Alumno: Maye Mamani Victor Raul

Tabla de contenido

Alumno: Maye Mamani Victor Raul

Actividad 5

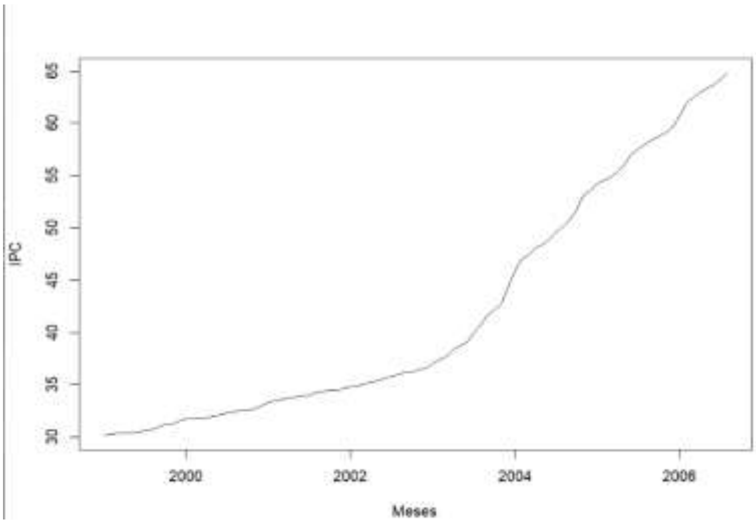
```
# Librerias necesaria
library(forecast) # Modelo ARIMA
library(tseries) # Para series de tiempo
library(TSA) # Para series de tiempo
library(urca) # Raiz Unitaria
library(ggplot2) # Para hacer gráficos
library(gridExtra)
library(dplyr) # Para la manipulación de datos
library(lmtest) # Inferencia para coeficientes estimados
library(MASS) # Transformacion de Box-Cox
library(nortest) # Pruebas de normalidad
library(strucchange) # Cambio estructural - Test de Chow
library(mFilter)
library(readxl)
library(fitdistrplus)
```

Caso 1

1 Identificación

Grafica inicial de la serie:

```
data <- read_excel("F:\\777--Programacion
repos\\Una\\r\\data\\actividad-05.xlsx",sheet = "1")
View(data)
# Gráfica de la serie
data_ts <- ts(data$interes, start = c(1999,1), frequency = 12)
plot(data_ts, xlab="Meses", ylab="IPC")
```



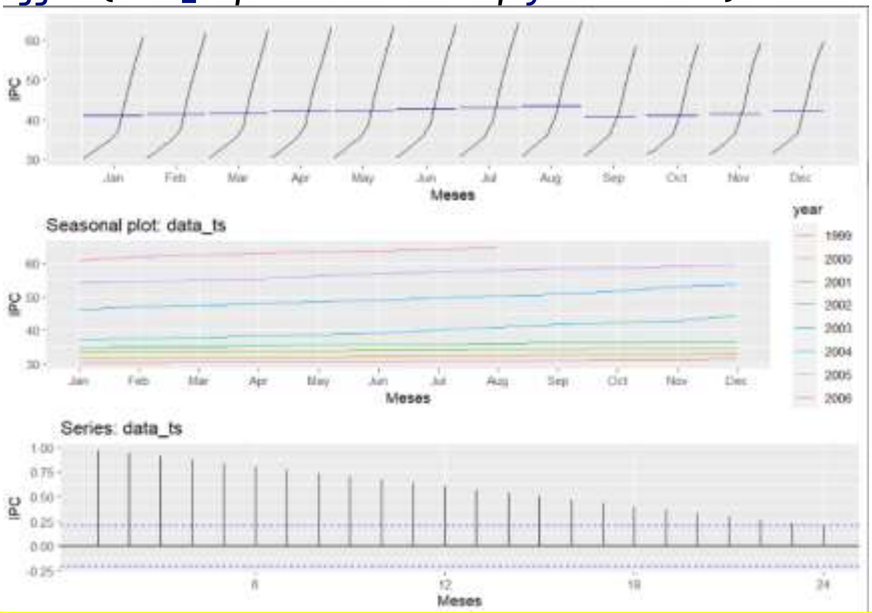
En primera instancia podemos visualizar una clara TENDENCIA creciente, la varianza no parece variar y no es estacionaria en media. Y tendremos que realizar un análisis de estacionalidad dado que la seri es trimestral

Alumno: Maye Mamani Victor Raul

1. 1 Análisis de la tendencia y la estacionalidad

1.1.1 Estacionalidad

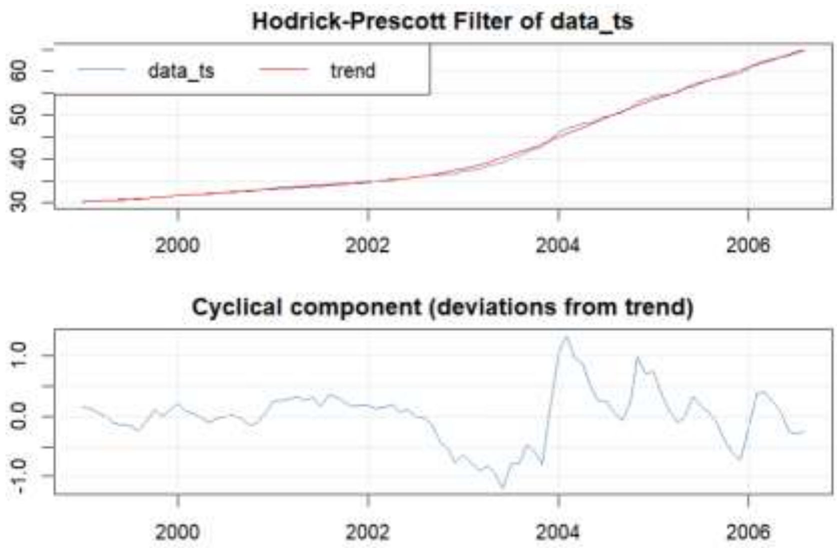
```
plot1 <- ggsubseriesplot(data_ts, xlab = "Meses", ylab = "IPC")
plot2 <- ggseasonplot(data_ts, xlab = "Meses", ylab = "IPC")
plot3 <- ggAcf(data_ts, xlab = "Meses", ylab = "IPC")
```



Vemos que la serie no presenta estacionalidad, dado que el primer grafico nos muestra se muestra a un mismo nivel, el segundo nos muestra que no hay un patrón claro y finalmente el correlograma decrece constantemente.

1.1.2 Análisis de tendencia

```
#Análisis de tendencia
lambda_hp <- 1600
data_hp <- hpfilter(data_ts, type="lambda", freq=lambda_hp)
plot(data_hp)
```



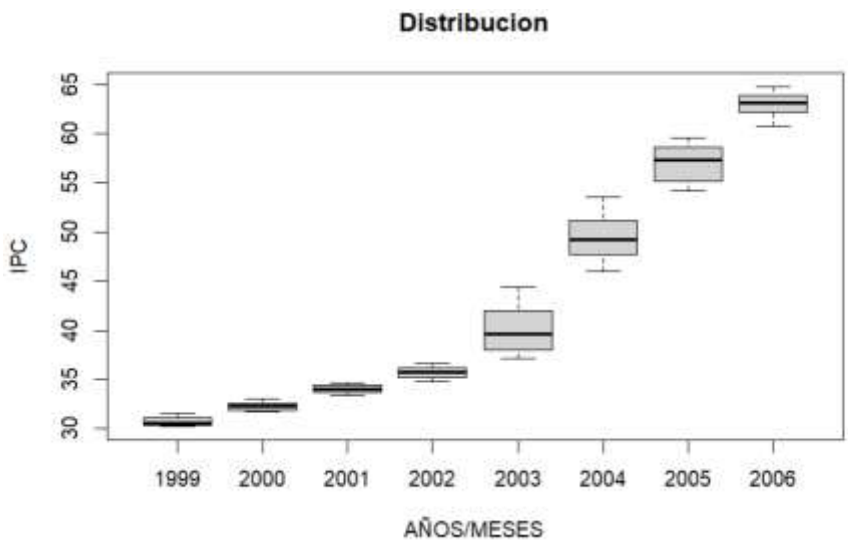
Hay una tendencia creciente y no parece haber algún patrón que se repita cíclicamente

Alumno: Maye Mamani Victor Raul

1.2 Análisis de estacionariedad

1.2.1 Estacionariedad en varianza

```
boxplot(data$IPC ~ data$Año, xlab = "AÑOS/MESES", ylab="IPC",
main="Distribucion" )
```

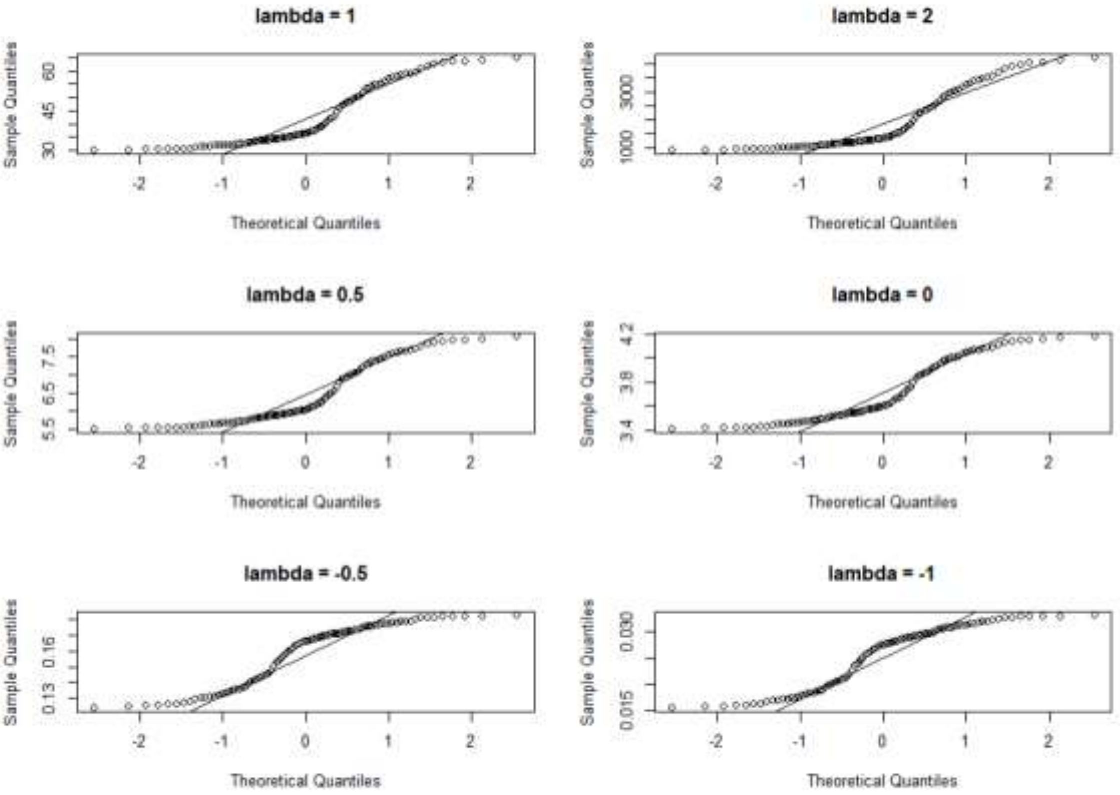


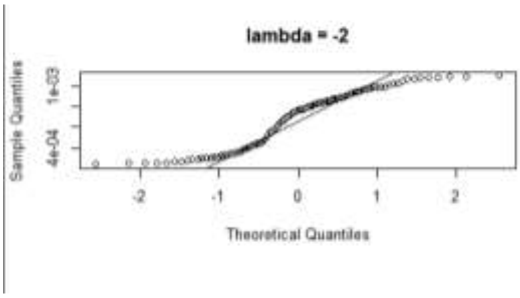
Tenemos ciertos indicios de estacionariedad en varianza, sin embargo a partir del año 2003 este presenta un cambio significativo en la serie lo que nos sugiere que no existe estacionariedad en varianza.

```
qqnorm(data_ts,main="lambda = 1")
qqline(data_ts)

t1.yt <- data_ts^2
qqnorm(t1.yt, main="lambda = 2")
qqline(t1.yt)

t3.yt <- sqrt(data_ts)
qqnorm(t3.yt, main="lambda = 0.5")
qqline(t3.yt)

t4.yt <- log(data_ts)
qqnorm(t4.yt, main="lambda = 0")
qqline(t4.yt)

t5.yt <- 1/sqrt(data_ts)
qqnorm(t5.yt, main="lambda = -0.5")
qqline(t5.yt)

t6.yt <- 1/data_ts
qqnorm(t6.yt, main="lambda = -1")
qqline(t6.yt)
```

Alumno: Maye Mamani Victor Raul

```
t7.yt <- 1/(Yt^2)
qqnorm(t7.yt, main="lambda = -2")
qqline(t7.yt)
```





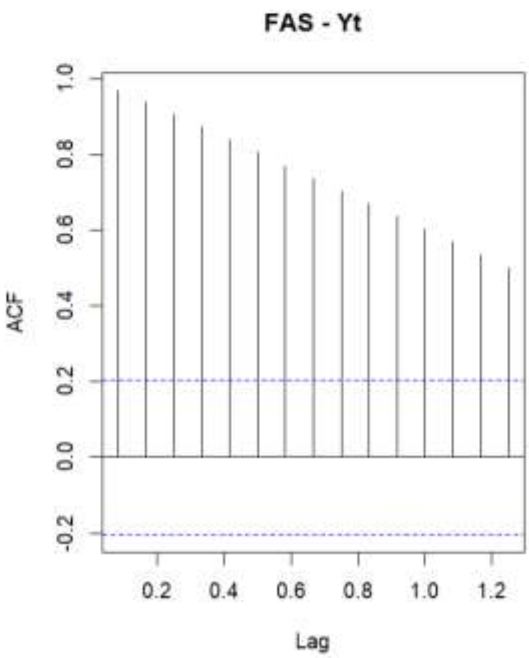
Se observa el ajuste a la distribución normal de la serie original y en las demás figuras se tiene todas las demás transformaciones, las cuales no son sensibles al valor que vaya a tomar lambda (λ)

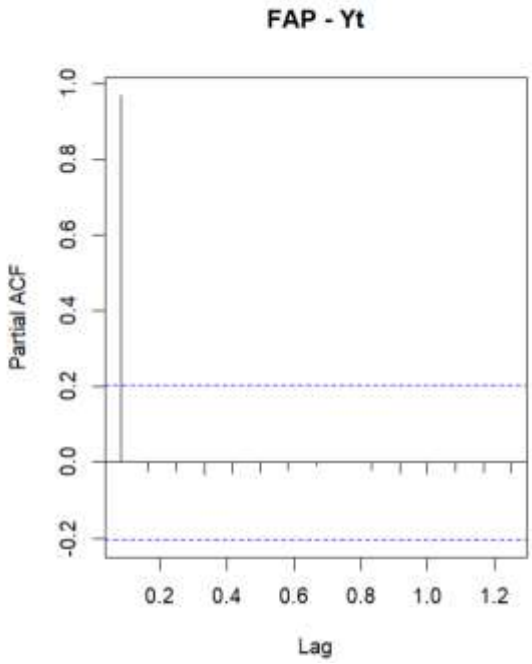
1.2.2 Estacionariedad en media

```
par(mfrow = c(1,2))
FAS <- acf(data_ts, lag.max = 15, main="FAS - Yt")
FAP <- pacf(data_ts, lag.max = 15, main="FAP - Yt")
```

FAP$acf[1]

Alumno: Maye Mamani Victor Raul






```
[1] 0.9689437
```

El FAS decrece lentamente y el primer FAP es significativo siendo mayor a 0.9, por lo que podríamos decir que la serie NO ES ESTACIONARIA

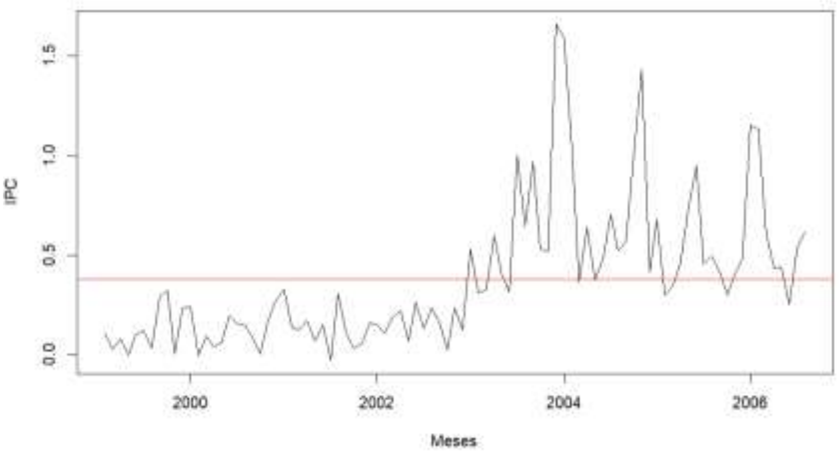
```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.194058   0.185783   1.045   0.2992    
z.lag.1     -0.008109   0.007524  -1.078   0.2842    
tt           0.007407   0.003211   2.306   0.0235 *  
z.diff.lag   0.485739   0.093634   5.188  1.4e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2499 on 86 degrees of freedom
Multiple R-squared:  0.5241,	Adjusted R-squared:  0.5075 
F-statistic: 31.57 on 3 and 86 DF,  p-value: 7.424e-14


Value of test-statistic is: -1.0777 7.7405 6.0705 

Critical values for test statistics: 
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Observamos que el T calculado (-1.0777) es MAYOR que el T critico (-3.45) por tanto se rechaza la hipótesis nula de la existencia de raíz unitaria, es **decir que la serie NO ES ESTACIONARIA**

Diferenciamos la serie

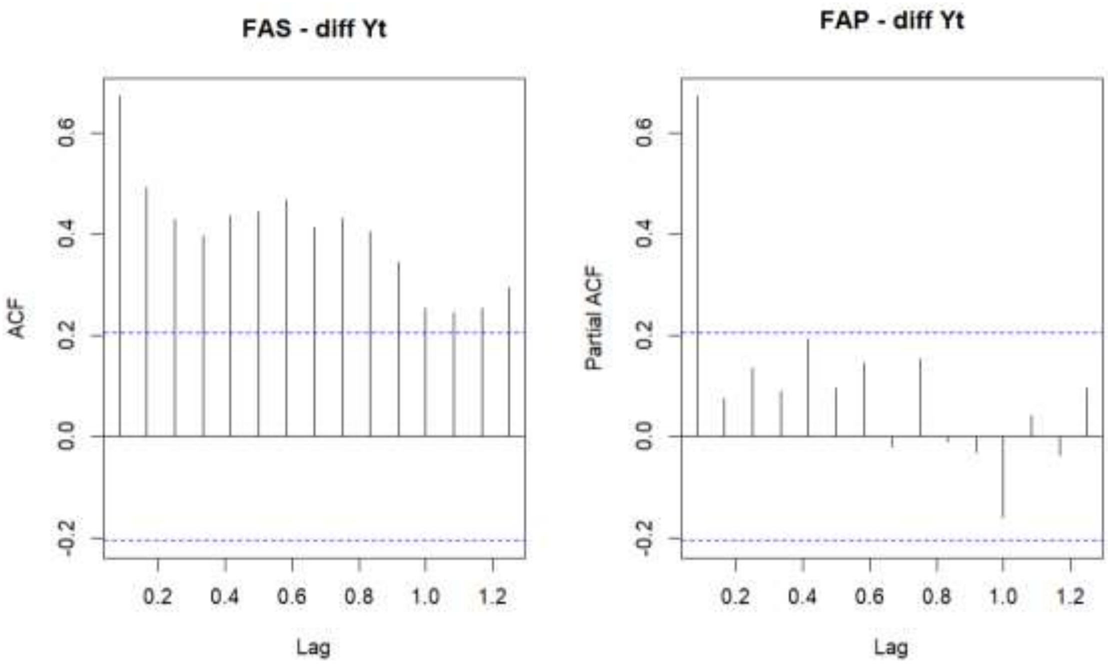
```
data_diff = diff(data_mod)
plot(data_diff, xlab="Meses", ylab="IPC")
abline(h = mean(data_diff), col = "red")
```

Alumno: Maye Mamani Victor Raul



Volvemos a analizar el correlograma

```
par(mfrow = c(1,2))
FAS <- acf(data_diff, lag.max = 15, main="FAS - diff Yt")
FAP <- pacf(data_diff, lag.max = 15, main="FAP - diff Yt")
FAP$acf[1]
```



Hay un decrecimiento lento en el FAS, sin embargó el valor del primer coeficiente de las FAP es menor a 0.9, por lo que existen indicios de que la serie NO es estacionaria en media. Es por que recurriremos a una confirmatoria para la evaluación de la estacionariedad del modelo.

