Mejorar Modelo GARCH

André Marx Puente Arévalo

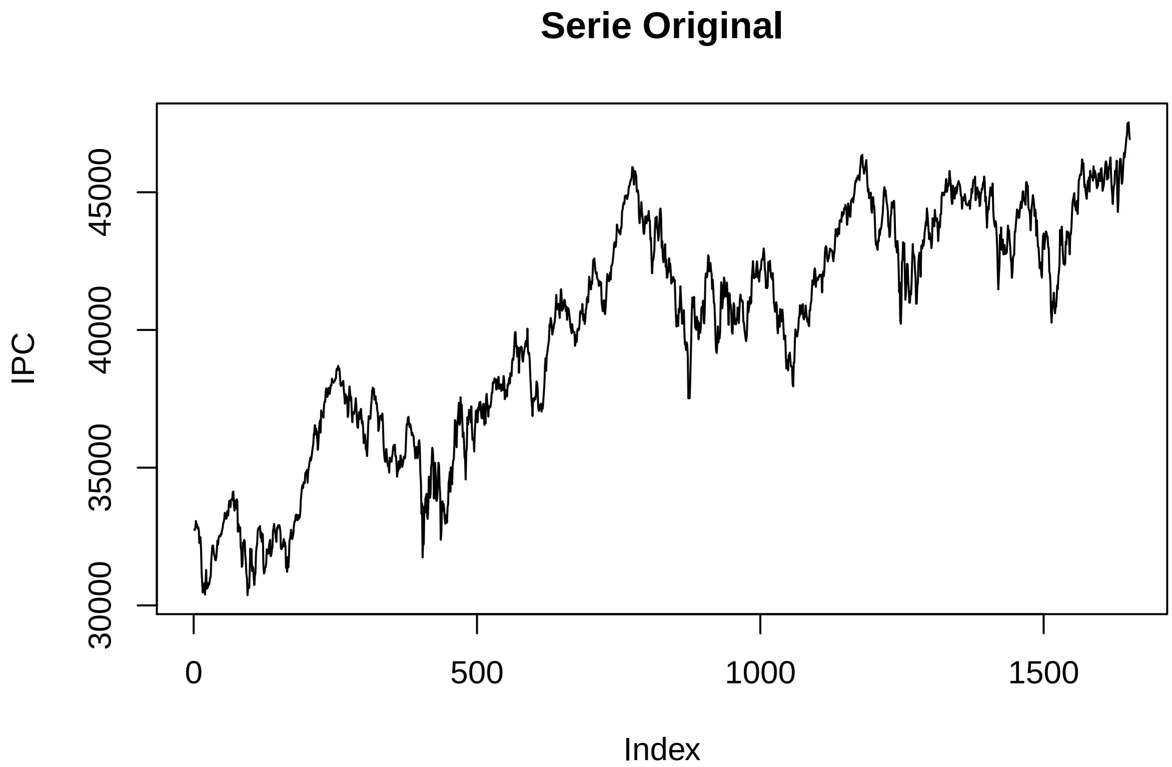
15/5/2020

```
# Cargamos las librerías a usar
library("PerformanceAnalytics")
library("quantmod")
library("car")
library("FinTS")
library("stats")
library("forecast")
library("fGarch")
library("TSA")
library("xtable")
library("tseries")
library("astsa")
library("TTR")
```

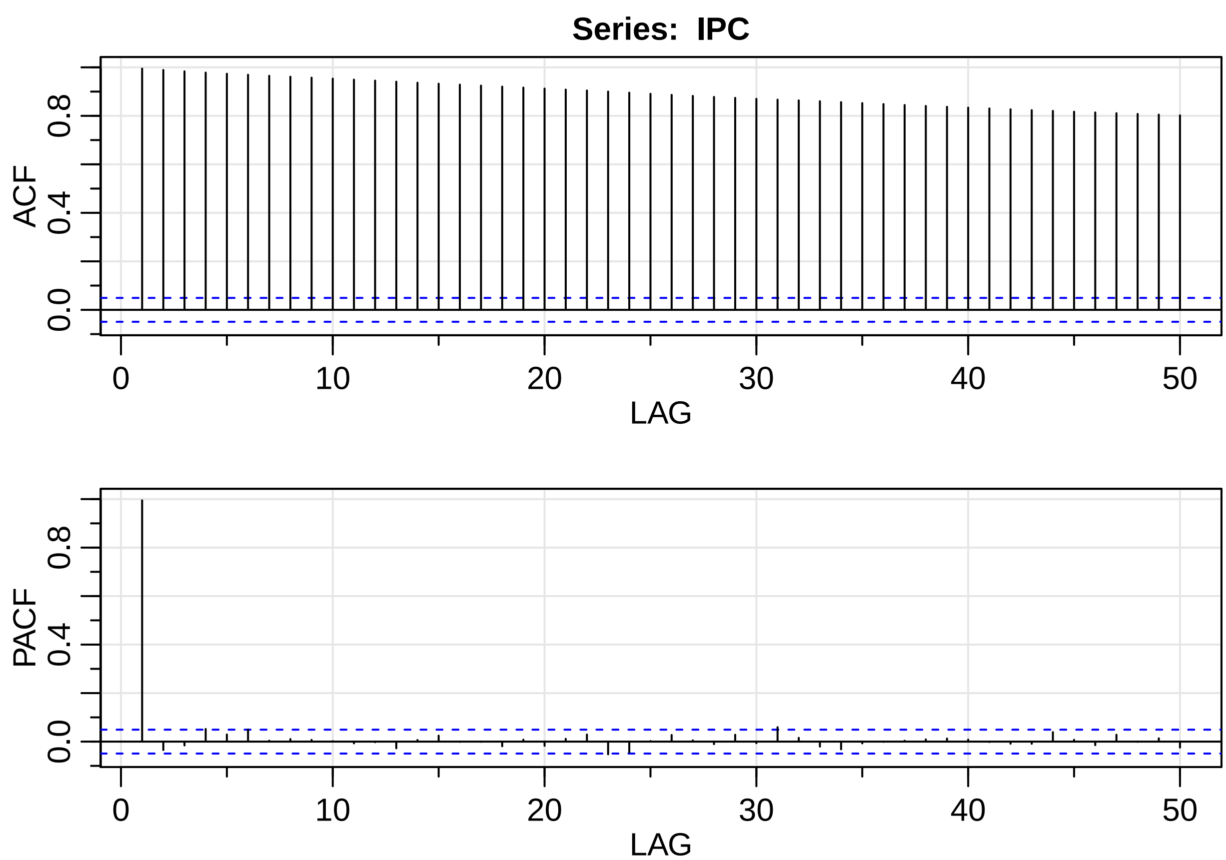
```
# Cargamos los datos
DATOS<-read.csv(file="/Users/andremarxpuentearevalo/Documents/FC/ST/Excel/IPC_MEX_2010_2016.csv",header=

# Extraemos los precios del cierre
IPC<-(DATOS)[,"CIERRE"]

# Obtenemos las estadísticas básicas de los datos
summary(IPC)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   30368   36863   40830   40008   43631   47537
```

```
# Graficamos los datos
plot(IPC,type="l", main = "Serie Original")
```

