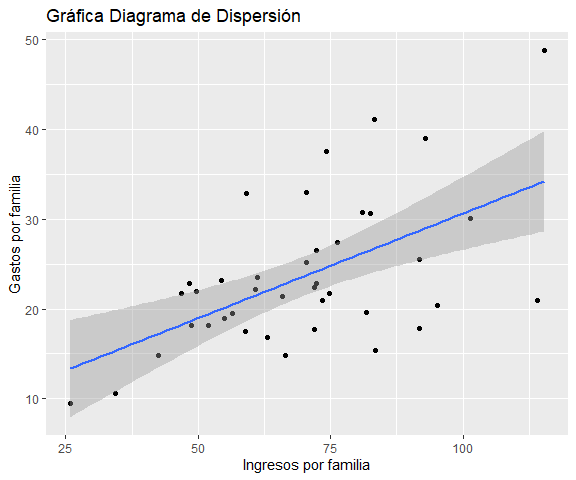
# Se busca estudiar la relacion que existe entre X e Y  
# Se quiere estimar el modelo Y\_i= beta\_1+beta\_2\*X\_i + U\_i con i=1.....40  
  
#install.packages("ggplot2")  
  
if (!require('ggplot2'))   
{  
 install.packages('ggplot2');  
 library(ggplot2);  
}

## Loading required package: ggplot2

## Warning: package 'ggplot2' was built under R version 3.6.3

library(ggplot2)  
ggplot(Datos\_2P, aes(x=X, y=Y)) + geom\_point() + ggtitle("Gráfica Diagrama de Dispersión") + xlab("Ingresos por familia") + ylab("Gastos por familia") + geom\_smooth(method=lm)

## `geom\_smooth()` using formula 'y ~ x'

 En la grafica de dispersion de ingresos Vs Gastos se puede observar una relacion lineal representada por la linea azul

## Estimacion Por Metodo Matricial OLS

X\_2P <- as.matrix(cbind(1,Datos\_2P[,3]))  
Y\_2P <- as.matrix(Datos\_2P[,2])  
#Reg2=lm(Y\_2P~Datos\_3[,3:5])  
  
XtX\_2 = t(X\_2P)%\*%X\_2P # Este es el resultado de multiplicar X traspuesta por X   
XtX\_2 # X Traspuesta por X (Xt\*X)

## [,1] [,2]  
## [1,] 40 2792.0  
## [2,] 2792 210206.2

XtX\_inv\_2 <- solve(XtX\_2) # LA FUNCION solve() halla la INVERSA de la matriz   
XtX\_inv\_2

## [,1] [,2]  
## [1,] 0.342922190 -4.554759e-03  
## [2,] -0.004554759 6.525443e-05

Xty\_2 <- t(X\_2P)%\*%Y\_2P # Resultado de X traspuesta por y ...   
Xty\_2

## [,1]  
## [1,] 943.48  
## [2,] 69417.39

b\_2 <- XtX\_inv\_2%\*%Xty\_2 # Resultado de los estimadores b = (Xt\*X)inv\*Xt\*y  
"Estos son los coeficientes estimados OLS"

## [1] "Estos son los coeficientes estimados OLS"

b\_2

## [,1]  
## [1,] 7.3607278  
## [2,] 0.2324681

Ye\_2 <- X\_2P%\*%b\_2 #Calcula yt estimado  
"Estos son los Y estimados por el modelo "

## [1] "Estos son los Y estimados por el modelo "

#OJO estos son los resultados estimados de y de acuerdo a los valores encontrados de los coeficientes   
plot(Y\_2P,Ye\_2, main = "Grafica Y Vs Y\_estimado", ) # Se grafica Y de los datos iniciales (y) contra Y estimado (yte)  
e\_2 <- Y\_2P - Ye\_2 # et es el vector de errores de estimacion (Ydado-Yestimado)   
(summary(e\_2)) # Aqui se muestra en consola un summary de los errores de estimacion (et)

## V1   
## Min. :-12.992   
## 1st Qu.: -3.606   
## Median : -1.071   
## Mean : 0.000   
## 3rd Qu.: 3.225   
## Max. : 14.509

hist(e\_2, prob=TRUE,main="Histogram e") #Histograma de los errores de estimacion   
curve(dnorm(x,mean = 0,sd=1),col=1,lty=2,lwd=3,add = TRUE) # Agrega una curva normal con media en cero y desviacion estandar 1   
boxplot(e\_2,main="Boxplot e") # grafica de caja de los errores de estimacion (et)  
"Prueba Jarque-bera sobre los errores"

## [1] "Prueba Jarque-bera sobre los errores"

jb.norm.test(e\_2) # Prueba de normalidad Jarque-Bera

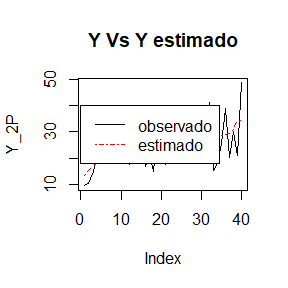
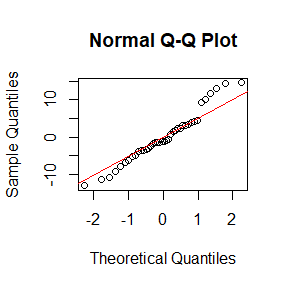
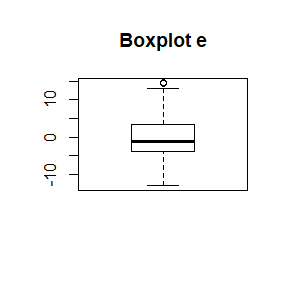
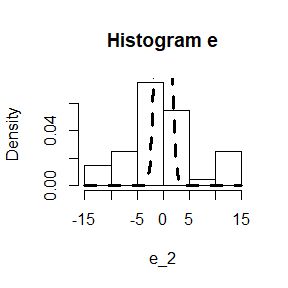
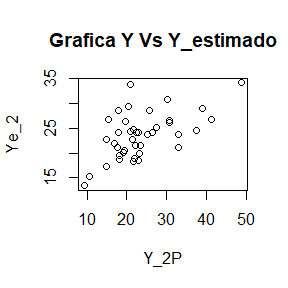
##   
## Jarque-Bera test for normality  
##   
## data: e\_2  
## JB = 1.0813, p-value = 0.4675

g1\_qq <- qqnorm(e\_2,col="black") # Genera una Q-Q plot que es un metodo para comparar distrubuciones de probabilidad de una muestra aleatoria y una de comparacion   
qqline(e\_2,col="red") # qqline agrega a la grafica anterior una linea de tendencia teorica de cuantiles de una distribucion normal   
  
# OTRA FORMA PARA YT vs YTE  
# De esta manera de grafica y y se agrega la linea LINES de yte estimado   
plot(Y\_2P,t="l",main="Y Vs Y estimado") # "l" significa linea, p significa puntos  
lines(Ye\_2,lty=4,col=2)  
legend(x=0,y=40,legend = c("observado","estimado"),col = c(1,2),lty = c(1,4) ) #"lty" significa 1=linea, 4= punto y linea   
  
"Estimacion con lm"

## [1] "Estimacion con lm"

Reg2lm=lm(X\_2P~Y\_2P)  
(summary(Reg2lm))

## Response Y1 :  
##   
## Call:  
## lm(formula = Y1 ~ Y\_2P)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -6.940e-15 5.500e-18 2.567e-16 3.933e-16 7.376e-16   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.000e+00 5.796e-16 1.725e+15 <2e-16 \*\*\*  
## Y\_2P 4.180e-17 2.325e-17 1.798e+00 0.0801 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.188e-15 on 38 degrees of freedom  
## Multiple R-squared: 0.4964, Adjusted R-squared: 0.4831   
## F-statistic: 37.45 on 1 and 38 DF, p-value: 3.902e-07  
##   
##   
## Response Y2 :  
##   
## Call:  
## lm(formula = Y2 ~ Y\_2P)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -24.692 -12.585 -1.147 7.793 47.997   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.6125 8.0971 4.645 4e-05 \*\*\*  
## Y\_2P 1.3646 0.3248 4.202 0.000154 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 16.59 on 38 degrees of freedom  
## Multiple R-squared: 0.3172, Adjusted R-squared: 0.2993   
## F-statistic: 17.66 on 1 and 38 DF, p-value: 0.0001544



Se observa que de los estimadores OLS del modelo tenemos un intercepto de 7.3607278 y una pendiente 0.2324681, siendo la pendiente Beta 2 se interpreta como la variacion de los gastos promedio semanales en dólares para hogares de tres personas (Y) freste a una variacion en los ingresos por lo que el signo es positivo, es el esperado ya que frente a un aumento en el ingreso se espera un aumento en el gasto. El test de jarque-Bera da P\_val>alpha 0,467 por lo que se acepta la Ho y se afirma la normalidad de los errores,

## Calculo de las sumas de cuadrados TSS, ESS y RSS

TSS\_2=t(Y\_2P)%\*%Y\_2P-length(Y\_2P)\*mean(Y\_2P)^2 # Suma de los cuadrados totales (y-ybarra)\*\*2  
#.........Tss= SUM Y\*\*2-n\*Y\_barra\*\*2  
"Este es el TSS"

## [1] "Este es el TSS"

TSS\_2

## [,1]  
## [1,] 2610.594

ESS\_2=t(b\_2)%\*% XtX\_2 %\*%b\_2-length(Y\_2P)\*mean(Y\_2P)^2 # Suma de los cuadrados explicados (yest-ybarra)\*\*2 .....ESS= b'\*(X'X)\*b-n\*Y\_barra\*\*2  
"Este es el ESS"

## [1] "Este es el ESS"

ESS\_2

## [,1]  
## [1,] 828.1646

RSS\_2=t(e\_2)%\*%e\_2 # suma de cuadrados residuales (yi-yest)\*\*2 ......(et\_traspuesto\*et)  
"Este es el RSS"

## [1] "Este es el RSS"

RSS\_2

## [,1]  
## [1,] 1782.43

TSS\_e16\_2=ESS\_2+RSS\_2 # las uma de cuadrados totales es igual al RSS + ESS  
"Esta es la suma TSS=ESS+RSS"

## [1] "Esta es la suma TSS=ESS+RSS"

TSS\_e16\_2

## [,1]  
## [1,] 2610.594

RC\_2<-ESS\_2[1,1]/TSS\_2[1,1] # calculo del coeficiente de determinacion R\_Cuadrado  
"Este es el coeficiente de determinacion R^2"

## [1] "Este es el coeficiente de determinacion R^2"

RC\_2

## [1] 0.3172322

Luego el coeficiente de determinacion R2 es bajo y explica el 31.72% de la varianza del modelo

