Alumno: Maye Mamani Victor Raul

Tabla de contenido

Caso 1 3
1 Identificación 3
1. 1 Análisis de la tendencia y la estacionalidad 4
1.1.1 Estacionalidad 4
1.1.2 Análisis de tendencia 4
1.2 Análisis de estacionariedad 5
1.2.1 Estacionariedad en varianza 5
1.2.2 Estacionariedad en media 9
1.3 Identificación del modelo estacionario 12
1.3.1 Identificación de las órdenes p y q 12
1.3.2 Inclusión del término independiente (δ) o intercepto 12
2 Estimación 13
3 Validación 13
3.1 Análisis de los coeficientes estimados 13
3.1.1 Significación de los coeficientes 13
3.1.2 Examen de la matriz de correlaciones entre los coeficientes 14
3.1.3 Condición de convergencia e invertibilidad 14
3.1.4 Análisis de la estabilidad 15
3.2 Análisis de los residuos 16
3.2.1 Media es igual a cero 16
3.2.2 Homocedasticidad o varianza constante 18
3.2.3 Ausencia de correlación serial 20
3.2.4 Contraste de normalidad 21
4 Pronostico 23
4.1Pronosticos de cada modelo 23
Modelo 1: ARIMA(2,1,0) 23
Modelo 2: ARIMA(0,1,1) 26
Modelo 3: ARIMA(1,1,2) 29
Métricas basadas en el error 32
Conclusión 33
CASO 2: PRODUCCION DE PAPA 34
1 Identificación 34
1. 1 Análisis de la tendencia y la estacionalidad 34
1.1.1 Estacionalidad 34
1.1.2 Análisis de tendencia 34
1.2 Análisis de estacionariedad 35

1.2.1 Estacionariedad en varianza 35
1.2.2 Estacionariedad en media 38
1.3 Identificación del modelo estacionario 41
1.3.1 Identificación de las órdenes p y q 41
1.3.2 Inclusión del término independiente (δ) o intercepto 41
2 Estimación 42
3 Validación 42
3.1 Análisis de los coeficientes estimados 42
3.1.1 Significación de los coeficientes 42
3.1.2 Examen de la matriz de correlaciones entre los coeficientes 43
3.1.3 Condición de convergencia e invertibilidad 43
3.1.4 Análisis de la estabilidad 45
3.2 Análisis de los residuos 45
3.2.1 Media es igual a cero 45
3.2.2 Homocedasticidad o varianza constante 47
3.2.3 Ausencia de correlación serial 49
3.2.4 Contraste de normalidad 50
4 Pronostico 52
4.1Pronosticos de cada modelo 52
Modelo 1: ARIMA (1,1,0) 52
Modelo 2: ARIMA (0,1,1) 56
Modelo 3: ARIMA (1,1,1) 58
Métricas basadas en el error 61
Conclusión 61

Alumno: Maye Mamani Victor Raul

Actividad 6

```r
# 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)
```

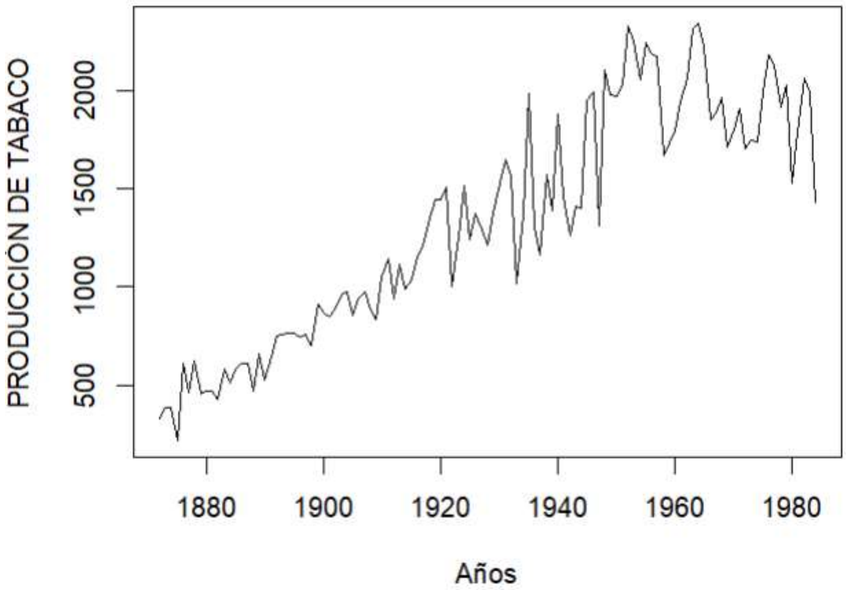
Caso 1

1 Identificación

Grafica inicial de la serie:

```r
data <- read_excel("F:\\777--Programacion
repos\\Una\\r\\data\\actividad-06.xlsx",sheet = "01")
View(data)
# Gráfica de la serie
data_ts <- ts(data$Yt, start = c(1872,1), frequency = 1)
plot(data_ts, xlab="Años", ylab="PRODUCCIÓN DE TABACO")
```

Alumno: Maye Mamani Victor Raul



En primera instancia podemos visualizar una clara TENDENCIA creciente, la varianza no parece variar mucho, y no es estacionaria en media. Y al ser una serie anual NO haremos un análisis de estacionalidad.

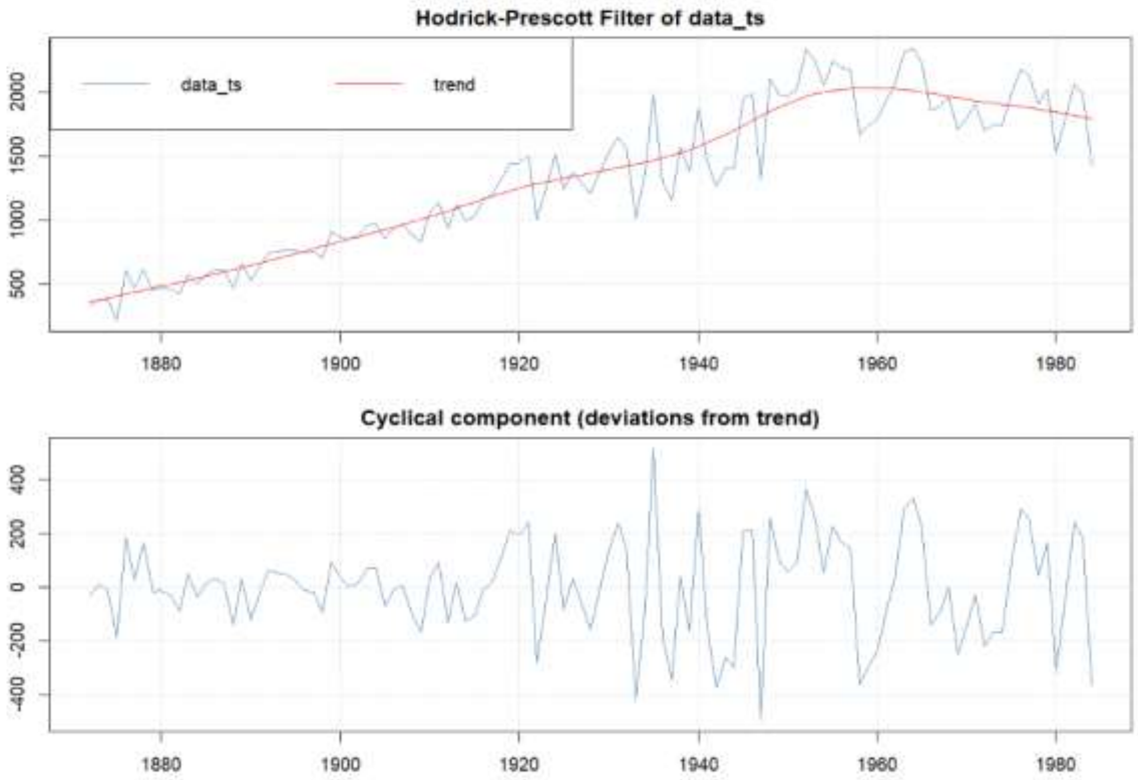
1. 1 Análisis de la tendencia y la estacionalidad

1.1.1 Estacionalidad

Debido a que se trata de un gráfico con un período anual, no se observa estacionalidad.

1.1.2 Análisis de tendencia

```
lambda_hp <- 1000
data_hp <- hpfilter(data_ts, type="lambda", freq=lambda_hp)
plot(data_hp)
```

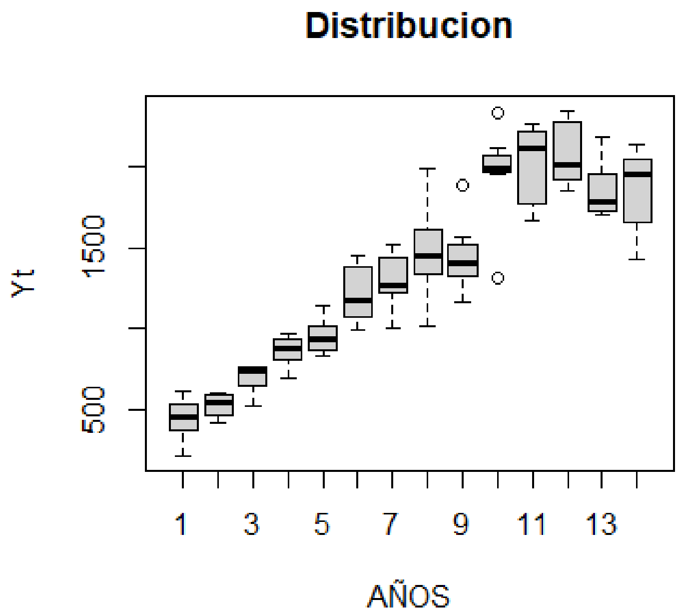


Hay una tendencia creciente hasta el año 1955 aproximadamente tras la cual se muestra un breve descenso en la misma. Y no parece haber algún patrón que se repita cíclicamente

1.2 Análisis de estacionariedad

1.2.1 Estacionariedad en varianza

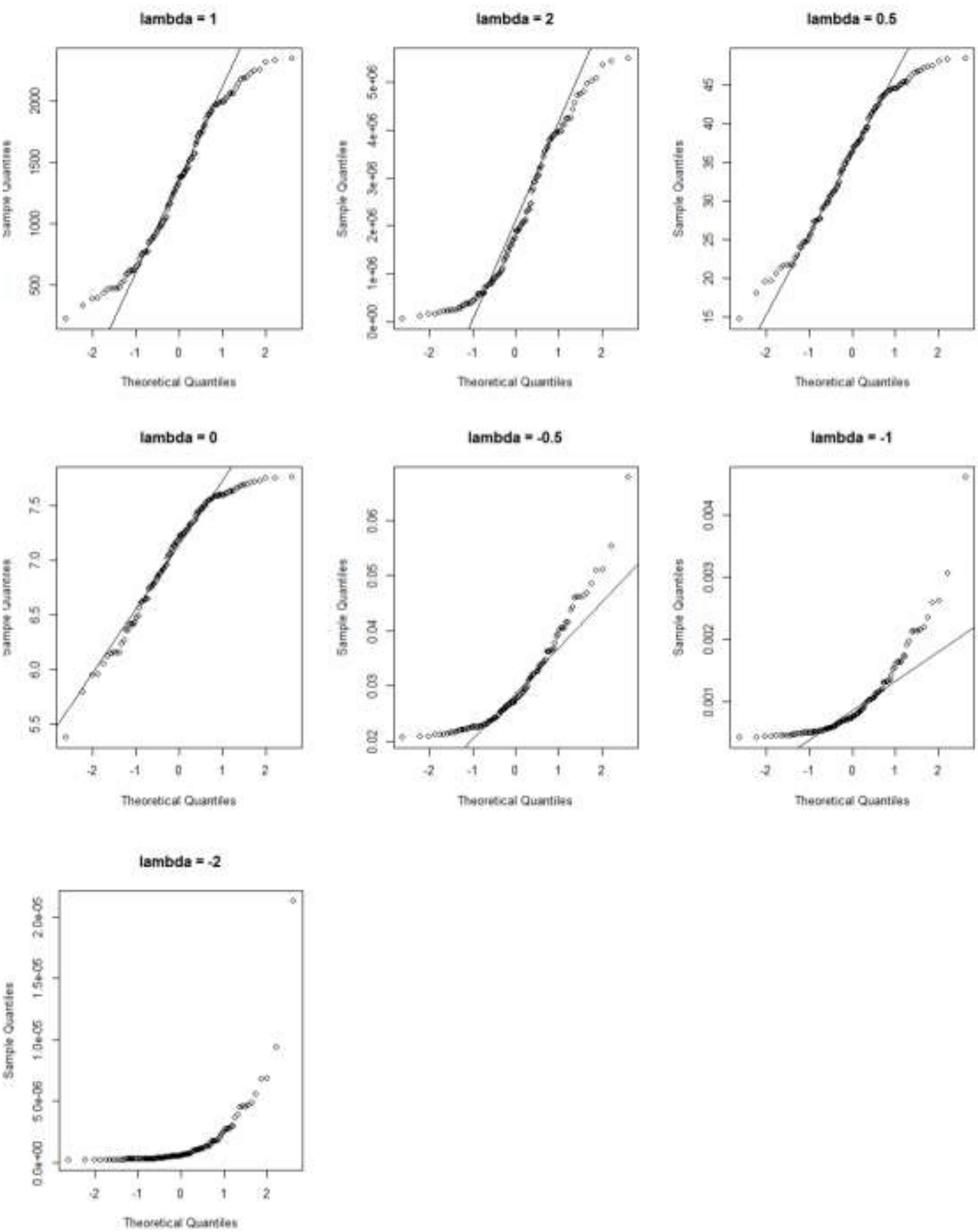
```
boxplot(data$IPC ~ data$Año, xlab = "AÑOS/MESES", ylab="IPC",
main="Distribucion" )
```



Alumno: Maye Mamani Victor Raul

Tenemos ciertos indicios de estacionariedad en varianza, sin embargo en los últimos años esta presenta un cambio significativo en la serie lo que nos sugiere que no existe estacionariedad en varianza.