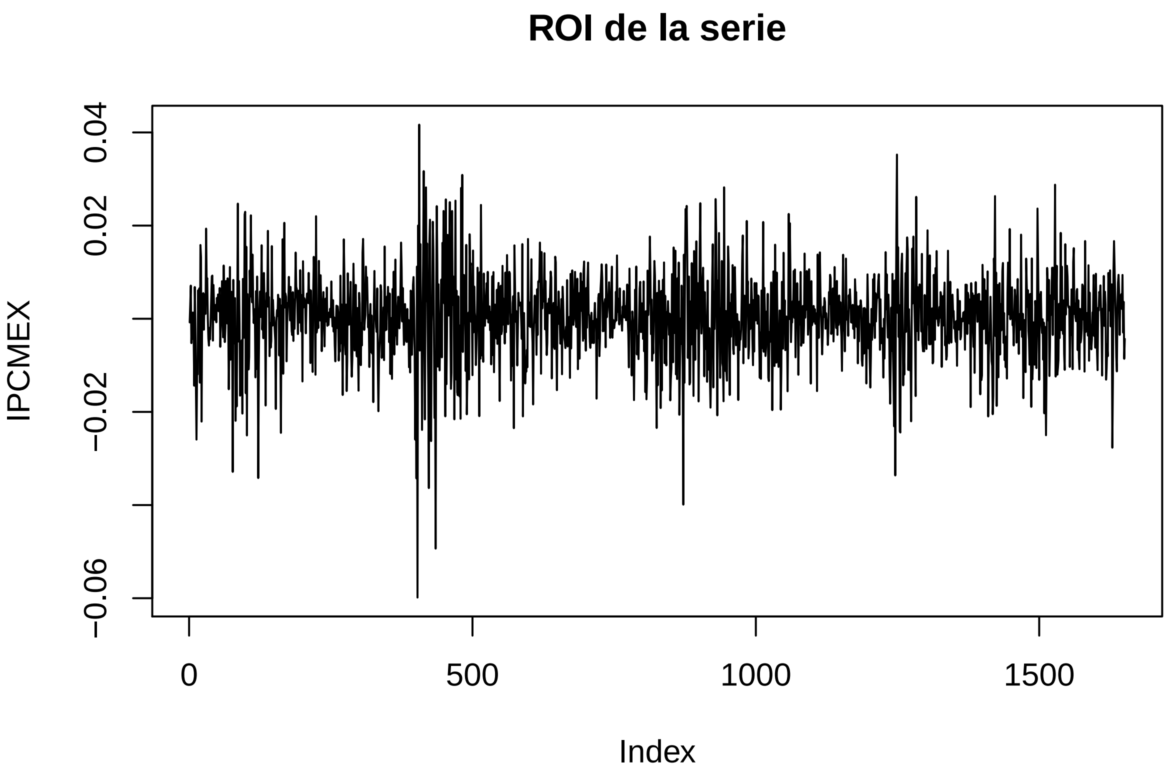



```
# Obtenemos y graficamos las autocorrelaciones
acf2(IPC,50)
```



Analizando las graficas obtenidas hasta el momento, nos damos cuenta que las autocorrelaciones no nos dicen mucho y por otro lado, notamos que es poque la serie presenta una varianza inestable y tiene tendencia, por lo que le sacaremos logaritmo para estabilizar la varianza y una diferencia para erradicar la tendencia.

```r
IPCMEX<-log(IPC)
IPCMEX<-diff(IPCMEX)

plot(IPCMEX,type="l", main = "ROI de la serie")
```



Ahora, aplicaré la prueba de hipótesis de **Ljung-Box**, la cual, tiene las siguientes hipótesis:

$H_0 : \rho_t = 0$ vs $H_1 : \rho_t \neq 0$

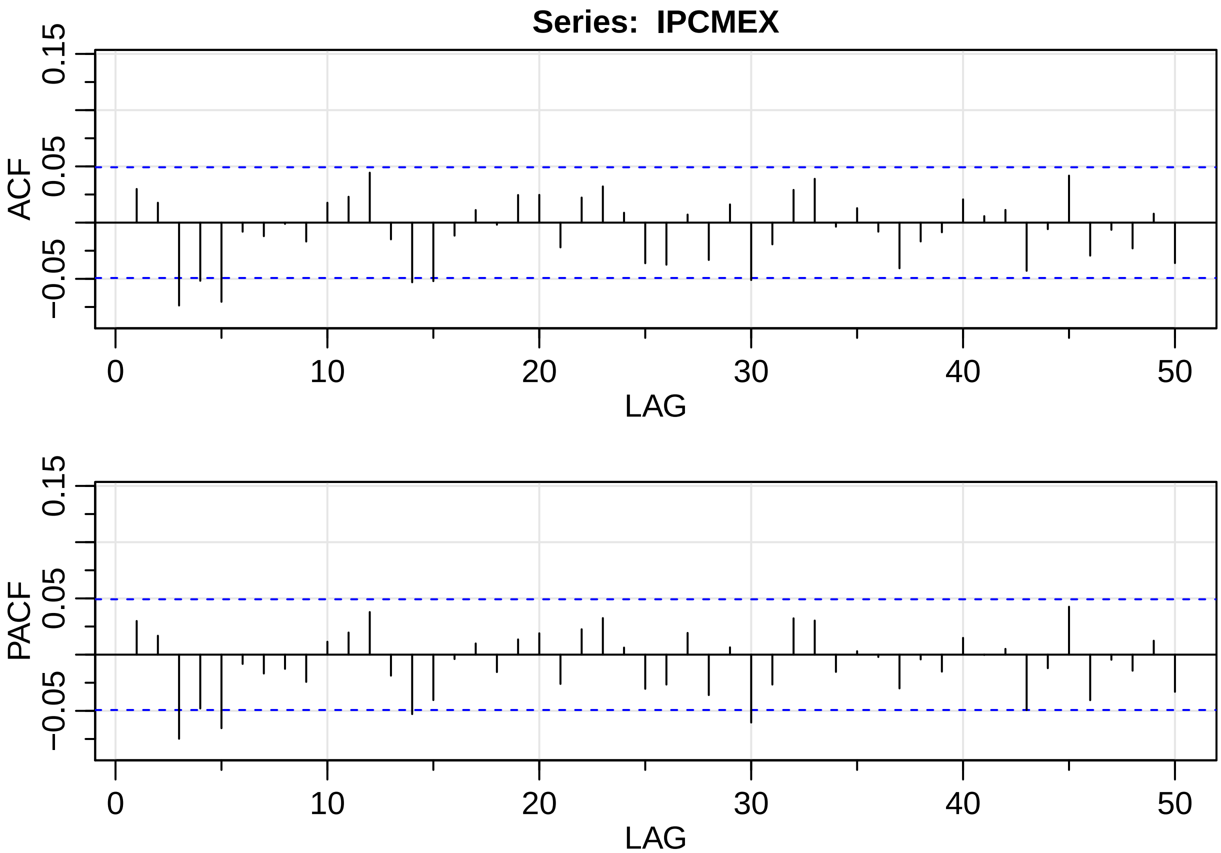
```r
Box.test(coredata(IPCMEX), type="Ljung-Box", lag=12)
```

```
##
##  Box-Ljung test
##
## data:  coredata(IPCMEX)
## X-squared = 29.144, df = 12, p-value = 0.003751
```

Dado que el p-value obtenido en la prueba es menor que el nivel de significancia, se rechaza la hipótesis nula, es decir, los datos no distribuyen en forma independiente.

Dado lo anterior, continuamos analizando la serie, calculamos las autocorrelaciones de la serie que ya tiene el logaritmo y la diferencia aplicada.

```r
acf2(IPCMEX,50)
```



Con base en lo observado en estas autocorrelaciones, comienzo a proponer modelos ARIMA.

```r
# Primero veremos los modelos propuestos por la profesora
arima110 <- arima(IPCMEX, order=c(1, 0, 0))
arima011 <- arima(IPCMEX, order=c(0, 0, 1))
arima111 <- arima(IPCMEX, order=c(1, 0, 1))
arima113 <- arima(IPCMEX, order=c(1, 0, 3))

# Los modelos que yo propongo observando las autocorrelaciones son
arima202 <- arima(IPCMEX, order=c(2, 0, 2))
arima303 <- arima(IPCMEX, order=c(3, 0, 3))
arima203 <- arima(IPCMEX, order=c(2, 0, 3))
arima003 <- arima(IPCMEX, order=c(0, 0, 3))
arima301 <- arima(IPCMEX, order=c(3, 0, 1))
arima302 <- arima(IPCMEX, order=c(3, 0, 2))

# Creamos una tabla con los AIC de los mdoelos
aicProfa = c(arima110$aic, arima011$aic, arima111$aic, arima113$aic, "NA", "NA")
ModelosProfa <- c("ARIMA(1,0,0)", "ARIMA(0,0,1)", "ARIMA(1,0,1)", "ARIMA(1,0,3)", "NA", "NA")
aicAndre <- c(arima202$aic, arima303$aic, arima203$aic, arima003$aic,
              arima301$aic, arima302$aic)
ModelosAndre <- c("ARIMA(2,0,2)", "ARIMA(3,0,3)", "ARIMA(2,0,3)", "ARIMA(0,0,3)",
                  "ARIMA(3,0,1)", "ARIMA(3,0,2)")
tabla <- data.frame(ModelosProfa,aicProfa,ModelosAndre,aicAndre)
tabla
```

```
##   ModelosProfa          aicProfa ModelosAndre  aicAndre
## 1 ARIMA(1,0,0) -10773.4361705346 ARIMA(2,0,2) -10782.93
## 2 ARIMA(0,0,1) -10773.3812211126 ARIMA(3,0,3) -10785.12
## 3 ARIMA(1,0,1) -10771.5092333319 ARIMA(2,0,3) -10784.91
## 4 ARIMA(1,0,3) -10786.1091262466 ARIMA(0,0,3) -10778.22
## 5           NA                NA ARIMA(3,0,1) -10787.03
## 6           NA                NA ARIMA(3,0,2) -10784.92
```