## Estimacion De La Varianza y Pruebas De Significancia Individual

# Recordar que b~N(beta,Sigma\*\*2(X'X)Inv) Donde b es el vector de estimadores hallados   
  
# Recordar que Var b\_estimado=s\*\*2\*(X'X)inv  
  
S2\_2<- RSS\_2[1,1]/38 # Aqui se hala la varianza S\*\*2estimado=RSS/(n-k)=et\*\*2/(n-k) ES UN ESCALAR  
Var\_b2=S2\_2\*XtX\_inv\_2 # Matriz de varianzas y covarianzas \_\_\_Los elementos de la diagonal de la matriz contienen las varianzas de las variables, mientras que los elementos que se encuentran fuera de la diagonal contienen las covarianzas entre todos los pares posibles de variables.  
S2b1\_2=Var\_b2[1,1]  
S2b2\_2=Var\_b2[2,2]  
  
sb1\_2=sqrt(S2b1\_2)  
sb2\_2=sqrt(S2b2\_2)  
  
colum1<-c("Coef","b1","b2")  
colum2<-c("Estimación",b\_2[1,1], b\_2[2,1])  
colum3<-c("Varianza",S2b1\_2, S2b2\_2)  
colum4<-c("Error\_Estandar",sb1\_2, sb2\_2)  
Tabla\_Coef\_Var\_2 <- cbind(colum1,colum2,colum3,colum4)  
kable(Tabla\_Coef\_Var\_2,caption = "Coeficientes y Varianzas ",align = "c")

Coeficientes y Varianzas

|  |  |  |  |
| --- | --- | --- | --- |
| colum1 | colum2 | colum3 | colum4 |
| Coef | Estimación | Varianza | Error\_Estandar |
| b1 | 7.36072783063958 | 16.0851229232876 | 4.0106262507603 |
| b2 | 0.232468082655594 | 0.00306082705693042 | 0.0553247418153074 |

""

## [1] ""

"Pruebas de Significancia Individual"

## [1] "Pruebas de Significancia Individual"

"|Ho: b\_j=0"

## [1] "|Ho: b\_j=0"

"|Ha: b\_j=/= 0"

## [1] "|Ha: b\_j=/= 0"

# Recordar que t\_sub\_b1=b1/Se(b1)=t\_sub\_c Donde Se(b1) es la desviacion estandar de b1 sqrt(var(b1))  
  
# Recordar que P\_Val= 2\*Prob[t\_sub\_n-k > abs(t\_sub\_b1)]  
  
# Aqui se hacen las pruebas de significancia individual de cada estimador   
tb1\_2=b\_2[1,1]/sb1\_2 # VALORES t   
pb1\_2=2\*(1-pt(abs(tb1\_2),38)) # IMPORTANTE con esta funcion pt() obtenemos la probabilidad de una distribucion t-Student con 47 grados de libertad VALORES p   
tb2\_2=b\_2[2,1]/sb2\_2  
pb2\_2=2\*(1-pt(abs(tb2\_2),38))  
  
colum1<-c("Coef","b1","b2")  
colum2<-c("Estimación",b\_2[1,1], b\_2[2,1])  
colum4<-c("Error\_Estandar",sb1\_2, sb2\_2)  
colum5<-c("Valores t",tb1\_2, tb2\_2)  
colum6<-c("Valores p",pb1\_2, pb2\_2)  
Tabla\_analisis\_reg\_2 <- cbind(colum1,colum2, colum4,colum5,colum6)  
kable(Tabla\_analisis\_reg\_2,caption = "Coeficientes y Varianzas ",align = "c")

Coeficientes y Varianzas

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| colum1 | colum2 | colum4 | colum5 | colum6 |
| Coef | Estimación | Error\_Estandar | Valores t | Valores p |
| b1 | 7.36072783063958 | 4.0106262507603 | 1.83530635128222 | 0.0742962908246338 |
| b2 | 0.232468082655594 | 0.0553247418153074 | 4.20188282905415 | 0.000154435397409225 |

## Modelo De Regresion Lineal En Desviaciones De La Media

I40 <- diag(40) # diag() nos arroja una matriz diagonal cuadrada con unos en la diagonal  
i\_2 <- rep(1,40) # Genera un vector de unos de tamaño 52  
iit\_2 <- i\_2%\*% t(i\_2) # Multiplica y da coo resultado una matriz i de unos 1 de orden 52len  
#vi1 <- matrix(nrow=52, ncol=1, rep(1, 52)) # Vector de unos   
#Miit1 <- matrix(nrow=52, ncol=52, rep(1,52\*52)) # Matriz de unos   
A <- I40-(1/40)\*iit\_2 # Matriz 52X52 donde a la diagonal principal se le resta 1/52 y el resto de entredas seran 0-1/52   
#X21 <- X0[1:52,2:5] # Define una matriz de 52x4  
Xa\_2 <- A%\*%X\_2P[,2] # Resulta una matriz de 40x1 A\*X2=A\*X21  
ya\_2 <- A%\*%Y\_2P[,1] # ya es un vector de 52 X 1 resultado de A\*y  
XatXa\_2 <- t(Xa\_2)%\*%Xa\_2   
XatXa\_2

## [,1]  
## [1,] 15324.63

XatXa\_2\_inv <- solve(XatXa\_2)  
XatXa\_2\_inv

## [,1]  
## [1,] 6.525443e-05

Xatya\_2 <- t(Xa\_2)%\*%ya\_2  
Xatya\_2

## [,1]  
## [1,] 3562.487

b2\_desv <- XatXa\_2\_inv%\*%Xatya\_2  
"b2\_desv es el vector de parametros estimados en en modelo en desviaciones"

## [1] "b2\_desv es el vector de parametros estimados en en modelo en desviaciones"

kable(b2\_desv, caption = "Parametros Estimados Mod-Desv",align = "c") # b2 es el vector de parametros estimados

## Warning in kable\_markdown(x, padding = padding, ...): The table should have a  
## header (column names)

Parametros Estimados Mod-Desv

|  |
| --- |
| 0.2324681 |

# OBTENEMOS EL INTERCEPTO  
# Ybarra-b2\*X2barra-b3\*X3barra-b4\*X4barra-b5\*X5barra  
Intercepto\_2=mean(Y\_2P)-b2\_desv[1,1]\*mean(X\_2P[,2])  
"Intercepto estimado" # Entre comillas --muestra el mensaje en la consola

## [1] "Intercepto estimado"

Intercepto\_2

## [1] 7.360728

yde\_2 <- Xa\_2%\*%b2\_desv # Vector de Y estimado   
ed\_2 <- ya\_2-yde\_2 # Vector de errores de estimacion   
TSSa\_2=t(ya\_2)%\*%ya\_2 # Suma de cuadrados totales   
"Suma total de cuadrados del modelo en desviaciones"

## [1] "Suma total de cuadrados del modelo en desviaciones"

TSSa\_2

## [,1]  
## [1,] 2610.594

ESSa\_2=t(b2\_desv)%\*% XatXa\_2 %\*%b2\_desv # SUma de cuadrados Estimados b2t\*XatXa\*b2  
"Suma de cuadrados explicados del modelo en desviaciones"

## [1] "Suma de cuadrados explicados del modelo en desviaciones"

ESSa\_2

## [,1]  
## [1,] 828.1646

RSSa\_2=t(ed\_2)%\*%ed\_2 # Suma de cuadrados residuales eta\*\*2  
"Suma de cuadrados residuales en mod de desviaciones"

## [1] "Suma de cuadrados residuales en mod de desviaciones"

RSSa\_2

## [,1]  
## [1,] 1782.43

c=c("Modelo en desviacioes")  
c1=c("TSS","ESS","RSS")  
c2=c(TSSa\_2,ESSa\_2,RSSa\_2)  
Sum\_cuad\_desv\_2=cbind(c,c1,c2)  
kable(Sum\_cuad\_desv\_2,caption = "Sumas de Cuadrados Mod-Desv",align = "c")

Sumas de Cuadrados Mod-Desv

|  |  |  |
| --- | --- | --- |
| c | c1 | c2 |
| Modelo en desviacioes | TSS | 2610.59424 |
| Modelo en desviacioes | ESS | 828.164615362709 |
| Modelo en desviacioes | RSS | 1782.42962463729 |

"TSS=ESS+RSS"

## [1] "TSS=ESS+RSS"

(TAAa\_2s=ESSa\_2+RSSa\_2)

## [,1]  
## [1,] 2610.594

"Prueba de normalidad Jarque-Bera"

## [1] "Prueba de normalidad Jarque-Bera"

jb.norm.test(ed\_2)

##   
## Jarque-Bera test for normality  
##   
## data: ed\_2  
## JB = 1.0813, p-value = 0.4735

Se puede determinar por la prueba de jarque bera que los errores del modelo 2 estimado se distribuyen de manera normal y

# Tercer Punto

Se tienen datos son anuales en el periodo 1947 – 2000 sobre las siguientes variables para la economía de US: C1 = Gastos de consumo real, Yd =Ingreso Real, Ri = Riqueza, y TI = Tasa de interés

Modelo P3

## Estimacion del modelo Beta2=Variacion enlos gastos con respecto a variacion en el ingreso real Beta3=Variacion en los gastos con respecto a una variacion en la riqueza disponible Beta4=Varacion en el gasto con respecto a una variacion en la tasa de interes Beta1=Intercepto en y= gasto autonomo

"Estimar el modelo C1\_t=B1+B2\*Yd\_t+B3\*RI\_t+B4\*TI\_t+U\_t"

## [1] "Estimar el modelo C1\_t=B1+B2\*Yd\_t+B3\*RI\_t+B4\*TI\_t+U\_t"

rm(list=ls()) # rm remueve objetos, ls = lista  
setwd('C://Users/f-pis/Desktop/Semestres/Archivos R') #Establece el workDirectory  
#install.packages("readxl") # instala el paquete para leer Excel   
library(readxl) #..library() Carga el paquete   
Datos\_3P<-read\_excel("Datos\_E3\_Consumo.xls") #read\_excel lee el excel  
Datos\_3P=data.frame(Datos\_3P)# Aqui se estable que Tiempo es una hoja de Excel  
Datos\_3P$c1<-as.numeric(Datos\_3P$c1)  
Datos\_3P$yd<-as.numeric(Datos\_3P$yd)  
Datos\_3P$Ri<-as.numeric(Datos\_3P$Ri)  
Datos\_3P$ti<-as.numeric(Datos\_3P$ti)  
attach(Datos\_3P) # attach() carga los datos

## The following object is masked from Datos\_1P:  
##   
## Date

kable(head(Datos\_3P),caption = "Head Datos 3P ",align = "c")

Head Datos 3P

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Date | c1 | yd | Ri | ti |
| 1947 | 975.5 | 1035.2 | 5166.8 | -10351.000 |
| 1948 | 998.1 | 1090.0 | 5280.8 | -4720.000 |
| 1949 | 1025.3 | 1095.6 | 5607.4 | 1044.000 |
| 1950 | 1090.9 | 1192.7 | 5759.5 | 0.407 |
| 1951 | 1107.1 | 1227.0 | 6086.1 | -5283.000 |
| 1952 | 1142.4 | 1266.8 | 6243.9 | -0.277 |

library(normtest) # carga el paquete de prueba de normalidad   
  
X\_3P <- as.matrix(cbind(1,Datos\_3P[,3:5]))  
Y\_3P <- as.matrix(Datos\_3P[,2])  
XtX\_3=t(X\_3P)%\*%X\_3P # Este es el resultado de multiplicar X traspuesta por X   
XtX\_3 # X Traspuesta por X..... (X'\*X)

## 1 yd Ri ti  
## 1 54.00 173654.4 833684 6.703646e+04  
## yd 173654.40 699882672.2 3423310436 3.479660e+08  
## Ri 833684.00 3423310436.2 16999040180 1.711348e+09  
## ti 67036.46 347965965.4 1711348276 4.758595e+08

XtX\_inv\_3 <- solve(XtX\_3) # LA FUNCION solve() halla la INVERSA de la matriz   
XtX\_inv\_3