```
data_adf <- ur.df(data_diff, type="trend", lags = 1)
summary(data_adf)
Residuals:
     Min       1Q   Median       3Q      Max
-0.58622 -0.14484 -0.02920  0.07888  1.15691
```

Alumno: Maye Mamani Victor Raul

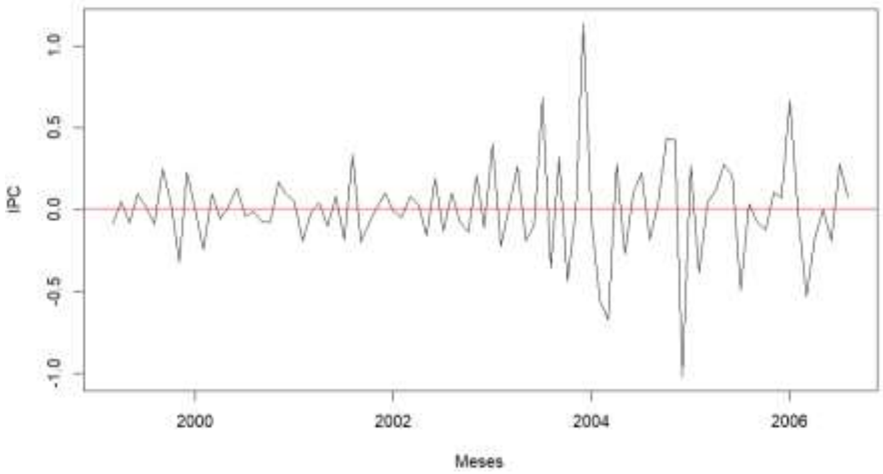
```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.008055   0.054982   0.147  0.88387
z.lag.1     -0.530904   0.110180  -4.819 6.25e-06 ***
tt           0.004355   0.001396   3.120  0.00247 **
z.diff.lag   0.035425   0.108660   0.326  0.74521
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2529 on 85 degrees of freedom
Multiple R-squared:  0.2566,   Adjusted R-squared:  0.2304
F-statistic: 9.782 on 3 and 85 DF,  p-value: 1.303e-05


Value of test-statistic is: -4.8185 7.7709 11.6148

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Observamos que el T calculado (-4.8185) es MENOR que el T critico (-3.45) por tanto se acepta la hipótesis nula de la existencia de raíz unitaria, es decir que la serie ES ESTACIONARIA

Teniendo en cuenta que el grafico de la serie y el correlograma nos indica que NO es estacionaria, pero la prueba de raíz unitaria indica que es estacionaria. Tomamos la decisión de volver a hacer una diferenciación mas para una mejor estabilización de la serie

Segunda Diferenciación

```
par(mfrow = c(1,1))
data_diff2 = diff(data_diff)
plot(data_diff2, xlab="Meses", ylab="IPC")
abline(h = mean(data_diff2), col = "red")
```

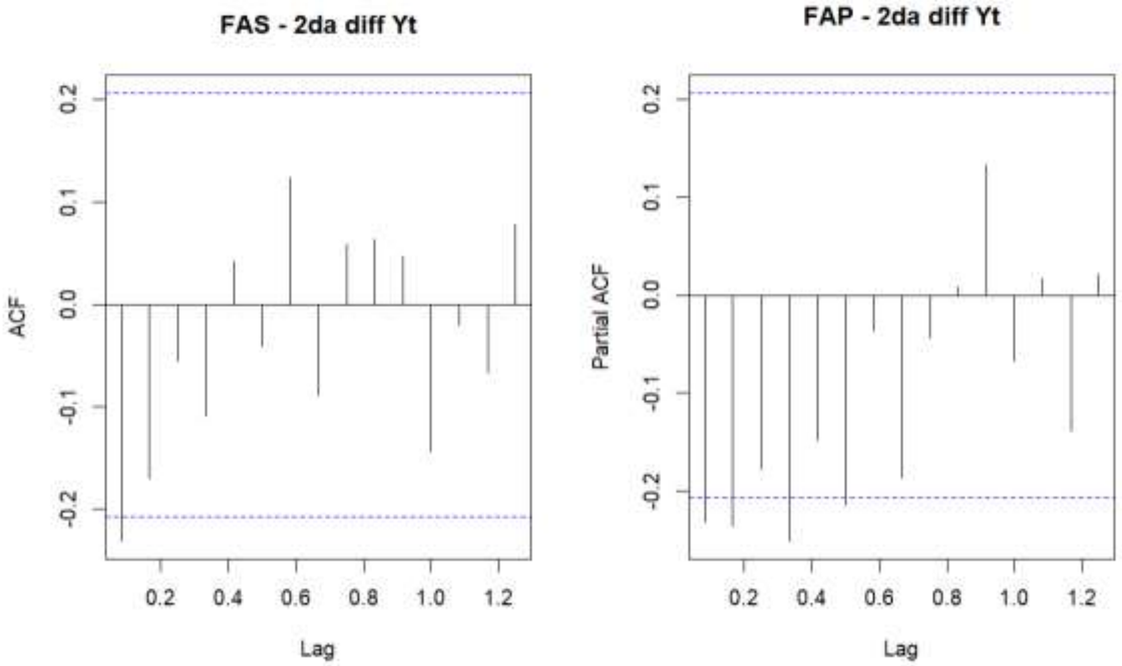


Análisis de estacionariedad

```
par(mfrow = c(1,2))
FAS <- acf(data_diff2, lag.max = 15, main="FAS - 2da diff Yt")
FAP <- pacf(data_diff2, lag.max = 15, main="FAP - 2da diff Yt")
```

Alumno: Maye Mamani Victor Raul

```r
FAP$acf[1]
```



El FAS decrece rápidamente a cero, y el primer coeficiente del FAP es menor a 0.9, por lo que la serie es estacionaria, pero de igual forma haremos la prueba de raíz unitaria.

```r
data_adf <- ur.df(data_diff2, type="drift", lags = 1)
summary(data_adf)
```

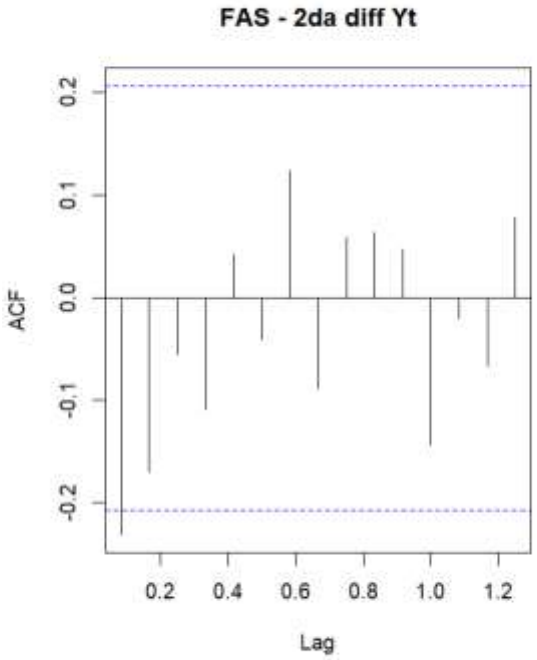
```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.008139   0.029542   0.276    0.784
z.lag.1     -1.527289   0.166203  -9.189 2.23e-14 ***
z.diff.lag   0.240127   0.105992   2.266    0.026 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2771 on 85 degrees of freedom
Multiple R-squared:  0.6372,	Adjusted R-squared:  0.6287 
F-statistic: 74.66 on 2 and 85 DF,  p-value: < 2.2e-16


Value of test-statistic is: -9.1893 42.2251 

Critical values for test statistics: 
      1pct  5pct 10pct
tau2 -3.51 -2.89 -2.58
phi1  6.70  4.71  3.86
```

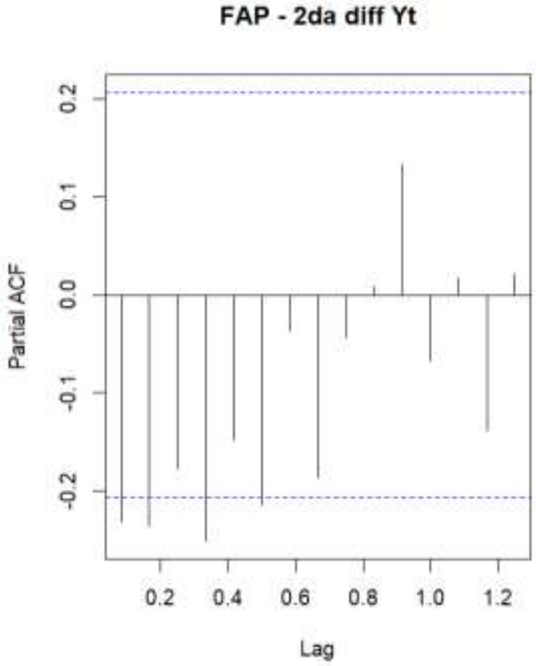
Observamos que el T calculado (-9.1893) es MENOR que el T critico (-2.89) por tanto se acepta la hipótesis nula de la existencia de raíz unitaria, es decir que la serie ES ESTACIONARIA

Alumno: Maye Mamani Victor Raul

1.3 Identificación del modelo estacionario

1.3.1 Identificación de las órdenes p y q





Observando el FAS vemos que decrece rápidamente y tiene un coeficiente significativo MA(1). Y el FAP decrece rápidamente a cero mostrando un patrón sinusoidal y tiene 2 coeficientes diferentes de cero por lo que también planteamos un AR(3). Adicionalmente planteamos un ARIMA(1,2).

1.3.2 Inclusión del término independiente (δ) o intercepto

```
Z <- mean(data_diff2)
Co <- var(data_diff2)
Tn <- length(data_diff2)
Ta <- Tn - 1
Sigma <- Co/Ta
t <- Z/Sigma
tt <- qt(1-0.05/2,Ta-1)
pruebaT <- c(t, tt)
names(pruebaT) <- c("t-calculado","t-critico")
pruebaT
```

```
t-calculado   t-critico
   6.085177    1.987290
```

Se rechaza la hipótesis nula y se incluye la constante.

Resumiendo, se proponen los siguientes modelos

ARIMA(3,2,0)

ARIMA(0,2,1)

ARIMA(1,2,2)

Alumno: Maye Mamani Victor Raul

2 Estimación

```
mod1 <- Arima(data_mod, order = c(3, 2, 0), include.constant = T)
coeftest(mod1)
```

```
z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ar1 -0.32564    0.10357 -3.1441 0.001666 **
ar2 -0.28464    0.10452 -2.7232 0.006466 **
ar3 -0.17365    0.10283 -1.6887 0.091283 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$\Delta^2 Y_t = -0.3256Y_{t-1} - 0.28464\Delta^2 Y_{t-2} - 0.1736\Delta^2 Y_{t-3} + a_t$$

```
mod2 <- Arima(data_mod, order = c(0, 2, 1), include.constant = T)
coeftest(mod2)
```

```
z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ma1 -0.67612    0.13061 -5.1767 2.259e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$\Delta^2 Y_t = a_t - 0.6761a_{t-1}$$

```
mod3 <- Arima(data_mod, order = c(1, 2, 2), include.constant = T)
coeftest(mod3)
```

```
z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ar1  0.15321    0.32614  0.4698  0.63852
ma1 -0.60038    0.32139 -1.8681  0.06175 .
ma2 -0.20687    0.23711 -0.8725  0.38294
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$\Delta^2 Y_t = 0.1532\Delta^2 Y_{t-1} + a_t - 0.6003a_{t-1} - 0.2068a_{t-2}$$

3 Validación

3.1 Análisis de los coeficientes estimados

3.1.1 Significación de los coeficientes

Para el modelo 1

$$AR(1): \phi_1 = -0.3256 \rightarrow p = 0.0000 < 0.01, altamente\ significativo$$

$$AR(2): \phi_2 = -0.28464 \rightarrow p = 0.0000 < 0.01, altamente\ significativo$$

$$AR(3): \phi_3 = -0.1736 \rightarrow p = 0.0912 > 0.01, NO\ es\ significativo$$

Para el modelo 2

$$MA(1): \theta_1 = -0.6761 \rightarrow p = 0.0000 < 0.01, altamente\ significativo$$

Para el modelo 3

$$AR(1): \phi_1 = 0.1532 \rightarrow p = 0.6385 < 0.01, altamente\ significativo$$

$$MA(1): \theta_1 = -0.6003 \rightarrow p = 0.0617 > 0.05, No\ es\ significativo$$

$$MA(2): \phi_1 = -0.2068 \rightarrow p = 0.3829 > 0.01, NO\ es\ significativo$$

3.1.2 Examen de la matriz de correlaciones entre los coeficientes

```
vcov(mod1)
```

```
            ar1         ar2         ar3
ar1 0.010726879 0.003143925 0.002536229
ar2 0.003143925 0.010925282 0.003018367
ar3 0.002536229 0.003018367 0.010574821
```

```
vcov(mod2)
```

```
           ma1
ma1 0.01705901
```

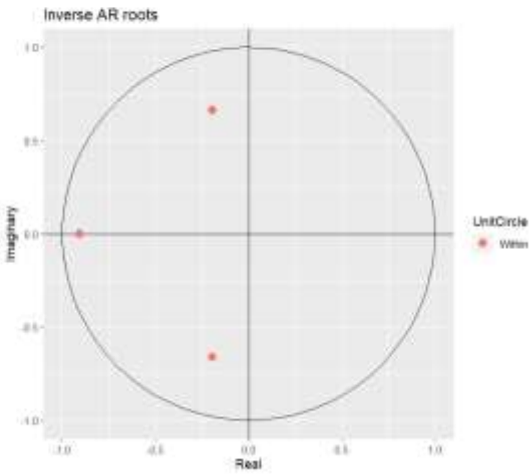
```
vcov(mod3)
```

```
            ar1         ma1         ma2
ar1  0.10636603 -0.09928611  0.07010027
ma1 -0.09928611  0.10329013 -0.07344415
ma2  0.07010027 -0.07344415  0.05621914
```

Se observa claramente que ningún coeficiente esta próximo ni cercano a 0.9, por tanto, podemos indicar que no hay problema de multicolinealidad en los modelos propuestos.

3.1.3 Condición de convergencia e invertibilidad

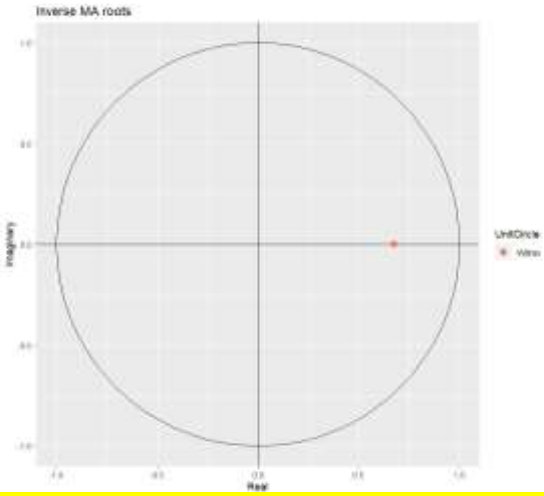
```
autoplot(mod1)
```



En la figura de raíces inversas de AR, se observa en el gráfico las raíces características se encuentran dentro del círculo, es decir que cumplen con la condición de estacionariedad para la parte autorregresiva.

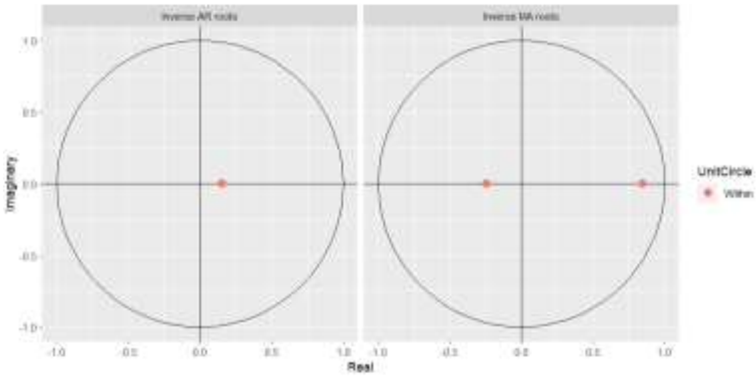
```
autoplot(mod2)
```

Alumno: Maye Mamani Victor Raul



En la figura de raíces inversas de MA, se observa en el gráfico las raíces características se encuentran dentro del círculo, es decir que cumplen con la condición de invertibilidad para la parte de media movíl.

```
autoplot(mod3)
```



Al estar los valores dentro de la circunferencia unitaria es un indicativo de que el modelo se ajusta correctamente. Tanto en su parte AR, como en su parte MA.

3.1.4 Análisis de la estabilidad

```
#Analisis de estabilidad
Chow_mod1 <- Fstats(mod1$fitted ~ 1, from = 0.60,to=0.70)
sctest(Chow_mod1)
supF test

data:  Chow_mod1
sup.F = 506.59, p-value < 2.2e-16
```

```
Chow_mod2 <- Fstats(mod2$fitted ~ 1, from = 0.60,to=0.70)
sctest(Chow_mod2)
supF test

data:  Chow_mod2
sup.F = 507.6, p-value < 2.2e-16
```

Alumno: Maye Mamani Victor Raul

```
Chow_mod3 <- Fstats(mod3$fitted ~ 1, from = 0.60,to=0.70)
sctest(Chow_mod3)
```

```
supF test

data:  Chow_mod3
sup.F = 507.35, p-value < 2.2e-16
```

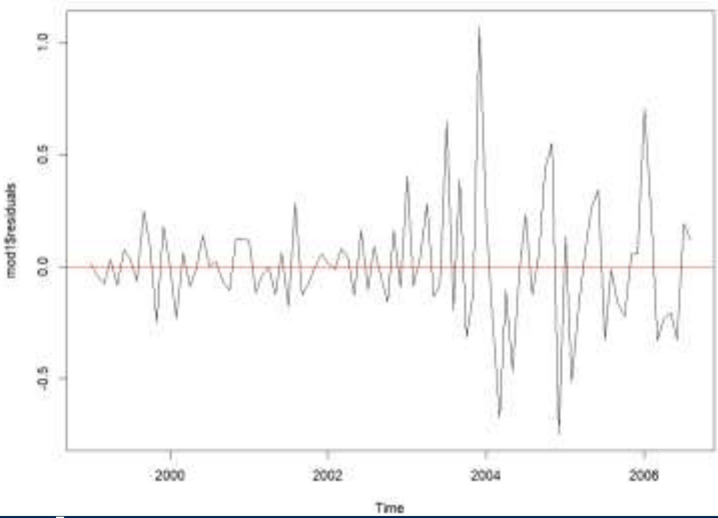
En las tres pruebas se rechaza la hipótesis nula ($p < \alpha = 0.05$), es decir, NO existe estabilidad de coeficientes

3.2 Análisis de los residuos

3.2.1 Media es igual a cero

Modelo 1

```
plot(mod1$residuals)
abline(h = 0, col = "red")
t.test(mod1$residuals, mu = 0)
```



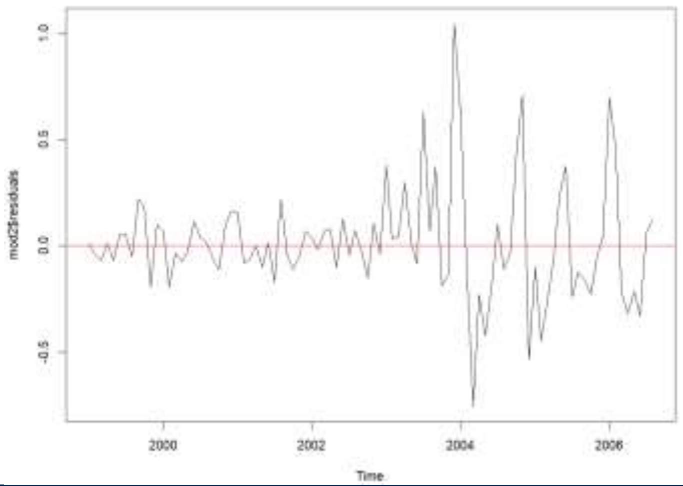