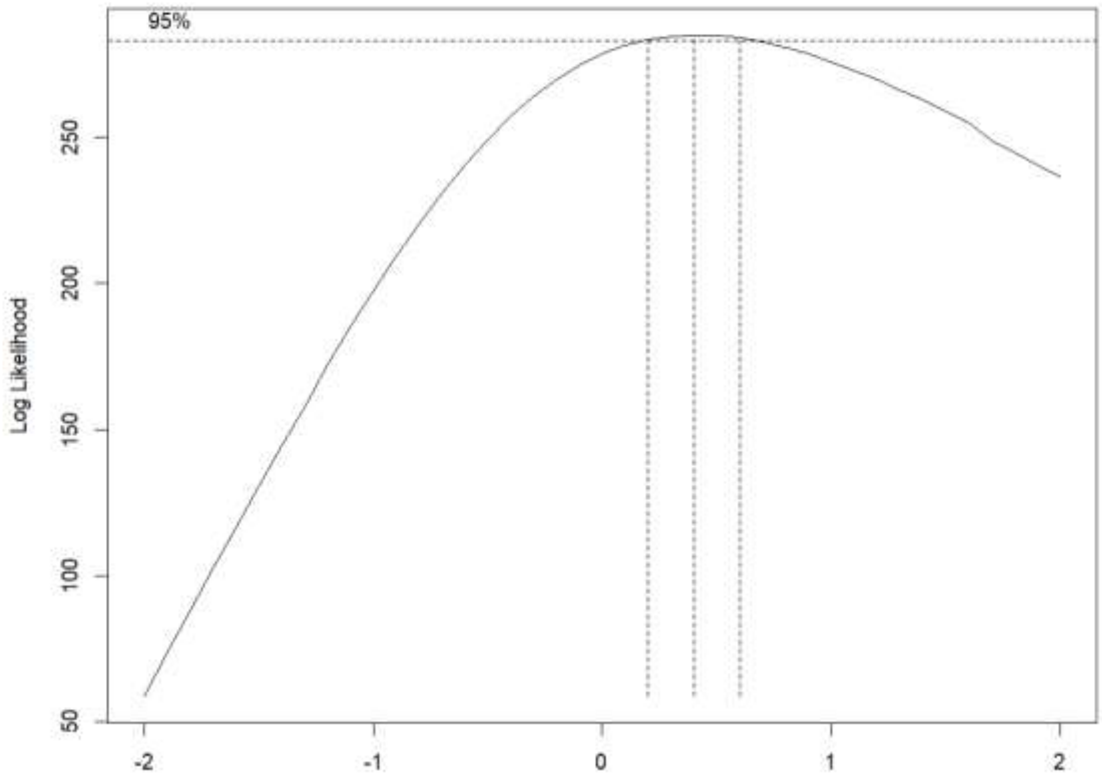
```
qqnorm(data_ts,main="lambda = 1")
qqline(data_ts)

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)

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



LAMBDA OPTIMO

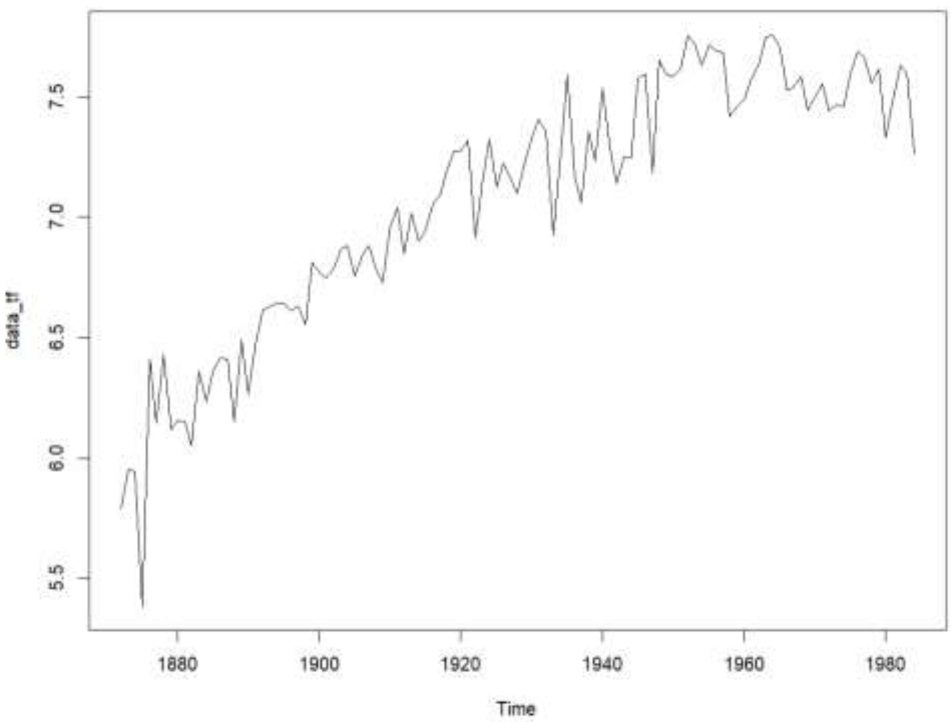
```
b <- BoxCox.ar(data_ts)
lambda <- b$mle
round(lambda,2)
```

```
[1] 0.4
```



Se observa una mejora en el ajuste a la distribución normal de la serie original al aplicar lambda 0.

```
data_tf <- log(data_ts)
plot(data_tf)
```



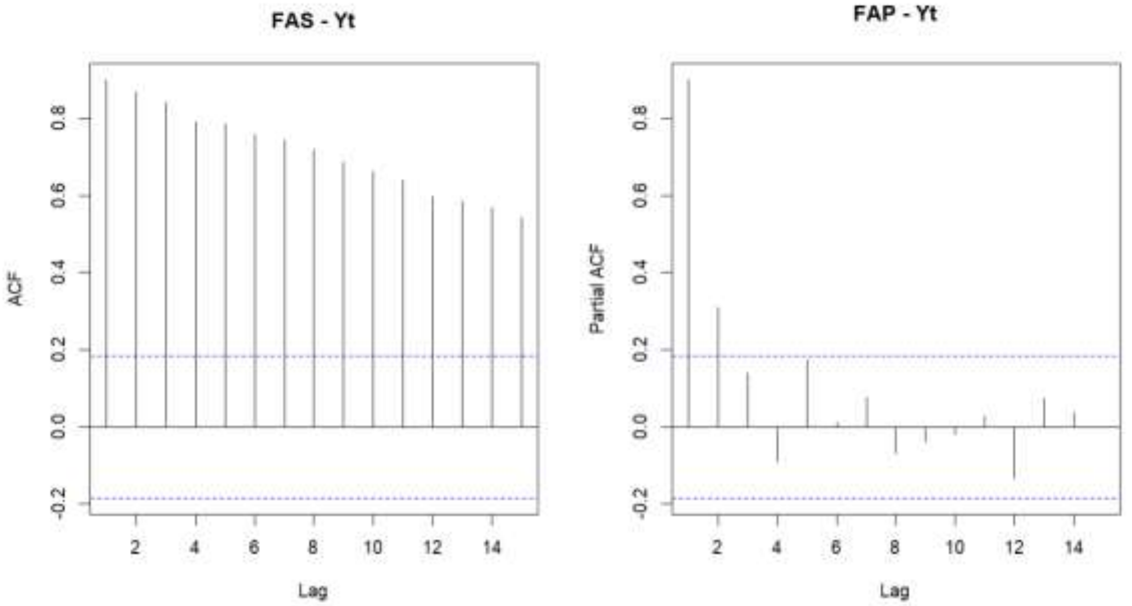
Se puede observar una mejora en la estacionariedad en varianza.

Alumno: Maye Mamani Victor Raul

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]




```
[1] 0.9689437
```

El FAS decrece lentamente y el primer FAP es significativo siendo muy cercano a 0.9, por lo que podríamos decir que la serie NO ES ESTACIONARIA, pero de igual forma haremos una prueba confirmatoria.

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

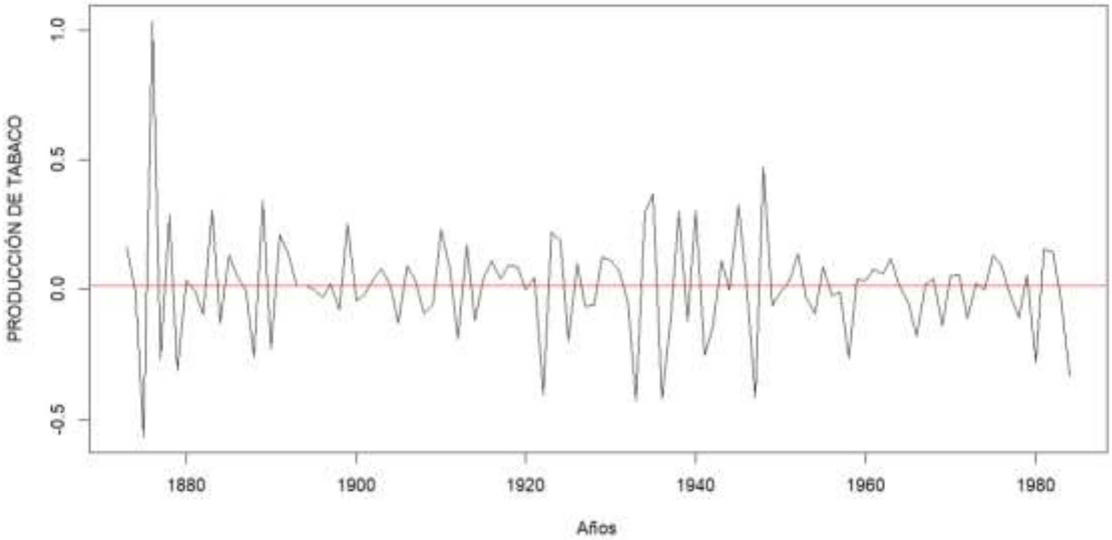
```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-0.71596 -0.08100  0.02031  0.09813  0.53372

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.885199   0.555697   3.392 0.000972 ***
z.lag.1     -0.294094   0.089150  -3.299 0.001319 **
tt           0.003753   0.001428   2.628 0.009844 **
z.diff.lag  -0.318628   0.093425  -3.411 0.000916 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1752 on 107 degrees of freedom
```

```
Multiple R-squared:  0.286,    Adjusted R-squared:  0.266
F-statistic: 14.29 on 3 and 107 DF,  p-value: 6.743e-08


Value of test-statistic is: -3.2989 4.5669 6.1851

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.99 -3.43 -3.13
phi2  6.22  4.75  4.07
phi3  8.43  6.49  5.47
```

Observamos que el T calculado (-3.2989) es MAYOR que el T critico (-3.43) por tanto se rechaza la hipótesis nula de la existencia de raíz unitaria, es **decir que la serie NO ES ESTACIONARIA**

Diferenciamos la serie

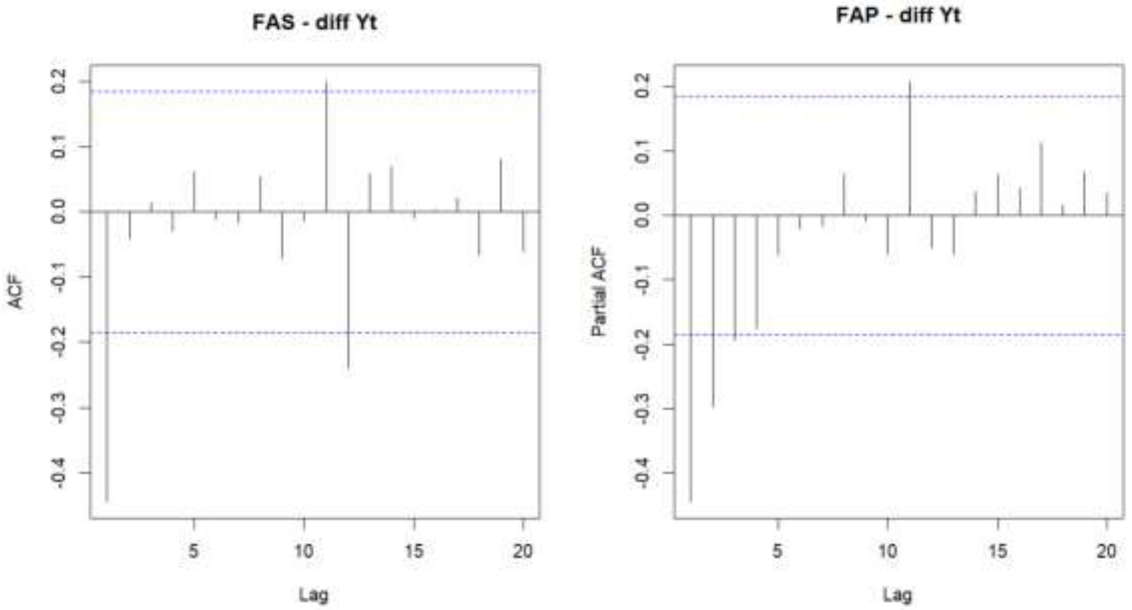
```
data_diff = diff(data_mod)
plot(data_diff, xlab="Meses", ylab="IPC")
abline(h = mean(data_diff), col = "red")
```



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]
```

Alumno: Maye Mamani Victor Raul



Hay un decrecimiento rapido en el FAS, y el valor del primer coeficiente de las FAP es menor a 0.9, por lo que existen indicios de que la serie ES ESTACIONARIA

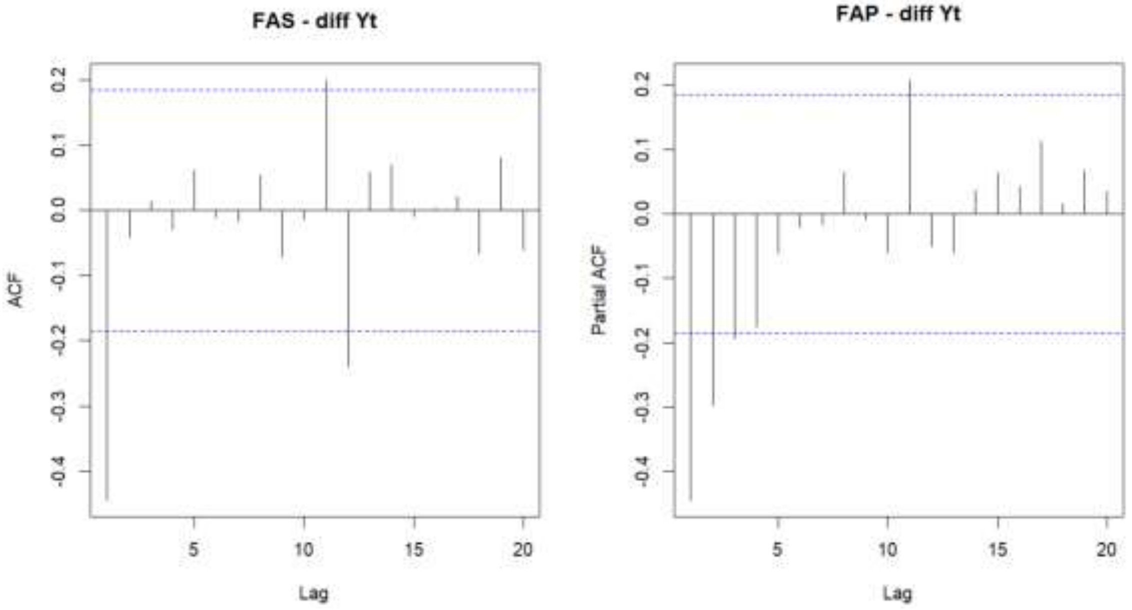
Prueba cofirmatoria

```
data_adf <- ur.df(data_diff, type="trend", lags = 1)
summary(data_adf)
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-0.54422 -0.09446 -0.00272  0.09511  0.66017

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.02644    0.01692   1.562 0.121155
z.lag.1     -1.92761    0.15958 -12.079  < 2e-16 ***
z.diff.lag   0.32145    0.09325   3.447 0.000811 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1755 on 107 degrees of freedom
Multiple R-squared:  0.7499,	Adjusted R-squared:  0.7453
F-statistic: 160.5 on 2 and 107 DF,  p-value: < 2.2e-16


Value of test-statistic is: -12.0794 72.986

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81
```

Observamos que el T calculado (-12.0767) 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.