## 1 yd Ri ti  
## 1 1.138203e-01 -6.480331e-05 6.758945e-06 7.044779e-06  
## yd -6.480331e-05 1.325795e-07 -2.301159e-08 -5.060596e-09  
## Ri 6.758945e-06 -2.301159e-08 4.331587e-09 2.969349e-10  
## ti 7.044779e-06 -5.060596e-09 2.969349e-10 3.741649e-09

Xty\_3 <- t(X\_3P)%\*%Y\_3P # Resultado de X traspuesta por y ...   
Xty\_3

## [,1]  
## 1 155998.3  
## yd 631417771.8  
## Ri 3097719840.4  
## ti 312918407.2

b\_3 <- XtX\_inv\_3%\*%Xty\_3 # Resultado de los estimadores b = (Xt\*X)inv\*Xt\*y  
"Estos son los coeficientes estimados OLS"

## [1] "Estos son los coeficientes estimados OLS"

b\_3

## [,1]  
## 1 -20.422436191  
## yd 0.736848710  
## Ri 0.035418543  
## ti -0.005724966

"Estos son los Y estimados por el modelo "

## [1] "Estos son los Y estimados por el modelo "

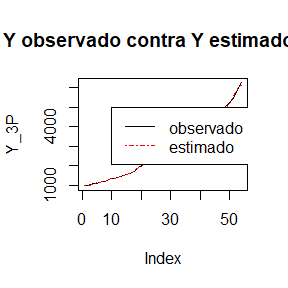
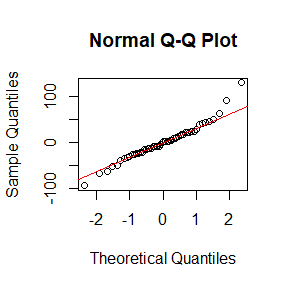
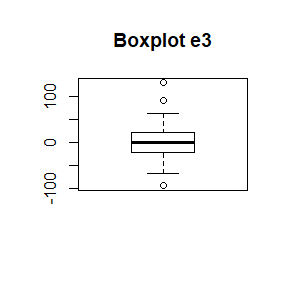
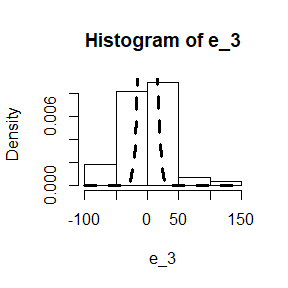
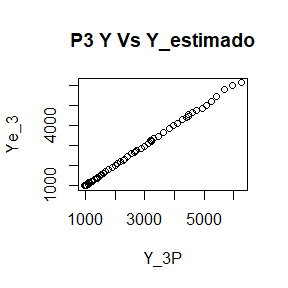
Ye\_3 <- X\_3P%\*%b\_3 #Calcula yt estimado  
#OJO estos son los resultados estimados de y de acuerdo a los valores encontrados de los coeficientes   
  
  
plot(Y\_3P,Ye\_3, main = "P3 Y Vs Y\_estimado") # Se grafica Y de los datos iniciales (y) contra Y estimado (yte)  
e\_3 <- Y\_3P - Ye\_3 # et es el vector de errores de estimacion (Ydado-Yestimado)   
summary(e\_3) # Aqui se muestra en consola un summary de los errores de estimacion (et)

## V1   
## Min. :-94.38719   
## 1st Qu.:-22.63371   
## Median : 0.00934   
## Mean : 0.00000   
## 3rd Qu.: 19.66640   
## Max. :128.44952

hist(e\_3, prob=TRUE ) #Histograma de los errores de estimacion   
curve(dnorm(x,mean = 0,sd=9),col=1,lty=2,lwd=3,add = TRUE) # Agrega una curva normal con media en cero y desviacion estandar 3,878   
boxplot(e\_3, main="Boxplot e3") # grafica de caja de los errores de estimacion (et)  
jb.norm.test(e\_3) # Prueba de normalidad Jarque-Bera

##   
## Jarque-Bera test for normality  
##   
## data: e\_3  
## JB = 11.193, p-value = 0.014

g1\_qq <- qqnorm(e\_3,col="black") # Genera una Q-Q plot que es un metodo para comparar distrubuciones de probabilidad de una muestra aleatoria y una de comparacion   
qqline(e\_3,col="red") # qqline agrega a la grafica anterior una linea de tendencia teorica de cuantiles de una distribucion normal   
  
# OTRA FORMA PARA YT vs YTE  
# De esta manera de grafica y y se agrega la linea LINES de yte estimado   
  
plot(Y\_3P,t="l", main = "Y observado contra Y estimado 3P") # "l" significa linea, p significa puntos  
lines(Ye\_3,lty=4,col=2)  
legend(x=10,y=5000,legend = c("observado","estimado"),lty = c(1,4),col = c(1,2)) #"lty" significa 1=linea, 4= punto y linea

 Observamos en este modelo que la prueba de jarque bera tiene un p\_val<alpha, sin embargo sacamos conclusiones a partir de los graficos en donde se observa que Y vs Ye tiene fuerte relacion lineal, su boxplot tiene media en cero y ademas los errores de estimacion en la qqplot se asemejan mucho a la distribucion normal

## Estimacion de la varianza y Intervalos de confianza

#x=seq(-3,3,length=100)  
#y=seq(-3,3,length=100)  
#parabola=function(x,y) x^3+y^2  
#z=outer(x, y, parabola)  
#persp(x,y,z,theta = 45)  
#image(x,y,z)  
#contour(x,y,z,add=T)  
  
  
# Recordar que b~N(beta,Sigma\*\*2(X'X)Inv) Donde b es el vector de estimadores hallados   
  
# Recordar que Var b\_estimado=s\*\*2\*(X'X)inv  
RSS\_3=t(e\_3)%\*%e\_3  
S2\_3<- RSS\_3[1,1]/51 # Aqui se hala la varianza S\*\*2estimado=RSS/(n-k)=et\*\*2/(n-k) ES UN ESCALAR  
Var\_b\_3=S2\_3\*XtX\_inv\_3 # Matriz de varianzas y covarianzas \_\_\_Los elementos de la diagonal de la matriz contienen las varianzas de las variables, mientras que los elementos que se encuentran fuera de la diagonal contienen las covarianzas entre todos los pares posibles de variables.  
S2b1\_3=Var\_b\_3[1,1]  
S2b2\_3=Var\_b\_3[2,2]  
S2b3\_3=Var\_b\_3[3,3]  
S2b4\_3=Var\_b\_3[4,4]  
  
sb1=sqrt(S2b1\_3)  
sb2=sqrt(S2b2\_3)  
sb3=sqrt(S2b3\_3)  
sb4=sqrt(S2b4\_3)  
  
  
# Recordar que para los intervalos de confianza se tiene que   
# Intervalo= b1 +- t\_(alfa/2)\*S\_b1 con n-2 grados de libertad  
ta2=qt(0.025,51,lower.tail = FALSE) # aqui se obtiene la funcion de quantiles de la t-Student con a/2=0.025 >>>  
ta2

## [1] 2.007584

lib1=b\_3[1,1]-ta2\*sb1 # Aqui se calcula el extremo superior e inferior del intervalo  
lib1

## 1   
## -46.47431

lsb1=b\_3[1,1]+ta2\*sb1  
lsb1

## 1   
## 5.629434

lib2=b\_3[2,1]-ta2\*sb2  
lib2

## yd   
## 0.7087318

lsb2=b\_3[2,1]+ta2\*sb2  
lsb2

## yd   
## 0.7649656

lib3=b\_3[3,1]-ta2\*sb3  
lib3

## Ri   
## 0.03033633

lsb3=b\_3[3,1]+ta2\*sb3  
lsb3

## Ri   
## 0.04050075

lib4=b\_3[4,1]-ta2\*sb4  
lib4

## ti   
## -0.01044843

lsb4=b\_3[4,1]+ta2\*sb4  
lsb4

## ti   
## -0.001001504

colum1<-c("Coef","b1","b2","b3","b4")  
colum2<-c("Estimación",b\_3[1,1], b\_3[2,1], b\_3[3,1], b\_3[4,1])  
colum3<-c("Varianza",S2b1\_3, S2b2\_3, S2b3\_3, S2b4\_3)  
colum4<-c("Error\_Estandar",sb1, sb2, sb3, sb4)  
colum5<-c("Int Inferior",lib1,lib2,lib3,lib4)  
colum6<-c("Int Superior",lsb1,lsb2,lsb3,lsb4)  
kable(Tabla\_Coef\_Var <- cbind(colum1,colum2,colum3,colum4,colum5,colum6))

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | colum1 | colum2 | colum3 | colum4 | colum5 | colum6 |
|  | Coef | Estimación | Varianza | Error\_Estandar | Int Inferior | Int Superior |
| 1 | b1 | -20.4224361911529 | 168.395494057141 | 12.9767289428862 | -46.474306608679 | 5.62943422637328 |
| yd | b2 | 0.736848709961494 | 0.000196149386803981 | 0.0140053342267859 | 0.70873182826995 | 0.764965591653038 |
| Ri | b3 | 0.0354185427496093 | 6.4085177347898e-06 | 0.00253150503353041 | 0.0303363343298208 | 0.0405007511693978 |
| ti | b4 | -0.00572496640457398 | 5.53571321024604e-06 | 0.00235280964173603 | -0.0104484288559659 | -0.00100150395318211 |

## PRUEBA DE RESTRICCIONES LINEALES: Ho:4B2 + 3B4 = 0 ; Ha:4B2 + 3B4 =! 0

Para esta restriccion lineal usamos la prueba F

# Se aplica la misma prueba F que se uso para la prueba de significancia de la regresion   
  
R=c(0,4,0,3)  
R=t(R)  
R # Matriz de coeficientes de las restricciones 2x5

## [,1] [,2] [,3] [,4]  
## [1,] 0 4 0 3

r=c(0) # vector fila r  
r

## [1] 0

# Aqui se vuelve a construir la estadistica de prueba F  
p1=R%\*%b\_3-r   
p1

## [,1]  
## [1,] 2.93022

p0=XtX\_inv\_3%\*%t(R)  
"XtX\_inv\*Rt"

## [1] "XtX\_inv\*Rt"

p0

## [,1]  
## 1 -2.380789e-04  
## yd 5.151362e-07  
## Ri -9.115554e-08  
## ti -9.017439e-09

p2=solve(R%\*%XtX\_inv\_3%\*%t(R)) # Inversa   
p2

## [,1]  
## [1,] 491764.8

p3=t(p1)%\*%p2%\*%p1   
p3

## [,1]  
## [1,] 4222385

fc=(p3/1)/S2\_3 # Estadistica F=[(Rb-r)'(R(X'X)^-1\*R')^-1\*(Rb-r)/#deRestc]/S2 S2=RSS/n-k  
"F\_Valor "

## [1] "F\_Valor "

fc

## [,1]  
## [1,] 2853.956

pv=pf(fc,1,51,lower.tail = F) # Recordar que el comando pf() nos da la probabilidad de una distribucion F   
"Probabilidad \_ F"

## [1] "Probabilidad \_ F"

pv

## [,1]  
## [1,] 1.917373e-46

# Hallamos los estimadores restringidos br  
  
br=b\_3+XtX\_inv\_3%\*%t(R)%\*%p2%\*%(r-R%\*%b\_3) # Donde p2=solve(R%\*%XtX\_inv%\*%t(R))  
"El estimador OLS restringido es"