```
        One Sample t-test

data:  mod1$residuals
t = 0.289, df = 91, p-value = 0.7732
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.04666254  0.06255243
sample estimates:
  mean of x
0.007944943
```

Parece indicar que un buen número de residuales están en torno a la media igual a cero. Para confirmar lo mencionado se realiza la prueba t: Como $p = 0.7732 > \alpha = 0.05$, se acepta Ho, es decir la media es igual a cero.

Modelo 2

```
plot(mod2$residuals)
abline(h = 0, col = "red")
t.test(mod2$residuals, mu = 0)
```

Alumno: Maye Mamani Victor Raul

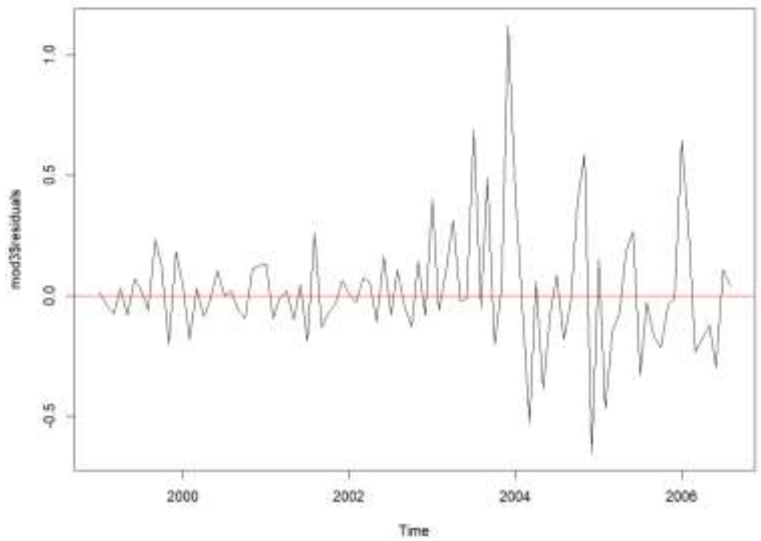


```
One Sample t-test

data:  mod2$residuals
t = 0.53477, df = 91, p-value = 0.5941
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.03987526  0.06925532
sample estimates:
 mean of x
0.01469003
```

Parece indicar que un buen número de residuales están en torno a la media igual a cero. Para confirmar lo mencionado se realiza la prueba t. Como p = 0.5941> α = 0.05, se acepta Ho, es decir la media es igual a cero.

Modelo 3

```
plot(mod3$residuals)
abline(h = 0, col = "red")
t.test(mod3$residuals, mu = 0)
```



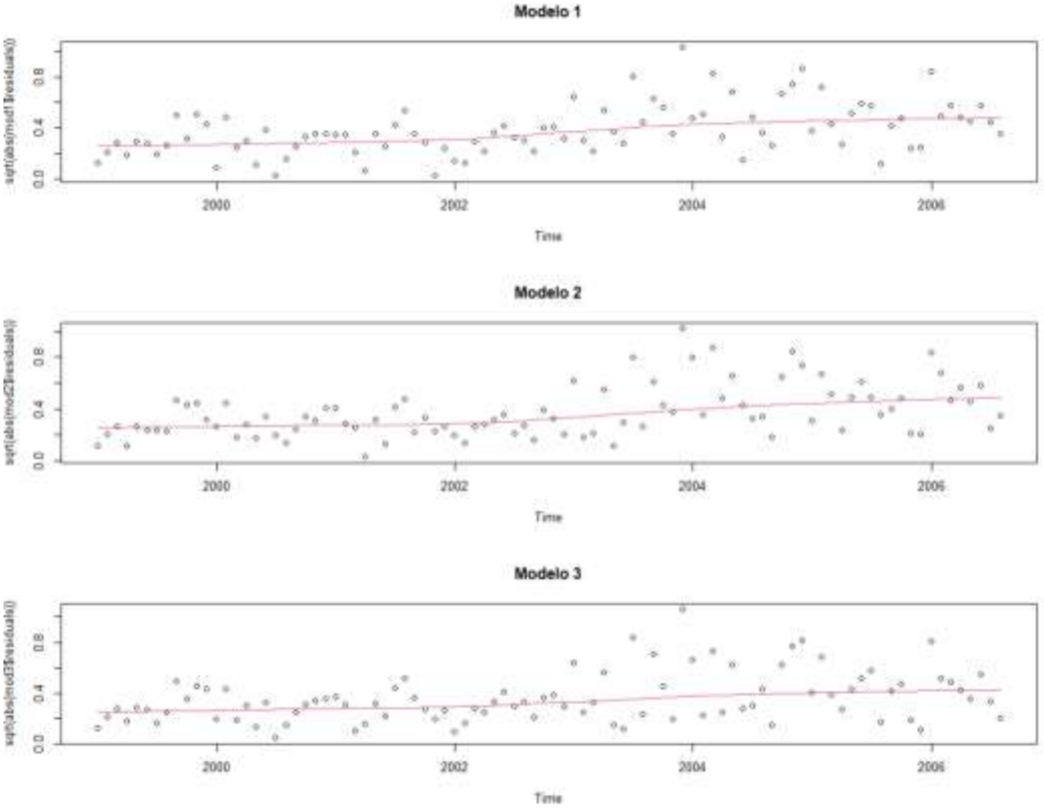
```
One Sample t-test

data:  mod3$residuals
t = 0.85858, df = 91, p-value = 0.3928
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
```

```
 -0.02927068  0.07383793
sample estimates:
 mean of x
0.02228363
```

Parece indicar que un buen número de residuales están en torno a la media igual a cero. Para confirmar lo mencionado se realiza la prueba t: Como p = 0.3928 > α = 0.05, se acepta Ho, es decir la media es igual a cero.

3.2.2 Homocedasticidad o varianza constante

```
par(mfrow = c(3,1))
scatter.smooth(sqrt(abs(mod1$residuals)), lpars=list(col=2), main =
"Modelo 1")
scatter.smooth(sqrt(abs(mod2$residuals)), lpars=list(col=2), main =
"Modelo 2")
scatter.smooth(sqrt(abs(mod3$residuals)), lpars=list(col=2), main =
"Modelo 3")
```



Se observa que los datos parecen no presentan una variabilidad considerable, por tanto, será necesario realizar la prueba de Breusch-Pagan para determinar finalmente si las varianzas constantes para los modelos

```
obs=get(mod1$series)
bptest(resid(mod1)~I(obs-resid(mod1)))
```

```
studentized Breusch-Pagan test

data:  resid(mod1) ~ I(obs - resid(mod1))
BP = 5.7253, df = 1, p-value = 0.01672
```

```
obs=get(mod2$series)
```

Alumno: Maye Mamani Victor Raul

```
bptest(resid(mod2)~I(obs-resid(mod2)))
studentized Breusch-Pagan test

data:  resid(mod2) ~ I(obs - resid(mod2))
BP = 6.3088, df = 1, p-value = 0.01201
```

```
obs=get(mod3$series)
bptest(resid(mod3)~I(obs-resid(mod3)))
studentized Breusch-Pagan test

data:  resid(mod3) ~ I(obs - resid(mod3))
BP = 3.3616, df = 1, p-value = 0.06673
```

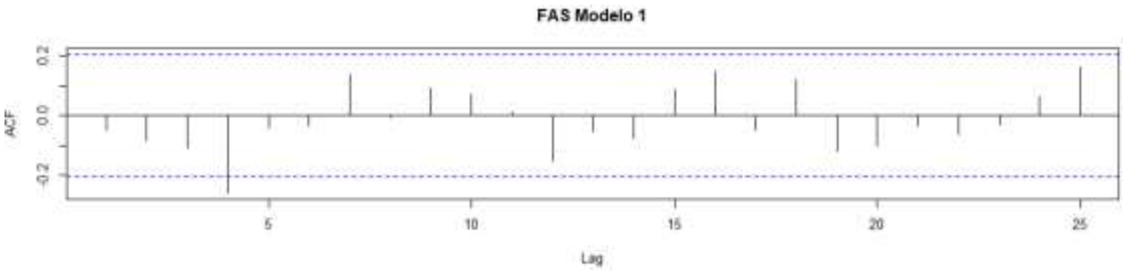
El valor de probabilidad (p-valor) asociado al estadístico BPasume un valor de 0.06673 para el MODELO 3, que es mayor a α=0.05. Por tanto, los residuales de estos modelos son constantes.

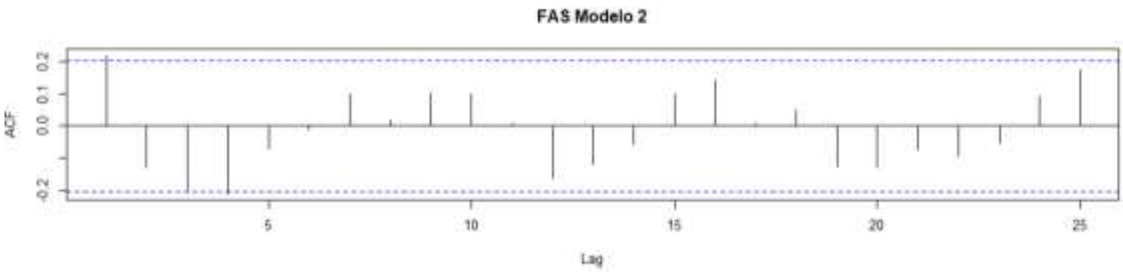
En cambio en el MODELO 1 y MODELO 2 tiene el valor de probabilidad de 0.1039 y 0.6629 respectivamente, que son menores a α=0.05, por lo cual podemos afirmar que los residuales de estos modelos no tiene varianza constante.

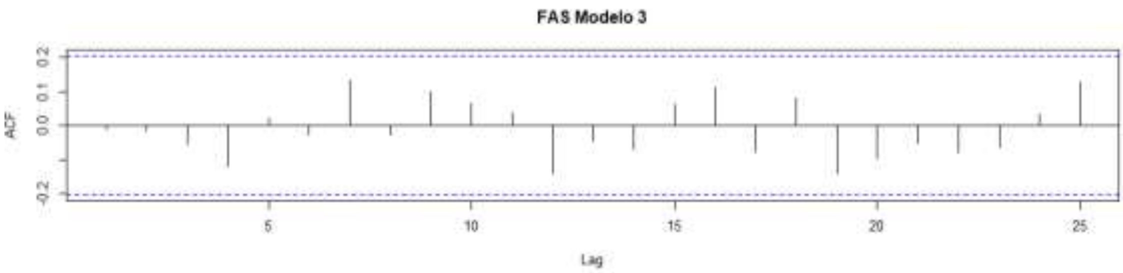
3.2.3 Ausencia de correlación serial

```
resid_m1 <- as.vector(mod1$residuals)
resid_m2 <- as.vector(mod2$residuals)
resid_m3 <- as.vector(mod3$residuals)
par(mfrow = c(1,1))
FAS_e.m1 <- acf(resid_m1, lag.max = 25,
                main="FAS Modelo 1")
FAS_e.m2 <- acf(resid_m2, lag.max = 25,
                main="FAS Modelo 2")
FAS_e.m3 <- acf(resid_m3, lag.max = 25,
                main="FAS Modelo 3")
```

Alumno: Maye Mamani Victor Raul







Se observa que casi la totalidad de los coeficientes del FAS para los modelos 1 y 3 se encuentran dentro de las bandas de no significación. Por el contrario, el modelo 2 presenta 3 coeficientes al borde de los límites de confianza indicando problemas de autocorrelación entre residuales. Por tanto, tenemos altos indicios de que los residuos de los modelos 1 y 3 sean ruido blanco

```
Box.test(resid_m1,type = "Ljung-Box")
Box-Ljung test

data:  resid_m1
X-squared = 0.21148, df = 1, p-value = 0.6456
```

```
Box.test(resid_m2,type = "Ljung-Box")
Box-Ljung test

data:  resid_m2
X-squared = 4.6216, df = 1, p-value = 0.03157
```

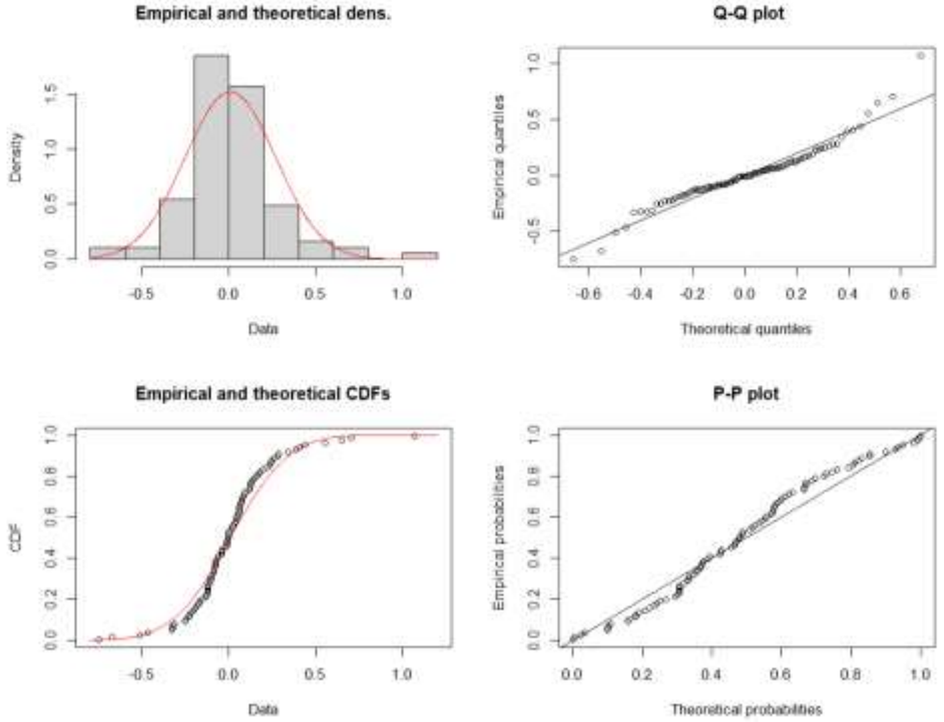
```
Box.test(resid_m3,type = "Ljung-Box")
Box-Ljung test

data:  resid_m3
X-squared = 0.012379, df = 1, p-value = 0.9114
```

Con los valores 0.6456 y 0.9114 en los MODELOS 1 Y 3 son mayores a α=0.05, por lo que se acepta la hipótesis nula de que los coeficientes de autocorrelación son cero; es decir, los residuos son independientes o están incorrelacionados. Pero el MODELO 2 con el valor de 0.03157 < 0.05

rechazamos la hipótesis nula indicándonos que los residuos están correlacionados.

3.2.4 Contraste de normalidad

```
ajuste_m1<-fitdist(data = resid_m1, distr="norm")
plot(ajuste_m1)
JB_m1 <- jarque.bera.test(resid_m1)
JB_m1
```

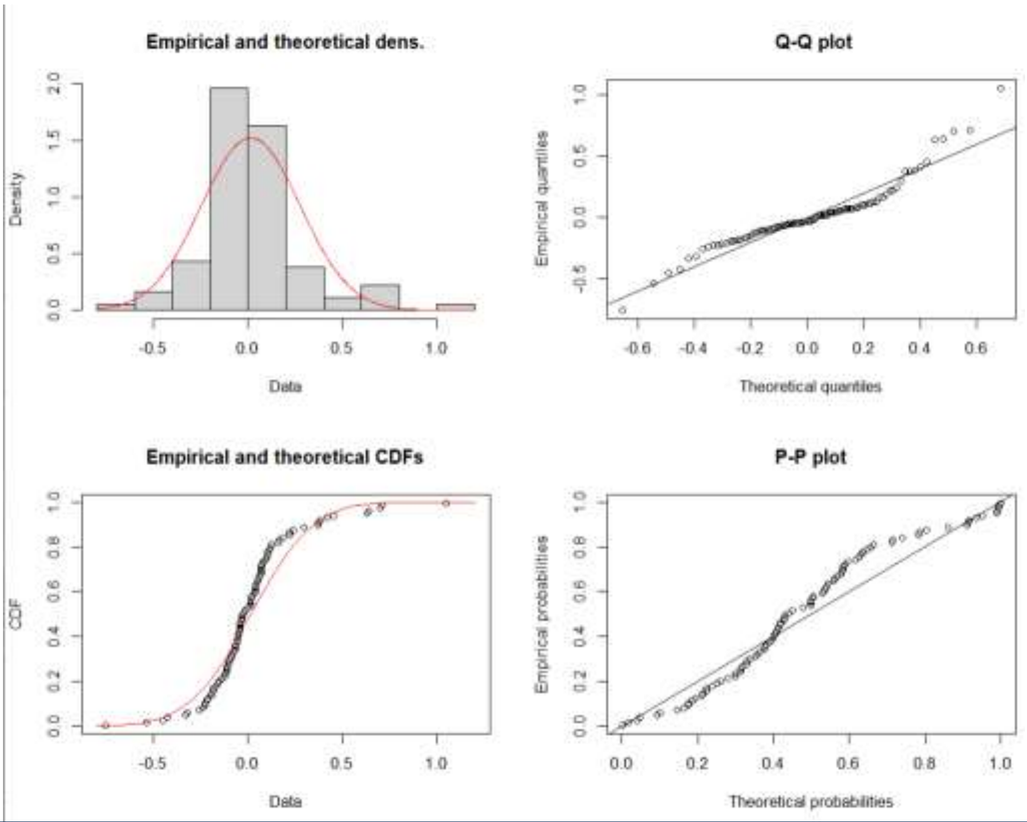



```
Jarque Bera Test

data:  resid_m1
X-squared = 41.778, df = 2, p-value = 8.471e-10
```

En las figuras se observa que los residuales del modelo 1 presentan indicios de ajuste hacia la distribución normal. En la prueba JB, como p = 0.0000 < 0.05, se rechaza la Ho, es decir, los residuos NO se aproximan a una distribución normal.

```
ajuste_m2<-fitdist(data = resid_m2, distr="norm")
plot(ajuste_m2)
JB_m2 <- jarque.bera.test(resid_m2)
JB_m2
```

Alumno: Maye Mamani Victor Raul

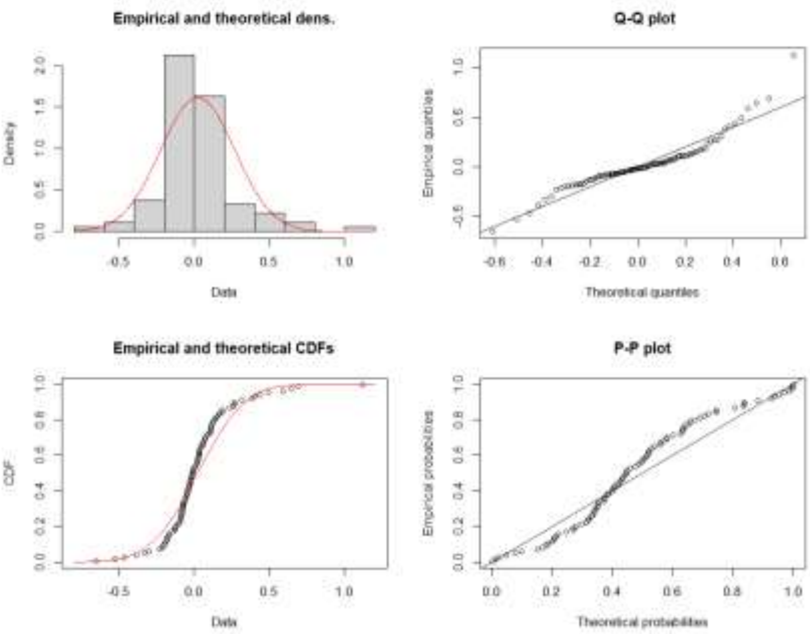



```
Jarque Bera Test

data:  resid_m2
X-squared = 47.925, df = 2, p-value = 3.919e-11
```

En las figuras se observa que los residuales del modelo 2 presentan indicios de ajuste hacia la distribución normal. En la prueba JB, como p = 0.00000 < 0.05, se rechaza la Ho, es decir, los residuos NO se aproximan a una distribución normal.