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 2 coeficientes significativos MA(1). Y el FAP decrecen forma de sinusoidal y tiene 2 coeficientes diferentes de cero por lo que también planteamos un AR(2). Adicionalmente planteamos un ARIMA(2,1).

1.3.2 Inclusión del término independiente (δ) o 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
  35.411314    1.981567
```

Se rechaza la hipótesis nula y se incluye la constante.

Resumiendo, se proponen los siguientes modelos

ARIMA(2,1,0)

ARIMA(0,1,1)

Alumno: Maye Mamani Victor Raul

ARIMA(1,1,2)

2 Estimación

```r
mod1 <- Arima(data_tf, order = c(2, 1, 0), include.constant = T)
coeftest(mod1)
```

```
z test of coefficients:

        Estimate Std. Error z value  Pr(>|z|)    
ar1   -0.6018932  0.0912564 -6.5956 4.235e-11 ***
ar2   -0.3173850  0.0906412 -3.5016 0.0004626 ***
drift  0.0141191  0.0085184  1.6575 0.0974214 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$\Delta Y_t = 0.0141 + -0.6018\Delta Y_{t-1} - 0.3173\Delta Y_{t-2} + a_t$$

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

```
z test of coefficients:

        Estimate Std. Error  z value  Pr(>|z|)    
ma1   -0.7242000  0.0614426 -11.7866 < 2.2e-16 ***
drift  0.0144458  0.0044053   3.2792  0.001041 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$\Delta\, Y_t = 0.01444 a_t - 0.7242 a_{t-1}$$

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

```
z test of coefficients:

       Estimate Std. Error  z value Pr(>|z|)    
ar1    0.977761   0.040107  24.3789   <2e-16 ***
ma1   -1.765500   0.074269 -23.7716   <2e-16 ***
ma2    0.784000   0.068231  11.4904   <2e-16 ***
drift  0.014212   0.010129   1.4032   0.1606    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$\Delta\, Y_t = 0.0142 + 0.9777\Delta\, Y_{t-1} + a_t - 1.7655 a_{t-1} - 0.7840 a_{t-2}$$

3 Validación

3.1 Análisis de los coeficientes estimados

3.1.1 Significación de los coeficientes

Para el modelo 1

$$AR(1): \phi_1 = -0.6018 \rightarrow p = 0.0000 < 0.01, altamente\ significativo$$

$$AR(2): \phi_2 = -0.3173 \rightarrow p = 0.0004 < 0.01, altamente\ significativo$$

Para el modelo 2

$$MA(1): \theta_1 = -0.7242 \rightarrow p = 0.0000 < 0.01, altamente\ significativo$$

Para el modelo 3

$$AR(1): \phi_1 = 0.9777 \rightarrow p = 0.0000 < 0.01, altamente\ significativo$$

Alumno: Maye Mamani Victor Raul

$$MA(1): \theta_1 = -1.7655 \rightarrow p = 0.0000 < 0.01, es\ altamente\ significativo$$

$$MA(2): \phi_1 = 0.7840 \rightarrow p = 0.0000 < 0.01, es\ altamente\ significativo$$

3.1.2 Examen de la matriz de correlaciones entre los coeficientes

```
vcov(mod1)
                ar1           ar2         drift
ar1    8.327732e-03  3.808960e-03 -6.492482e-06
ar2    3.808960e-03  8.215831e-03 -7.907764e-06
drift -6.492482e-06 -7.907764e-06  7.256258e-05
```

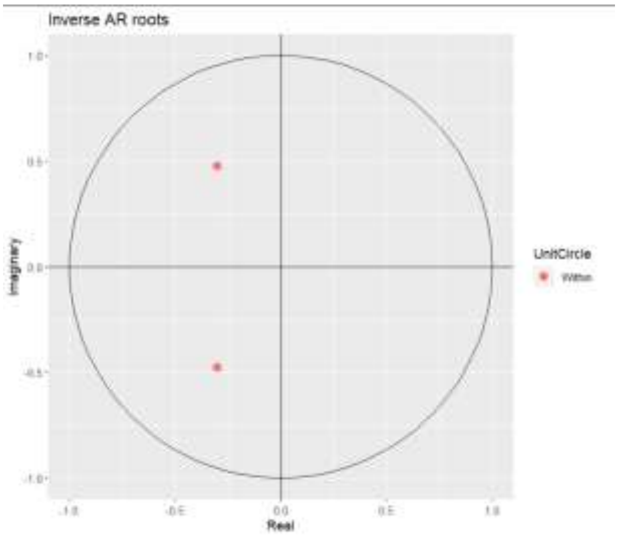
```
vcov(mod2)
                ma1         drift
ma1    3.775192e-03 -1.116834e-06
drift -1.116834e-06  1.940628e-05
```

```
vcov(mod3)
                ar1           ma1           ma2         drift
ar1    1.608558e-03 -1.610558e-03  1.134299e-03 -1.624084e-05
ma1   -1.610558e-03  5.515904e-03 -4.980805e-03  2.648729e-05
ma2    1.134299e-03 -4.980805e-03  4.655448e-03 -2.125514e-05
drift -1.624084e-05  2.648729e-05 -2.125514e-05  1.025866e-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

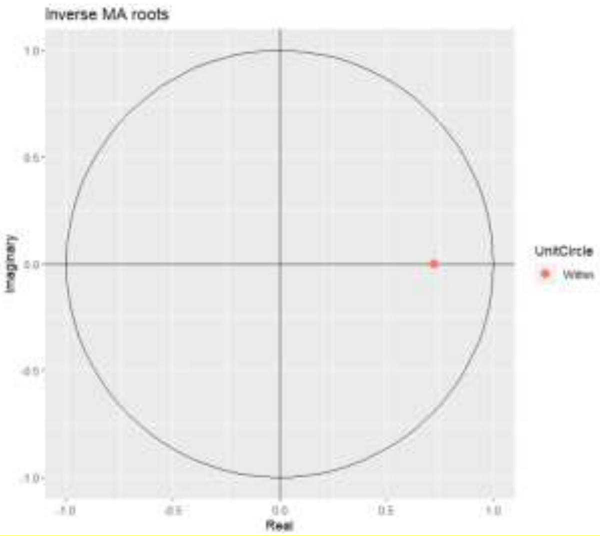
```
autoplot(mod1)
```



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.

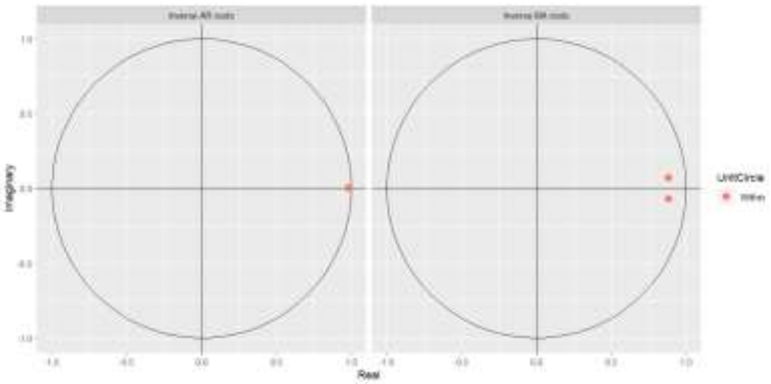
```
autoplot(mod2)
```

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 = 324.15, 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 = 340.91, 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 = 320.61, p-value < 2.2e-16
```

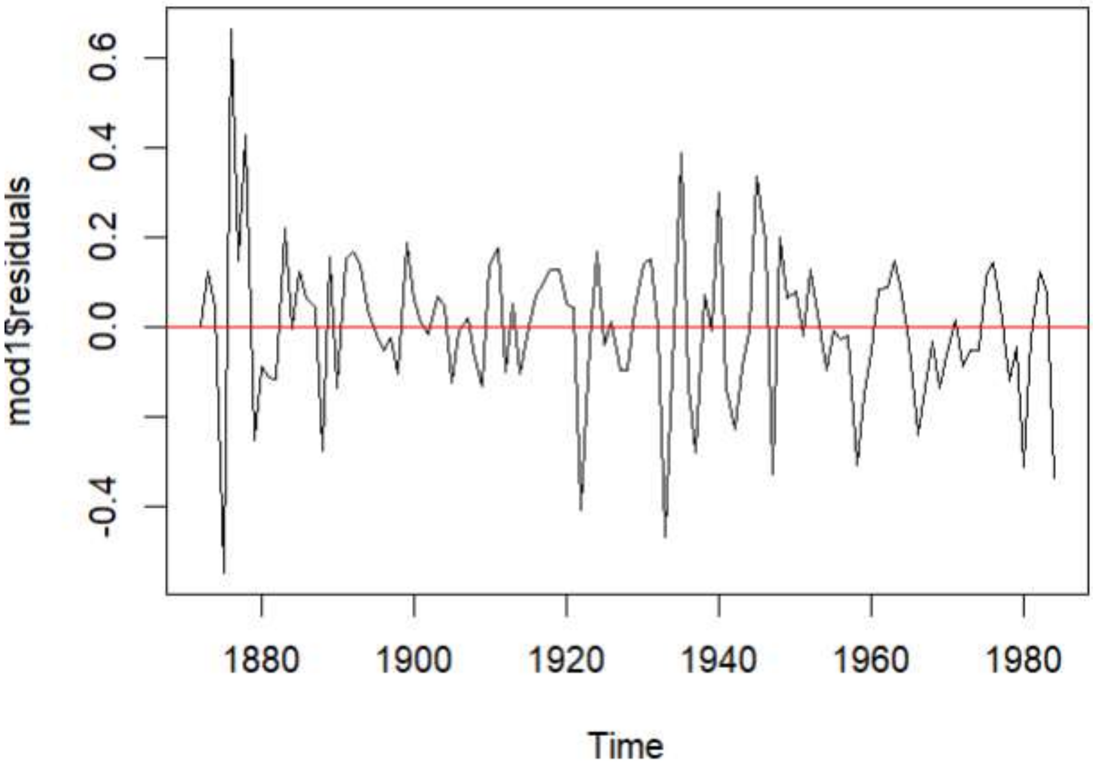
En las tres pruebas se rechaza la hipótesis nula ($p < \alpha = 0.05$), es decir, NO existe estabilidad de coeficientes

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 = 0.048446, df = 112, p-value = 0.9614
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.03127590  0.03284367
sample estimates:
   mean of x
0.0007838851
```

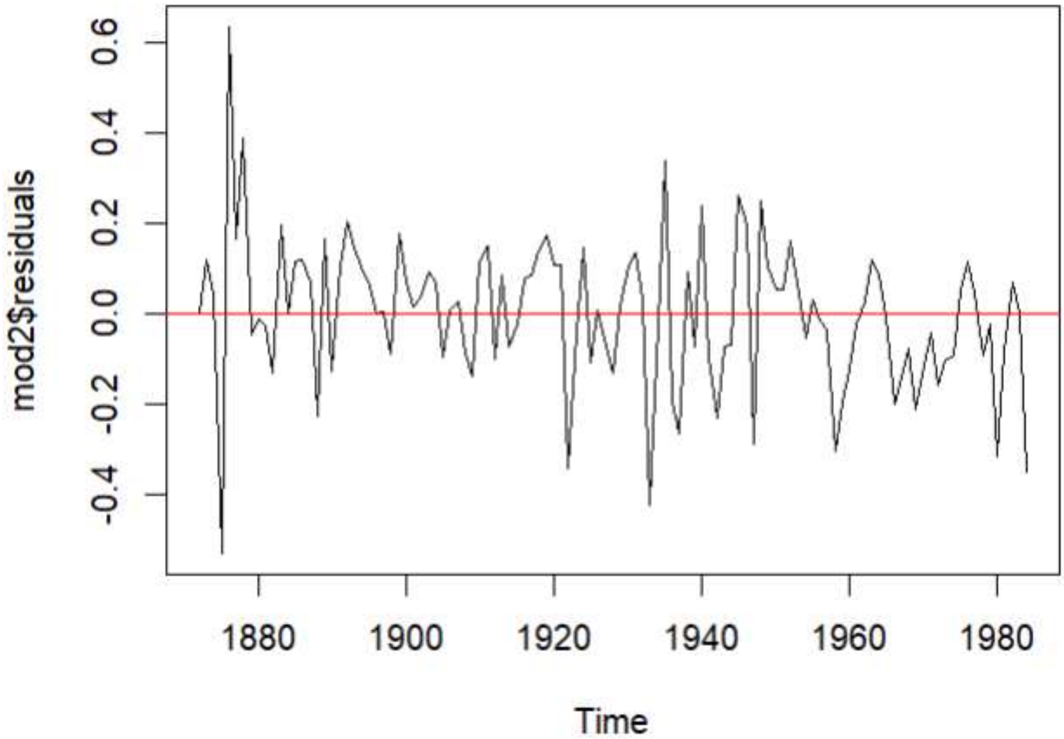
Parece indicar que un buen número de residuales están en torno a la media igual a cero. Para confirmar lo mencionado se realiza la prueba t:

Alumno: Maye Mamani Victor Raul

Como p = 0.9614 > α = 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)
```

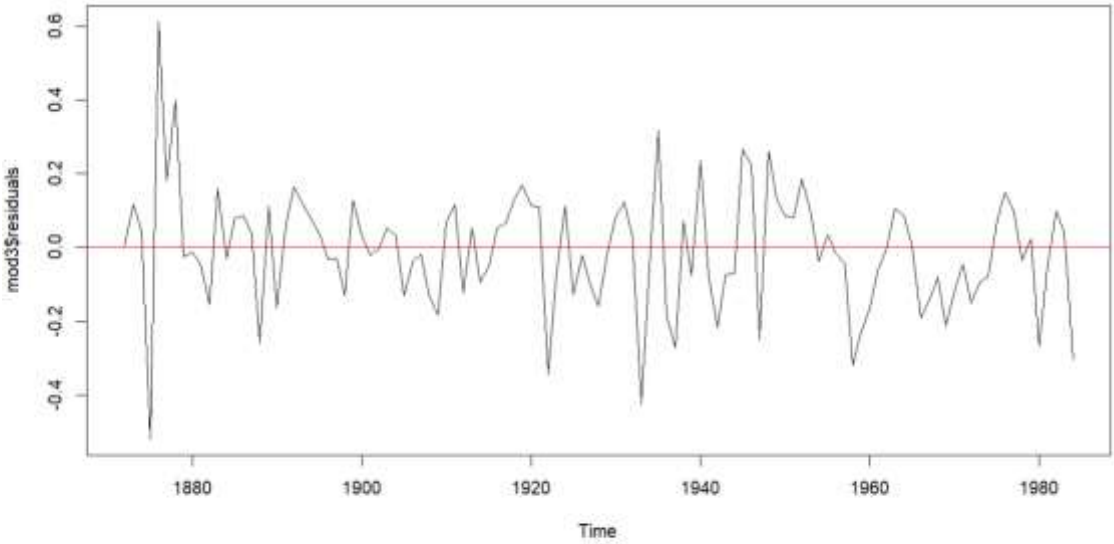



```
One Sample t-test

data:  mod2$residuals
t = 0.138, df = 112, p-value = 0.8905
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.02860809  0.03289153
sample estimates:
  mean of x
0.002141719
```