## [1] "El estimador OLS restringido es"

br

## [,1]  
## 1 322.644249696  
## yd -0.005451729  
## Ri 0.166771752  
## ti 0.007268972

prueba=4\*br[2,1]+3\*br[4,1]  
"Vefificación de la restricción 1"

## [1] "Vefificación de la restricción 1"

prueba

## yd   
## -2.844947e-16

ye\_r=X\_3P%\*%br # Vector de valores y\_estimados\_restringidos Se usa la matriz X para hallar el intercepto   
et\_r=Y\_3P-ye\_r # Vector de errores de estimacion restringidos   
RSSr=t(et\_r)%\*%et\_r # RSS= e^2  
"El RSSr es"

## [1] "El RSSr es"

RSSr

## [,1]  
## [1,] 4297839

fc1=((RSSr-RSS\_3)/1)/S2\_3 # Valor de F calculado para el modelo restringido   
"El Fc de la ecuación restringida (44) es"

## [1] "El Fc de la ecuación restringida (44) es"

fc1

## [,1]  
## [1,] 2853.956

ye\_r <- X\_3P%\*%br # Vector de Y estimado   
etr <- Y\_3P-ye\_r # Vector de errores de estimacion   
TSSa=t(Y\_3P)%\*%Y\_3P # Suma de cuadrados totales   
"TSS con restricciones lineales"

## [1] "TSS con restricciones lineales"

TSSa

## [,1]  
## [1,] 570074234

ESSa=t(br)%\*% XtX\_3 %\*%br # SUma de cuadrados Estimados b2t\*XatXa\*b2  
"ESS con restricciones lineales"

## [1] "ESS con restricciones lineales"

ESSa

## [,1]  
## [1,] 565776395

RSSa=t(et\_r)%\*%et\_r # Suma de cuadrados residuales eta\*\*2  
"RSS con restricciones lineales"

## [1] "RSS con restricciones lineales"

RSSa

## [,1]  
## [1,] 4297839

# Cuarto Punto

* CAR: Stock de vehículos en USA
* QMG: Consumo de gasolina en miles de galones
* PMG: Precio de la gasolina en dólares
* POP: población en miles
* RGNP: PIB real en miles de millones de dólares de 1982
* PGNP: deflactor del PIB

## Estimacion del modelo 4.1 y 4.2 con los datos 1950-1972=n1

rm(list=ls()) # rm remueve objetos, ls = lista  
setwd('C://Users/f-pis/Desktop/Semestres/Archivos R') #Establece el workDirectory  
#install.packages("readxl") # instala el paquete para leer Excel   
library(readxl) #..library() Carga el paquete   
Datos\_4PT<-read\_excel("Datos\_E4.xls") #read\_excel lee el excel  
Datos\_4P=data.frame(Datos\_4PT[1:23,])  
  
lnQMG=log(Datos\_4P$QMG)  
lnCAR=log(Datos\_4P$CAR)  
lnPOP=log(Datos\_4P$POP)  
lnRGNP=log(Datos\_4P$RGNP)  
lnPGNP=log(Datos\_4P$PGNP)  
lnPMG=log(Datos\_4P$PMG)  
  
Reg\_4.1<-lm(lnQMG~lnCAR+lnPOP+lnRGNP+lnPGNP+lnPMG)  
summary(Reg\_4.1)

##   
## Call:  
## lm(formula = lnQMG ~ lnCAR + lnPOP + lnRGNP + lnPGNP + lnPMG)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.041097 -0.005890 -0.000394 0.010531 0.025579   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.6801 2.7936 0.601 0.55549   
## lnCAR 0.3635 0.5152 0.706 0.48994   
## lnPOP 1.0539 0.9048 1.165 0.26019   
## lnRGNP -0.3114 0.1625 -1.916 0.07231 .   
## lnPGNP 0.1250 0.1580 0.791 0.44000   
## lnPMG 1.0481 0.2682 3.907 0.00113 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.01872 on 17 degrees of freedom  
## Multiple R-squared: 0.9952, Adjusted R-squared: 0.9937   
## F-statistic: 699.8 on 5 and 17 DF, p-value: < 2.2e-16

TSS\_4.1=sum((lnQMG)^2)-length(lnQMG)\*mean(lnQMG)^2  
TSS\_4.1

## [1] 1.232235

e\_4.1=resid(Reg\_4.1)  
RSS\_4.1=t(e\_4.1)%\*%e\_4.1  
RSS\_4.11=deviance(Reg\_4.1)  
Ye\_4.1<-fitted(Reg\_4.1)  
ESS\_4.1=sum((Ye\_4.1)^2)-length(Ye\_4.1)\*mean(lnQMG)^2  
  
col1=c("ESS Modelo 4 Rest",ESS\_4.1)  
col2=c("TSS Modelo 4 Rest",TSS\_4.1)  
col3=c("RSS Modelo 4 Rest",RSS\_4.1)  
tab4=cbind(col1,col2,col3)  
kable(tab4)

|  |  |  |
| --- | --- | --- |
| col1 | col2 | col3 |
| ESS Modelo 4 Rest | TSS Modelo 4 Rest | RSS Modelo 4 Rest |
| 1.22627695174197 | 1.23223511095148 | 0.00595815921018831 |

plot(Reg\_4.1)  
  
if (!require('corrplot'))   
{  
 install.packages('corrplot');  
 library(corrplot);  
}

## Loading required package: corrplot

## Warning: package 'corrplot' was built under R version 3.6.3

## corrplot 0.84 loaded

library("corrplot")  
col4 <- colorRampPalette(c("#7F0000", "red", "#FF7F00", "yellow", "#7FFF7F",  
 "cyan", "#007FFF", "blue", "#00007F"))  
MD<-cor(Datos\_4P[,2:7])  
corrplot(MD,method="number", title = "Grafica de Correlaciones",col = col4(20),tl.col = "black")  
pairs(lnQMG~lnCAR+lnPOP+lnRGNP+lnPGNP+lnPMG,data = Datos\_4P)  
  
library(GGally)

## Warning: package 'GGally' was built under R version 3.6.3

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2

ggpairs(Datos\_4P[,2:7], lower = list(continuous = "smooth"),  
 diag = list(continuous = "bar"), axisLabels = "none")

## Warning in check\_and\_set\_ggpairs\_defaults("diag", diag, continuous =  
## "densityDiag", : Changing diag$continuous from 'bar' to 'barDiag'

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

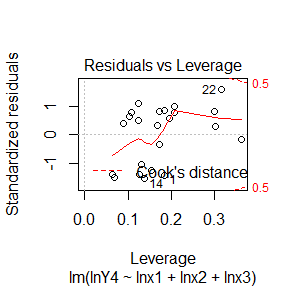
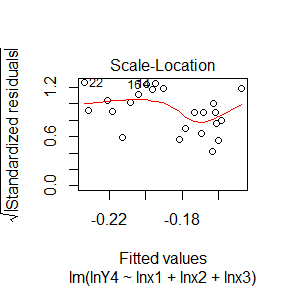
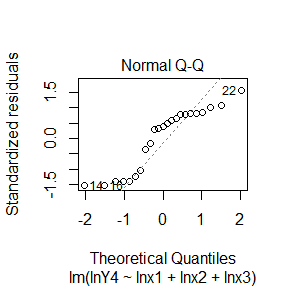
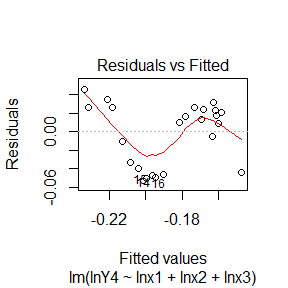
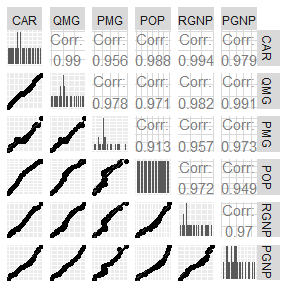
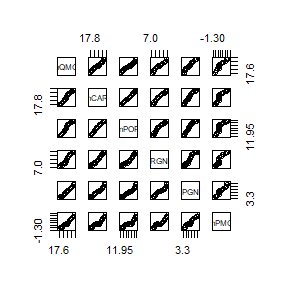
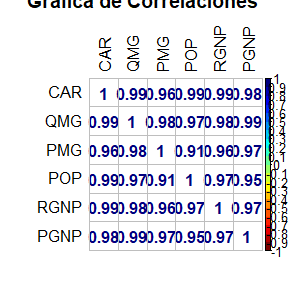
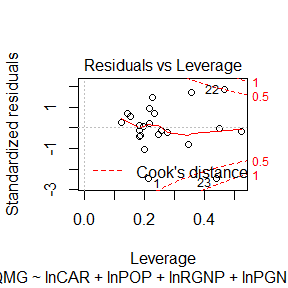
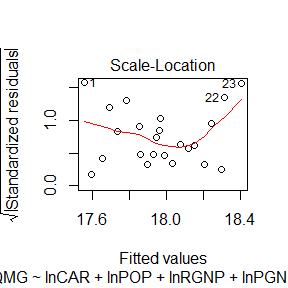
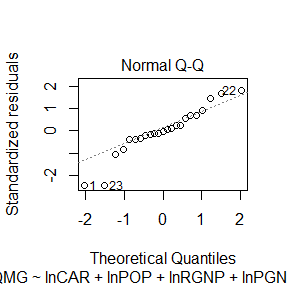
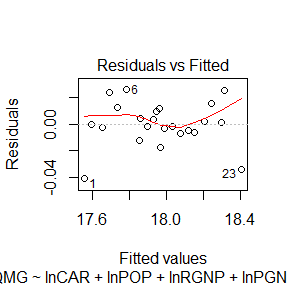
# Estimacion del modelo 4.2  
  
lnY4=log(Datos\_4P$QMG/Datos\_4P$CAR)  
lnx1=log(Datos\_4P$RGNP/Datos\_4P$POP)  
lnx2=log(Datos\_4P$CAR/Datos\_4P$POP)  
lnx3=log(Datos\_4P$PMG/Datos\_4P$PGNP)  
  
Reg\_4.2<-lm(lnY4~lnx1+lnx2+lnx3)  
summary(Reg\_4.2)

##   
## Call:  
## lm(formula = lnY4 ~ lnx1 + lnx2 + lnx3)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.05079 -0.03639 0.01337 0.02515 0.04526   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -0.30653 2.37844 -0.129 0.899  
## lnx1 -0.13972 0.23851 -0.586 0.565  
## lnx2 0.05446 0.28276 0.193 0.849  
## lnx3 0.18527 0.27882 0.664 0.514  
##   
## Residual standard error: 0.03478 on 19 degrees of freedom  
## Multiple R-squared: 0.389, Adjusted R-squared: 0.2925   
## F-statistic: 4.032 on 3 and 19 DF, p-value: 0.02238

TSS\_4.2=sum((lnY4)^2)-length(lnY4)\*mean(lnY4)^2  
e\_4.2=resid(Reg\_4.2)  
RSS\_4.2=t(e\_4.2)%\*%e\_4.2  
RSS\_4.21=deviance(Reg\_4.2)  
Ye\_4.2<-fitted(Reg\_4.2)  
ESS\_4.2=sum((Ye\_4.2)^2)-length(Ye\_4.2)\*mean(lnY4)^2  
  
col1=c("ESS Modelo 4 Rest",ESS\_4.2)  
col2=c("TSS Modelo Rest",TSS\_4.2)  
col3=c("RSS Modelo 4 Rest",RSS\_4.2)  
tab4.2=cbind(col1,col2,col3)  
kable(tab4.2)

|  |  |  |
| --- | --- | --- |
| col1 | col2 | col3 |
| ESS Modelo 4 Rest | TSS Modelo Rest | RSS Modelo 4 Rest |
| 0.0146323457937422 | 0.0376153514771659 | 0.0229830056834237 |

plot(Reg\_4.2)



lnRGNP -0.3114 0.1625 -1.916 0.07231 . lnPMG 1.0481 0.2682 3.907 0.00113 \*\* Estas son las variables significativas del modelo 3 por lo que podemos eliminarlas otras variables explicativas no significativas, podemos observar que el R2 es muy alto, explica el 99,52% de la varianza de modelo, y el p\_valor < alpha, por lo tanto concluimos que la regresion es significativa.