```
ajuste_m3<-fitdist(data = resid_m3, distr="norm")
plot(ajuste_m3)
JB_m3 <- jarque.bera.test(resid_m3)
JB_m3
```

Alumno: Maye Mamani Victor Raul



```
Jarque Bera Test

data:  resid_m3
X-squared = 87.238, df = 2, p-value < 2.2e-16
```

En las figuras se observa que los residuales del modelo 3 presentan indicios de ajuste hacia la distribución normal. En la prueba JB, como p = 0.00000 < 0.05, se acepta Ho, es decir, los residuos NO se aproximan a una distribución normal.

4 Pronostico

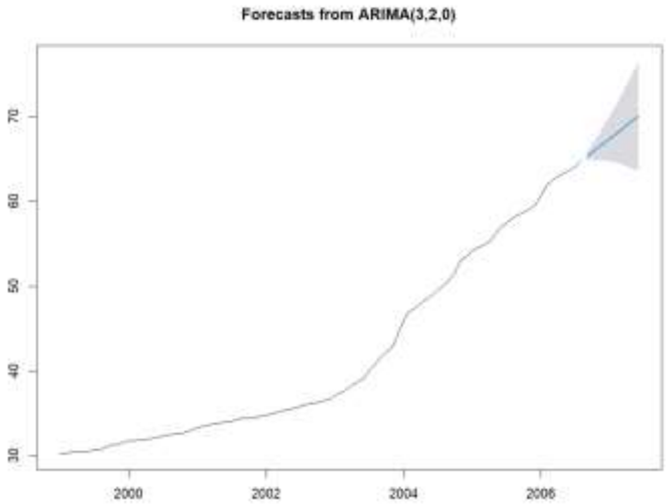
4.1Pronosticos de cada modelo

Modelo 1: ARIMA(3,2,0)

```
Pron1 <- forecast(mod1,level=c(95),h=10)
plot(Pron1)
summary(Pron1)
```

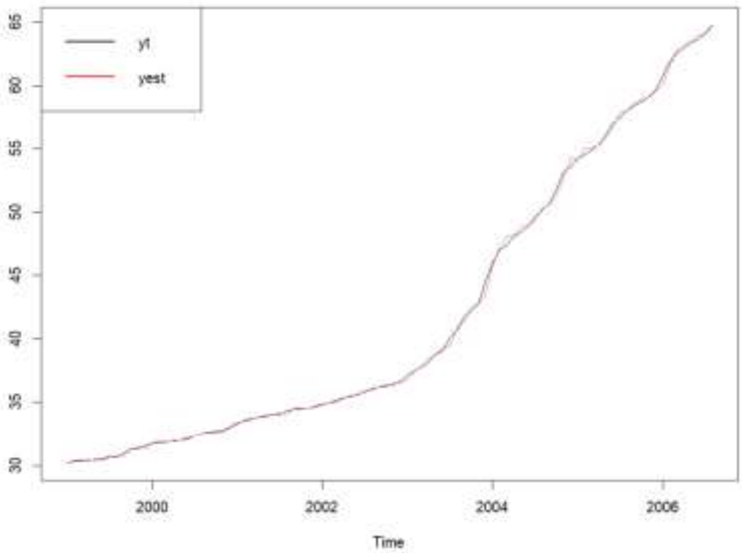
```
Forecast method: ARIMA(3,2,0)

Model Information:
Series: data_ts
ARIMA(3,2,0)

Coefficients:
          ar1      ar2      ar3
      -0.3256  -0.2846  -0.1737
s.e.   0.1036   0.1045   0.1028

sigma^2 = 0.07279:  log likelihood = -8.39
AIC=24.78   AICc=25.25   BIC=34.78

Error measures:
                      ME      RMSE      MAE        MPE      MAPE
MASE        ACF1
Training set 0.007944943 0.2623677 0.183105 0.02477163 0.4173242 0.039
46547 -0.04717314

Forecasts:
         Point Forecast    Lo 95    Hi 95
Sep 2006       65.32976 64.80096 65.85856
Oct 2006       65.82460 64.79331 66.85590
Nov 2006       66.34229 64.79941 67.88517
```

```
Dec 2006       66.87907 64.79726 68.96089
Jan 2007       67.41145 64.71834 70.10456
Feb 2007       67.93587 64.56976 71.30197
Mar 2007       68.46081 64.37547 72.54615
Apr 2007       68.98861 64.14337 73.83386
May 2007       69.51672 63.86925 75.16419
Jun 2007       70.04382 63.55289 76.53475
```



PRONÓSTICO DE SERIE ORIGINAL Y VALORES ESTIMADOS

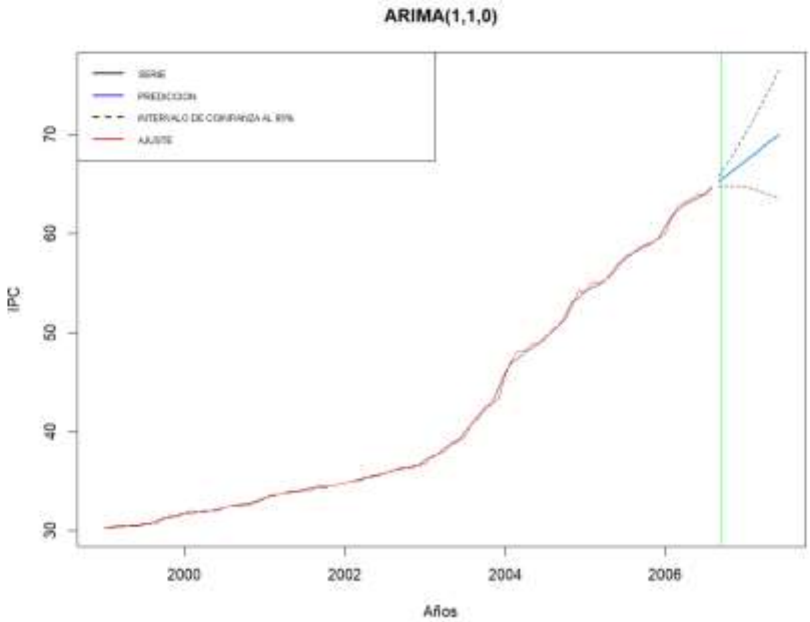
```
data_modelo1 <- (mod1$fitted)
grafico_comparativo <- cbind(data_ts,data_modelo1)
ts.plot(grafico_comparativo, col=c(1,2), lwd = 1)
legend("topleft",c("yt","yest"),lty = c(1,1), lwd = 2,col=c("black",
"red"))
```



GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

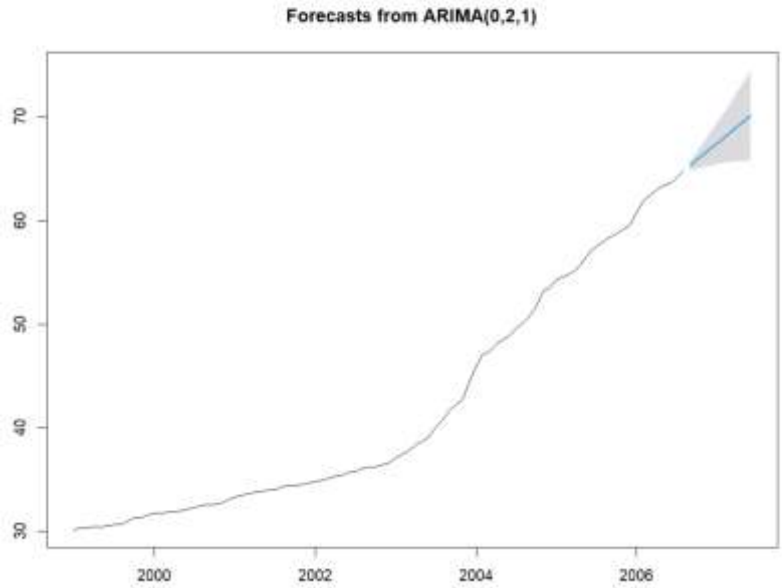
```
plot(Pron1, shaded = FALSE, xlab = "Años", ylab = "IPC",main =
"ARIMA(1,1,0)")
lines(Pron1$fitted, col = "red")
```

```
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE
COINFIANZA AL 95%", "AJUSTE"),col=c("black", "blue", "black", "red"),
lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=2006.7, lwd = 1, col="green")
```



Modelo 2: ARIMA(0,2,1)

```
Pron2 <- forecast(mod2,level=c(95),h=10)
plot(Pron2)
summary(Pron2)
```

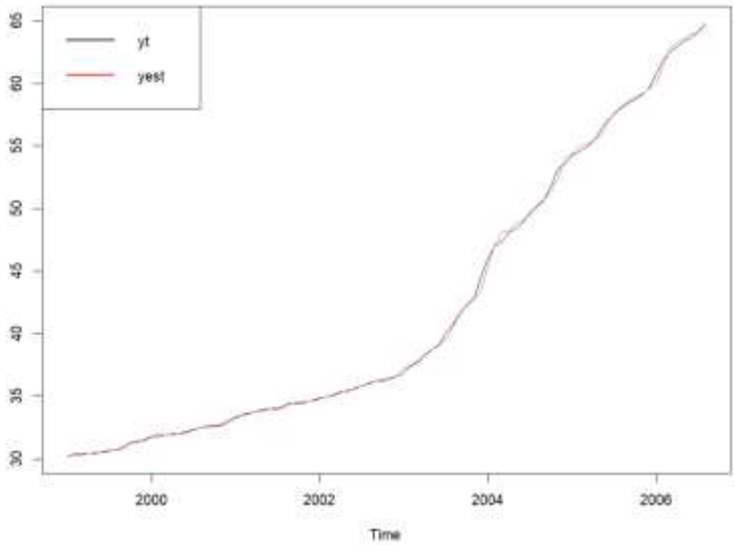



```
Forecast method: ARIMA(0,2,1)

Model Information:
Series: data_ts
ARIMA(0,2,1)

Coefficients:
          ma1
```

```
      -0.6761
s.e.   0.1306

sigma^2 = 0.07121:  log likelihood = -8.6
AIC=21.19   AICc=21.33   BIC=26.19

Error measures:
                     ME      RMSE       MAE        MPE      MAPE
MASE      ACF1
Training set 0.01469003 0.2624562 0.1774635 0.04408597 0.4000181 0.038
24954 0.2205244

Forecasts:
         Point Forecast    Lo 95    Hi 95
Sep 2006       65.32250 64.79950 65.84551
Oct 2006       65.85801 64.99029 66.72573
Nov 2006       66.39351 65.17057 67.61646
Dec 2006       66.92902 65.32936 68.52868
Jan 2007       67.46452 65.46446 69.46458
Feb 2007       68.00003 65.57578 70.42428
Mar 2007       68.53553 65.66385 71.40721
Apr 2007       69.07104 65.72944 72.41263
May 2007       69.60654 65.77332 73.43976
Jun 2007       70.14204 65.79627 74.48782
```

PRONÓSTICO DE SERIE ORIGINAL Y VALORES ESTIMADOS

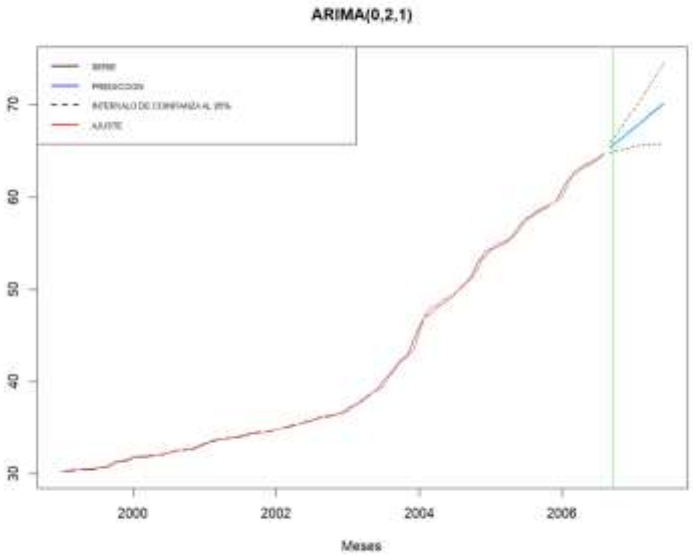
```
data_modelo2 <- (mod2$fitted)
grafico_comparativo <- cbind(data_ts,data_modelo2)
ts.plot(grafico_comparativo, col=c(1,2), lwd = 1)
legend("topleft",c("yt","yest"),lty = c(1,1), lwd = 2,col=c("black",
"red"))
```



GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

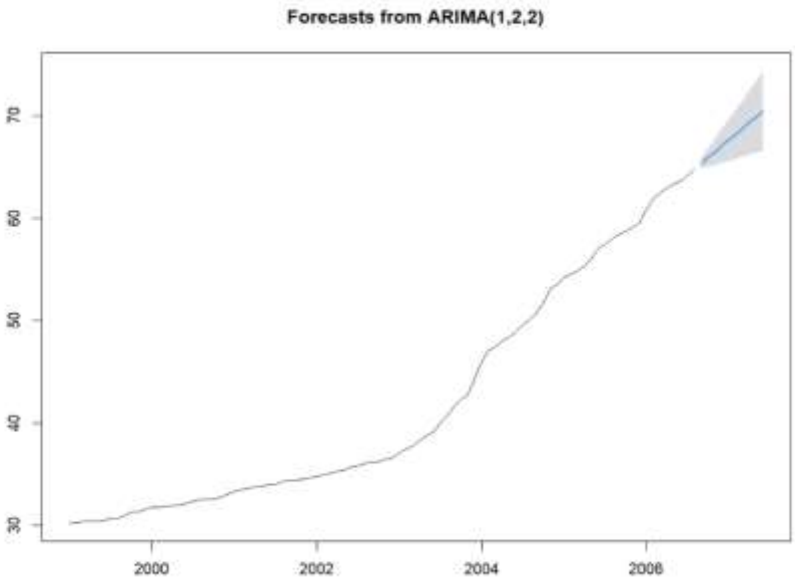
```
plot(Pron2, shaded = FALSE, xlab = "Meses", ylab = "",main =
"ARIMA(0,1,5)")
lines(Pron2$fitted, col = "red")
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE
COINFIANZA AL 95%", "AJUSTE"),col=c("black", "blue", "black", "red"),
lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=2006.7, lwd = 1, col="green")
```

Alumno: Maye Mamani Victor Raul



Modelo 3: ARIMA(1,2,2)

```
Pron3 <- forecast(mod3,level=c(95),h=10)
plot(Pron3)
summary(Pron3)
```



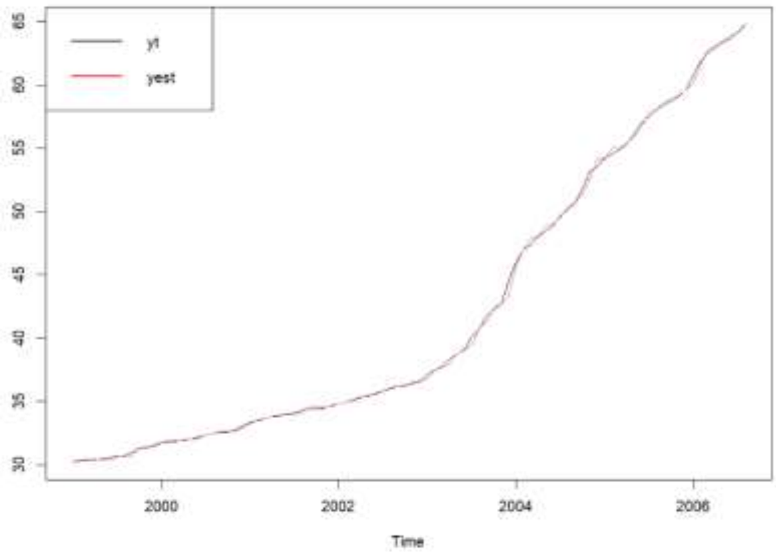

```
Forecast method: ARIMA(1,2,2)

Model Information:
Series: data_ts
ARIMA(1,2,2)

Coefficients:
         ar1      ma1      ma2
      0.1532  -0.6004  -0.2069
s.e.  0.3261   0.3214   0.2371

sigma^2 = 0.06535:  log likelihood = -3.78
AIC=15.56   AICc=16.03   BIC=25.56

Error measures:
                     ME      RMSE       MAE        MPE      MAPE
MASE        ACF1
Training set 0.02228363 0.2485855 0.1641209 0.06164755 0.3777026 0.035
37375 -0.01141304
```

Alumno: Maye Mamani Victor Raul

```
Forecasts:
         Point Forecast     Lo 95     Hi 95
Sep 2006       65.36983 64.86881 65.87086
Oct 2006       65.93916 65.01379 66.86454
Nov 2006       66.50643 65.20365 67.80921
Dec 2006       67.07338 65.40958 68.73717
Jan 2007       67.64028 65.61801 69.66255
Feb 2007       68.20717 65.82256 70.59178
Mar 2007       68.77406 66.02007 71.52805
Apr 2007       69.34095 66.20890 72.47300
May 2007       69.90784 66.38814 73.42755
Jun 2007       70.47473 66.55732 74.39215
```

PRONÓSTICO DE SERIE ORIGINAL Y VALORES ESTIMADOS

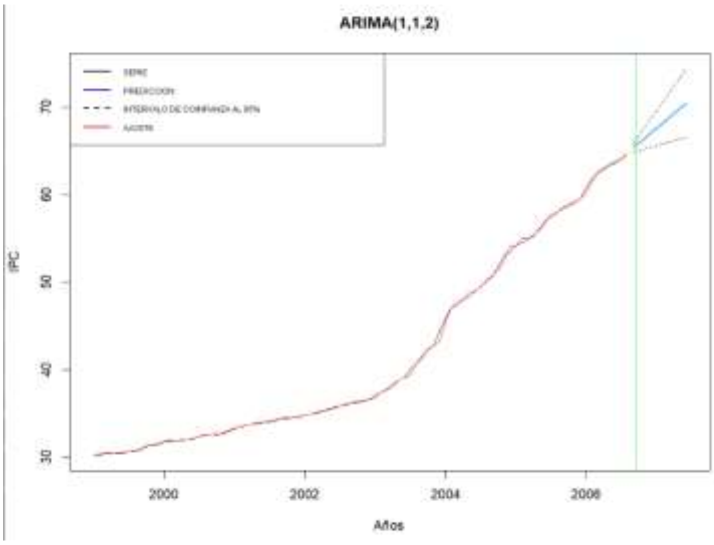
```
data_modelo3 <- (mod3$fitted)
grafico_comparativo <- cbind(data_ts,data_modelo3)
ts.plot(grafico_comparativo, col=c(1,2), lwd = 1)
legend("topleft",c("yt","yest"),lty = c(1,1), lwd = 2,col=c("black",
"red"))
```



GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

```
plot(Pron3, shaded = FALSE, xlab = "Años", ylab = "IPC",main =
"ARIMA(1,1,2)")
lines(Pron3$fitted, col = "red")
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE
COINFIANZA AL 95%", "AJUSTE"),col=c("black", "blue", "black", "red"),
lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=2006.7, lwd = 1, col="green")
```

Alumno: Maye Mamani Victor Raul



Métricas basadas en el error

```
accuracy(Pron1)
ME          RMSE         MAE          MPE         MAPE          MASE          ACF1
Training set 0.007944943 0.2623677 0.183105 0.02477163 0.4173242 0.039
46547 -0.04717314
```

```
accuracy(Pron2)
  ME         RMSE          MAE          MPE         MAPE          MASE         ACF1
Training set 0.01469003 0.2624562 0.1774635 0.04408597 0.4000181 0.038
24954 0.2205244
```