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 = 0.8905> 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)
```



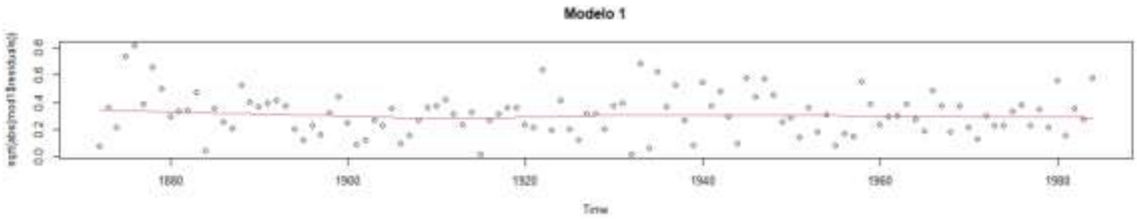

```
One Sample t-test

data:  mod3$residuals
t = -0.56544, df = 112, p-value = 0.5729
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.03888680  0.02161961
sample estimates:
   mean of x
-0.008633598
```

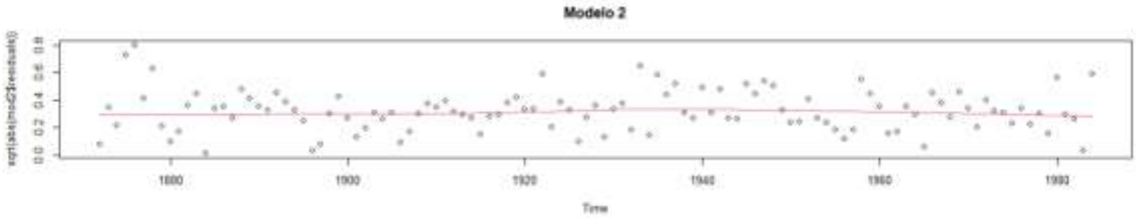
Parece indicar que un buen número de residuales están en torno a la media igual a cero. Para confirmar lo mencionado se realiza la prueba t: Como p = 0.5729 > α = 0.05, se acepta Ho, es decir la media es igual a cero.

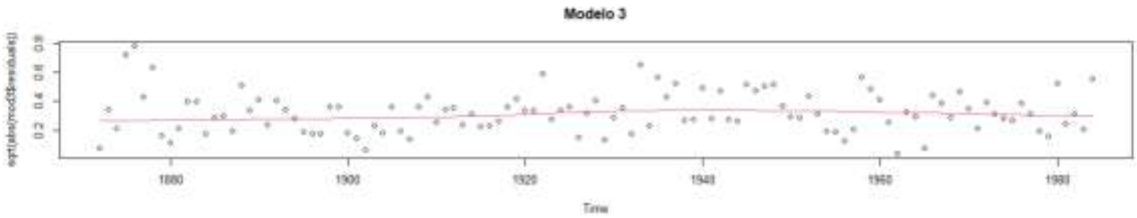
3.2.2 Homocedasticidad o varianza constante

```
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 presentar 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 = 5.5028, df = 1, p-value = 0.01899
```

```
obs=get(mod2$series)
bptest(resid(mod2)~I(obs-resid(mod2)))
studentized Breusch-Pagan test

data:  resid(mod2) ~ I(obs - resid(mod2))
BP = 4.1873, df = 1, p-value = 0.04073
```

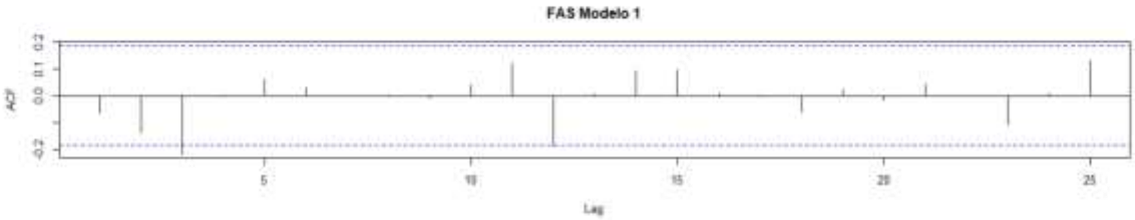
```
obs=get(mod3$series)
bptest(resid(mod3)~I(obs-resid(mod3)))
studentized Breusch-Pagan test

data:  resid(mod3) ~ I(obs - resid(mod3))
BP = 4.3006, df = 1, p-value = 0.0381
```

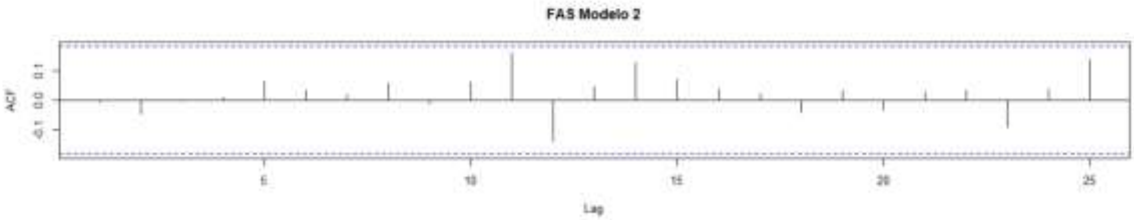
El valor de probabilidad (p-valor) asociado al estadístico BP asume los valores de probabilidad de 0.01899, 0.04073 y 0.0381 para los modelos 1,2 y 3 respectivamente, que son menores a α=0.05, por lo cual podemos afirmar que los residuales de estos modelos NO son constantes.

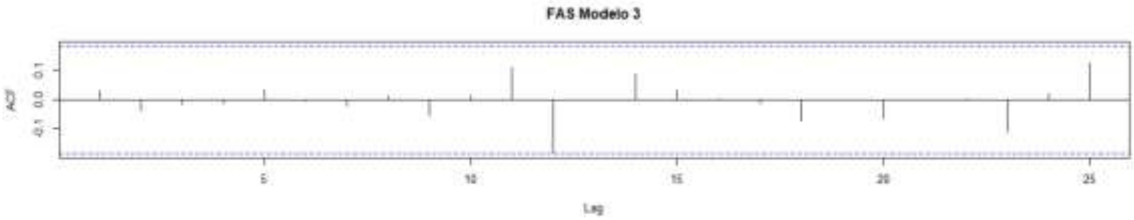
Alumno: Maye Mamani Victor Raul

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")
```







Se observa que casi la totalidad de los coeficientes del FAS para los 3 modelos se encuentran dentro de las bandas de no significación. Por tanto, tenemos altos indicios de que los residuos de los modelos 1 ,2 y 3 sean ruido blanco

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

data:  resid_m1
X-squared = 0.48086, df = 1, p-value = 0.488
```

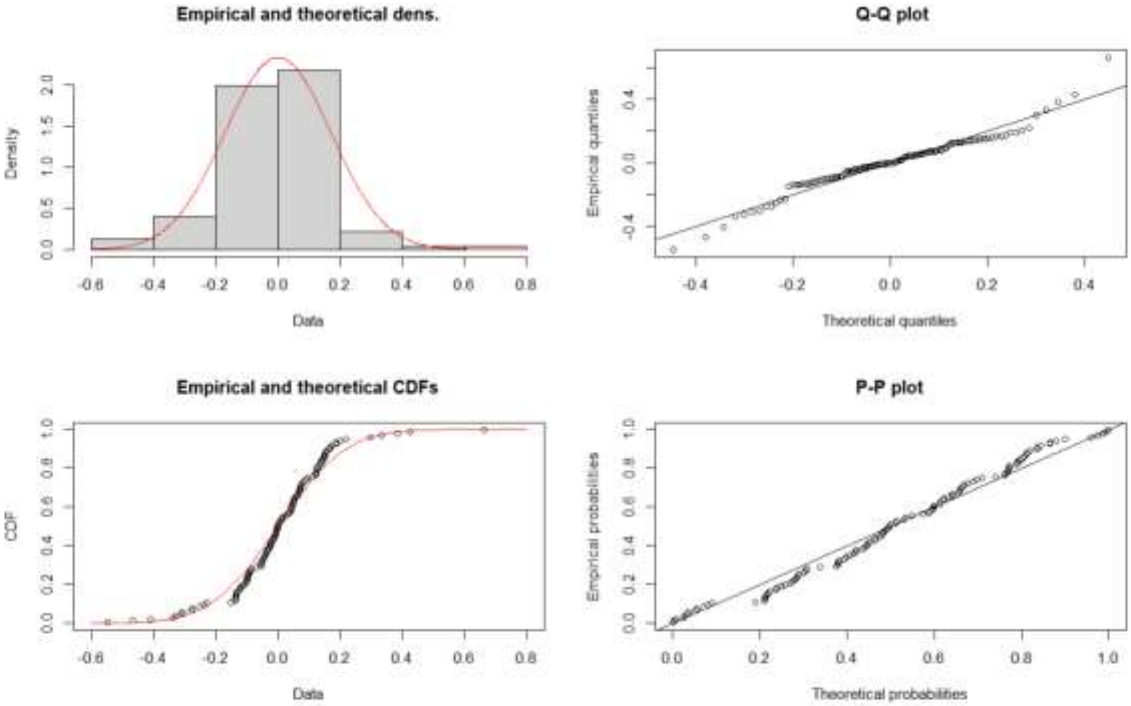
```
Box.test(resid_m2,type = "Ljung-Box")
Box-Ljung test

data:  resid_m2
X-squared = 0.0028218, df = 1, p-value = 0.9576
```

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

data:  resid_m3
X-squared = 0.1021, df = 1, p-value = 0.7493
```

Teniendo los valores 0.488, 0.9576 y 0.7493 en los MODELOS 1, 2 Y 3 respectivamente, los cuales 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
```

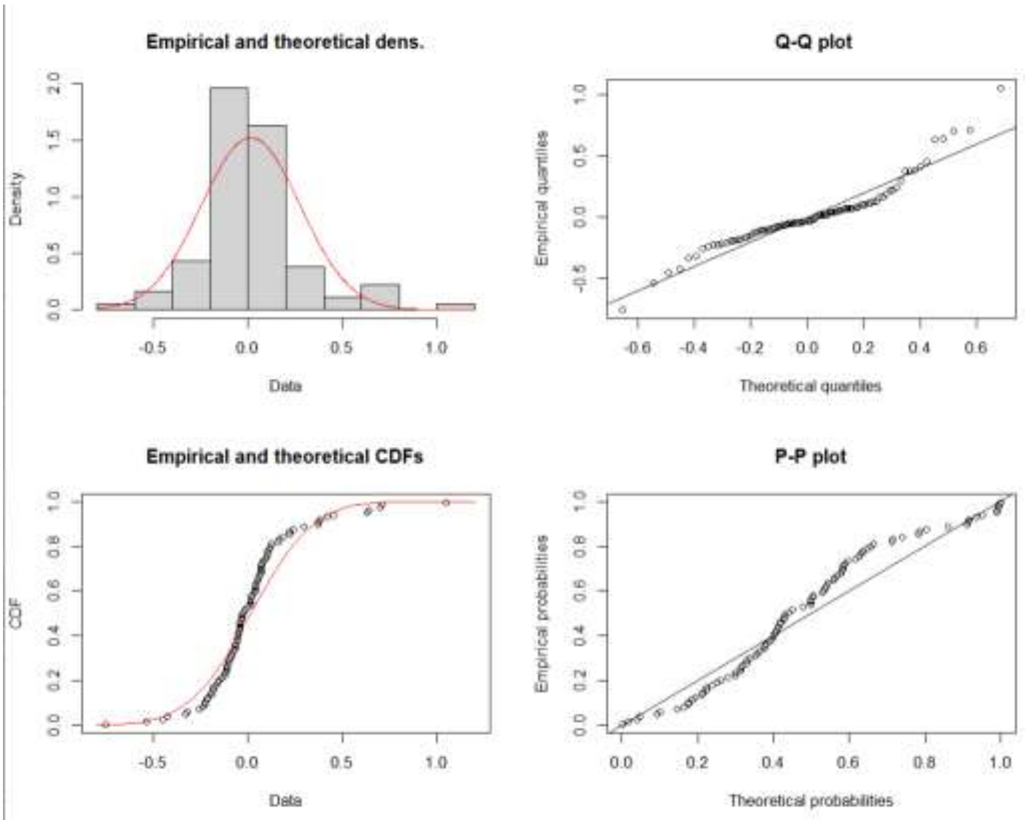



```
Jarque Bera Test

data:  resid_m1
X-squared = 24.485, df = 2, p-value = 4.821e-06
```

En las figuras se observa que los residuales del modelo 1 presentan indicios de ajuste hacia la distribución normal. Pero 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

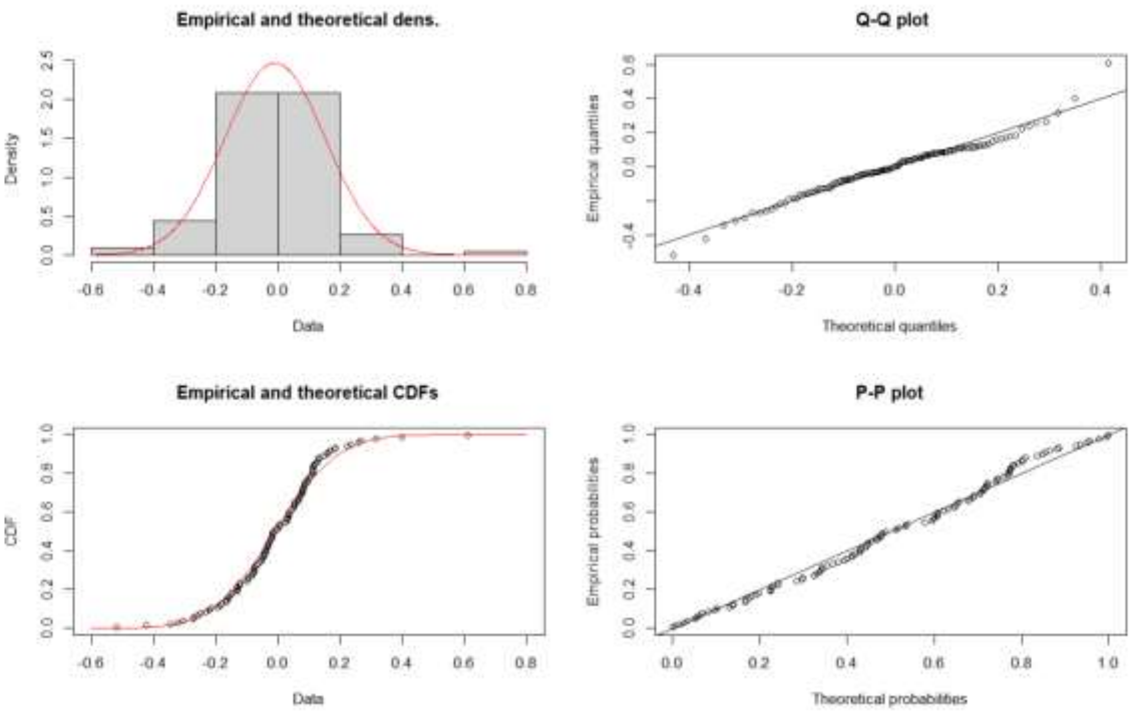



```
Jarque Bera Test

data:  resid_m2
X-squared = 18.485, df = 2, p-value = 9.685e-05
```

En las figuras se observa que los residuales del modelo 2 presentan indicios de ajuste hacia la distribución normal. Pero enn 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 = 15.299, df = 2, p-value = 0.0004763
```

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 = 0.00047 < 0.05, se acepta Ho, es decir, los residuos NO se aproximan a una distribución normal.