(Intercept) -0.30653 2.37844 -0.129 0.899 lnx1 -0.13972 0.23851 -0.586 0.565 lnx2 0.05446 0.28276 0.193 0.849 lnx3 0.18527 0.27882 0.664 0.514 Observamos que en el segundo modelo no nay variables significativas, por lo que este modelo a su vez solo esplica un R2 de 0,389 lo que es un 38% de la varianza del modelo, en la grafica qqplot se observan serias desviaciones respecto de la normal teorica.

## Reestimacion de los modelos usando la muestra completa 1950-1987

# Estimacion del modelo 4.1  
"Estimacion del modelo 4.1"

## [1] "Estimacion del modelo 4.1"

lnQMGt=log(Datos\_4PT$QMG)  
lnCARt=log(Datos\_4PT$CAR)  
lnPOPt=log(Datos\_4PT$POP)  
lnRGNPt=log(Datos\_4PT$RGNP)  
lnPGNPt=log(Datos\_4PT$PGNP)  
lnPMGt=log(Datos\_4PT$PMG)  
X4.1=cbind(lnCARt,lnPOPt,lnRGNPt,lnPGNPt,lnPMGt)  
  
  
Reg\_4.1T<-lm(lnQMGt~lnCARt+lnPOPt+lnRGNPt+lnPGNPt+lnPMGt)  
summary(Reg\_4.1T)

##   
## Call:  
## lm(formula = lnQMGt ~ lnCARt + lnPOPt + lnRGNPt + lnPGNPt + lnPMGt)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.059240 -0.020056 -0.001106 0.027460 0.041927   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 9.98698 2.62062 3.811 0.000594 \*\*\*  
## lnCARt 2.55992 0.22813 11.221 1.26e-12 \*\*\*  
## lnPOPt -2.87808 0.45344 -6.347 3.99e-07 \*\*\*  
## lnRGNPt -0.42927 0.14838 -2.893 0.006811 \*\*   
## lnPGNPt -0.17887 0.06336 -2.823 0.008115 \*\*   
## lnPMGt -0.14111 0.04340 -3.252 0.002704 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.02825 on 32 degrees of freedom  
## Multiple R-squared: 0.9927, Adjusted R-squared: 0.9915   
## F-statistic: 868.8 on 5 and 32 DF, p-value: < 2.2e-16

plot(Reg\_4.1T)  
# Estimacion del modelo 4.2  
"Estimacion del modelo 4.2"

## [1] "Estimacion del modelo 4.2"

lnY4t=log(Datos\_4PT$QMG/Datos\_4PT$CAR)  
lnx1t=log(Datos\_4PT$RGNP/Datos\_4PT$POP)  
lnx2t=log(Datos\_4PT$CAR/Datos\_4PT$POP)  
lnx3t=log(Datos\_4PT$PMG/Datos\_4PT$PGNP)  
X4=cbind(lnx1t,lnx2t,lnx3t)  
  
  
Reg\_4.2T<-lm(lnY4t~lnx1t+lnx2t+lnx3t)  
summary(Reg\_4.2T)

##   
## Call:  
## lm(formula = lnY4t ~ lnx1t + lnx2t + lnx3t)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.109139 -0.045609 0.000628 0.051406 0.110555   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -5.85398 3.10248 -1.887 0.0677 .  
## lnx1t -0.69046 0.29337 -2.354 0.0245 \*  
## lnx2t 0.28874 0.27723 1.041 0.3050   
## lnx3t -0.14313 0.07488 -1.911 0.0644 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.0605 on 34 degrees of freedom  
## Multiple R-squared: 0.7415, Adjusted R-squared: 0.7187   
## F-statistic: 32.51 on 3 and 34 DF, p-value: 4.249e-10

(RSS\_R=deviance(Reg\_4.1T))

## [1] 0.02553086

"Regresion con la muestra 2 n2"

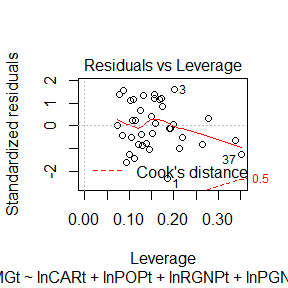
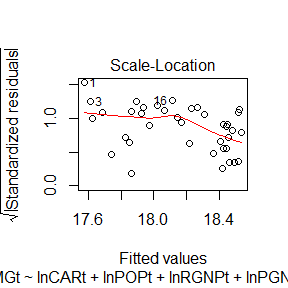
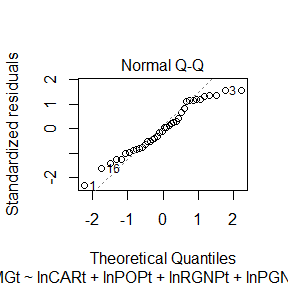
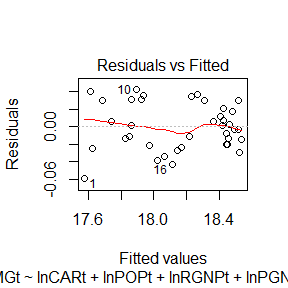
## [1] "Regresion con la muestra 2 n2"

Reg\_4.2F<-lm(lnY4t[23:38]~lnx1t[23:38]+lnx2t[23:38]+lnx3t[23:38])  
summary(Reg\_4.2F)

##   
## Call:  
## lm(formula = lnY4t[23:38] ~ lnx1t[23:38] + lnx2t[23:38] + lnx3t[23:38])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.064850 -0.029484 -0.008218 0.012563 0.107417   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 4.05308 5.68468 0.713 0.489  
## lnx1t[23:38] -0.29873 0.45431 -0.658 0.523  
## lnx2t[23:38] -0.94261 0.56801 -1.660 0.123  
## lnx3t[23:38] -0.09151 0.07782 -1.176 0.262  
##   
## Residual standard error: 0.05183 on 12 degrees of freedom  
## Multiple R-squared: 0.8274, Adjusted R-squared: 0.7843   
## F-statistic: 19.18 on 3 and 12 DF, p-value: 7.149e-05

(RSS4F=deviance(Reg\_4.2F))

## [1] 0.0322316

 Se observa que (Intercept) 9.98698 2.62062 3.811 0.000594  ***lnCARt 2.55992 0.22813 11.221 1.26e-12***  lnPOPt -2.87808 0.45344 -6.347 3.99e-07 \* **lnRGNPt -0.42927 0.14838 -2.893 0.006811**  lnPGNPt -0.17887 0.06336 -2.823 0.008115  **lnPMGt -0.14111 0.04340 -3.252 0.002704**  Ahora cuando tomamos toda la muestra y estimamos el modelo, ahora todos los coeficientes son significativos, por lo que se observa un mejor ajuste, tambien debido a que el R2 incremento y ahora es de 0,9927 un 99,27%de la varianza total explicada, muy buen auste, finalmente el p\_val de la prueba es < alpha por lo que la regresion es significativa

No con la misma suerte observamos que el modelo 2 cuando se toman todos los datos para la estimacion los coeficientes estimados no se vuelven significativos, sin embargo tiene un buen R2 0,8274 (Intercept) 4.05308 5.68468 0.713 0.489 lnx1t[23:38] -0.29873 0.45431 -0.658 0.523 lnx2t[23:38] -0.94261 0.56801 -1.660 0.123 lnx3t[23:38] -0.09151 0.07782 -1.176 0.262

## CAMBIO ESTRUCTURAL CON VARIABLES DUMMY

pruebe si la demanda de gasolina por automóvil sufrió un cambio permanente luego del embargo petrolero de 1973.

Usamos el modelo 4.2 logaritmo de la demanda de gasolina por automobil, tener en cuenta que el modelo restringido es el modelo original con todos los datos

#Acontinuacion procedemos a estimar el modelo 4.2 con variables Dummy  
library(readxl)   
Datos\_4PT<-read\_excel("Datos\_E4.xls") #read\_excel lee el excel  
d1=cbind(rep(0,23))  
d2=cbind(rep(1,15))  
d=rbind(d1,d2)  
  
  
eq1\_mnr\_d=lm(lnY4t~d+X4+d\*X4)  
summary(eq1\_mnr\_d)

##   
## Call:  
## lm(formula = lnY4t ~ d + X4 + d \* X4)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.050788 -0.032354 0.007884 0.025500 0.085824   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.30653 2.67525 -0.115 0.90954   
## d 10.75911 5.68382 1.893 0.06805 .   
## X4lnx1t -0.13972 0.26828 -0.521 0.60634   
## X4lnx2t 0.05446 0.31804 0.171 0.86519   
## X4lnx3t 0.18527 0.31361 0.591 0.55910   
## d:X4lnx1t 0.25268 0.46632 0.542 0.59192   
## d:X4lnx2t -1.74034 0.61306 -2.839 0.00805 \*\*  
## d:X4lnx3t -0.32025 0.31955 -1.002 0.32427   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.03912 on 30 degrees of freedom  
## Multiple R-squared: 0.9046, Adjusted R-squared: 0.8824   
## F-statistic: 40.65 on 7 and 30 DF, p-value: 1.341e-13

"Suma de cuadrados de los errores del modelo no restringido con DUMMY"

## [1] "Suma de cuadrados de los errores del modelo no restringido con DUMMY"

(RSS\_NRD2=deviance(eq1\_mnr\_d))

## [1] 0.04591107

"prueba de cambio estructural con variables Dummy"

## [1] "prueba de cambio estructural con variables Dummy"

"La estadística F de la ecuación (22) calculada"

## [1] "La estadística F de la ecuación (22) calculada"

(FC\_22=((RSS\_R-RSS\_NRD2)/5)/(RSS\_NRD2/28))

## [1] -2.485875

"verificación de coeficientes"

## [1] "verificación de coeficientes"

(PV\_F\_22=pf(FC\_22,5,28,lower.tail = F))

## [1] 1

"El valor p de la prueba nos indica que debemos rechazar H0 si p\_val<<alpha; por tanto, concluimos que hay cambio estructural en la función a partir del año 1972."