```
accuracy(Pron3)
ME          RMSE          MAE          MPE         MAPE          MASE          ACF1
Training set 0.02228363 0.2485855 0.1641209 0.06164755 0.3777026 0.035
37375 -0.01141304
```

Basado en las métricas MAE y RMSE, el modelo Pron3 es el mejor entre los tres, ya que tiene los valores más bajos(MAE = 0.1641 Y RMSE=0.2485) en ambas métricas. Esto indica que sus predicciones son más precisas y con menos variabilidad en los errores grandes.

Conclusión

Modelo ARIMA(3,2,0)

1. **Estimación y Validación:**
    - Los coeficientes estimados para AR(1) y AR(2) fueron significativos, lo que indica que estos términos contribuyen de manera significativa al modelo, pero AR(3) no fue significativo.
    - La prueba de estabilidad de coeficientes (Chow test) indicó inestabilidad, lo que podría sugerir que el modelo podría no ser robusto en el tiempo.
2. **Análisis de Residuos:**
    - Los residuos mostraron una media cercana a cero y no presentaron homocedasticidad, como indicado por la prueba de Breusch-Pagan.

- No se observaron problemas significativos de autocorrelación en los residuos, como se confirmó mediante el análisis del correlograma de los residuos.

Modelo ARIMA(0,2,1)

1. **Estimación y Validación:**
    - El coeficiente estimado para MA(1) fue significativo, lo que indica que este término es importante para modelar la serie.
    - La prueba de estabilidad de coeficientes mostró inestabilidad, similar al primer modelo.
2. **Análisis de Residuos:**
    - Los residuos mostraron una media cercana a cero y no presentaron homocedasticidad, según la prueba de Breusch-Pagan.
    - Se observaron problemas de autocorrelación en los residuos, sugiriendo que el modelo podría no capturar completamente la estructura de la serie.

Modelo ARIMA(1,2,2)

1. **Estimación y Validación:**
    - El coeficiente AR(1) , MA(1) y MA(2) no fueron significativos , lo que indica que estos términos pueden no ser necesarios para el modelo.
    - La prueba de estabilidad de coeficientes mostró inestabilidad, similar a los otros modelos.
2. **Análisis de Residuos:**
    - Los residuos mostraron una media cercana a cero y no presentaron homocedasticidad, según la prueba de Breusch-Pagan.
    - Se observaron problemas de autocorrelación en los residuos, lo que indica que el modelo puede no ser adecuado para capturar toda la variabilidad de la serie.

Considerando la significancia de los coeficientes, la ausencia de autocorrelación en los residuos y la estabilidad de los coeficientes, el *ARIMA(0,2,1)* ***emerge como el modelo preferido.***

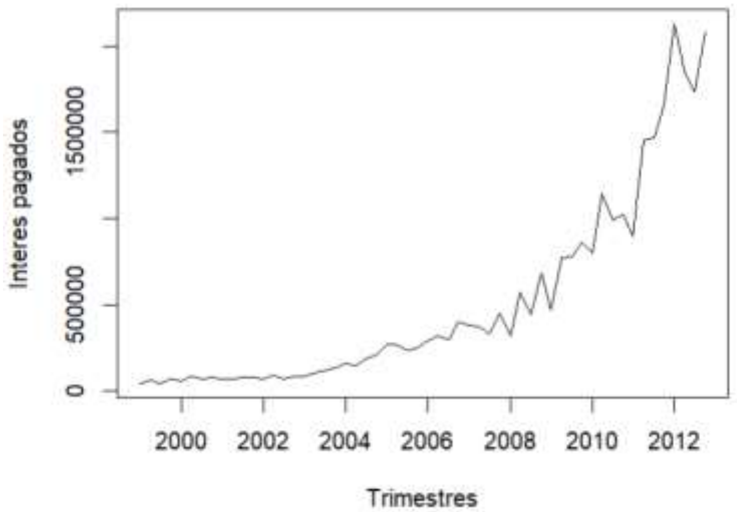
Pero basado en las métricas MAE (Error Absoluto Medio) y RMSE (Error Cuadrático Medio), se puede concluir que el modelo 3 ARIMA(1,2,2) se destaca como el mejor entre los tres evaluados. Con un MAE de 0.1641 y un RMSE de 0.2486, el modelo 3 exhibe los errores más bajos en sus predicciones, indicando una mayor precisión y menor variabilidad en comparación con los modelos 1 y 2. Estas métricas sugieren que el modelo 3 produce predicciones más cercanas a los valores reales y con una consistencia superior, haciendo de este modelo la opción preferida para este análisis.

En este caso eligiremos el modelo 3 ya que sus predicciones son mejores

Alumno: Maye Mamani Victor Raul

CASO 2: INTERESES PAGADOS AL EXTERIOR (IPE)

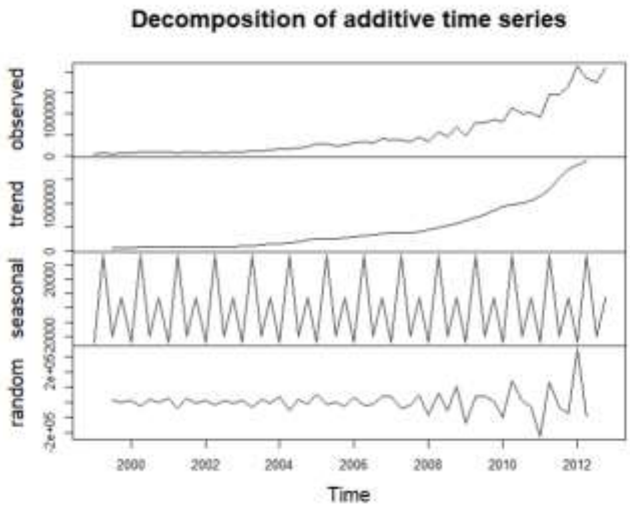
```
data <- read_excel("F:\\777--Programacion
repos\\Una\\r\\data\\actividad-05.xlsx",sheet = "2")
View(data)
# Gráfica de la serie
data_ts <- ts(data$interes, start = c(1999,1), frequency = 4)
plot(data_ts, xlab="Trimestres", ylab="Interes pagados")
```



En primera instancia podemos visualizar una clara TENDENCIA creciente, la varianza no parece variar mucho, y no es estacionaria en media. Por lo que tendremos que realizar un análisis de estacionalidad dado que la serie es trimestral

1 Identificación

```
#Descomposición de la serie
data_des <- decompose(data_ts, type = "multiplicative")
plot(data_des,type="l")
```

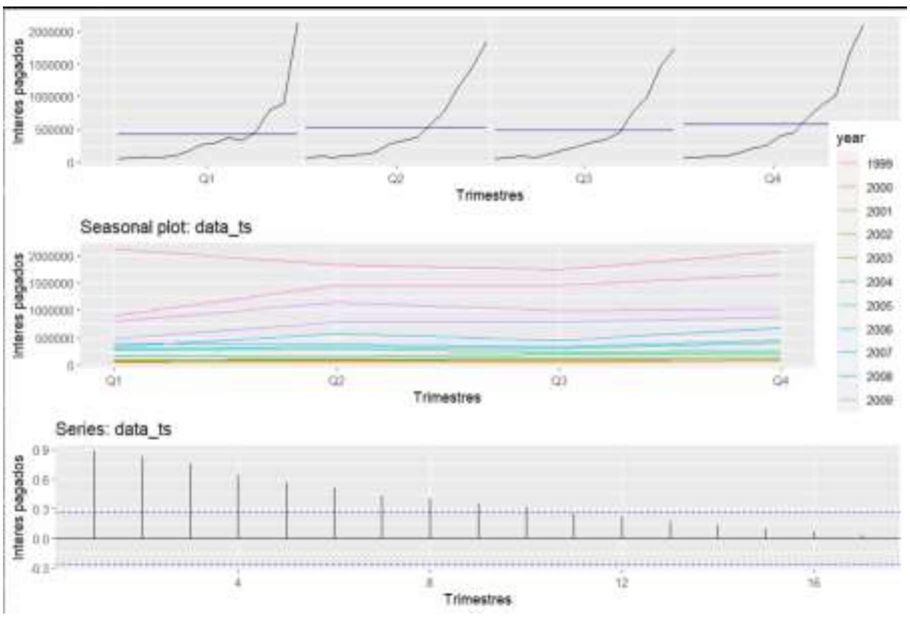


Se aprecia una tendencia creciente

1. 1 Análisis de la tendencia y la estacionalidad

1.1.1 Estacionalidad

```
#Graficos de la serie para identificar estacionalidad
plot1 <- ggsubseriesplot(data_ts, xlab = "Trimestres", ylab =
"Interes pagados")
plot2 <- ggseasonplot(data_ts, xlab = "Trimestres", ylab = "Interes
pagados")
plot3 <- ggAcf(data_ts, xlab = "Trimestres", ylab = "Interes
pagados")

grid.arrange(plot1, plot2, plot3, ncol = 1)
```



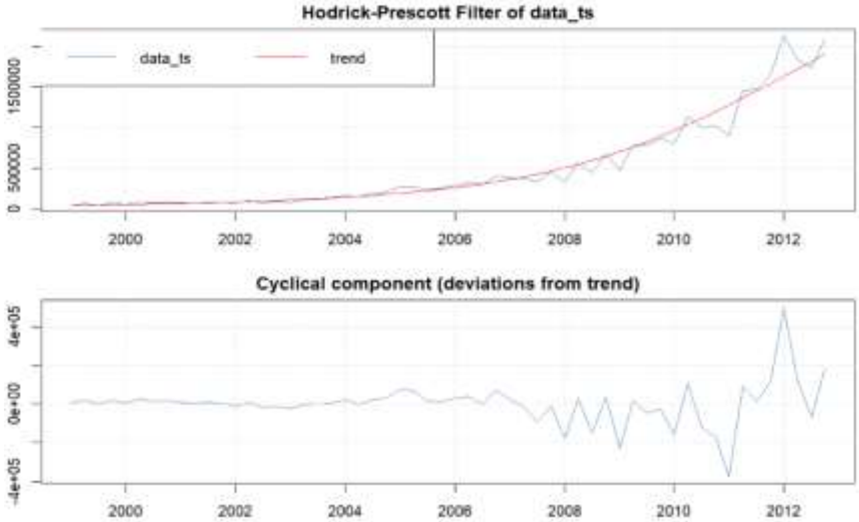
Vemos que la serie no presenta estacionalidad, dado que el primer grafico nos muestra se muestra a un mismo nivel, el segundo nos muestra que

Alumno: Maye Mamani Victor Raul

no hay un patrón claro y finalmente el correlograma decrece constantemente.

1.1.2 Análisis de tendencia

```
#Análisis de tendencia
lambda_hp <- 1600
data_hp <- hpfilter(data_ts, type="lambda", freq=lambda_hp)
plot(data_hp)
```

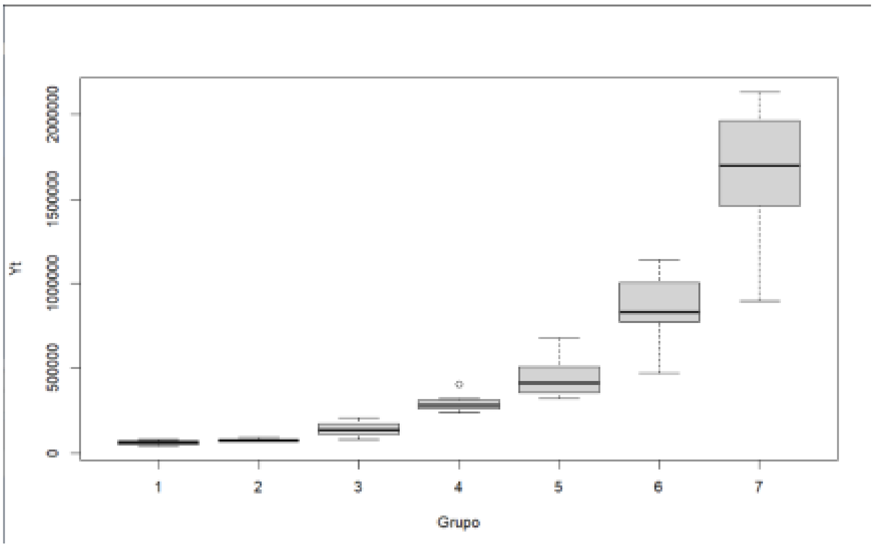


Hay una tendencia creciente y no parece haber algún patrón que se repita cíclicamente

1.2 Análisis de estacionariedad

1.2.1 Estacionariedad en varianza

```
#Estacionaridad en varianza
Grupo <- rep(1:7, each = 8)
boxplot(data$interes ~ Grupo, xlab = "Grupo", ylab="Yt")
```

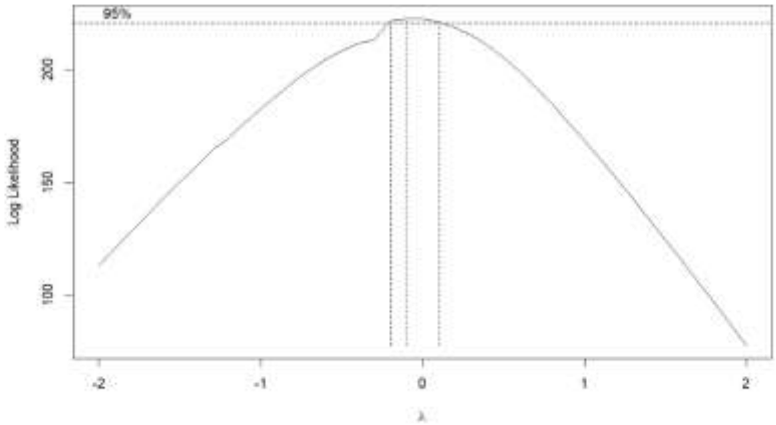


La varianza va creciendo, por lo que diremos que no es constante

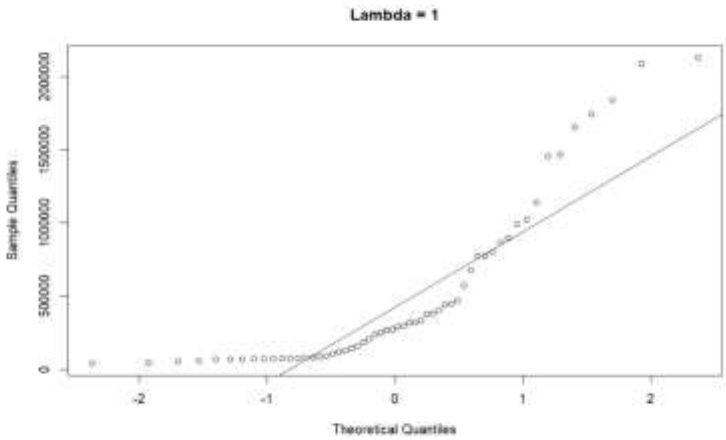
Transformación de la serie para que sea estacionaria

Alumno: Maye Mamani Victor Raul

```
b <- BoxCox.ar(data_ts)
lambda <- b$mle
round(lambda,1)
[1] -0.1
```

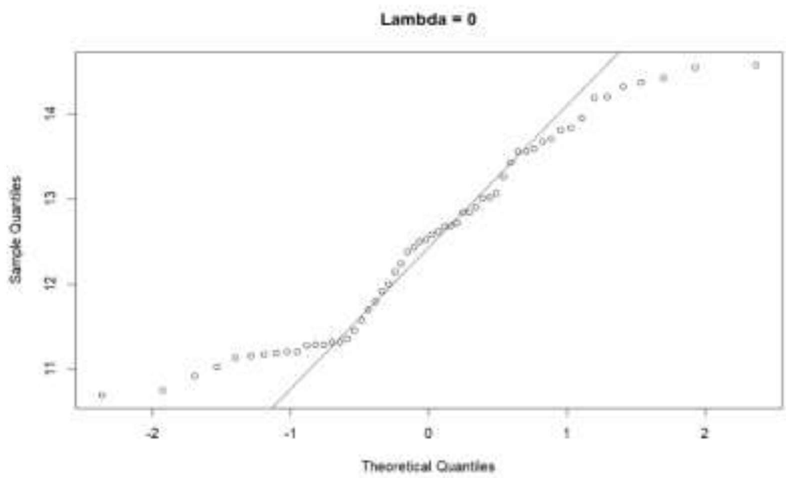



```
qqnorm(data_ts, main = "Lambda = 1") # Yt: Original
qqline(data_ts)
```

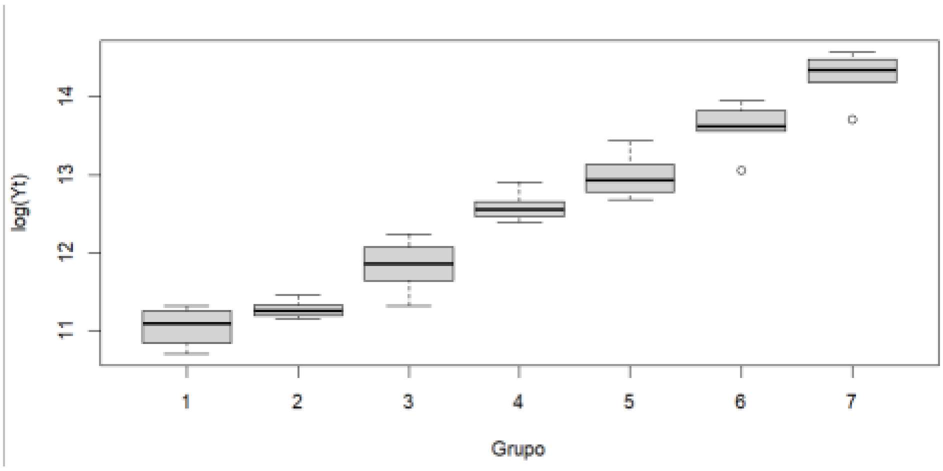



```
data_mod = log(data_ts)
qqnorm(data_mod, main = "Lambda = 0")
qqline(data_mod)
```

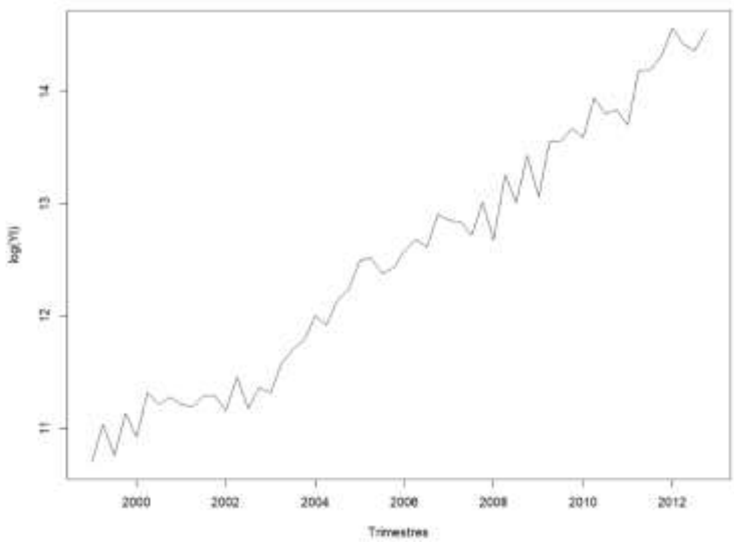
Alumno: Maye Mamani Victor Raul




```
boxplot(data_mod ~ Grupo, xlab = "Grupo", ylab="log(Yt)")
```




```
plot(data_mod, xlab="Trimestres", ylab="log(Yt)")
```

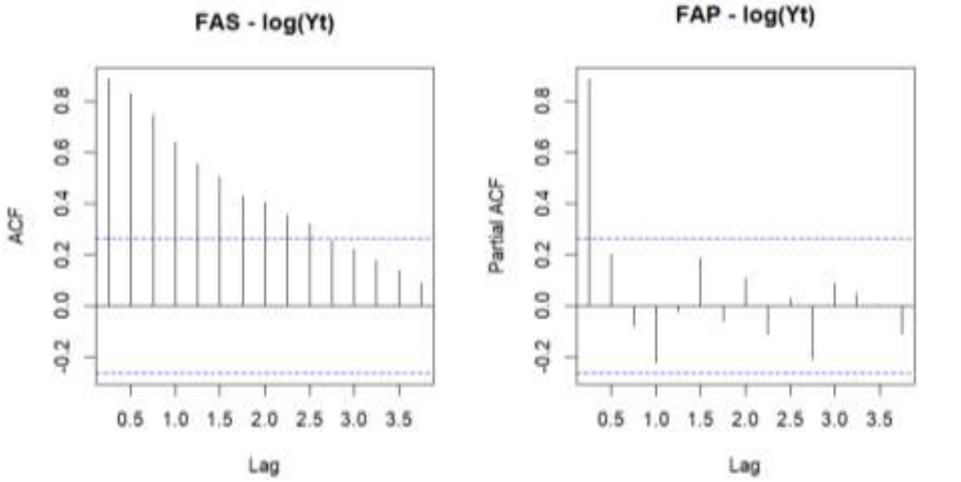


Alumno: Maye Mamani Victor Raul

Después de haber transformado la serie podemos notar que la varianza se ha estabilizado y aparenta ser constante

1.2.2 Estacionariedad en media

```
par(mfrow = c(1,2))
FAS <- acf(data_ts, lag.max = 15, main="FAS - log(Yt)")
FAP <- pacf(data_ts, lag.max = 15, main="FAP - log(Yt)")
```



El FAS decrece lentamente y el primer FAP es significativo siendo casi 0.9, por lo que podríamos decir que la serie NO ES ESTACIONARIA