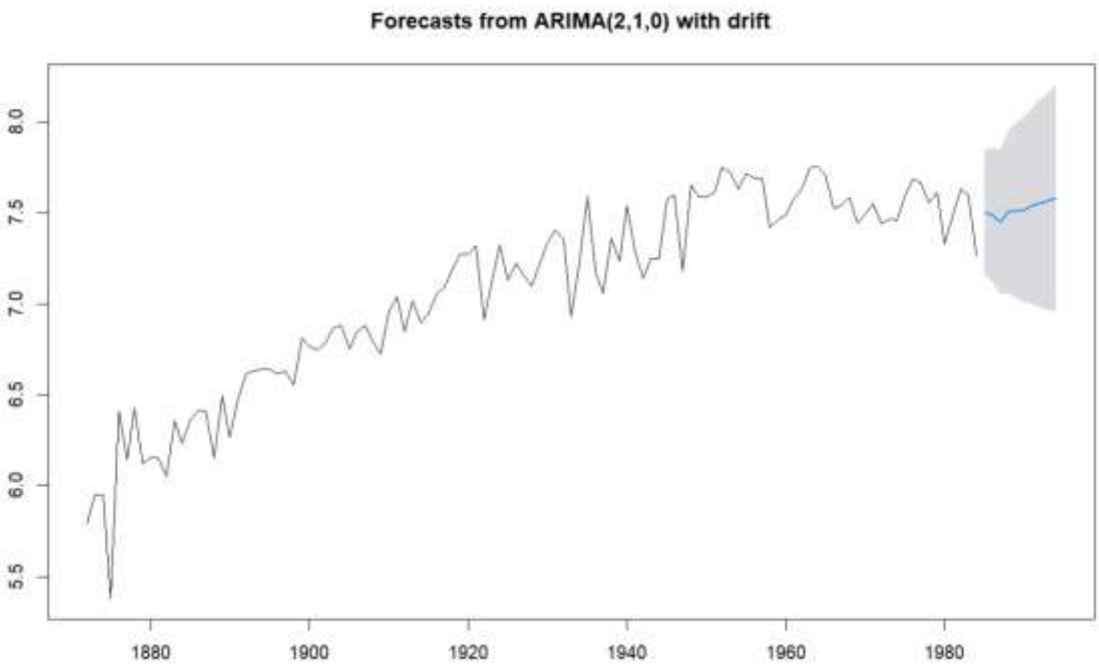
4 Pronostico

4.1Pronosticos de cada modelo

Modelo 1: ARIMA(2,1,0)

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

Alumno: Maye Mamani Victor Raul



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

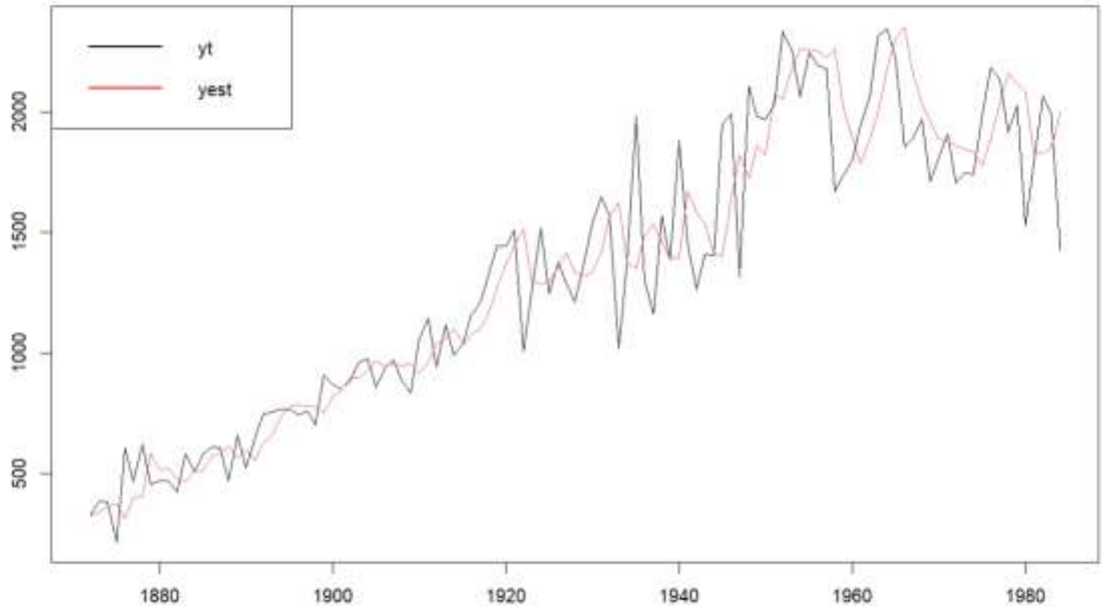
Los datos proyectados para los siguientes 10 años son:

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

Model Information:
Series: data_tf
ARIMA(2,1,0) with drift

Coefficients:
          ar1      ar2   drift
      -0.6019  -0.3174  0.0141
s.e.   0.0913   0.0906  0.0085

sigma^2 = 0.0304:  log likelihood = 38
AIC=-68.01   AICc=-67.63   BIC=-57.13

Error measures:
                       ME      RMSE       MAE         MPE    MAPE
MASE        ACF1
Training set 0.0007838851 0.1712412 0.1233828 -0.01076704 1.78273 0.88
24134 -0.06437716

Forecasts:
     Point Forecast    Lo 95    Hi 95
1985       7.503311 7.161581 7.845040
1986       7.492552 7.124738 7.860366
1987       7.450404 7.052655 7.848153
1988       7.506286 7.057696 7.954876
1989       7.513126 7.033818 7.992435
1990       7.518371 7.008363 8.028380
1991       7.540142 6.998211 8.082072
1992       7.552472 6.982641 8.122303
1993       7.565239 6.968384 8.162094
1994       7.580740 6.957633 8.203846
```

Alumno: Maye Mamani Victor Raul

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

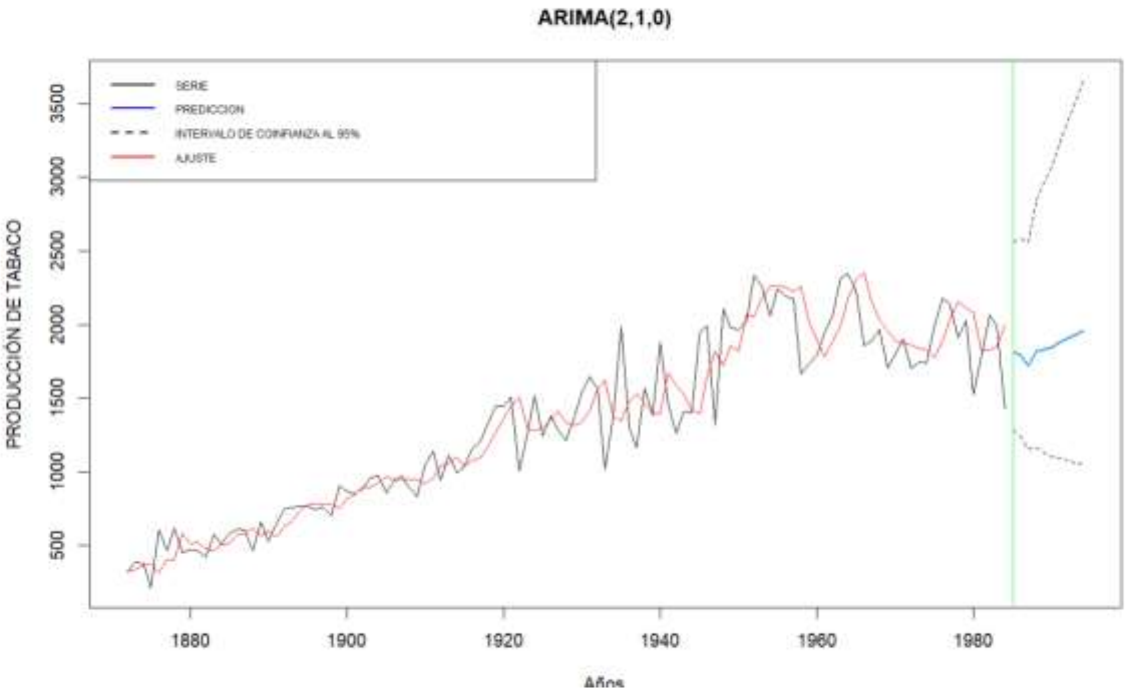
```
Pron1$mean <- exp(Pron1$mean)
Pron1$lower <- exp(Pron1$lower)
Pron1$upper <- exp(Pron1$upper)
Pron1$x <- exp(Pron1$x)
Pron1$fitted <- exp(Pron1$fitted)
Pron1$residuals <- exp(Pron1$residuals)
summary(Pron1)
```

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

Model Information:
Series: data_tf
ARIMA(2,1,0) with drift

Coefficients:
          ar1      ar2   drift
      -0.6019  -0.3174  0.0141
s.e.   0.0913   0.0906  0.0085

sigma^2 = 0.0304:  log likelihood = 38
AIC=-68.01   AICc=-67.63   BIC=-57.13

Error measures:
                    ME     RMSE      MAE       MPE     MAPE      MASE
ACF1
Training set -4.832891 216.5697 153.8088 -1.405363 12.55187 0.9204205
0.0930709
```

```
Forecasts:
     Point Forecast    Lo 95    Hi 95
1985       1814.038 1288.948 2553.040
1986       1794.625 1242.322 2592.468
1987       1720.558 1155.923 2561.001
1988       1819.443 1161.765 2849.434
1989       1831.932 1134.353 2958.491
1990       1841.566 1105.843 3066.768
1991       1882.097 1094.673 3235.933
1992       1905.447 1077.761 3368.770
1993       1929.931 1062.505 3505.521
1994       1960.078 1051.143 3654.981
```

GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

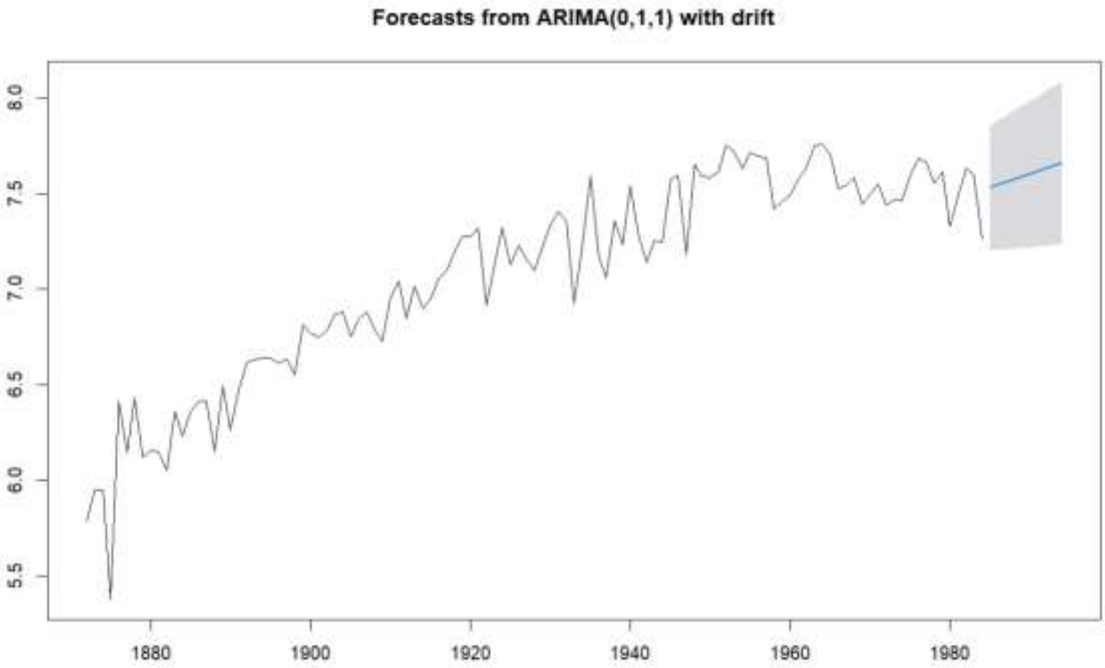
```
plot(Pron1, shaded = FALSE, xlab = "Años", ylab = "PRODUCCIÓN DE
TABACO",main = "ARIMA(2,1,0)")
lines(Pron1$fitted, col = "red")
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE
COINFIANZA AL 95%", "AJUSTE"),col=c("black", "blue", "black", "red"),
lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=1985, lwd = 1, col="green")
```



Modelo 2: ARIMA(0,1,1)