## [1] "El valor p de la prueba nos indica que debemos rechazar H0 si p\_val<<alpha; por tanto, concluimos que hay cambio estructural en la función a partir del año 1972."

bnrd=cbind(coef(eq1\_mnr\_d))  
bnrd

## [,1]  
## (Intercept) -0.30652784  
## d 10.75910981  
## X4lnx1t -0.13971549  
## X4lnx2t 0.05446182  
## X4lnx3t 0.18527038  
## d:X4lnx1t 0.25267628  
## d:X4lnx2t -1.74033781  
## d:X4lnx3t -0.32024822

b11nrd=bnrd[1]  
delta1=bnrd[2]  
b12nrd=b11nrd+delta1  
b21nrd=bnrd[3]  
delta2=bnrd[4]  
b22nrd=b21nrd+delta2

Mediante el uso de las variables ficticias Dummy podemos concluir que hay o no un cambio estructural en el modelo a partir de cierto punto. El valor p de la prueba nos indica que debemos rechazar H0 si p\_val<<alpha; por tanto, concluimos que hay cambio estructural en la función a partir del año 1972. Esto explica el cambio en los estimadores encontrados en cada una de las muestras

## Estimacion del modelo 4.1 por variables Dummy

# i) Estimacion del modelo 4.1 por variables Dummy  
# Construya una regresión con variable dummy para probar si la elasticidad precio cambió luego de 1973.  
  
  
Reg\_4.1D<-lm(lnQMGt~d+X4.1+d\*X4.1)  
summary(Reg\_4.1D)

##   
## Call:  
## lm(formula = lnQMGt ~ d + X4.1 + d \* X4.1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.041097 -0.006074 0.000566 0.008870 0.025579   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.6801 2.4793 0.678 0.503967   
## d 60.4078 20.9419 2.885 0.007775 \*\*   
## X4.1lnCARt 0.3635 0.4572 0.795 0.433738   
## X4.1lnPOPt 1.0539 0.8030 1.312 0.200846   
## X4.1lnRGNPt -0.3114 0.1442 -2.159 0.040252 \*   
## X4.1lnPGNPt 0.1250 0.1403 0.891 0.381128   
## X4.1lnPMGt 1.0481 0.2381 4.403 0.000163 \*\*\*  
## d:X4.1lnCARt 1.3176 0.6643 1.984 0.057965 .   
## d:X4.1lnPOPt -7.4647 1.7777 -4.199 0.000278 \*\*\*  
## d:X4.1lnRGNPt 0.4959 0.2560 1.938 0.063602 .   
## d:X4.1lnPGNPt 0.3918 0.3523 1.112 0.276264   
## d:X4.1lnPMGt -1.2737 0.2443 -5.214 1.91e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.01662 on 26 degrees of freedom  
## Multiple R-squared: 0.9979, Adjusted R-squared: 0.9971   
## F-statistic: 1147 on 11 and 26 DF, p-value: < 2.2e-16

"RSS No Restringido Dummy Lm"

## [1] "RSS No Restringido Dummy Lm"

(RSS\_NRDlm=deviance(Reg\_4.1D))

## [1] 0.007177586

X1=cbind(1,X4.1[1:23,])  
X2=cbind(1,X4.1[24:38,])  
O23=matrix(nrow=23,ncol=6,0)  
O15=matrix(nrow=15,ncol=6,0)  
A1=rbind(X1,O15)  
A2=rbind(O23,X2)  
"MATRIZ XNR"

## [1] "MATRIZ XNR"

XNR=cbind(A1,A2)  
  
"LA MATRIZ XNRtXNR"

## [1] "LA MATRIZ XNRtXNR"

XNRtXNR=t(XNR)%\*%XNR  
"LA MATRIZ XNRtXNR inversa"

## [1] "LA MATRIZ XNRtXNR inversa"

XNRtXNR\_inv=solve(XNRtXNR)  
"EL VECTOR XNRty"

## [1] "EL VECTOR XNRty"

XNRty=t(XNR)%\*%lnQMGt  
"EL ESTIMADOR OLS DEL MODELO NO RESTRINGIDO ES"

## [1] "EL ESTIMADOR OLS DEL MODELO NO RESTRINGIDO ES"

(bnr=XNRtXNR\_inv%\*%XNRty)

## [,1]  
## 1.6801434  
## lnCARt 0.3635327  
## lnPOPt 1.0539307  
## lnRGNPt -0.3113880  
## lnPGNPt 0.1249568  
## lnPMGt 1.0481449  
## 62.0879904  
## lnCARt 1.6811160  
## lnPOPt -6.4108036  
## lnRGNPt 0.1845502  
## lnPGNPt 0.5168058  
## lnPMGt -0.2255728

etnr\_D=lnQMGt-XNR%\*%bnr  
"SUMA DE CUADRADOS DE ERRROES NR modelo Dummy Matriz"

## [1] "SUMA DE CUADRADOS DE ERRROES NR modelo Dummy Matriz"

(RSS\_NRDM=t(etnr\_D)%\*%etnr\_D)

## [,1]  
## [1,] 0.007177586

"SUMA DE CUADRADOS EXPLICADA NR modelo Dummy Matriz "

## [1] "SUMA DE CUADRADOS EXPLICADA NR modelo Dummy Matriz "

(ESS\_NRD=(t(bnr)%\*%XNRtXNR%\*%bnr)[1:1]-38\*mean(lnQMGt)^2)

## [1] 3.484047

"LA SUMA TOTAL CUADRADOS NR modelo Dummy Matriz"

## [1] "LA SUMA TOTAL CUADRADOS NR modelo Dummy Matriz"

(TSS\_NRD=(t(lnQMGt)%\*%lnQMGt)[1:1]-38\*mean(lnQMGt)^2)

## [1] 3.491195

"PRUEBA DE SUMAS DE CUADRADOS EN EL MODELO NR"

## [1] "PRUEBA DE SUMAS DE CUADRADOS EN EL MODELO NR"

(TSS\_NRD\_PRUEBA=ESS\_NRD+RSS\_NRDM)

## [,1]  
## [1,] 3.491225

"ESTADÍSTICA F DE LA ECUACION (20)"

## [1] "ESTADÍSTICA F DE LA ECUACION (20)"

"EL VALOR CALCULADO DE LA ESTADÍSTICA ES"

## [1] "EL VALOR CALCULADO DE LA ESTADÍSTICA ES"

(FC\_20=(((RSS\_R-RSS\_NRDM)/5)/(RSS\_NRDM/28)))

## [,1]  
## [1,] 14.31935

"EL P VALOR ES"

## [1] "EL P VALOR ES"

(PV\_F\_20=pf(FC\_20,5,28,lower.tail = F))

## [,1]  
## [1,] 5.466818e-07

"El valor p de la prueba nos indica que debemos rechazar H0 si p\_val<<alpha; por tanto, concluimos que hay cambio estructural en la función a partir del año 1972."

## [1] "El valor p de la prueba nos indica que debemos rechazar H0 si p\_val<<alpha; por tanto, concluimos que hay cambio estructural en la función a partir del año 1972."

El valor p de la prueba nos indica que debemos rechazar H0 si p\_val<<alpha; por tanto, concluimos que hay cambio estructural en la función a partir del año 1972 en el modelo 2 Verificamos esto haciendo la comparacion de los estimadores no restringidos del modelo Dummy con los estimadores restringidos, observamos que las pendientes y el intercepto son los mismos, pero en el modelo NR se le suman valores delta diferentes de cero, por lo que la hipotesis del cambio estructural se acepta

# Quinto Punto

En un paper de 1963, Marc Nerlove analizó una función de costos para 145 compañías eléctricas estadounidenses. El archivo “Datos\_E3\_Nerlove” contiene la información para

* TC: costo total en millones de dólares
* Q: producción en miles de millones de kilovatios hora
* PL: precio del trabajo
* PF: precio de combustibles
* PK: precio del capital

## Lectura de Datos 5P

rm(list=ls()) # rm remueve objetos, ls = lista  
setwd('C://Users/f-pis/Desktop/Semestres/Archivos R') #Establece el workDirectory  
#install.packages("readxl") # instala el paquete para leer Excel   
library(readxl) #..library() Carga el paquete   
Datos\_5P<-read\_excel("Datos\_E5\_Nerlove.xls") #read\_excel lee el excel  
Datos\_5P=data.frame(Datos\_5P)  
  
kable(head(Datos\_5P))

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TC | Q | PL | PF | PK |
| 0.082 | 2 | 2.09 | 17.9 | 183 |
| 0.661 | 3 | 2.05 | 35.1 | 174 |
| 0.990 | 4 | 2.05 | 35.1 | 171 |
| 0.315 | 4 | 1.83 | 32.2 | 166 |
| 0.197 | 5 | 2.12 | 28.6 | 233 |
| 0.098 | 9 | 2.12 | 28.6 | 195 |

## Estimacion Modelo 5

Estime un modelo en el que se tenga como variable dependiente los costos y como variables explicativas una constante y el resto de variables listadas. Con esta estimación obtenga las elasticidades de los costos a las distintas variables explicativas. Interprete los principales resultados.

Y5=Datos\_5P[,1] # en este caso Y5 es el vector de costos   
  
logx2=log(Datos\_5P[,2])  
logx3=log(Datos\_5P[,3])  
logx4=log(Datos\_5P[,4])  
logx5=log(Datos\_5P[,5])  
lnY5=log(Y5)  
  
  
Reg\_5<-lm(lnY5~logx2+logx3+logx4+logx5)  
summary(Reg\_5)

##   
## Call:  
## lm(formula = lnY5 ~ logx2 + logx3 + logx4 + logx5)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.97784 -0.23817 -0.01372 0.16031 1.81751   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -3.52650 1.77437 -1.987 0.0488 \*   
## logx2 0.72039 0.01747 41.244 < 2e-16 \*\*\*  
## logx3 0.43634 0.29105 1.499 0.1361   
## logx4 0.42652 0.10037 4.249 3.89e-05 \*\*\*  
## logx5 -0.21989 0.33943 -0.648 0.5182   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3924 on 140 degrees of freedom  
## Multiple R-squared: 0.926, Adjusted R-squared: 0.9238   
## F-statistic: 437.7 on 4 and 140 DF, p-value: < 2.2e-16

(RSS\_NR5=deviance(Reg\_5))

## [1] 21.55201

plot(Reg\_5)  
  
library(normtest)  
e5=resid(Reg\_5)  
jb.norm.test(e5)

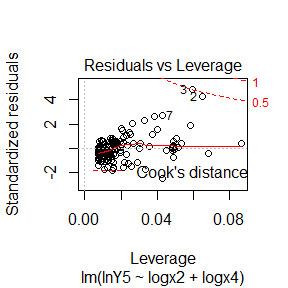
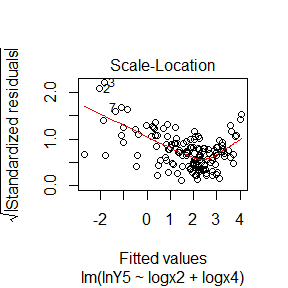
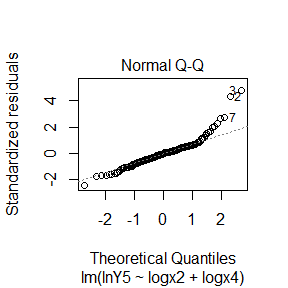
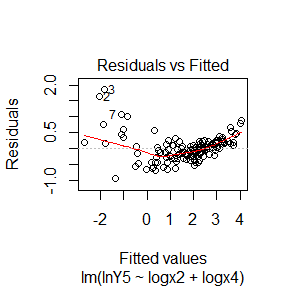
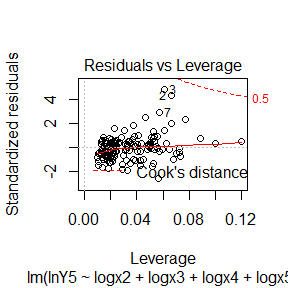
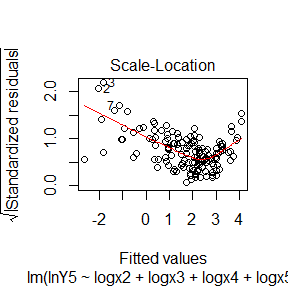
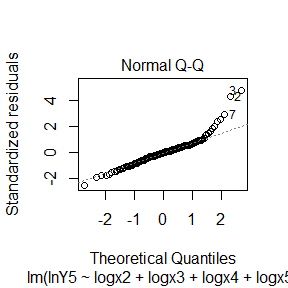
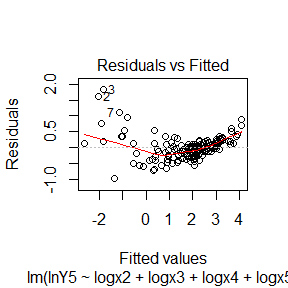
##   
## Jarque-Bera test for normality  
##   
## data: e5  
## JB = 175.7, p-value < 2.2e-16