Observando la tabla generada con los modelos, nos damos cuenta que el que tiene el menor AIC es el modelo ARIMA(3, 0, 1), propuesto por mi, minimizando más el AIC del ARIMA(1, 0, 3) propuesto por la profesora.

Obtendremos el resumen del mejor modelo:

```r
# Resumen de mi emjor modelo
summary(arima301)
```

```
##
## Call:
## arima(x = IPCMEX, order = c(3, 0, 1))
##
## Coefficients:
##          ar1      ar2      ar3      ma1  intercept
##       0.7202  -0.0012  -0.0889  -0.6976      2e-04
## s.e.  0.1182   0.0305   0.0265   0.1173      2e-04
##
## sigma^2 estimated as 8.46e-05:  log likelihood = 5398.52,  aic = -10787.03
##
## Training set error measures:

## Warning in trainingaccuracy(f, test, d, D): test elements must be within
## sample

##               ME RMSE MAE MPE MAPE
## Training set NaN  NaN NaN NaN  NaN
```

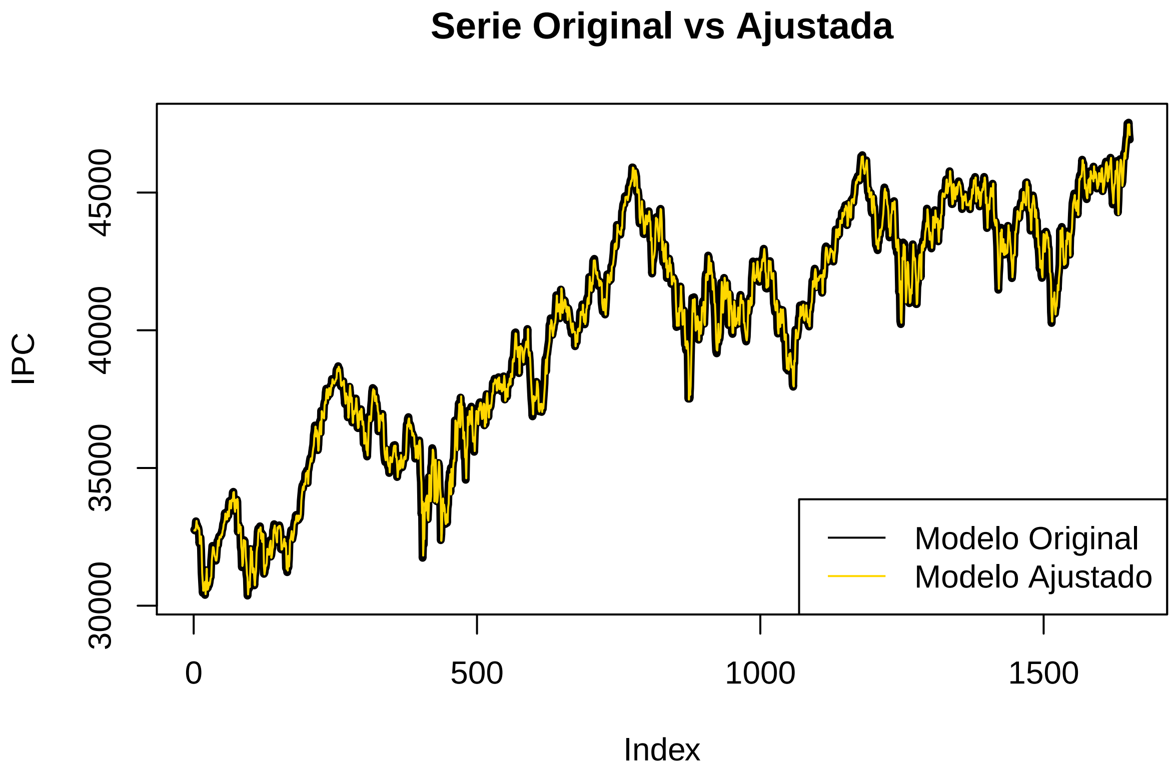
```r
# Resumen del mejor modelo de la profesora
summary(arima113)
```

```
##
## Call:
## arima(x = IPCMEX, order = c(1, 0, 3))
##
## Coefficients:
##          ar1      ma1      ma2      ma3  intercept
##       0.6831  -0.6613  -0.0023  -0.0850      2e-04
## s.e.  0.1190   0.1197   0.0296   0.0257      2e-04
##
## sigma^2 estimated as 8.465e-05:  log likelihood = 5398.05,  aic = -10786.11
##
## Training set error measures:

## Warning in trainingaccuracy(f, test, d, D): test elements must be within
## sample

##               ME RMSE MAE MPE MAPE
## Training set NaN  NaN NaN NaN  NaN
```

En el resumen del modelo, podemos darnos cuenta que el modelo estima una $\hat{\sigma}^2$ muy pequeña y podemos ver los coeficientes del modelo.