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

Alumno: Maye Mamani Victor Raul



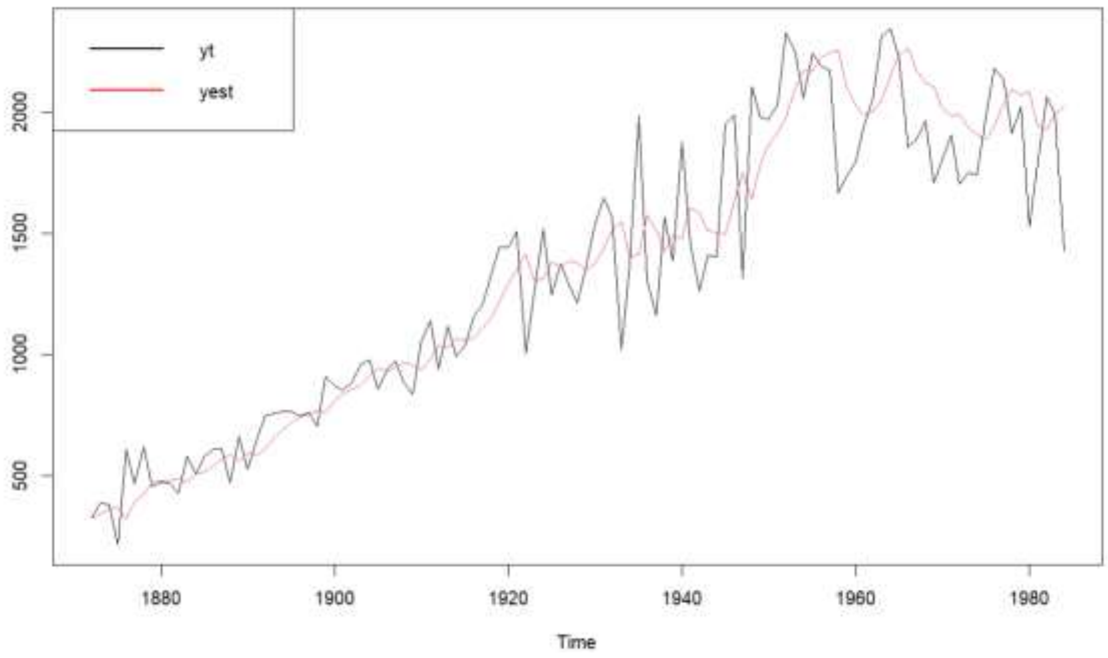

```
Forecast method: ARIMA(0,1,1) with drift

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

Coefficients:
          ma1   drift
      -0.7242  0.0144
s.e.   0.0614  0.0044

sigma^2 = 0.02772:  log likelihood = 42.52
AIC=-79.04   AICc=-78.82   BIC=-70.88

Error measures:
                      ME      RMSE       MAE      MPE     MAPE       MA
SE         ACF1
Training set 0.002141719 0.1642564 0.1229315 0.024795 1.767541 0.87918
64 -0.004931528

Forecasts:
     Point Forecast    Lo 95    Hi 95
1985       7.530439 7.204142 7.856736
1986       7.544885 7.206405 7.883365
1987       7.559331 7.209092 7.909570
1988       7.573777 7.212161 7.935392
1989       7.588222 7.215577 7.960868
1990       7.602668 7.219310 7.986026
1991       7.617114 7.223335 8.010893
1992       7.631560 7.227628 8.035492
1993       7.646006 7.232171 8.059841
1994       7.660452 7.236945 8.083959
```

PRONÓSTICO DE LA SERIE ORIGINAL Y VALORES ESPERADOS

Alumno: Maye Mamani Victor Raul

Como en el anterior modelo para obtener los verdaderos valores de pronóstico, se tendrán que transformar aplicando la función exp().

```
data_modelo2 <- exp(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"))
```



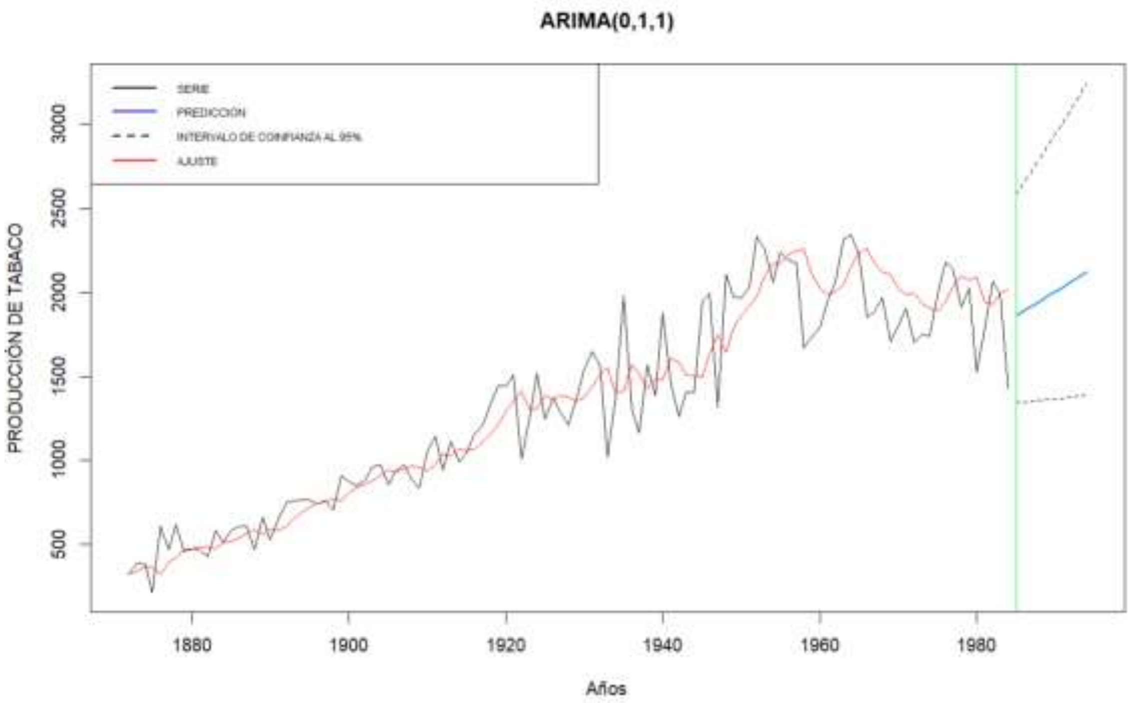

```
Pron2$mean <- exp(Pron2$mean)
Pron2$lower <- exp(Pron2$lower)
Pron2$upper <- exp(Pron2$upper)
Pron2$x <- exp(Pron2$x)
Pron2$fitted <- exp(Pron2$fitted)
Pron2$residuals <- exp(Pron2$residuals)
summary(Pron2)
```

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

Coefficients:
          ma1   drift
      -0.7242  0.0144
s.e.   0.0614  0.0044

sigma^2 = 0.02772:  log likelihood = 42.52
AIC=-79.04   AICc=-78.82   BIC=-70.88

Error measures:
                    ME     RMSE      MAE       MPE    MAPE      MASE
ACF1
Training set -10.17216 212.6267 156.7662 -1.149227 12.4767 0.9381179 0
.1640972
```

```
Forecasts:
     Point Forecast    Lo 95    Hi 95
1985       1863.924 1344.990 2583.076
1986       1891.045 1348.037 2652.783
1987       1918.561 1351.664 2723.218
1988       1946.477 1355.819 2794.455
1989       1974.800 1360.458 2866.560
1990       2003.535 1365.547 2939.592
1991       2032.687 1371.054 3013.608
1992       2062.264 1376.953 3088.657
1993       2092.272 1383.222 3164.786
1994       2122.716 1389.841 3242.042
```

GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

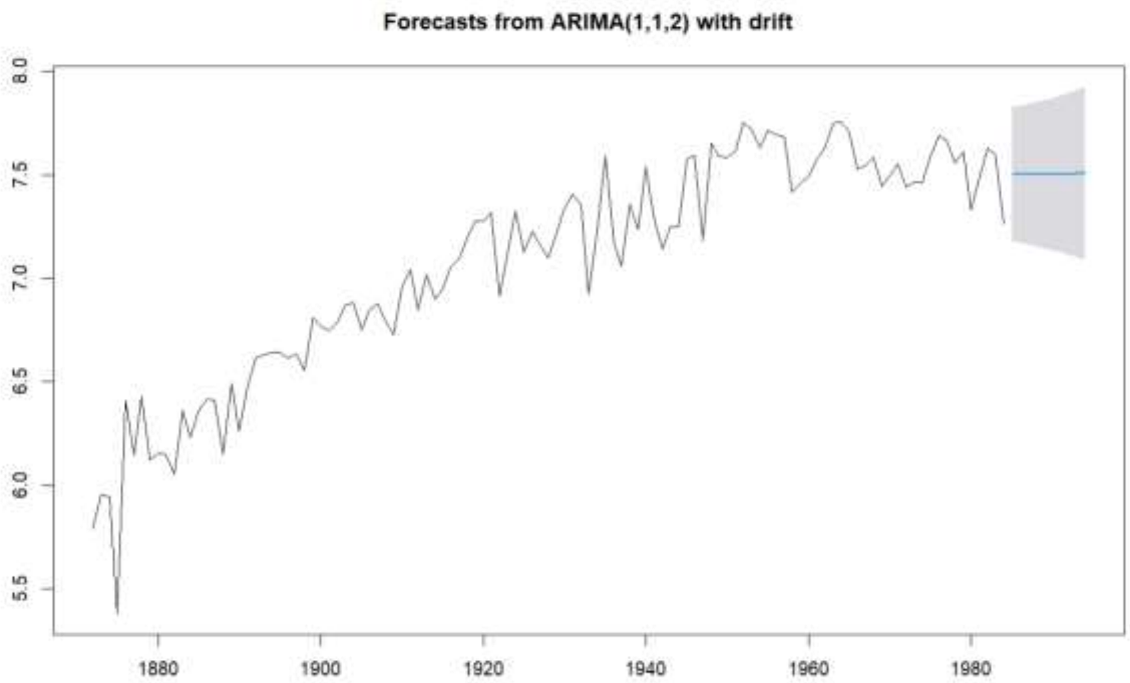
```
plot(Pron2, shaded = FALSE, xlab = "Años", ylab = "PRODUCCIÓN DE
TABACO",main = "ARIMA(0,1,1)")
lines(Pron2$fitted, col = "red")
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE
COINFIANZA AL 95%", "AJUSTE"),col=c("black", "blue", "black", "red"),
lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=2020, lwd = 1, col="green")
```



Modelo 3: ARIMA(1,1,2)

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

Alumno: Maye Mamani Victor Raul



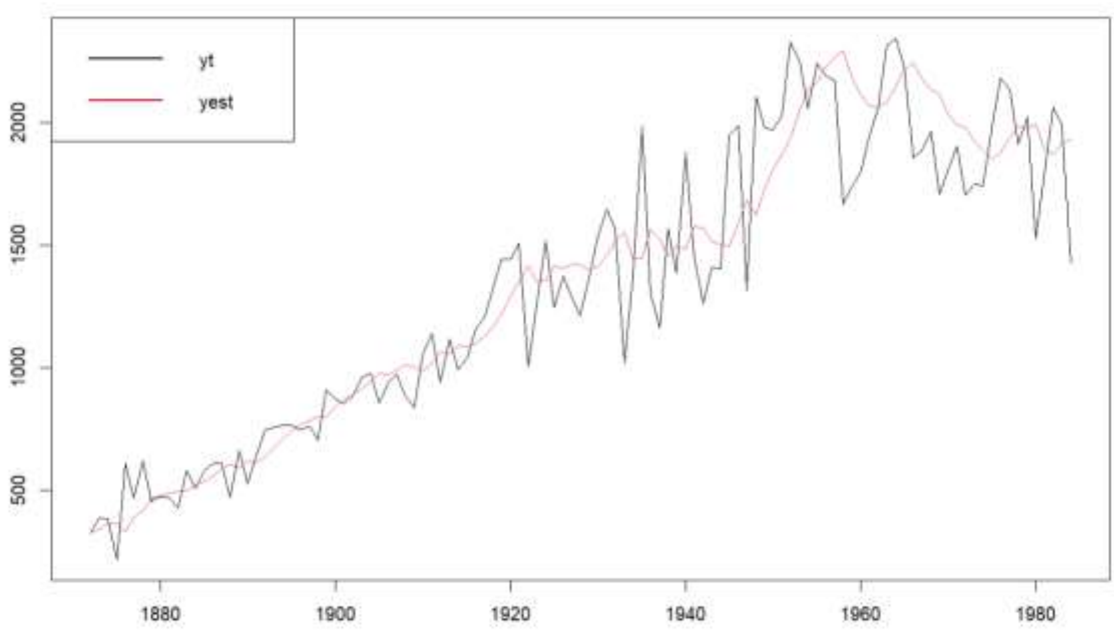

```
Forecast method: ARIMA(1,1,2) with drift

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

Coefficients:
         ar1      ma1     ma2   drift
      0.9778  -1.7655  0.7840  0.0142
s.e.  0.0401   0.0743  0.0682  0.0101

sigma^2 = 0.0274:  log likelihood = 43.87
AIC=-77.74   AICc=-77.17   BIC=-64.14

Error measures:
                       ME      RMSE       MAE        MPE     MAPE
MASE      ACF1
Training set -0.008633598 0.1618205 0.1223072 -0.1396587 1.755203 0.87
47209 0.0296646

Forecasts:
     Point Forecast    Lo 95    Hi 95
1985       7.505582 7.181161 7.830002
1986       7.504563 7.172914 7.836211
1987       7.503883 7.164223 7.843542
1988       7.503534 7.155099 7.851968
1989       7.503509 7.145560 7.861457
1990       7.503800 7.135629 7.871971
1991       7.504401 7.125332 7.883471
1992       7.505305 7.114697 7.895913
1993       7.506505 7.103756 7.909254
1994       7.507994 7.092538 7.923450
```

PRONÓSTICO DE SERIE ORIGINAL Y VALORES ESTIMADOS

Como en el anterior modelo para obtener los verdaderos valores de pronóstico, se tendrán que transformar aplicando la función exp().

Alumno: Maye Mamani Victor Raul

```
data_modelo3 <- exp(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"))
```



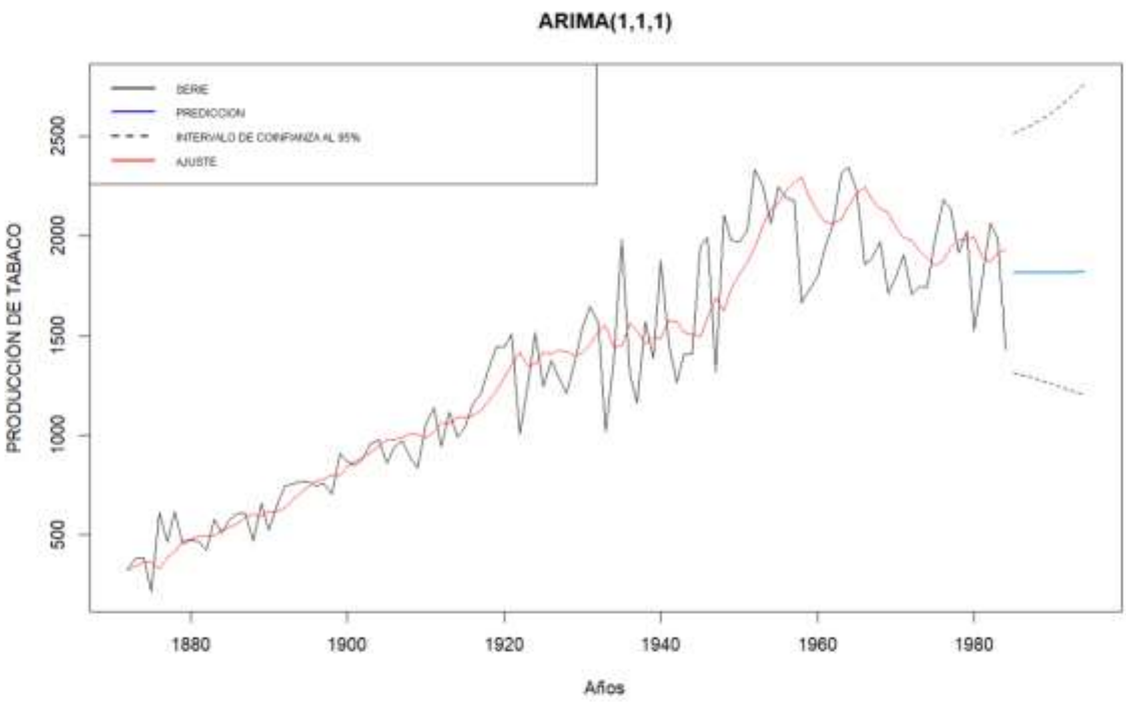