Reg\_5a<-lm(lnY5~logx2+logx4)  
summary(Reg\_5a)

##   
## Call:  
## lm(formula = lnY5 ~ logx2 + logx4)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.95874 -0.21080 -0.00947 0.16515 1.84465   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -4.53676 0.33825 -13.412 < 2e-16 \*\*\*  
## logx2 0.72415 0.01742 41.562 < 2e-16 \*\*\*  
## logx4 0.47163 0.09286 5.079 1.17e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3942 on 142 degrees of freedom  
## Multiple R-squared: 0.9242, Adjusted R-squared: 0.9231   
## F-statistic: 865.6 on 2 and 142 DF, p-value: < 2.2e-16

plot(Reg\_5a)  
  
library(normtest)  
e5a=resid(Reg\_5a)  
jb.norm.test(e5a)

##   
## Jarque-Bera test for normality  
##   
## data: e5a  
## JB = 182.63, p-value < 2.2e-16

 Observamos en las graficas que los residuales se distribuyen al rededor de cero en la grafica residuals-fitted, se observan algunos valores outliers, en la grafica de cuantiles observamos que los residuales se asemejan a los cuantiles teoricos de la normal, sin embargo en la grafica de los residuales estandarizados se observa cierto comportamiento no lineal, al parecer cuadratico Obseravamos que al eliminar los estimadores no significativos y hacer de nuevo la regresion el modelo no aumenta su R2, aunque ya es alto 98% no varia y por lo tanto procedemos a hacer el test de jarque bera el cual arroja un p-valor<alpha por lo que se rechaza ho y se concluye que no hay normalidad en los errores y por tanto en la dist prob

(Intercept) -4.53676 0.33825 -13.412 < 2e-16  ***logx2 0.72415 0.01742 41.562 < 2e-16***  logx4 0.47163 0.09286 5.079 1.17e-06 \*\*\*

##Modelo 5 Restringido Restriccion de Homogeneidad beta\_3+beta\_4+beta\_5=1

"ESTIMACIÓN DEL MODELO RESTINGIDO CON LA FUNCIÓN lm()"

## [1] "ESTIMACIÓN DEL MODELO RESTINGIDO CON LA FUNCIÓN lm()"

#if (!require("restriktor")) install.packages("restriktor")  
# Se usa el paquete "restriktor" para imponer la restricción beta3=0 sobre la ecuación estimada del modelo no restringido  
  
  
if (!require("restriktor"))   
{  
 install.packages('restriktor');  
 library(restriktor);  
}

## Loading required package: restriktor

## Warning: package 'restriktor' was built under R version 3.6.3

## This is restriktor 0.2-800

## restriktor is BETA software! Please report any bugs.

library(restriktor)  
#options(digits=3) # Para este ejemplo solo se utilizan 3 digitos decimales  
Restricciones <- '  
logx3+logx4+logx5 ==1;'  
Reg\_5R = restriktor(Reg\_5, constraints = Restricciones)  
summary(Reg\_5R)

##   
## Call:  
## conLM.lm(object = object, constraints = constraints)  
##   
## Restriktor: restricted linear model:  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.012004 -0.217588 -0.007524 0.160480 1.819218   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -4.6907891 0.8848713 -5.3011 4.379e-07 \*\*\*  
## logx2 0.7206875 0.0174357 41.3340 < 2.2e-16 \*\*\*  
## logx3 0.5929096 0.2045722 2.8983 0.004357 \*\*   
## logx4 0.4144715 0.0989512 4.1886 4.940e-05 \*\*\*  
## logx5 -0.0073811 0.1907356 -0.0387 0.969186   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.39176 on 140 degrees of freedom  
## Standard errors: standard   
## Multiple R-squared reduced from 0.926 to 0.926   
##   
## Generalized order-restricted information criterion:   
## Loglik Penalty goric   
## -67.838 5.000 145.677

"Estos son los estimadores OLS del modelo 5 restringido"

## [1] "Estos son los estimadores OLS del modelo 5 restringido"

(b5R=coef(Reg\_5R))

## (Intercept) logx2 logx3 logx4 logx5   
## -4.690789123 0.720687524 0.592909608 0.414471455 -0.007381064

sR5\_mr=summary(Reg\_5R)  
"LA VARIANZA ESTIMADA DEL MODELO RESTRINGIDO ES"

## [1] "LA VARIANZA ESTIMADA DEL MODELO RESTRINGIDO ES"

S2\_mr=sR5\_mr$s2  
S2\_mr

## [1] 0.1534774

" EL ERROR ESTÁNDAR DEL MODELO RESTRINGIDO ES"

## [1] " EL ERROR ESTÁNDAR DEL MODELO RESTRINGIDO ES"

s\_mr=(sR5\_mr$s2)^0.5  
s\_mr

## [1] 0.391762

"LA SUMA DE CUADRADOS DE LOS ERROES DEL MODELO RESTRINGIDO ES: RSS\_R=S2\_mr\*(n-(k-1))"

## [1] "LA SUMA DE CUADRADOS DE LOS ERROES DEL MODELO RESTRINGIDO ES: RSS\_R=S2\_mr\*(n-(k-1))"

RSS\_R5=141\*S2\_mr  
RSS\_R5

## [1] 21.64032

(Intercept) -4.6907891 0.8848713 -5.3011 4.379e-07  ***logx2 0.7206875 0.0174357 41.3340 < 2.2e-16***  logx3 0.5929096 0.2045722 2.8983 0.004357 \*\* logx4 0.4144715 0.0989512 4.1886 4.940e-05 \*\*\*

Se observa que en el modelo con la restriccion de homogeneidad el modelo se ajusta mas y todos sus estimadores a excepcion de logx5 son significativos y tiene un r2 del 92% de explicacion

##Prueba de wald

"ESTADÍSTICA F DE WALD DE LA ECUACIÓN F (20)"

## [1] "ESTADÍSTICA F DE WALD DE LA ECUACIÓN F (20)"

"El valor calculado de esta estadística es:"

## [1] "El valor calculado de esta estadística es:"

"FC\_20"

## [1] "FC\_20"

(FC\_20=((RSS\_R5-RSS\_NR5)/5)/(RSS\_NR5/145-2\*5))

## [1] -0.001792867

"EL VALOR P PARA LA PRUEBA ES"

## [1] "EL VALOR P PARA LA PRUEBA ES"

(pv\_wald=pchisq(FC\_20,df=1,lower.tail = F))

## [1] 1

"El Test de Wald es un contraste de hipótesis donde se trata de ver la coherencia de afirmar un valor concreto de un parámetro de un modelo probabilístico una vez tenemos ya un modelo previamente seleccionado y ajustado."

## [1] "El Test de Wald es un contraste de hipótesis donde se trata de ver la coherencia de afirmar un valor concreto de un parámetro de un modelo probabilístico una vez tenemos ya un modelo previamente seleccionado y ajustado."

"El p-valor de la F es"

## [1] "El p-valor de la F es"

(PV\_F\_20=pf(FC\_20,5,135,lower.tail = F))

## [1] 1

#Es este caso el anterior es es mismo estadistico de la prueba de Chow

## Secto Punto

El siguiente modelo puede ser utilizado para estudiar si los gastos de una campaña afectan los resultados de la elección:

Vote A = porcentaje de votos del candidato A expendA= gastos de campaña candidato A expendB= gastos de campaña candidato B prtystrA= porcentaje de votos obtenido por el pardito de A en la elección más reciente

#Estimacion del Modelo

rm(list=ls()) # rm remueve objetos, ls = lista  
setwd('C://Users/f-pis/Desktop/Semestres/Archivos R') #Establece el workDirectory  
#install.packages("readxl") # instala el paquete para leer Excel   
library(readxl) #..library() Carga el paquete   
Datos\_6P<-read\_excel("Datos\_E6\_Votacion.xls") #read\_excel lee el excel  
Datos\_6P=data.frame(Datos\_6P)  
kable(head(Datos\_6P))

|  |  |  |  |
| --- | --- | --- | --- |
| voteA | expendA | expendB | prtystrA |
| 68 | 328.30 | 8.74 | 41 |
| 62 | 626.38 | 402.48 | 60 |
| 73 | 99.61 | 3.07 | 55 |
| 69 | 319.69 | 26.28 | 64 |
| 75 | 159.22 | 60.05 | 66 |
| 69 | 570.16 | 21.39 | 46 |

voteA=Datos\_6P[,1]  
logExpendA=log(Datos\_6P[,2])  
logExpendB=log(Datos\_6P[,3])  
prtystrA=Datos\_6P[,4]  
  
Reg\_6=lm(voteA~logExpendA+logExpendB+prtystrA)  
summary(Reg\_6)

##   
## Call:  
## lm(formula = voteA ~ logExpendA + logExpendB + prtystrA)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -20.3990 -5.4184 -0.8737 4.9563 26.0575   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 45.08788 3.92680 11.482 <2e-16 \*\*\*  
## logExpendA 6.08136 0.38211 15.915 <2e-16 \*\*\*  
## logExpendB -6.61563 0.37889 -17.461 <2e-16 \*\*\*  
## prtystrA 0.15201 0.06203 2.451 0.0153 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 7.713 on 169 degrees of freedom  
## Multiple R-squared: 0.7925, Adjusted R-squared: 0.7888   
## F-statistic: 215.2 on 3 and 169 DF, p-value: < 2.2e-16

RSS\_NR6=deviance(Reg\_6)  
plot(Reg\_6)  
plot(lm.influence(Reg\_6)$hat)  
  
if (!require("olsrr"))   
{  
 install.packages('olsrr');  
 library(rolsrr);  
}