```
# Verificación con la prueba de Raíz unitaria de Dickey-Fuller
Aumentada.
data_adf <- ur.df(data_mod, type="trend", lags = 1)
summary(data_adf)
```

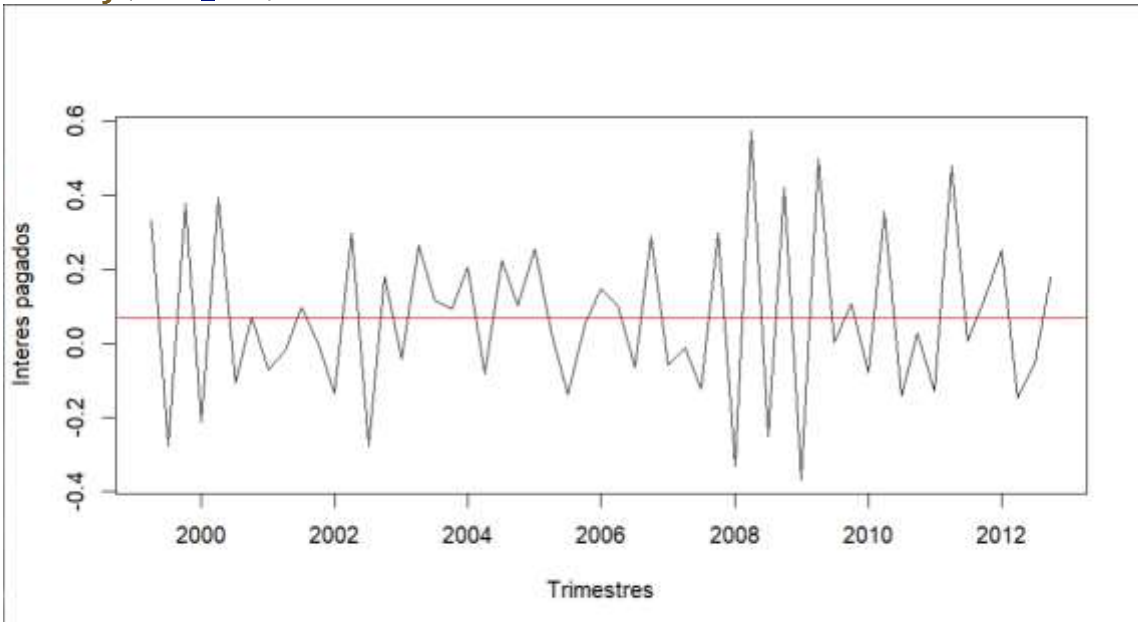
```
Residual standard error: 0.148 on 50 degrees of freedom
Multiple R-squared:  0.5748,   Adjusted R-squared:  0.5493
F-statistic: 22.53 on 3 and 50 DF,   p-value: 2.279e-09


Value of test-statistic is: -2.4952 12.0028 3.5166

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Observamos que el T calculado (-2.4952) es MAYOR que el T critico (-3.45) por tanto se acepta la hipótesis nula de la existencia de raíz unitaria, es decir que la serie NO ES ESTACIONARIA

Diferenciamos debido a que la serie no es estacionaria

```
data_diff = diff(data_mod)
plot(data_diff, xlab="Trimestres", ylab="Interes pagados")
abline(h = mean(data_diff), col = "red")
```

```
data_adf <- ur.df(data_diff, type="drift", lags = 1)
summary(data_adf)
```



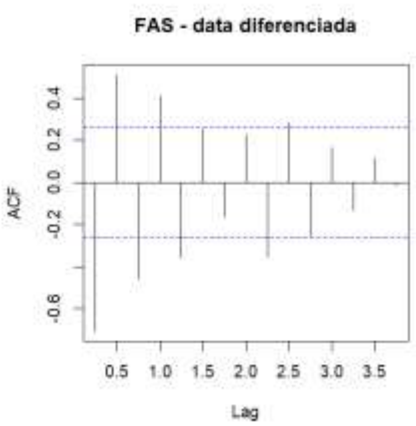
Y volvemos a aplicar el test de Dickey Fuller

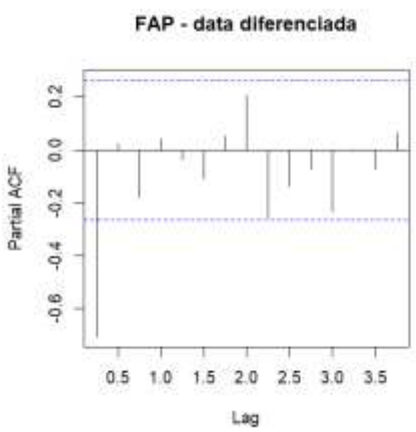
```
Residual standard error: 0.1564 on 50 degrees of freedom
Multiple R-squared:  0.8575,   Adjusted R-squared:  0.8518
F-statistic: 150.5 on 2 and 50 DF,  p-value: < 2.2e-16


Value of test-statistic is: -6.2316 19.4227

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.51 -2.89 -2.58
phi1  6.70  4.71  3.86
```

Ahora vemos que el T calculado (-6.23)16 es MENOR que el T critico (-2.89) por tanto se RECHAZA la hipótesis nula de la existencia de raíz unitaria, es decir que la serie ES ESTACIONARIA

1.3 Identificación del modelo estacionario

1.3.1 Identificación de las órdenes p y q





Alumno: Maye Mamani Victor Raul

Observando el FAS vemos que decrece en forma sinusoidal y tiene 5 coeficientes significativos por lo que proponemos un MA(5). Y el FAP decrece rápidamente a cero mostrando un patrón sinusoidal y tiene el primer coeficiente diferente de cero por lo que también planteamos un AR(1) aunque no esta muy claro. Adicionalmente planteamos un ARMA(1,2).

1.3.2 Inclusión del término independiente (δ) o intercepto

```
#incluir el intercepto
Z <- mean(data_diff)
Co <- var(data_diff)
Tn <- length(data_diff)
Ta <- Tn - 1
Sigma <- Co/Ta
t <- Z/Sigma
tt <- qt(1-0.05/2,Ta-1)
pruebaT <- c(t, tt)
names(pruebaT) <- c("t-calculado","t-critico")
pruebaT
```

```
t-calculado   t-critico
  77.157356     2.00574
```

Al tener un T calculado (77.157356) MAYOR al T critico (2.00574) se acepta la hipótesis alterna por lo que incluimos la constante en el modelo

2 Estimación

```
mod1 <- Arima(data_mod, order = c(1, 1, 0), include.constant = T)
coeftest(mod1)
```

```
z test of coefficients:

       Estimate Std. Error z value  Pr(>|z|)
ar1   -0.715918   0.092892 -7.7070 1.288e-14 ***
drift  0.067188   0.012105  5.5505 2.849e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Con los parámetros estimados del modelo tenemos lo siguiente

$$\Delta Y_t = 0.0671 - 0.7159\Delta Y_{t-1} + a_t$$

```
mod2 <- Arima(data_mod, order = c(0, 1, 5), include.constant = T)
coeftest(mod2)
```

```
z test of coefficients:

        Estimate Std. Error z value  Pr(>|z|)
ma1   -0.4485415  0.1576038 -2.8460  0.004427 **
ma2    0.2633784  0.2748540  0.9582  0.337938
ma3   -0.6407278  0.1198853 -5.3445 9.066e-08 ***
ma4    0.5254175  0.2013839  2.6090  0.009080 **
ma5   -0.6995139  0.2720900 -2.5709  0.010144 *
drift  0.0694389  0.0033106 20.9749 < 2.2e-16 ***
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Con los parámetros estimados del modelo tenemos lo siguiente

$$\Delta Y_t = 0.6994 + a_t - 0.4485 + 0.2633a_{t-2} - 0.6407a_{t-3} + 0.5254a_{t-4} - 0.6995a_{t-5}$$

Modelo ARIMA (1,1,2)

```
mod3 <- Arima(data_mod, order = c(1, 1, 2), include.constant = T)
coeftest(mod3)
```

```
z test of coefficients:

       Estimate Std. Error z value  Pr(>|z|)    
ar1   -0.906507   0.093423 -9.7032 < 2.2e-16 ***
ma1    0.243719   0.171342  1.4224    0.1549    
ma2   -0.239670   0.163010 -1.4703    0.1415    
drift  0.067546   0.010765  6.2745 3.508e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Con los parámetros estimados del modelo tenemos lo siguiente

$$\Delta Y_t = 0.0675 - 0.9065\Delta Y_{t-1} + a_t + 0.2437a_{t-1} - 0.2396a_{t-2}$$

3 Validación

3.1 Análisis de los coeficientes estimados

3.1.1 Significación de los coeficientes

Para el modelo AR(1)

```
z test of coefficients:

       Estimate Std. Error z value  Pr(>|z|)    
ar1   -0.715918   0.092892 -7.7070 1.288e-14 ***
drift  0.067188   0.012105  5.5505 2.849e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$AR(1): \phi_1 = -0.7159 \rightarrow p = 0.00000 < 0.01, altamente\ significativo$$

Para el modelo MA(5)

```
z test of coefficients:

        Estimate Std. Error z value  Pr(>|z|)    
ma1   -0.4485415  0.1576038 -2.8460  0.004427 ** 
ma2    0.2633784  0.2748540  0.9582  0.337938    
ma3   -0.6407278  0.1198853 -5.3445 9.066e-08 ***
ma4    0.5254175  0.2013839  2.6090  0.009080 ** 
ma5   -0.6995139  0.2720900 -2.5709  0.010144 *  
drift  0.0694389  0.0033106 20.9749 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$MA(1): \theta_1 = -0.4485 \rightarrow p = 0.0044 < 0.01, altamente\ significativo$$

$$MA(2): \theta_2 = 0.2633 \rightarrow p = 0.3379 > 0.05, NO\ es\ significativo$$

$$MA(3): \theta_3 = -0.6407 \rightarrow p = 0.0000 < 0.01, altamente\ significativo$$

$$MA(4): \theta_4 = 0.5254 \rightarrow p = 0.0090 < 0.01, altamente\ significativo$$

$$MA(5): \theta_5 = -0.6995 \rightarrow p = 0.0101 < 0.05, es\ significativo$$

```
z test of coefficients:

       Estimate Std. Error z value  Pr(>|z|)
ar1   -0.906507   0.093423 -9.7032 < 2.2e-16 ***
ma1    0.243719   0.171342  1.4224    0.1549
ma2   -0.239670   0.163010 -1.4703    0.1415
drift  0.067546   0.010765  6.2745 3.508e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$AR(1): \phi_1 = -0.9065 \rightarrow p = 0.0000 < 0.01, altamente\ significativo$$

$$MA(1): \theta_1 = 0.2437 \rightarrow p = 0.1549 > 0.05, NO\ es\ significativo$$

$$MA(2): \theta_2 = -0.2396 \rightarrow p = 0.1415 > 0.05, NO\ es\ significativo$$

3.1.2 Examen de la matriz de correlaciones entre los coeficientes

```
vcov(mod1)
```

```
                ar1        drift
ar1   8.628880e-03 2.049485e-05
drift 2.049485e-05 1.465264e-04
```

```
vcov(mod2)
```

```
ma1     2.483896e-02  2.580993e-02  3.811391e-03 -0.0182930187 -0.02455
07250   1.659998e-05
ma2     2.580993e-02  7.554474e-02  1.455536e-02 -0.0419279506 -0.06756
89521  -8.637586e-06
ma3     3.811391e-03  1.455536e-02  1.437248e-02 -0.0138970967 -0.00937
73322  -8.930819e-06
ma4    -1.829302e-02 -4.192795e-02 -1.389710e-02  0.0405554824  0.03557
91858  -1.317850e-05
ma5    -2.455073e-02 -6.756895e-02 -9.377332e-03  0.0355791858  0.07403
29643   1.413850e-05
drift   1.659998e-05 -8.637586e-06 -8.930819e-06 -0.0000131785  0.00001
41385   1.095990e-05
```

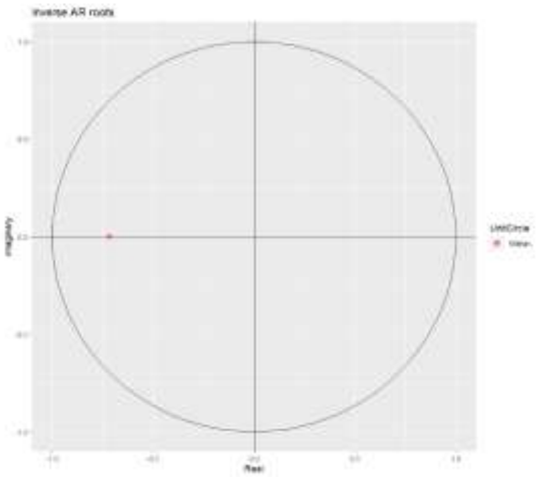
```
vcov(mod3)
```

```
                 ar1           ma1           ma2         drift
ar1     8.727900e-03 -0.0103213052  8.607613e-03  1.139592e-06
ma1    -1.032131e-02  0.0293581486 -3.515797e-03 -1.031090e-05
ma2     8.607613e-03 -0.0035157970  2.657211e-02 -4.151149e-05
drift   1.139592e-06 -0.0000103109 -4.151149e-05  1.158904e-04
```

Se observa claramente que ningún coeficiente esta próximo ni cercano a 0.9, por tanto, podemos indicar que no hay problema de multicolinealidad en los modelos propuestos.

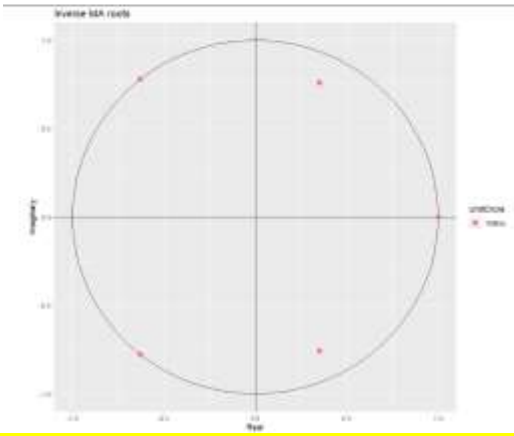
3.1.3 Condición de convergencia e invertibilidad

```
autoplot(mod1)
```



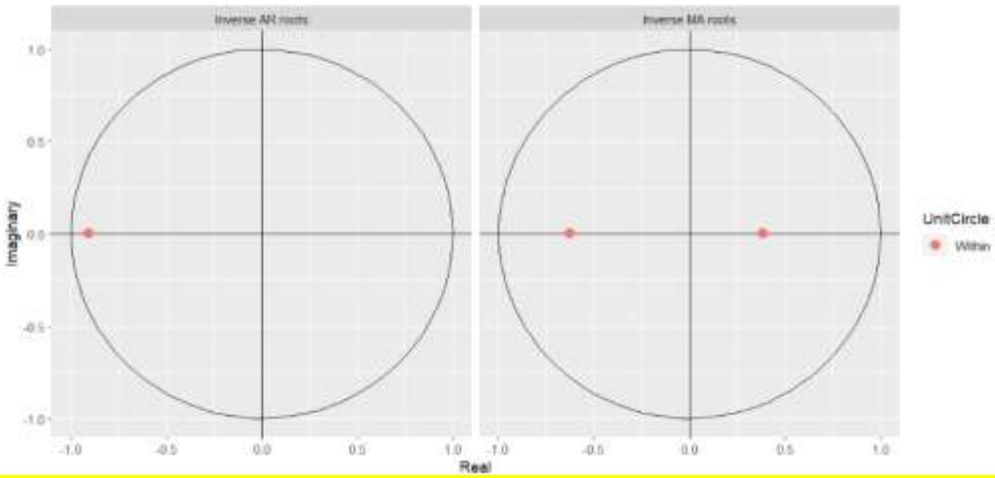
En la figura de raíces inversas de AR, se observa en el gráfico las raíces características se encuentran dentro del círculo, es decir que cumplen con la condición de estacionariedad para la parte autorregresiva.

```
autoplot(mod2)
```



En la figura de raíces inversas de MA, se observa en el gráfico las raíces características se encuentran dentro del círculo, es decir que cumplen con la condición de invertibilidad para la parte de media movíl.

```
autoplot(mod3)
```



Al estar los valores dentro de la circunferencia unitaria es un indicativo de que el modelo se ajusta correctamente. Tanto en su parte AR, como en su parte MA.

Alumno: Maye Mamani Victor Raul

3.1.4 Análisis de la estabilidad

```
Chow_mod1 <- Fstats(mod1$fitted ~ 1, from = 0.67)
sctest(Chow_mod1)
```

```
supF test

data:  Chow_mod1
sup.F = 160.43, p-value < 2.2e-16
```

```
Chow_mod2 <- Fstats(mod2$fitted ~ 1, from = 0.67)
sctest(Chow_mod2)
```

```
supF test

data:  Chow_mod2
sup.F = 156.99, p-value < 2.2e-16
```

```
Chow_mod3 <- Fstats(mod3$fitted ~ 1, from = 0.67)
sctest(Chow_mod3)
```

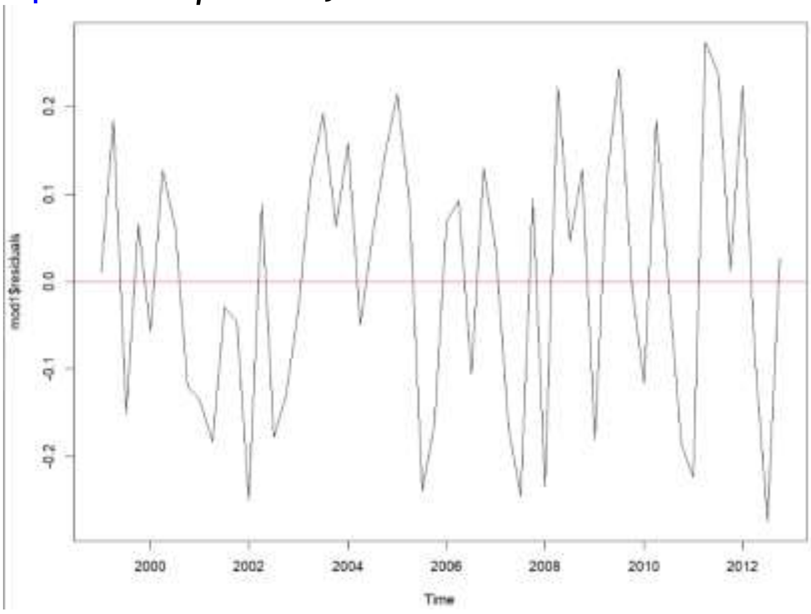
```
supF test

data:  Chow_mod3
sup.F = 157.84, p-value < 2.2e-16
```

En las tres pruebas se rechaza la hipótesis nula ($p < \alpha = 0.05$), es decir, NO existe estabilidad de coeficientes.