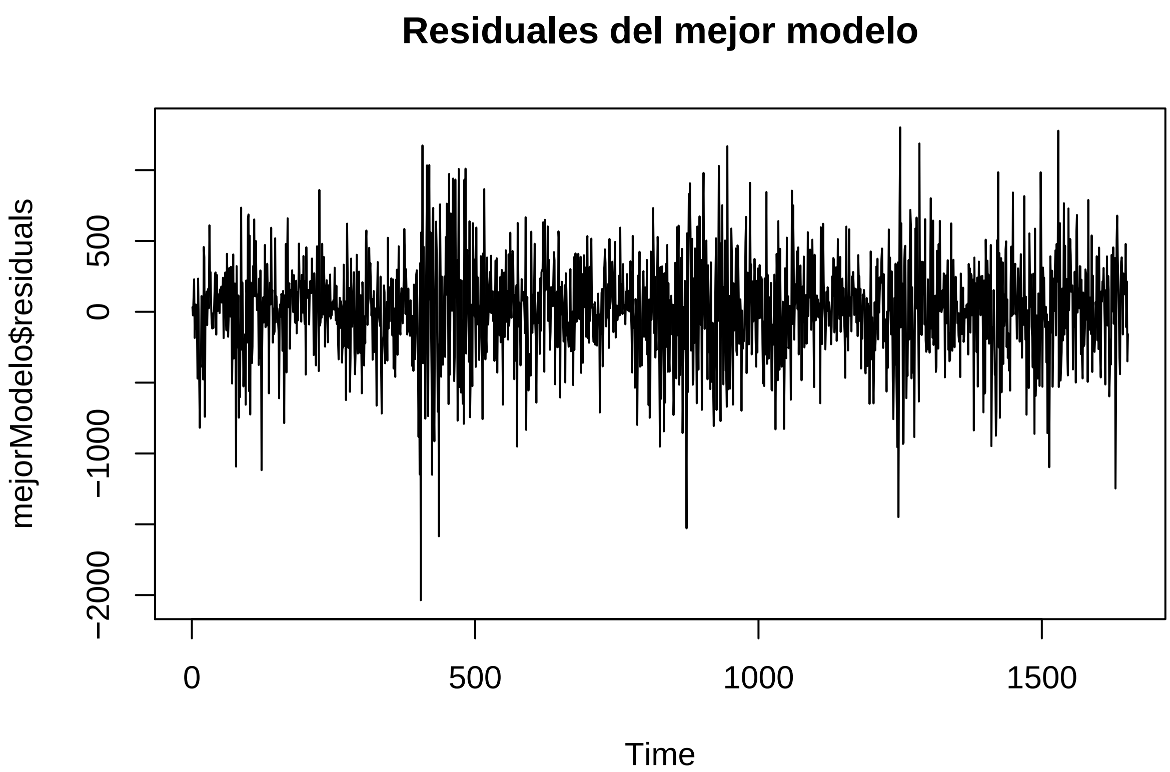
Ahora, para ajustar sobre los datos de la serie original, la cual, no tiene logaritmo ni diferencias, ajusto el modelo ARIMA(3, 1, 1) donde el 1 de en medio, hace referencia a que se le hace una diferencia a la serie.

```
mejorModelo <- arima(IPC, order = c(3,1,1))

# Obtengo los valores ajustados
ajuste <- fitted.values(mejorModelo)

# Finalmente grafico la serie original vs la ajustada
plot(IPC, type = "l", main = "Serie Original vs Ajustada", lwd=4)
lines(ajuste, col = "gold", lwd=1, lty = 1)
legend("bottomright", legend = c("Modelo Original", "Modelo Ajustado"), col = c("black", "gold"), lty =
```



Obtendremos la gráfica de los residuales del mejor modelo, para ver como se comportan.

```
plot(mejorModelo$residuals, main = "Residuales del mejor modelo")
```



Dado que tiene unos picos muy extraños, me hace pensar que se podría modelar estos mediante su varianza.

Por lo que aplicaremos la prueba de hipótesis de Multiplicadores de Lagrange, la cual, tiene las siguientes hipótesis:

H_0 : La serie no presenta efecto ARCH vs H_1 : La serie presenta efecto ARCH

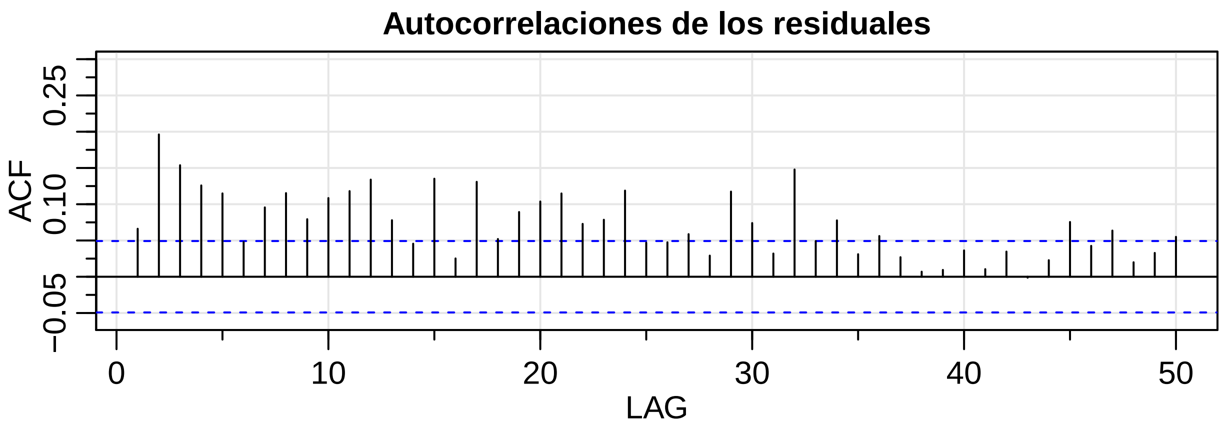
```
ArchTest(IPCMEX)
```

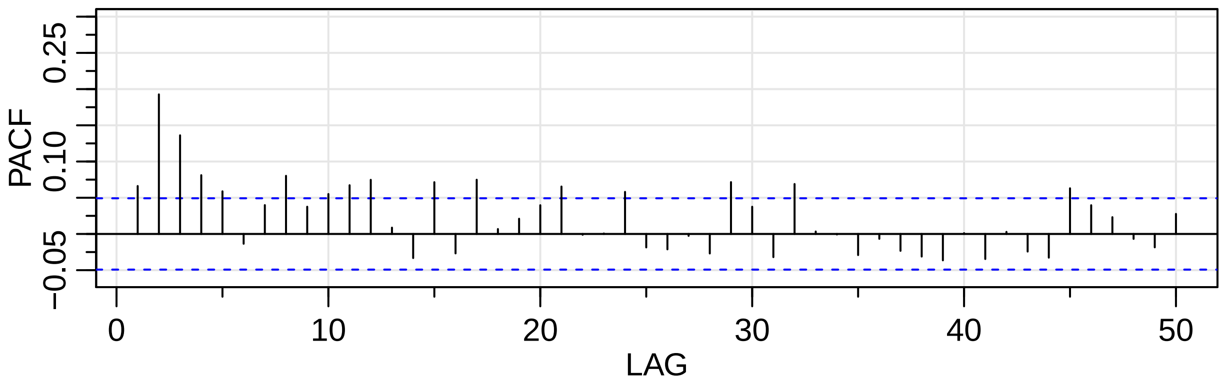
```
##
##  ARCH LM-test; Null hypothesis: no ARCH effects
##
## data:  IPCMEX
## Chi-squared = 175.9, df = 12, p-value < 2.2e-16
```

Dado que el p-value es menor que el nivel de significacia $\alpha = 0.05$, se rechaza la hipótesis nula, es decir, la serie sí presenta efecto ARCH.

A continuación obtendré las autocorrelaciones de los residuales del modelo ARIMA(3,0,1), esto porque ahora quiero modelar el ROI de la serie original.

```
# Obtengo los resuales del modelo y los elevo al cuadrado
res_arima301 <- arima301$res
res_arima301_2 <- arima301$res^2

acf2(res_arima301_2,50, main = "Autocorrelaciones de los residuales")
```






```r
# Propongo mis modelos
GARCH01 <- garch(res_arima301,order=c(0,1),trace=F)
GARCH02 <- garch(res_arima301,order=c(0,2),trace=F)
GARCH03 <- garch(res_arima301,order=c(0,3),trace=F)
GARCH11 <- garch(res_arima301,order=c(1,1),trace=F)
GARCH15 <- garch(res_arima301,order=c(1,5),trace=F)
GARCH117 <- garch(res_arima301,order=c(11,7),trace=F)

# Obtengo los AIC de los modelos
aicGARCH01 <- AIC(GARCH01)
aicGARCH02 <- AIC(GARCH02)
aicGARCH03 <- AIC(GARCH03)
aicGARCH11 <- AIC(GARCH11)
aicGARCH15 <- AIC(GARCH15)
aicGARCH117 <- AIC(GARCH117)
```

```r
# Obtengo los residuales del mejor modelo de la profesora
res_arima111 <- arima111$res
res_arima111_2 <- arima111$res^2

# Modelos propuestos por la profesora
arch07=garch(res_arima111,order=c(0,7),trace=F)
arch14=garch(res_arima111,order=c(0,14),trace=F)
garch11=garch(res_arima111,order=c(1,1),trace=F)
garch77=garch(res_arima111,order=c(7,7),trace=F)
```

```
## Warning in garch(res_arima111, order = c(7, 7), trace = F): singular
## information
```

```r
garch714=garch(res_arima111,order=c(7,14),trace=F)

# Obtiene sus aic
aicarch07=AIC(arch07)
aicarch14=AIC(arch14)
aicgarch11=AIC(garch11)
aicgarch77=AIC(garch77)
aicgarch714=AIC(garch714)

# Creamos una tabla para comparar los mejores modelos
GARCHdeProfesora <-c("aicarch07","aicarch14","aicgarch11",
                     "aicgarch77","aicgarch714", "NA")
aicProfesora <- c(aicarch07,aicarch14,aicgarch11,
                             aicgarch77,aicgarch714, "NA")

GARCHdeAndre <- c("GARCH01", "GARCH02", "GARCH03",
                  "GARCH11","GARCH15", "GARCH117")
aicAndre <- c(aicGARCH01,aicGARCH02,aicGARCH03,
              aicGARCH11,aicGARCH15,aicGARCH117)

tabla2 <- data.frame(GARCHdeProfesora, aicProfesora, GARCHdeAndre, aicAndre)
tabla2
```

```
##   GARCHdeProfesora      aicProfesora GARCHdeAndre  aicAndre
## 1        aicarch07  -10909.306977598      GARCH01 -10807.81
## 2        aicarch14 -10897.0450221852      GARCH02 -10881.90
## 3       aicgarch11 -10993.9935129654      GARCH03 -10892.30
## 4       aicgarch77 -10933.6300844178      GARCH11 -10996.66
## 5      aicgarch714  3034.31630405909      GARCH15 -10928.78
## 6               NA                NA     GARCH117 -10890.13
```