```
Pron3$mean <- exp(Pron3$mean)
Pron3$lower <- exp(Pron3$lower)
Pron3$upper <- exp(Pron3$upper)
Pron3$x <- exp(Pron3$x)
Pron3$fitted <- exp(Pron3$fitted)
Pron3$residuals <- exp(Pron3$residuals)
summary(Pron3)
```

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

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

Coefficients:
         ar1      ma1     ma2   drift
      0.9778  -1.7655  0.7840  0.0142
s.e.  0.0401   0.0743  0.0682  0.0101

sigma^2 = 0.0274:  log likelihood = 43.87
AIC=-77.74   AICc=-77.17   BIC=-64.14

Error measures:
                    ME     RMSE      MAE       MPE     MAPE      MASE
ACF1
Training set -16.43305 213.2257 158.9872 -2.192223 12.49731 0.9514089
0.2194535

Forecasts:
     Point Forecast    Lo 95    Hi 95
1985       1818.162 1314.433 2514.936
```

```
1986         1816.311 1303.638 2530.599
1987         1815.076 1292.357 2549.218
1988         1814.443 1280.620 2570.789
1989         1814.397 1268.462 2595.298
1990         1814.926 1255.927 2622.730
1991         1816.017 1243.060 2653.064
1992         1817.659 1229.911 2686.280
1993         1819.842 1216.527 2722.358
1994         1822.554 1202.957 2761.280
```

GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

```
plot(Pron3, shaded = FALSE, xlab = "Años", ylab = "PRODUCCIÓN DE
TABACO",main = "ARIMA(1,1,1)")
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=2020, lwd = 1, col="green")
```



Métricas basadas en el error

```
accuracy(Pron1)
```

```
                    ME     RMSE      MAE       MPE     MAPE      MASE
ACF1
Training set -4.832891 216.5697 153.8088 -1.405363 12.55187 0.9204205
0.0930709
```

```
accuracy(Pron2)
```

```
                    ME     RMSE      MAE       MPE    MAPE      MASE
ACF1
Training set -10.17216 212.6267 156.7662 -1.149227 12.4767 0.9381179 0
.1640972
```

Alumno: Maye Mamani Victor Raul

```
accuracy(Pron3)
                    ME      RMSE       MAE       MPE     MAPE      MASE
ACF1
Training set -16.43305 213.2257 158.9872 -2.192223 12.49731 0.9514089
0.2194535
```

Basado en las métricas MAE y RMSE, el modelo ARIMA(0,1,1) es el mejor entre los tres, ya que tiene los mejores valores (MAE = 156.76 Y RMSE=212.62). Esto indica que sus predicciones son más precisas y con menos variabilidad en los errores grandes.

Conclusión

Se evaluaron tres modelos ARIMA para el conjunto de datos:

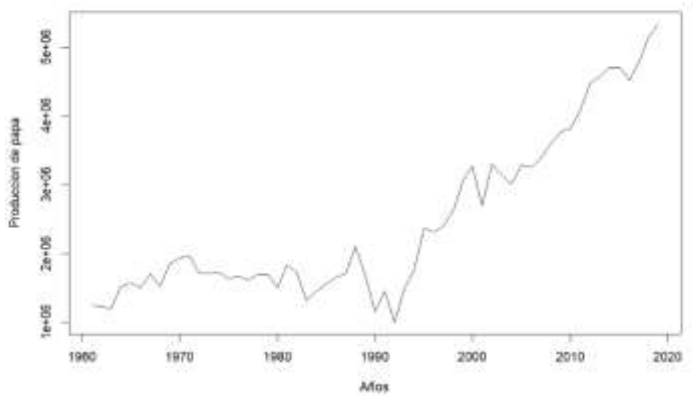
1. **ARIMA(2,1,0)**
    - Los coeficientes estimados son significativos para ar1, ar2 y drift.
    - Los residuos parecen seguir un comportamiento de ruido blanco según las pruebas de autocorrelación y heterocedasticidad.
    - Sin embargo, el p-valor del estadístico BP sugiere que los residuos no son constantes.
2. **ARIMA(0,1,1)**
    - El coeficiente ma1 es significativo.
    - Los residuos también muestran comportamiento de ruido blanco según las pruebas realizadas.
    - El p-valor del estadístico BP indica que los residuos no son constantes.
3. **ARIMA(1,1,2)**
    - Los coeficientes ar1, ma1 y ma2 son significativos.
    - Los residuos muestran comportamiento de ruido blanco, confirmado por las pruebas de autocorrelación y heterocedasticidad.
    - El p-valor del estadístico BP indica que los residuos no son constantes.

El modelo seleccionado es** *ARIMA(1,1,2)*. **Aunque todos los modelos mostraron buenos indicadores de ajuste y comportamiento de los residuos como ruido blanco, ARIMA(1,1,2) presenta coeficientes significativos tanto en los términos autorregresivos como en los de medias móviles. Además, sus métricas de precisión (RMSE, MAE) son comparables con los otros modelos evaluados. Este modelo proporciona un equilibrio adecuado entre complejidad y ajuste, capturando mejor las características de la serie temporal analizada.

Alumno: Maye Mamani Victor Raul

CASO 2: PRODUCCION DE PAPA

```
data <- read_excel("F:\\777--Programacion
repos\\Una\\r\\data\\actividad-06.xlsx",sheet = "02")
View(data)
# Gráfica de la serie
data_ts <- ts(data$Producción, start = c(1961,1), frequency = 1)
plot(data_ts, xlab="Años", ylab="Produccion de papa")
```



En primera instancia podemos visualizar una clara TENDENCIA creciente, la varianza no parece variar mucho, y no es estacionaria en media. Y al ser una serie anual NO haremos un análisis de estacionalidad

1 Identificación

1. 1 Análisis de la tendencia y la estacionalidad

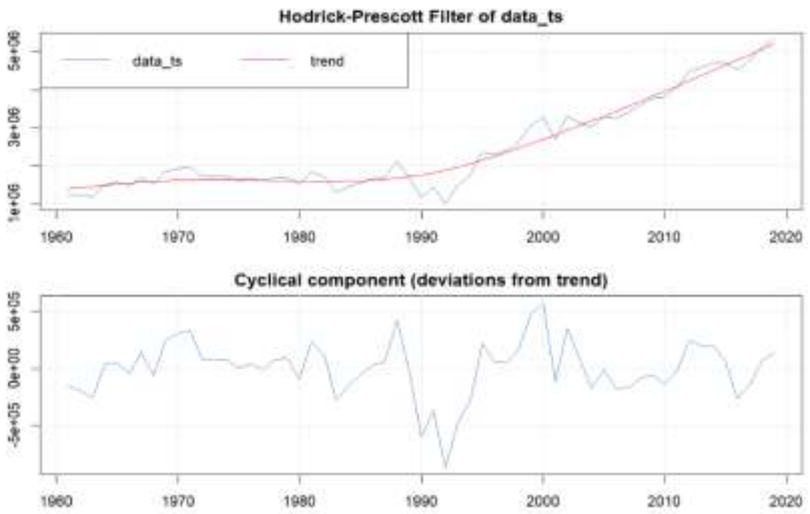
1.1.1 Estacionalidad

Debido a que se trata de un gráfico con un período anual, no se observa estacionalidad.

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)
```

Alumno: Maye Mamani Victor Raul

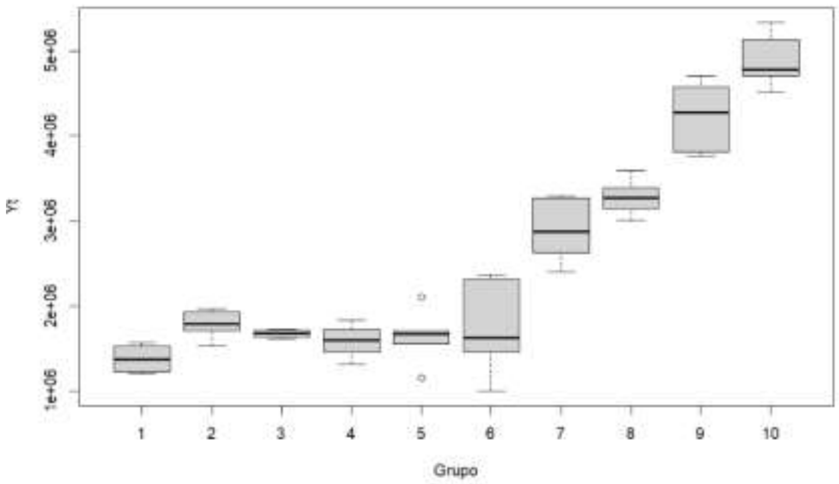


Hay una tendencia creciente y no parece haber algún patrón que se repita cíclicamente

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")
```



Se aprecia que la varianza no es constante, por lo que afirmamos que la serie no es estacionaria en varianza.

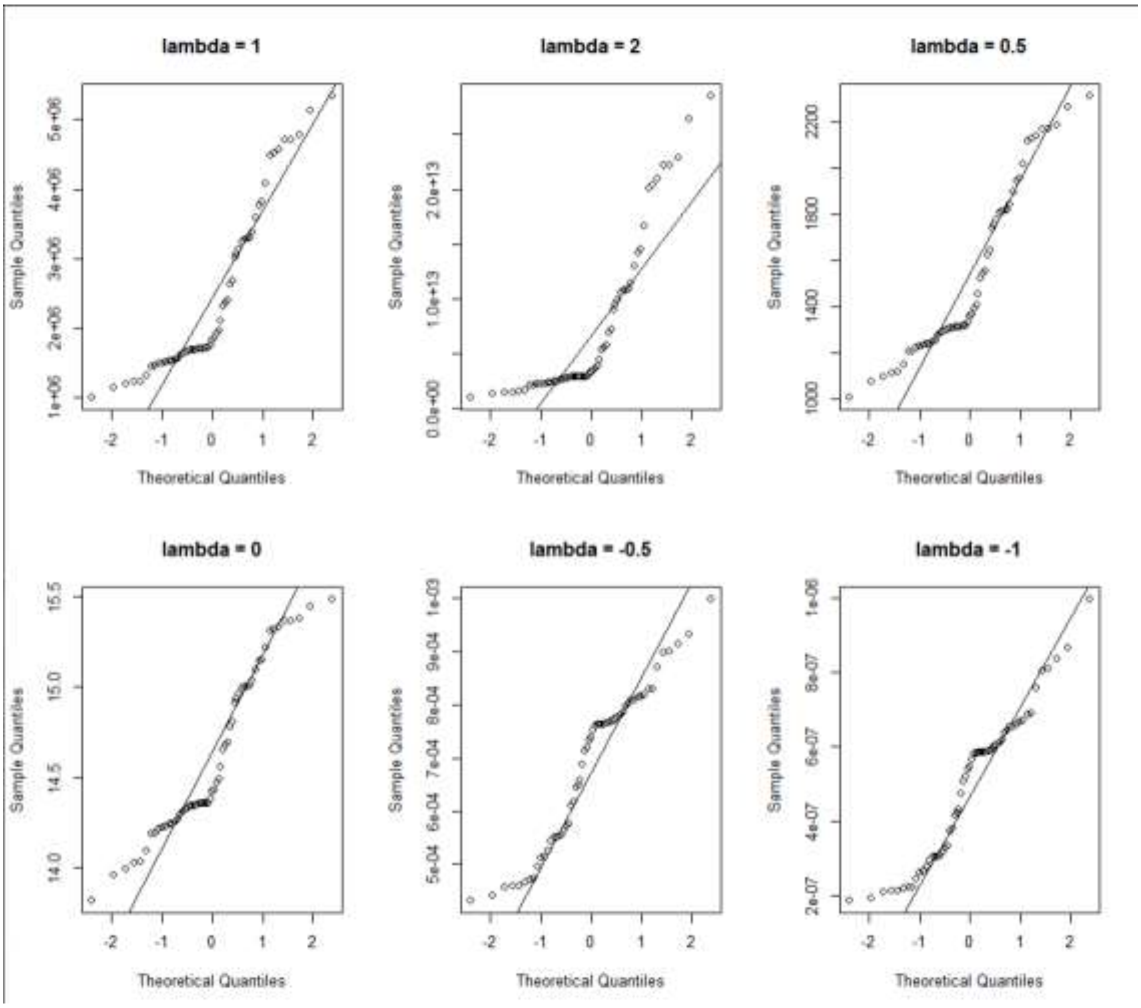
Transformación de la serie para que sea estacionaria

```
par(mfrow = c(2,3))
#lambda  = 1 <- yt
qqnorm(data_ts,main="lambda = 1")
qqline(data_ts)


#lambda =2 <- yt^2
```