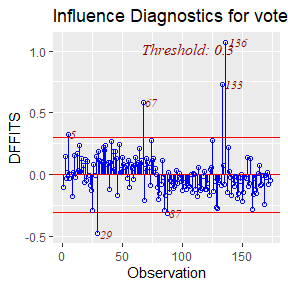
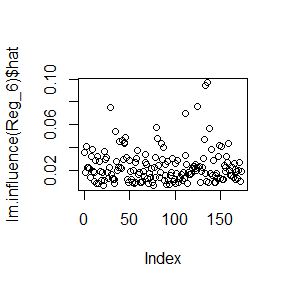
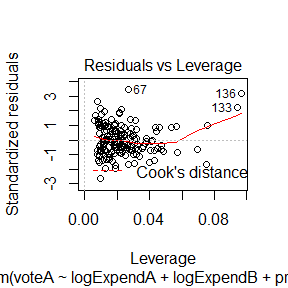
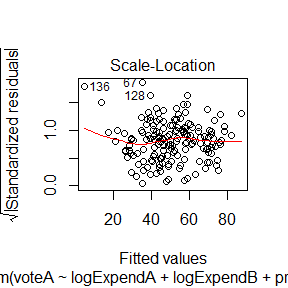
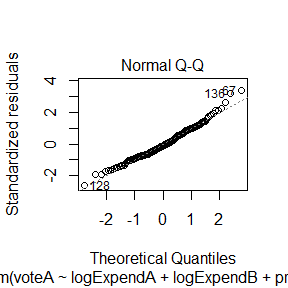
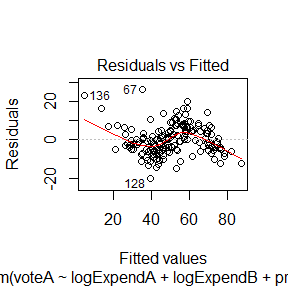
## Loading required package: olsrr

## Warning: package 'olsrr' was built under R version 3.6.3

##   
## Attaching package: 'olsrr'

## The following object is masked from 'package:datasets':  
##   
## rivers

library(olsrr)  
require(olsrr)  
ols\_plot\_dffits(Reg\_6)



## Prueba de Hipotesis

Pruebe la hipótesis según la cual un aumento de 1% en los gastos de A se compensa por un aumento de 1% en los gastos de B. Utilice las pruebas lr,Wald y LM

Ho: b1-b2=0 Ha: b1-b2!=0

### Estimacion del modelo 6 restringido

"ESTIMACIÓN DEL MODELO 6 RESTINGIDO CON LA FUNCIÓN lm()"

## [1] "ESTIMACIÓN DEL MODELO 6 RESTINGIDO CON LA FUNCIÓN lm()"

if (!require("restriktor"))   
{  
 install.packages('restriktor');  
 library(restriktor);  
}  
library(restriktor)  
  
Restricciones <- '  
logExpendA-logExpendB ==0;'  
Reg\_6R = restriktor(Reg\_6, constraints = Restricciones)  
summary(Reg\_6R)

##   
## Call:  
## conLM.lm(object = object, constraints = constraints)  
##   
## Restriktor: restricted linear model:  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -33.5332 -14.1190 0.7192 14.0651 34.6726   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 24.44182 7.85032 3.1135 0.002172 \*\*   
## logExpendA -0.31997 0.54688 -0.5851 0.559278   
## logExpendB -0.31997 0.54688 -0.5851 0.559278   
## prtystrA 0.58788 0.12138 4.8434 2.868e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 15.826 on 169 degrees of freedom  
## Standard errors: standard   
## Multiple R-squared reduced from 0.793 to 0.121   
##   
## Generalized order-restricted information criterion:   
## Loglik Penalty goric   
## -721.73 4.00 1451.45

"Estos son los estimadores OLS del modelo 5 restringido"

## [1] "Estos son los estimadores OLS del modelo 5 restringido"

(b6R=coef(Reg\_6R))

## (Intercept) logExpendA logExpendB prtystrA   
## 24.4418217 -0.3199659 -0.3199659 0.5878761

sR6\_mr=summary(Reg\_6R)  
"LA VARIANZA ESTIMADA DEL MODELO RESTRINGIDO ES"

## [1] "LA VARIANZA ESTIMADA DEL MODELO RESTRINGIDO ES"

S2\_mr6=sR6\_mr$s2  
S2\_mr6

## [1] 250.4518

" EL ERROR ESTÁNDAR DEL MODELO RESTRINGIDO ES"

## [1] " EL ERROR ESTÁNDAR DEL MODELO RESTRINGIDO ES"

s\_mr6=(sR6\_mr$s2)^0.5  
s\_mr6

## [1] 15.82567

"LA SUMA DE CUADRADOS DE LOS ERROES DEL MODELO RESTRINGIDO ES: RSS\_R=S2\_mr\*(n-(k-1))"

## [1] "LA SUMA DE CUADRADOS DE LOS ERROES DEL MODELO RESTRINGIDO ES: RSS\_R=S2\_mr\*(n-(k-1))"

RSS\_R6=170\*S2\_mr6  
RSS\_R6

## [1] 42576.8

## Prueba LR Razon de Verosimilitud

#PRUEBA LR: RAZÓN DE VEROSIMILITUD  
"PRUEBA LR: RAZÓN DE VEROSIMILITUD"

## [1] "PRUEBA LR: RAZÓN DE VEROSIMILITUD"

"LA SUMA DE LOS CUADRADOS DE LOS ERRORES DEL MODELO NO RESTRINGIDO ES"

## [1] "LA SUMA DE LOS CUADRADOS DE LOS ERRORES DEL MODELO NO RESTRINGIDO ES"

"RSS\_NR6"

## [1] "RSS\_NR6"

RSS\_NR6

## [1] 10054.83

npi=pi  
n=173  
k=3  
"El numero e"

## [1] "El numero e"

e=exp(1)  
  
(LMR6=((2\*npi\*e/n)^(-n/2))\*((RSS\_R6)^(-n/2)))

## [1] 0

"El logaritmo de la verosimilitud restringida es"

## [1] "El logaritmo de la verosimilitud restringida es"

(lmr6=log(LMR6))

## [1] -Inf

"LA SUMA DE LOS CUADRADOS DE LOS ERRORES DEL MODELO RESTRINGIDO ES"

## [1] "LA SUMA DE LOS CUADRADOS DE LOS ERRORES DEL MODELO RESTRINGIDO ES"

"RSS\_R6"

## [1] "RSS\_R6"

RSS\_R6

## [1] 42576.8

"LA VEROSIMILITUD DEL MODELO NO RESTRINGIDO ES:"

## [1] "LA VEROSIMILITUD DEL MODELO NO RESTRINGIDO ES:"

(LMNR6=((2\*npi\*e/n)^(-n/2))\*((RSS\_NR6)^(-n/2)))

## [1] 0

"El logaritmo de la verosimilitud no restringido es"

## [1] "El logaritmo de la verosimilitud no restringido es"

(lmnr6=log(LMNR6))

## [1] -Inf

"LA ESTADÍSTICA CALCULADA DE LA ECUACIÓN 17 ES"

## [1] "LA ESTADÍSTICA CALCULADA DE LA ECUACIÓN 17 ES"

"LR\_C\_17"

## [1] "LR\_C\_17"

(LRC\_17=-2\*(lmnr6-lmr6)) ##LR=2\*ln(lambda)

## [1] NaN

"Valores grandes de LR nos indican que lambda tiende a ceroy por esta razón debemos rechazar Ho y decimos que logExpendA-logExpendB !==0 "

## [1] "Valores grandes de LR nos indican que lambda tiende a ceroy por esta razón debemos rechazar Ho y decimos que logExpendA-logExpendB !==0 "

"LA ESTADÍSTICA CALCULADA DE LA ECUACIÓN 18 ES"

## [1] "LA ESTADÍSTICA CALCULADA DE LA ECUACIÓN 18 ES"

"LR\_C\_18"

## [1] "LR\_C\_18"

(LRC\_18=173\*log(RSS\_R6/RSS\_NR6)) #LR=n\*(ln(RSS\_R)-ln(RSS\_NR))

## [1] 249.6834

"EL VALOR P PARA LA PRUEBA ES"

## [1] "EL VALOR P PARA LA PRUEBA ES"

(pv\_valor=pchisq(LRC\_18,df = 1, lower.tail = F))

## [1] 3.044172e-56

Valores grandes de LR nos indican que lambda tiende a ceroy por esta razón debemos rechazar Ho y decimos que logExpendA-logExpendB !==0

##Prueba de wald

"ESTADÍSTICA F DE WALD DE LA ECUACIÓN F (20)"

## [1] "ESTADÍSTICA F DE WALD DE LA ECUACIÓN F (20)"

"El valor calculado de esta estadística es:"

## [1] "El valor calculado de esta estadística es:"

"FC\_20"

## [1] "FC\_20"

(FC\_20W=n\*(RSS\_R6-RSS\_NR6)/RSS\_NR6)

## [1] 559.562

"EL VALOR P PARA LA PRUEBA ES"

## [1] "EL VALOR P PARA LA PRUEBA ES"

(pv\_wald=pchisq(FC\_20W,df=1,lower.tail = F))

## [1] 1.046861e-123

## Prueba LM

Multiplicadores de Lagrange

#options(digits=4)  
"REGRESIÓN AUXILIAR PARA LA PRUEBA LM"

## [1] "REGRESIÓN AUXILIAR PARA LA PRUEBA LM"

"Errores de estimación etr del modelo restingido vs X2t, X2t"

## [1] "Errores de estimación etr del modelo restingido vs X2t, X2t"

etr6=resid(Reg\_6R)  
Reg\_6aux=lm(etr6~logExpendA+logExpendB+prtystrA)  
(sReg\_6aux\_lm=summary(Reg\_6aux))

##   
## Call:  
## lm(formula = etr6 ~ logExpendA + logExpendB + prtystrA)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -20.3990 -5.4184 -0.8737 4.9563 26.0575   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 20.64606 3.92680 5.258 4.36e-07 \*\*\*  
## logExpendA 6.40132 0.38211 16.752 < 2e-16 \*\*\*  
## logExpendB -6.29566 0.37889 -16.616 < 2e-16 \*\*\*  
## prtystrA -0.43586 0.06203 -7.027 4.95e-11 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 7.713 on 169 degrees of freedom  
## Multiple R-squared: 0.7638, Adjusted R-squared: 0.7597   
## F-statistic: 182.2 on 3 and 169 DF, p-value: < 2.2e-16

"El R2\_aux para la prueba LM es el R2 de la regresión auxiliar"

## [1] "El R2\_aux para la prueba LM es el R2 de la regresión auxiliar"

"El R2\_aux es"

## [1] "El R2\_aux es"

R2\_aux=0.7638  
  
R2\_aux

## [1] 0.7638

(LM6\_22=n\*R2\_aux)

## [1] 132.1374

(LM6\_23=n\*(RSS\_R6- RSS\_NR6)/(RSS\_R6))

## [1] 132.1448

(pv\_LM6=pchisq(LM6\_22,df=1,lower.tail = F))

## [1] 1.396139e-30

(pv\_LM6\_23=pchisq(LM6\_23,df=1,lower.tail = F))

## [1] 1.390974e-30

"Ya que P\_val <<alpha concluimos que se rechaza Ho y por lo tanto logExpendA-logExpendB!==0 "

## [1] "Ya que P\_val <<alpha concluimos que se rechaza Ho y por lo tanto logExpendA-logExpendB!==0 "

Ya que P\_val <<alpha concluimos que se rechaza Ho y por lo tanto logExpendA-logExpendB!==0

Las tres pruebas nos hacen concluir que se debe rechazar la hipotesis nula y por lo tanto un cambio del 1% del gasto de A no es compensado por un gasto del 1% del gasto de B, ya que esto se representa como elasticidades tiene sentido hablar de porcentajes, luego logExpendA-logExpendB es diferente de cero y los cambios no son iguales por conclusion.