El modelo con el que la profesra se quedó fue con un GARCH(7,14), el cual tiene el siguiente resumen:

```r
summary(garch714)
```

```
##
## Call:
## garch(x = res_arima111, order = c(7, 14), trace = F)
##
## Model:
## GARCH(7,14)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.060664 -0.004875  0.000109  0.005363  0.041798
##
## Coefficient(s):
##      Estimate  Std. Error  t value Pr(>|t|)
## a0  9.880e-01   1.985e+00    0.498    0.619
## a1  5.002e-02   2.802e+02    0.000    1.000
## a2  5.002e-02   2.801e+02    0.000    1.000
## a3  5.002e-02   2.822e+02    0.000    1.000
## a4  5.002e-02   2.847e+02    0.000    1.000
## a5  5.002e-02   2.849e+02    0.000    1.000
## a6  5.001e-02   2.852e+02    0.000    1.000
```

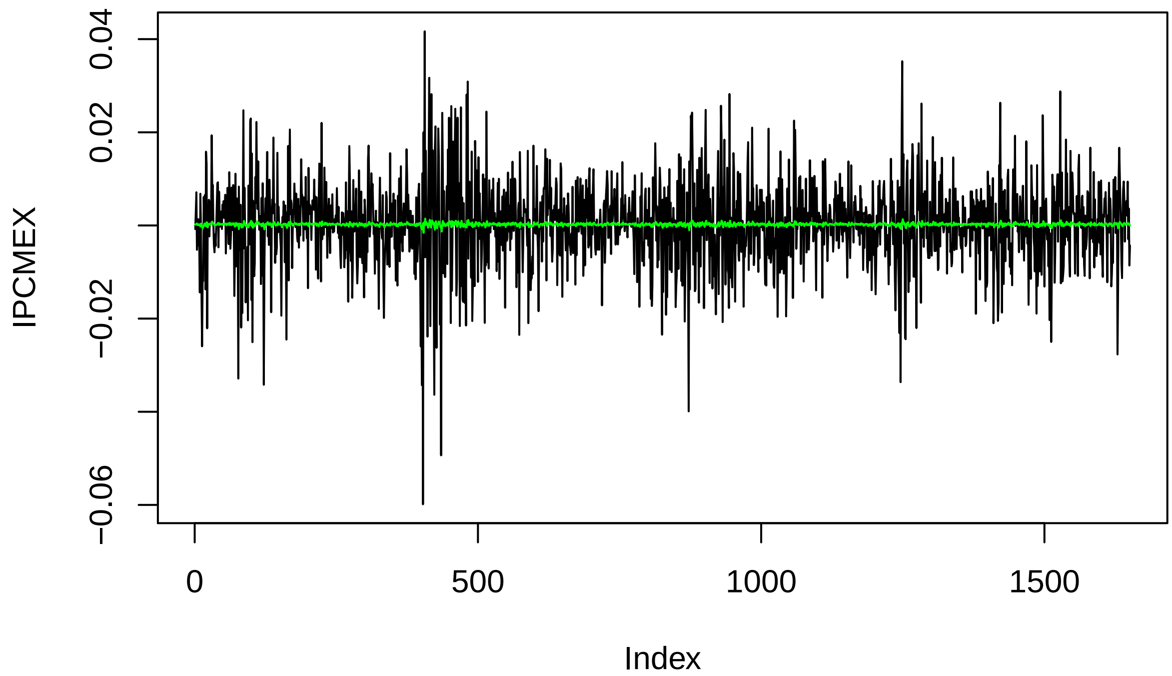
```
## a7  5.002e-02   2.849e+02    0.000    1.000
## a8  5.001e-02   2.849e+02    0.000    1.000
## a9  5.002e-02   2.841e+02    0.000    1.000
## a10 5.001e-02   2.837e+02    0.000    1.000
## a11 5.002e-02   2.834e+02    0.000    1.000
## a12 5.002e-02   2.809e+02    0.000    1.000
## a13 5.001e-02   2.789e+02    0.000    1.000
## a14 5.001e-02   2.789e+02    0.000    1.000
## b1  4.548e-15   2.838e+00    0.000    1.000
## b2  3.824e-05   2.838e+00    0.000    1.000
## b3  7.521e-05   2.838e+00    0.000    1.000
## b4  1.113e-04   2.839e+00    0.000    1.000
## b5  1.464e-04   2.840e+00    0.000    1.000
## b6  1.815e-04   2.841e+00    0.000    1.000
## b7  2.155e-04   2.012e+00    0.000    1.000
##
## Diagnostic Tests:
##  Jarque Bera Test
##
## data:  Residuals
## X-squared = 519.15, df = 2, p-value < 2.2e-16
##
##
##  Box-Ljung test
##
## data:  Squared.Residuals
## X-squared = 10.456, df = 1, p-value = 0.001222
```

Notemos que los coeficientes obtenidos en su modelo son todos no significativos.

El ajuste que ella propuso es el siguiente:

```
ht.garch714=garch714$fit[,1]^2

estimaciones_arima111=fitted.values(arima111)

inf_garch714= estimaciones_arima111-1.96*sqrt(ht.garch714)
sup_garch714= estimaciones_arima111+1.96*sqrt(ht.garch714)


plot(IPCMEX, main = "Grafico de los retornos de la serie con los intervalos de confianza del modelo Arc
lines(inf_garch714,col='red')
lines(sup_garch714,col='blue')
lines(estimaciones_arima111,col='green')
```

fico de los retornos de la serie con los intervalos de confianza del mod



En esta gráfica, podemos observar que el ajuste propuesto por la profesora fue muy bueno, ya que el intervalo en el que intenta atrapar a la serie es tan grande que no se muestra en la gráfica.