3.2 Análisis de los residuos

3.2.1 Media es igual a cero

```
plot(mod1$residuals)
abline(h = 0, col = "red")
t.test(mod1$residuals, mu = 0)
```



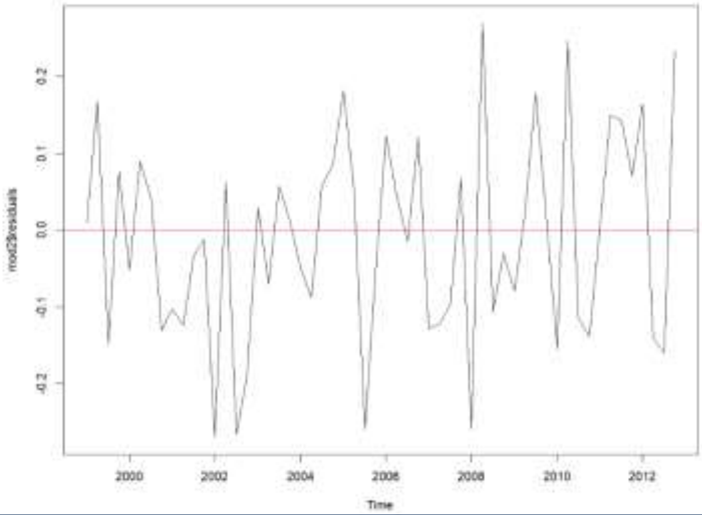

```
One Sample t-test

data:  mod1$residuals
t = 0.10559, df = 55, p-value = 0.9163
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.03876991  0.04308281
sample estimates:
  mean of x
```

Alumno: Maye Mamani Victor Raul

```
0.002156447
```

Parece indicar que un buen número de residuales están en torno a la media igual a cero. Para confirmar lo mencionado se realiza la prueba t: Como p = 0.9163 > α = 0.05, se acepta Ho, es decir la media es igual a cero.

```
plot(mod2$residuals)
abline(h = 0, col = "red")
t.test(mod2$residuals, mu = 0)
```

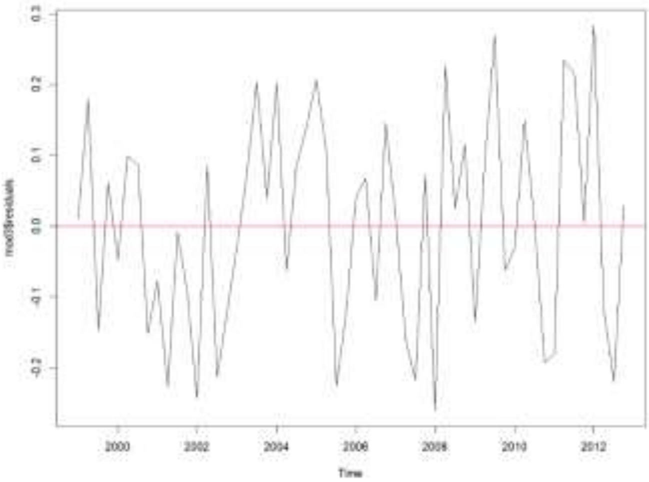



```
One Sample t-test

data:  mod2$residuals
t = -0.61655, df = 55, p-value = 0.5401
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.04695583  0.02486117
sample estimates:
  mean of x
-0.01104733
```

Parece indicar que un buen número de residuales están en torno a la media igual a cero. Para confirmar lo mencionado se realiza la prueba t: Como p = 0.5401 > α = 0.05, se acepta Ho, es decir la media es igual a cero.

```
plot(mod3$residuals)
abline(h = 0, col = "red")
t.test(mod3$residuals, mu = 0)
```

Alumno: Maye Mamani Victor Raul



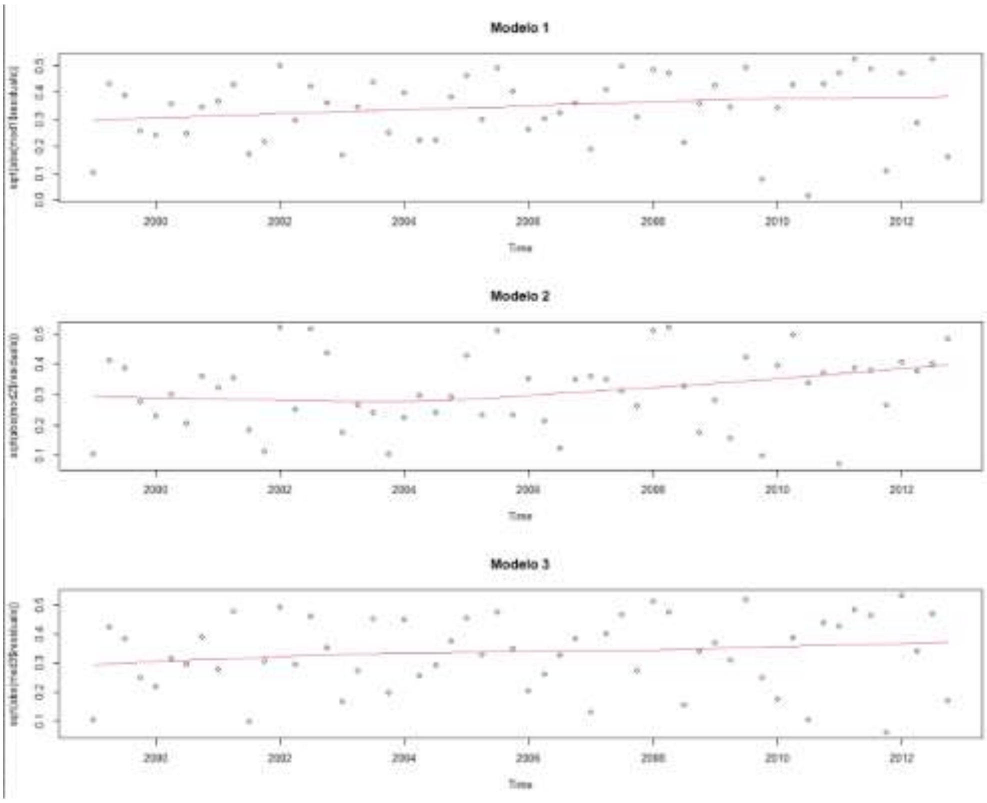

```
One Sample t-test

data:  mod3$residuals
t = 0.10756, df = 55, p-value = 0.9147
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.03789693  0.04219582
sample estimates:
  mean of x
0.002149444
```

Parece indicar que un buen número de residuales están en torno a la media igual a cero. Para confirmar lo mencionado se realiza la prueba t: Como p = 0.9147 > α = 0.05, se acepta Ho, es decir la media es igual a cero.

3.2.2 Homocedasticidad o varianza constante

```
par(mfrow = c(3,1))
scatter.smooth(sqrt(abs(mod1$residuals)), lpars=list(col=2), main =
"Modelo 1")
scatter.smooth(sqrt(abs(mod2$residuals)), lpars=list(col=2), main =
"Modelo 2")
scatter.smooth(sqrt(abs(mod3$residuals)), lpars=list(col=2), main =
"Modelo 3")
```

Alumno: Maye Mamani Victor Raul



Se observa que los datos parecen no presentan una variabilidad considerable, por tanto, será necesario realizar la prueba de Breusch-Pagan para determinar finalmente si las varianzas constantes para los modelos.

Prueba de Breusch – Pagan

```
obs=get(mod1$series)
bptest(resid(mod1)~I(obs-resid(mod1)))
```

```
studentized Breusch-Pagan test

data:  resid(mod1) ~ I(obs - resid(mod1))
BP = 4.9887, df = 1, p-value = 0.02551
```

```
obs=get(mod2$series)
bptest(resid(mod2)~I(obs-resid(mod2)))
```

```
studentized Breusch-Pagan test

data:  resid(mod2) ~ I(obs - resid(mod2))
BP = 1.8583, df = 1, p-value = 0.1728
```

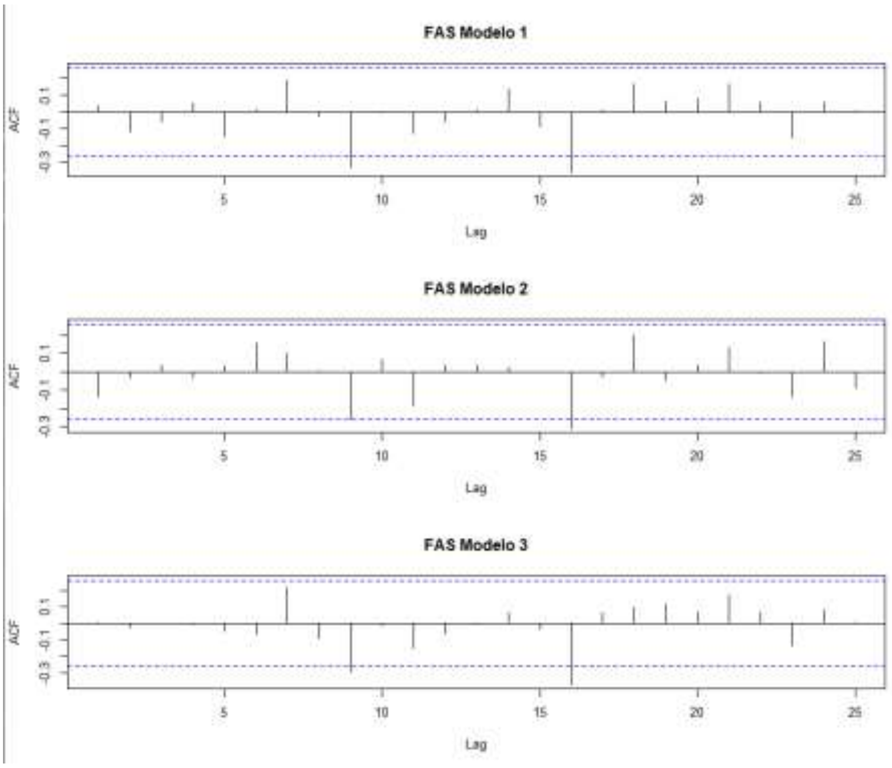
```
obs=get(mod3$series)
bptest(resid(mod3)~I(obs-resid(mod3)))
```

```
studentized Breusch-Pagan test

data:  resid(mod3) ~ I(obs - resid(mod3))
BP = 2.5607, df = 1, p-value = 0.1096
```

El valor de probabilidad (p-valor) asociado al estadístico BPasume un valor de 0.02551 para el MODELO 1, que es menor a α=0.05. Por tanto, estos residuos no tiene varianza constante.

En cambio en el MODELO 2 y MODELO3 tiene el valor de probabilidad de 0.1728 y 0.1096 respectivamente, que son mayores a α=0.05, por lo cual podemos afirmar que los residuales de estos modelos son constantes.

3.2.3 Ausencia de correlación serial



Se observa que casi la totalidad de los coeficientes del FAS para los modelos 1 , 2 y 3 se encuentran dentro de las bandas de no significación, sobre todo los de los primeros retardos.

Por tanto, tenemos altos indicios de que los residuos de los 3 modelos sean ruido blanco.

```
Box.test(resid_m1,type = "Ljung-Box")
Box-Ljung test

data:  resid_m1
X-squared = 0.079631, df = 1, p-value = 0.7778
```

```
Box.test(resid_m2,type = "Ljung-Box")
Box-Ljung test

data:  resid_m2
X-squared = 1.1497, df = 1, p-value = 0.2836
```

```
Box.test(resid_m3,type = "Ljung-Box")

        Box-Ljung test

data:  resid_m3
```

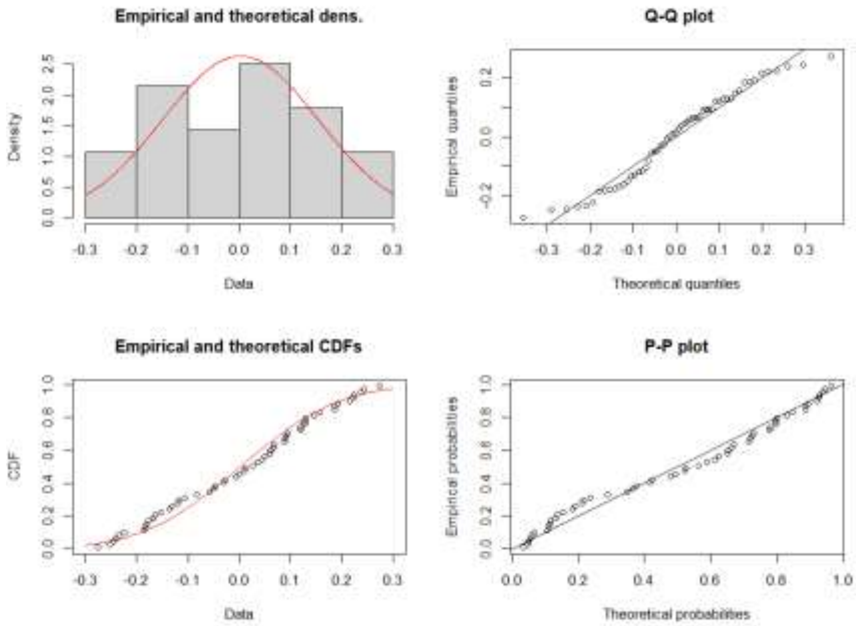
Alumno: Maye Mamani Victor Raul

```
X-squared = 0.0026241, df = 1, p-value = 0.9591
```

Con los valores 0.7778, 0.2836 y 0.9591 son mayores a α=0.05, por lo que se acepta la hipótesis nula de que los coeficientes de autocorrelación son cero; es decir, los residuos son independientes o están incorrelacionados.

3.2.4 Contraste de normalidad

```
ajuste_m1<-fitdist(data = resid_m1, distr="norm")
plot(ajuste_m1)
JB_m1 <- jarque.bera.test(resid_m1)
JB_m1
```

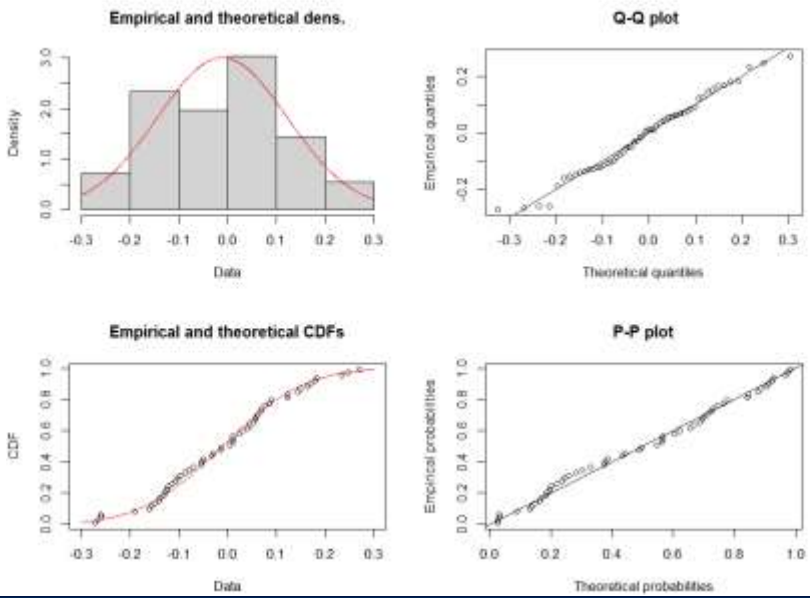



```
Jarque Bera Test

data:  resid_m1
X-squared = 3.0068, df = 2, p-value = 0.2224
```

En las figuras se observa que los residuales del modelo 1 presentan indicios de ajuste hacia la distribución normal. En la prueba JB, como p = 0.2224 > 0.05, se acepta Ho, es decir, los residuos se aproximan a una distribución normal.

```
ajuste_m2<-fitdist(data = resid_m2, distr="norm")
plot(ajuste_m2)
JB_m2 <- jarque.bera.test(resid_m2)
JB_m2
```

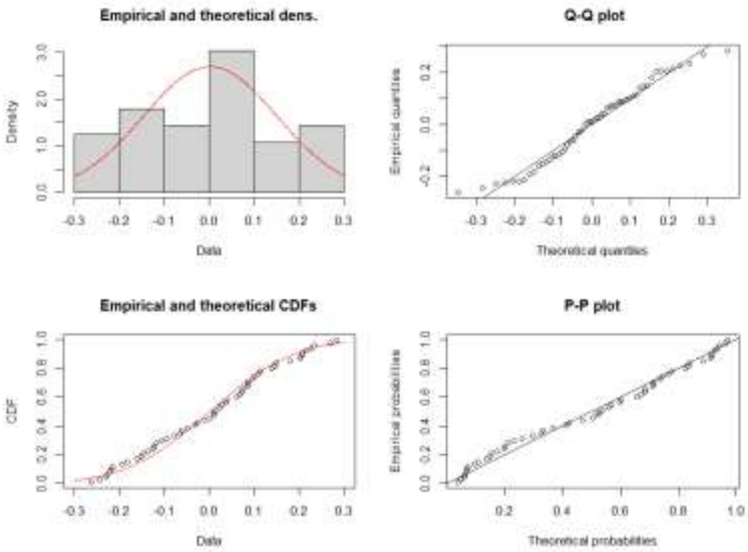
Alumno: Maye Mamani Victor Raul




```
Jarque Bera Test

data:  resid_m2
X-squared = 0.88085, df = 2, p-value = 0.6438
```

En las figuras se observa que los residuales del modelo 2 presentan indicios de ajuste hacia la distribución normal. En la prueba JB, como p = 0.6438 > 0.05, se acepta Ho, es decir, los residuos se aproximan a una distribución normal.

```
ajuste_m3<-fitdist(data = resid_m3, distr="norm")
plot(ajuste_m3)
JB_m3 <- jarque.bera.test(resid_m3)
JB_m3
```




```
Jarque Bera Test

data:  resid_m3
X-squared = 2.4313, df = 2, p-value = 0.2965
```

Alumno: Maye Mamani Victor Raul

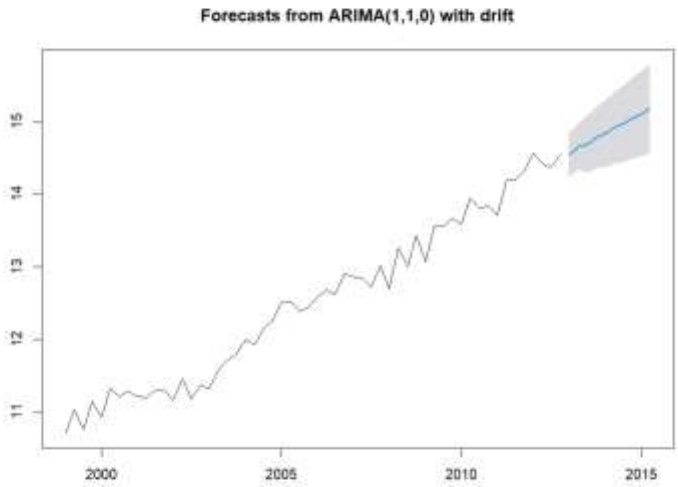
En las figuras se observa que los residuales del modelo 3 presentan indicios de ajuste hacia la distribución normal. En la prueba JB, como p = 0.2965 > 0.05, se acepta Ho, es decir, los residuos se aproximan a una distribución normal.

4 Pronostico

4.1Pronosticos de cada modelo

Modelo 1: ARIMA (1,1,0)

```
Pron1 <- forecast(mod1,level=c(95),h=10)
plot(Pron1)
summary(Pron1)
```



En la figura se puede observar el comportamiento de la función de predicción por punto y por intervalo.

Los datos proyectados para los siguientes 10 trimestres son:

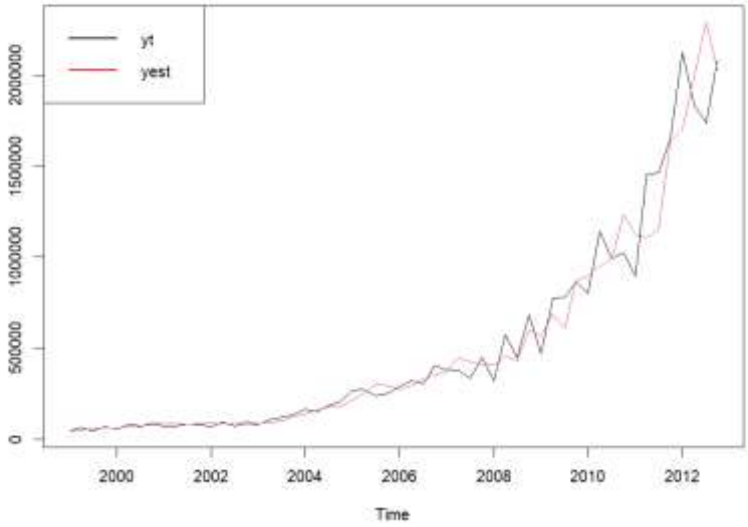
```
Forecast method: ARIMA(1,1,0) with drift

Model Information:
Series: data_mod
ARIMA(1,1,0) with drift

Coefficients:
          ar1   drift
      -0.7159  0.0672
s.e.   0.0929  0.0121

sigma^2 = 0.02424:  log likelihood = 24.91
AIC=-43.82   AICc=-43.35   BIC=-37.8

Error measures:
                      ME      RMSE       MAE          MPE     MAPE
MASE       ACF1
Training set 0.002156447 0.1514681 0.1302622 -0.001002706 1.037708 0.4
450654 0.03672105

Forecasts:
        Point Forecast    Lo 95    Hi 95
2013 Q1       14.53611 14.23095 14.84127
2013 Q2       14.66134 14.34410 14.97857
2013 Q3       14.68697 14.28731 15.08664
2013 Q4       14.78391 14.36328 15.20453
2014 Q1       14.82980 14.35909 15.30051
2014 Q2       14.91223 14.41701 15.40746
2014 Q3       14.96851 14.43628 15.50074
```

Alumno: Maye Mamani Victor Raul

```
2014 Q4        15.04351 14.48612 15.60090
2015 Q1        15.10510 14.51730 15.69290
2015 Q2        15.17630 14.56397 15.78862
```

Es importante mencionar que, los valores pronosticados son logaritmos, por tanto, para obtener los verdaderos valores de pronóstico, se tendrán que transformar aplicando la función exp().

```
data_modelo1 <- exp(mod1$fitted)
grafico_comparativo <- cbind(data_ts,data_modelo1)
ts.plot(grafico_comparativo, col=c(1,2), lwd = 1)
legend("topleft",c("yt","yest"),lty = c(1,1), lwd = 2,col=c("black",
"red"))
```