Finalmente, analizando mis modelos propuestos, me quedo con el GARCH(0,1)=ARCH(1) que es el que tiene un AIC más cercano a 0 y todos sus coeficientes son significativos, como se muestra en el resumen siguiente.

Nota: De los modelos que aparecen en mi tabla el de menor AIC es el GARCH(11, 7), pero este modelo genera NA en los coeficientes, es decir, no es bueno. Mientras que modelos como el GARCH(1, 1) presentan un AIC que su distancia al 0 es mayor que el modelo que yo propongo.

Vemos el resumen del mejor modelo que propongo:

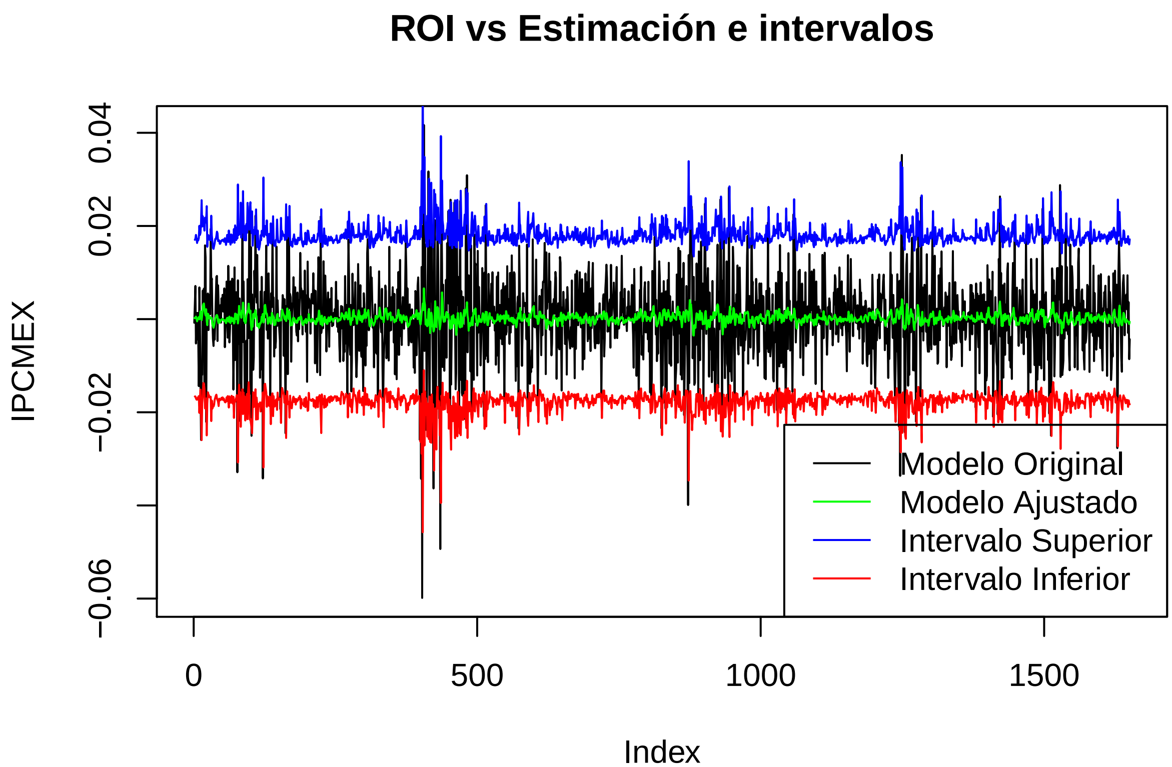
```r
summary(GARCH01)
```

```
##
## Call:
## garch(x = res_arima301, order = c(0, 1), trace = F)
##
## Model:
## GARCH(0,1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.6548 -0.5449  0.0325  0.5776  3.9108
##
## Coefficient(s):
##     Estimate  Std. Error  t value Pr(>|t|)
## a0 7.345e-05   2.672e-06   27.494   < 2e-16 ***
## a1 1.436e-01   3.073e-02    4.673 2.96e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Diagnostic Tests:
##  Jarque Bera Test
##
## data:  Residuals
## X-squared = 503.35, df = 2, p-value < 2.2e-16
##
##
##  Box-Ljung test
##
## data:  Squared.Residuals
## X-squared = 1.0219, df = 1, p-value = 0.3121
```

Ajustaré el modelo al ROI de la serie y graficaré el ROI original con sus intervalos de confianza y el ajuste obtenido.

```
# Obtengo el ajuste
ht.garch01=GARCH01$fit[,1]^2

estimaciones_arima301=fitted.values(arima301)

# Obtengo los intervalos
inf_garch01= estimaciones_arima301-1.96*sqrt(ht.garch01)
sup_garch01= estimaciones_arima301+1.96*sqrt(ht.garch01)

# Grafico los valores
plot(IPCMEX, main = "ROI vs Estimación e intervalos",type="l")
lines(estimaciones_arima301,col='green')
lines(inf_garch01,col='red')
lines(sup_garch01,col='blue')
legend("bottomright", legend = c("Modelo Original", "Modelo Ajustado", "Intervalo Superior", "Intervalo
       col = c("black", "green", "blue", "red"), lty = c(1,1,1,1))
```



Analizando ésta gráfica, nos damos cuenta que los intervalos de confianza no distan mucho de la serie original, la encierran de manera adecuada y el ajuste del arima ajusta de manera asecuada cayendo dentro del intervalo.

Por lo tanto el modelo que yo propongo modela mejor la varianza de la serie que el de la profesora.