PRONÓSTICO DE LA SERIE ORIGINAL

Deshaciendo la transformación:

```
Pron1$mean <- exp(Pron1$mean)
Pron1$lower <- exp(Pron1$lower)
Pron1$upper <- exp(Pron1$upper)
Pron1$x <- exp(Pron1$x)
Pron1$fitted <- exp(Pron1$fitted)
Pron1$residuals <- exp(Pron1$residuals)
summary(Pron1)
```

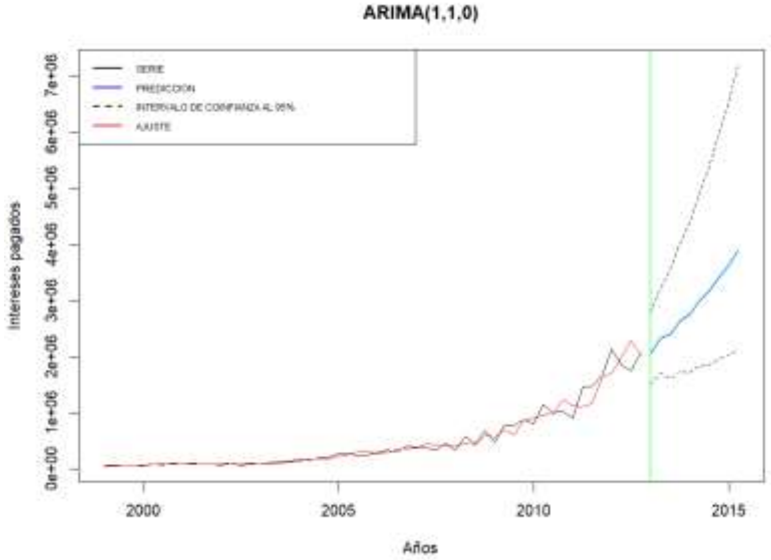
```
Forecast method: ARIMA(1,1,0) with drift

Model Information:
Series: data_mod
ARIMA(1,1,0) with drift

Coefficients:
          ar1   drift
      -0.7159  0.0672
s.e.   0.0929  0.0121

sigma^2 = 0.02424:  log likelihood = 24.91
AIC=-43.82   AICc=-43.35   BIC=-37.8

Error measures:
```

Alumno: Maye Mamani Victor Raul

```
                  ME      RMSE       MAE        MPE      MAPE       MASE
ACF1
Training set 6211.225 132113.9 71919.14 -0.9399903 13.13878 0.4836564
0.0547565

Forecasts:
        Point Forecast   Lo 95   Hi 95
2013 Q1        2055669 1515042 2789213
2013 Q2        2329903 1696545 3199708
2013 Q3        2390408 1602876 3564876
2013 Q4        2633723 1729399 4010930
2014 Q1        2757406 1722164 4414961
2014 Q2        2994341 1824848 4913330
2014 Q3        3167672 1860355 5393673
2014 Q4        3414391 1955433 5961884
2015 Q1        3631306 2017363 6536445
2015 Q2        3899254 2113747 7192999
```

GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

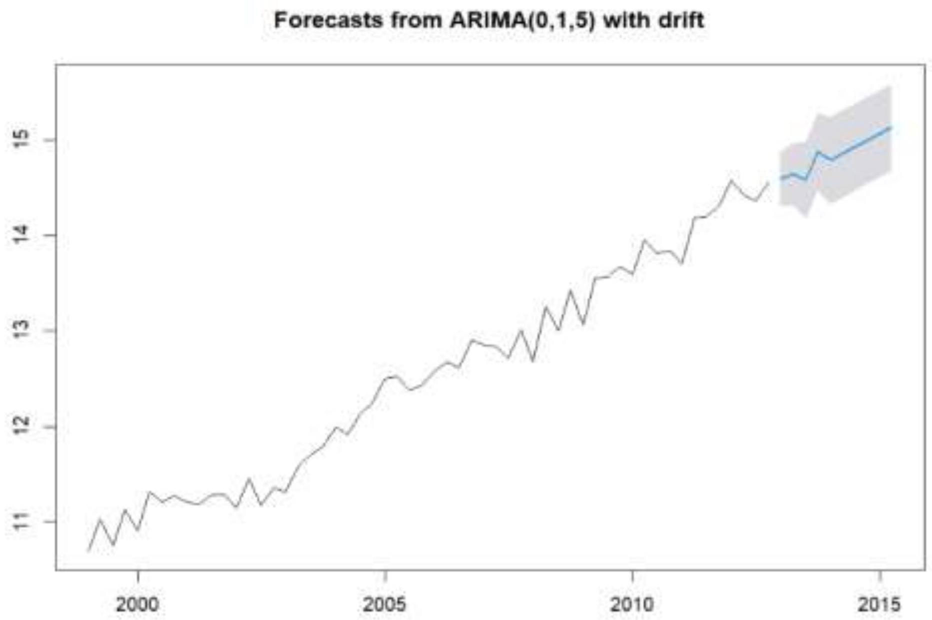
```
plot(Pron1, shaded = FALSE, xlab = "Años", ylab = "Intereses
pagados",main = "ARIMA(1,1,0)")
lines(Pron1$fitted, col = "red")
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE
COINFIANZA AL 95%", "AJUSTE"),col=c("black", "blue", "black", "red"),
lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=2013, lwd = 1, col="green")
```



Modelo 2: ARIMA (0,1,5)

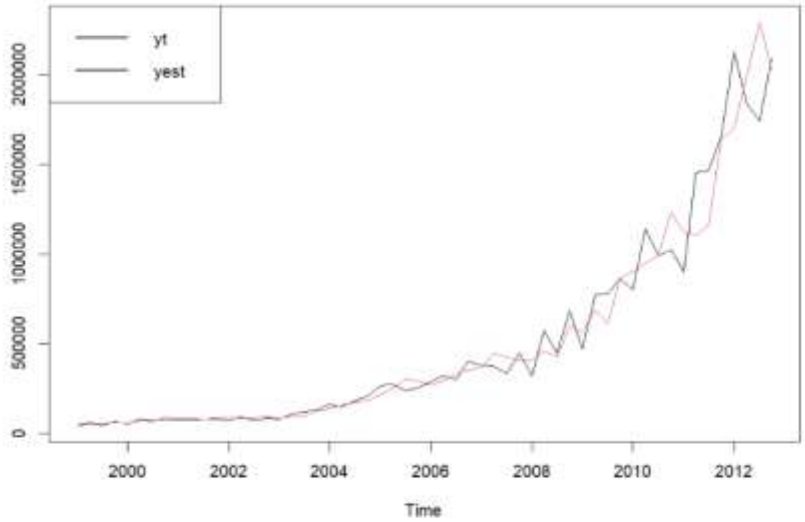
```
Pron2 <- forecast(mod2,level=c(95),h=10)
plot(Pron2)
summary(Pron2)
```

Alumno: Maye Mamani Victor Raul



PRONÓSTICO DE LA SERIE ORIGINAL

```
yt_arima2 <- exp(mod2$fitted)
grafico_comparativo <- cbind(Yt,yt_arima2)
ts.plot(grafico_comparativo, col=c(1,2), lwd = 1)
legend("topleft",c("yt","yest"),lty = c(1,1), lwd = 2)
```



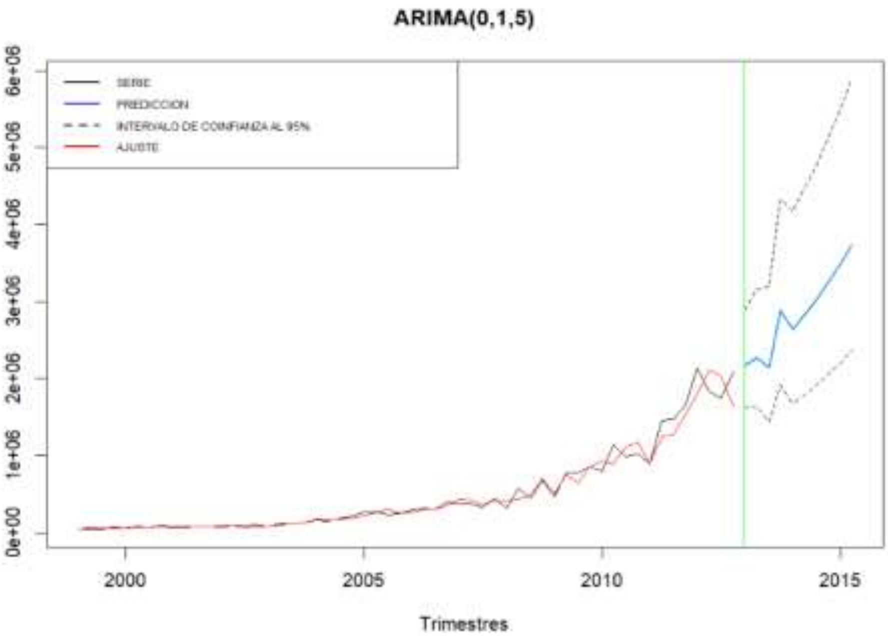

```
Pron2$mean <- exp(Pron2$mean)
Pron2$lower <- exp(Pron2$lower)
Pron2$upper <- exp(Pron2$upper)
Pron2$x <- exp(Pron2$x)
Pron2$fitted <- exp(Pron2$fitted)
Pron2$residuals <- exp(Pron2$residuals)
summary(Pron2)
```

```
Forecast method: ARIMA(0,1,5) with drift

Model Information:
```

```
Series: data_mod
ARIMA(0,1,5) with drift

Coefficients:
          ma1     ma2      ma3     ma4      ma5   drift
      -0.4485  0.2634  -0.6407  0.5254  -0.6995  0.0694
s.e.   0.1576  0.2749   0.1199  0.2014   0.2721  0.0033

sigma^2 = 0.02032:  log likelihood = 28.17
AIC=-42.33   AICc=-39.95   BIC=-28.28

Error measures:
                  ME     RMSE      MAE       MPE     MAPE      MASE
ACF1
Training set 9162.53 114974.8 64506.76 -2.005822 11.21007 0.4338081 -0
.1393904

Forecasts:
        Point Forecast   Lo 95   Hi 95
2013 Q1        2165890 1626312 2884490
2013 Q2        2276487 1642389 3155400
2013 Q3        2144462 1433980 3206961
2013 Q4        2886901 1921031 4338399
2014 Q1        2638674 1672330 4163414
2014 Q2        2828412 1792581 4462791
2014 Q3        3031794 1921480 4783696
2014 Q4        3249800 2059647 5127675
2015 Q1        3483482 2207749 5496389
2015 Q2        3733968 2366501 5891615
```

GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

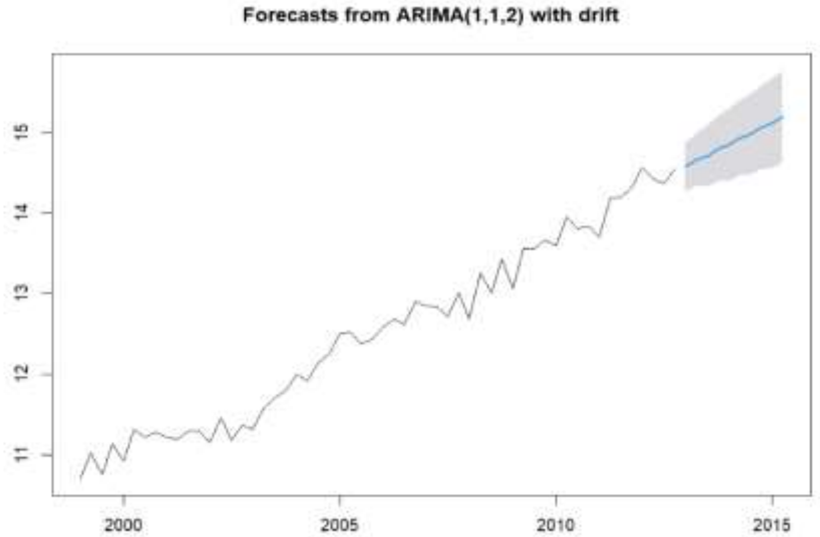
```
plot(Pron2, shaded = FALSE, xlab = "Trimestres", ylab = "",main =
"ARIMA(0,0,2)")
lines(Pron2$fitted, col = "red")
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE
COINFIANZA AL 95%", "AJUSTE"),col=c("black", "blue", "black", "red"),
lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=1930, lwd = 1, col="green")
```



Alumno: Maye Mamani Victor Raul

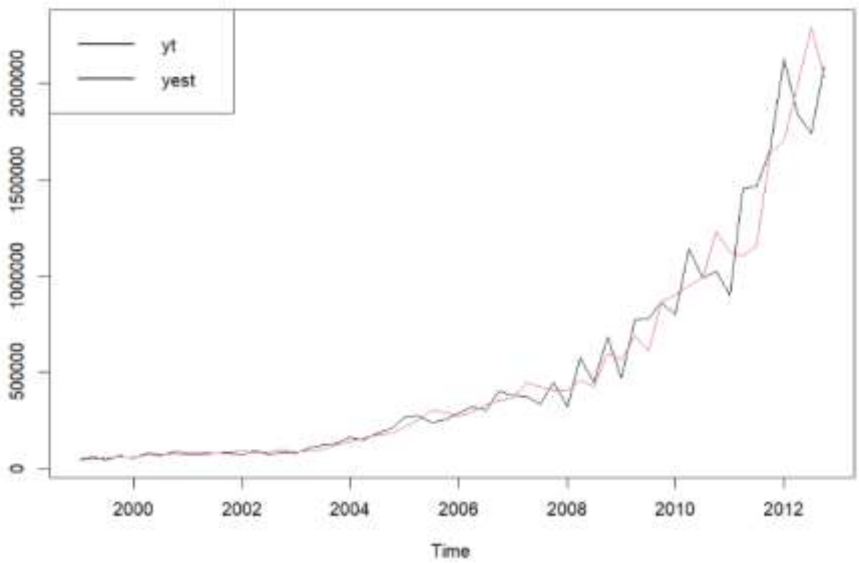
Modelo 3: ARIMA (1,1,2)

```
Pron3 <- forecast(modelo3,level=c(95),h=10)
plot(Pron3)
summary(Pron3)
```



PRONÓSTICO DE LA SERIE ORIGINAL

```
yt_arima3 <- exp(modelo3$fitted)
grafico_comparativo <- cbind(Yt,yt_arima1)
ts.plot(grafico_comparativo, col=c(1,2), lwd = 1)
legend("topleft",c("yt","yest"),lty = c(1,1), lwd = 2)
```

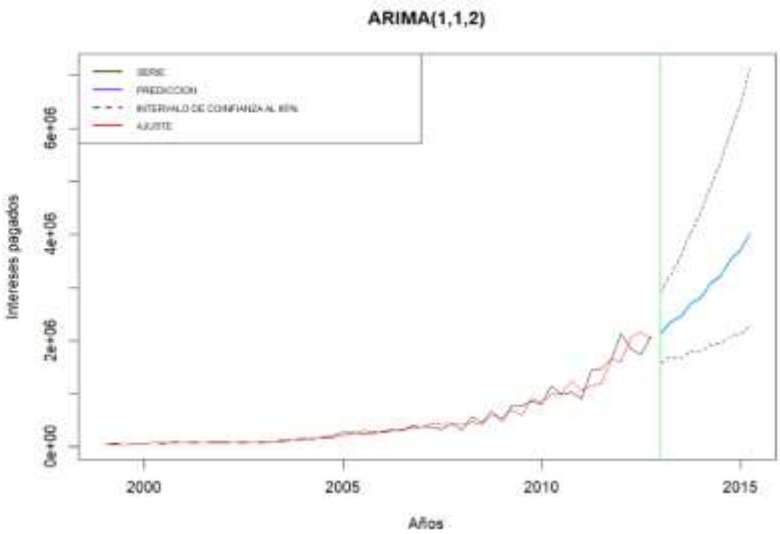


GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

```
Pron3$mean <- exp(Pron3$mean)
Pron3$lower <- exp(Pron3$lower)
Pron3$upper <- exp(Pron3$upper)
Pron3$x <- exp(Pron3$x)
Pron3$fitted <- exp(Pron3$fitted)
Pron3$residuals <- exp(Pron3$residuals)
summary(Pron3)
```

```
Forecast method: ARIMA(1,1,2) with drift

Model Information:
Series: data_mod
ARIMA(1,1,2) with drift

Coefficients:
          ar1     ma1      ma2   drift
      -0.9065  0.2437  -0.2397  0.0675
s.e.   0.0934  0.1713   0.1630  0.0108

sigma^2 = 0.02412:  log likelihood = 26.03
AIC=-42.06   AICc=-40.84   BIC=-32.03

Error measures:
                   ME     RMSE   MAE         MPE     MAPE       MASE
ACF1
Training set 8205.438 126569.1 68489 -0.8861896 12.64779 0.4605887 0.0
06992194

Forecasts:
        Point Forecast   Lo 95   Hi 95
2013 Q1        2136975 1576165 2897326
2013 Q2        2360379 1711867 3254569
2013 Q3        2453377 1669058 3606260
2013 Q4        2694501 1803620 4025422
2014 Q1        2815128 1795102 4414761
2014 Q2        3077400 1931037 4904303
2014 Q3        3228810 1947797 5352310
2014 Q4        3516103 2087372 5922749
2015 Q1        3701954 2125720 6446973
2015 Q2        4018652 2271042 7111081
```

```
plot(Pron3, shaded = FALSE, xlab = "Años", ylab = "N° DE PIELES",main
= "ARIMA(1,1,2)")
lines(Pron3$fitted, col = "red")
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE
COINFIANZA AL 95%", "AJUSTE"),col=c("black", "blue", "black", "red"),
lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=1930, lwd = 1, col="green")
```



Métricas basadas en el error

```
accuracy(Pron1)
```

Alumno: Maye Mamani Victor Raul

```
                        ME     RMSE      MAE        MPE     MAPE      MASE
ACF1
Training set 6211.225 132113.9 71919.14 -0.9399903 13.13878 0.4836564
0.0547565
```

```
accuracy(Pron2)
```

```
                       ME     RMSE      MAE       MPE     MAPE      MASE
ACF1
Training set 9162.53 114974.8 64506.76 -2.005822 11.21007 0.4338081 -0
.1393904
```

```
accuracy(Pron3)
```

```
                        ME     RMSE   MAE        MPE     MAPE      MASE
ACF1
Training set 8205.438 126569.1 68489 -0.8861896 12.64779 0.4605887 0.0
06992194
```

Basado en las métricas MAE y RMSE, el modelo Pron2 es el mejor entre los tres, ya que tiene los valores más bajos(MAE = 64506 Y RMSE=114974) en ambas métricas. Esto indica que sus predicciones son más precisas y con menos variabilidad en los errores.

Conclusión

Sobre los modelos:

ARIMA (1,1,0)

- **Coeficientes Significativos**: Todos los coeficientes son altamente significativos.
- **Residuos**: Los residuos tienen una media cercana a cero y no presentan autocorrelación.
- **Estabilidad**: Las pruebas de Chow indicaron que no hay estabilidad de coeficientes, lo cual es un punto en contra.
- **Condición de Invertibilidad y Convergencia**: Cumple con las condiciones.

ARIMA (0,1,5)

- **Coeficientes Significativos**: Tres de los cinco coeficientes de media móvil son significativos.
- **Residuos**: Los residuos tienen una media cercana a cero y no presentan autocorrelación.
- **Estabilidad**: Al igual que el modelo anterior, no muestra estabilidad de coeficientes en las pruebas de Chow.
- **Condición de Invertibilidad y Convergencia**: Cumple con las condiciones.

ARIMA (1,1,2)

- **Coeficientes Significativos**: El coeficiente AR es significativo, pero los coeficientes MA no lo son.
- **Residuos**: Los residuos tienen una media cercana a cero y no presentan autocorrelación.
- **Estabilidad**: Similar a los otros dos modelos, no presenta estabilidad de coeficientes.
- **Condición de Invertibilidad y Convergencia**: Cumple con las condiciones.

Alumno: Maye Mamani Victor Raul

Teniendo en cuenta todos los aspectos analizados, el** modelo ARIMA (1,1,0)* ***es el más adecuado para la serie de datos evaluada. Esto se debe a que todos sus coeficientes son altamente significativos, los residuos cumplen con los supuestos necesarios, y cumple con las condiciones de invertibilidad y convergencia. Aunque no presenta estabilidad en los coeficientes, este aspecto es una limitación compartida con los otros modelos evaluados. Pero si lo que mas nos importa es el error el modelo ARIMA(0,1,5) tiene menor error y cumple con los mismo supuestos del modelo ARIMA(1,1,0)