Alumno: Maye Mamani Victor Raul

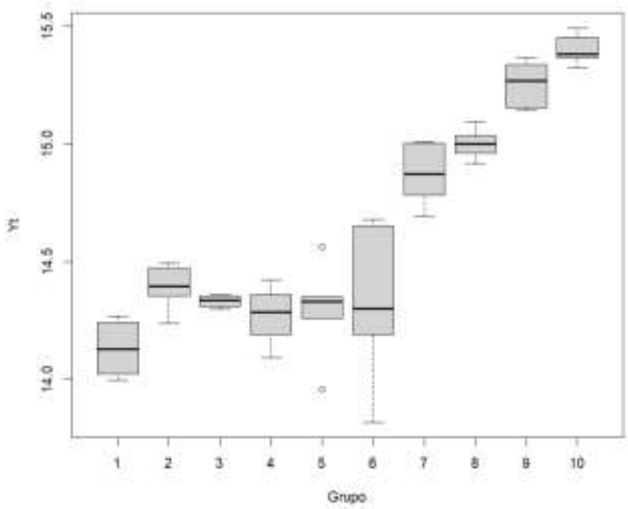
```r
t1.yt <- data_ts^2
qqnorm(t1.yt, main="lambda = 2")
qqline(t1.yt)

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

#lambda  =  0 <- log(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)

#lambda  =  -1 <- 1/Yt

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


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


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

Alumno: Maye Mamani Victor Raul



Elegimos lambda 0

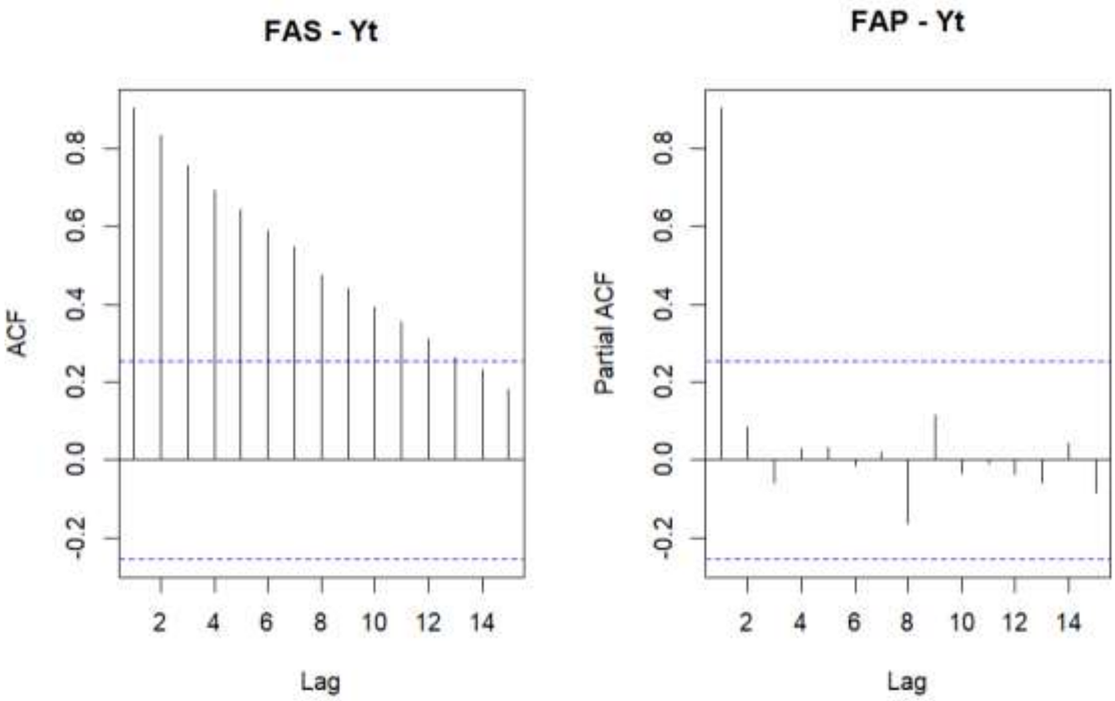
```
data_tf <- log(data_ts)
boxplot(data_mod ~ Grupo, xlab = "Grupo", ylab="log(Yt)")
plot(data_tf)
```



Vemos que hay una leve mejora al transformar la serie, por lo que tomamos la decisión de transformar la serie.

1.2.2 Estacionariedad en media

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



El FAS decrece lentamente y el primer FAP es significativo siendo 0.917, por lo que podríamos decir que la serie NO ES ESTACIONARIA

Alumno: Maye Mamani Victor Raul

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

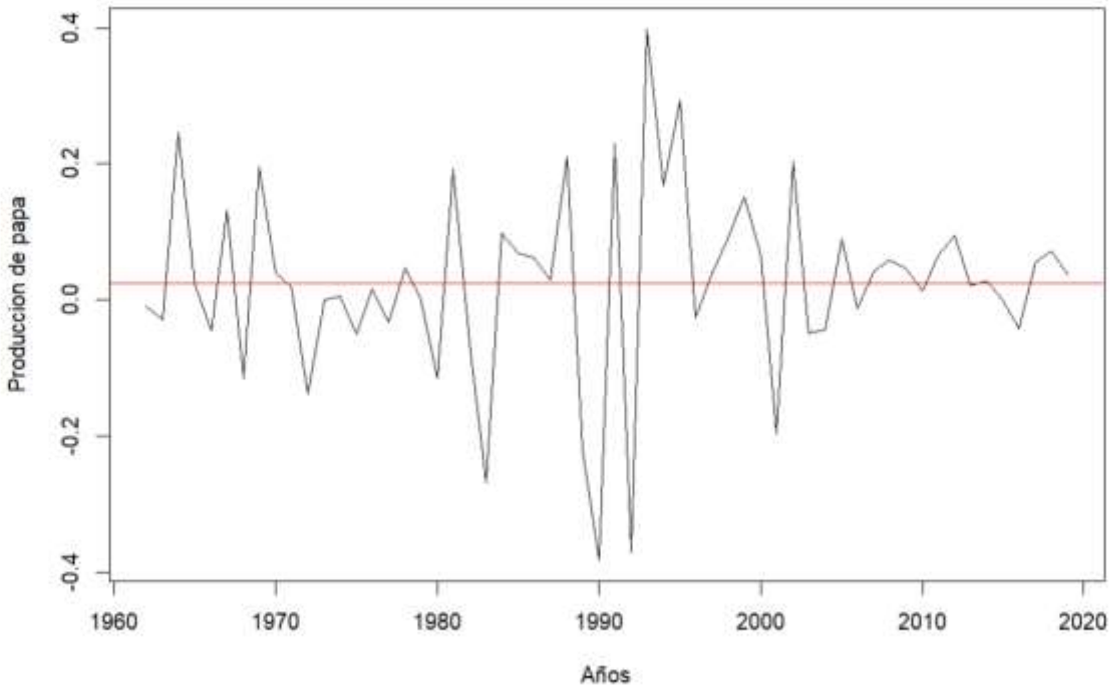
```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.48953 -0.03199  0.01448  0.05906  0.24742 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)  
(Intercept)  2.009477   1.163248   1.727   0.0899 .
z.lag.1     -0.143329   0.083471  -1.717   0.0918 .
tt           0.003831   0.002161   1.773   0.0820 .
z.diff.lag  -0.192866   0.135898  -1.419   0.1617  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1369 on 53 degrees of freedom
Multiple R-squared:  0.1234,	Adjusted R-squared:  0.07373 
F-statistic: 2.486 on 3 and 53 DF,  p-value: 0.07056


Value of test-statistic is: -1.7171 2.1159 1.6428 

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 (-1.7121) 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

```r
data_diff = diff(data_tf)
plot(data_diff, xlab="Años", ylab="Produccion de papa")
abline(h = mean(data_diff), col = "red")
```

Alumno: Maye Mamani Victor Raul



Y volvemos a aplicar el test de Dickey Fuller

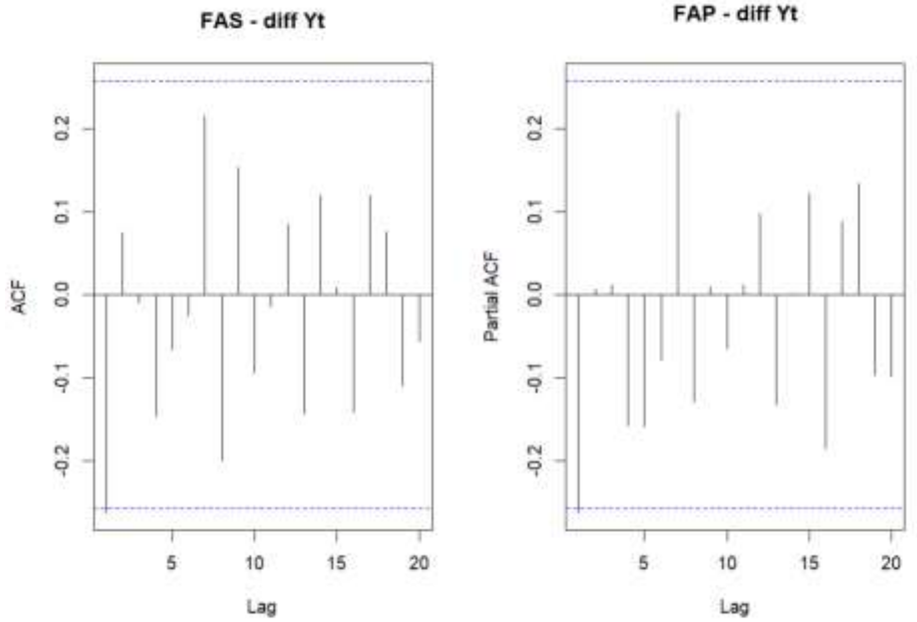
```r
data_adf <- ur.df(data_diff, type="drift", lags = 1)
summary(data_adf)
```

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-0.47435 -0.04849  0.00947  0.06474  0.30152

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.033224   0.019578   1.697   0.0956 .
z.lag.1     -1.256728   0.218038  -5.764 4.29e-07 ***
z.diff.lag  -0.006175   0.137193  -0.045   0.9643
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1408 on 53 degrees of freedom
Multiple R-squared:  0.6329,	Adjusted R-squared:  0.6191
F-statistic: 45.69 on 2 and 53 DF,  p-value: 2.926e-12


Value of test-statistic is: -5.7638 16.617

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 (-5.7638) 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



Observando el FAS vemos que decrece rápidamente y solo tiene un coeficiente ligeramente significativo por lo que proponemos un MA(1). Y el FAP decrece rápidamente a cero con el primer coeficiente diferente de cero por lo que también planteamos un AR(1) aunque no esta muy claro. Adicionalmente planteamos un ARMA(1,1).

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
  71.848700     2.003241
```

Al tener un T calculado (71.8487) MAYOR al T critico (2.003241) se acepta la hipótesis alterna por lo que incluimos la constante en el modelo

Alumno: Maye Mamani Victor Raul

2 Estimación

Modelo ARIMA(1,1,0)

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

```
z test of coefficients:

       Estimate Std. Error z value Pr(>|z|)
ar1   -0.258410   0.125577 -2.0578  0.03961 *
drift  0.025171   0.014133  1.7809  0.07492 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Con los parámetros estimados del modelo tenemos lo siguiente

$$\Delta Z_t = 0.02517 - 0.2584\Delta Z_{t-1} + a_t$$

Modelo ARIMA(0,1,1)

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

```
z test of coefficients:

       Estimate Std. Error z value Pr(>|z|)
ma1   -0.249715   0.125657 -1.9873  0.04689 *
drift  0.025133   0.013402  1.8753  0.06075 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Con los parámetros estimados del modelo tenemos lo siguiente

$$\Delta Z_t = 0.02513 + a_t - 0.2497a_{t-1}$$

Modelo ARIMA (1,1,1)

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

```
z test of coefficients:

       Estimate Std. Error z value Pr(>|z|)
ar1   -0.276737   0.446566 -0.6197  0.53545
ma1    0.019661   0.463039  0.0425  0.96613
drift  0.025163   0.014202  1.7717  0.07644 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Con los parámetros estimados del modelo tenemos lo siguiente

$$\Delta Z_t = 70260 - 0.2767\Delta Z_{t-1} + a_t + 0.0196a_{t-1}$$

3 Validación

3.1 Análisis de los coeficientes estimados

3.1.1 Significación de los coeficientes

Para el modelo ARIMA(1,1,0)

```
z test of coefficients:

       Estimate Std. Error z value Pr(>|z|)
ar1   -0.258410   0.125577 -2.0578  0.03961 *
drift  0.025171   0.014133  1.7809  0.07492 .
```

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

$$AR(1): \phi_1 = -0.2584 \rightarrow p = 0.0396 < 0.05, es\ significativo$$

Para el modelo ARIMA (0,1,1)

```
z test of coefficients:

       Estimate Std. Error z value Pr(>|z|)
ma1   -0.249715   0.125657 -1.9873  0.04689 *
drift  0.025133   0.013402  1.8753  0.06075 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$MA(1): \theta_1 = -0.2497 \rightarrow p = 0.0468 < 0.05, es\ significativo$$

Para el modelo ARIMA(1,1,1)

```
 test of coefficients:

       Estimate Std. Error z value Pr(>|z|)
ar1   -0.276737   0.446566 -0.6197  0.53545
ma1    0.019661   0.463039  0.0425  0.96613
drift  0.025163   0.014202  1.7717  0.07644 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$AR(1): \phi_1 = -0.2767 \rightarrow p = 0.5354 > 0.05, NO\ es\ significativo$$

$$MA(1): \theta_1 = -0.0196 \rightarrow p = 0.9661 > 0.05, NO\ es\ significativo$$

3.1.2 Examen de la matriz de correlaciones entre los coeficientes

```
vcov(mod1)
```

```
                ar1         drift
ar1    1.576965e-02 -3.699575e-06
drift -3.699575e-06  1.997519e-04
```

```
vcov(mod2)
```

```
               ma1        drift
ma1   1.578965e-02 4.544012e-06
drift 4.544012e-06 1.796084e-04
```

```
vcov(mod3)
```

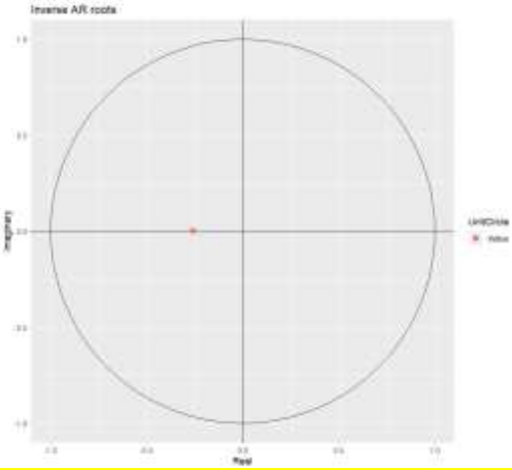
```
                ar1           ma1         drift
ar1    1.994215e-01 -1.985281e-01  2.399729e-05
ma1   -1.985281e-01  2.144049e-01 -2.993031e-05
drift  2.399729e-05 -2.993031e-05  2.017064e-04
```

Se observa claramente que NO hay coeficientes que esten cerca o sean superiores 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

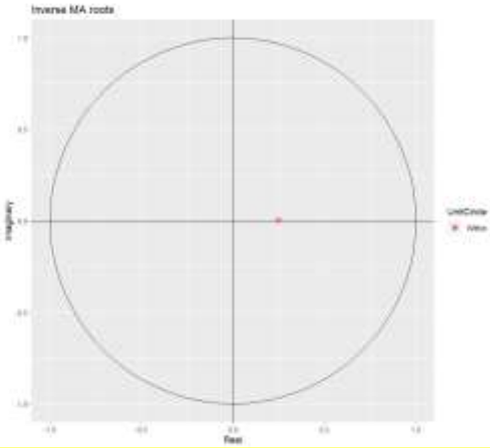
```
autoplot(mod1)
```

Alumno: Maye Mamani Victor Raul



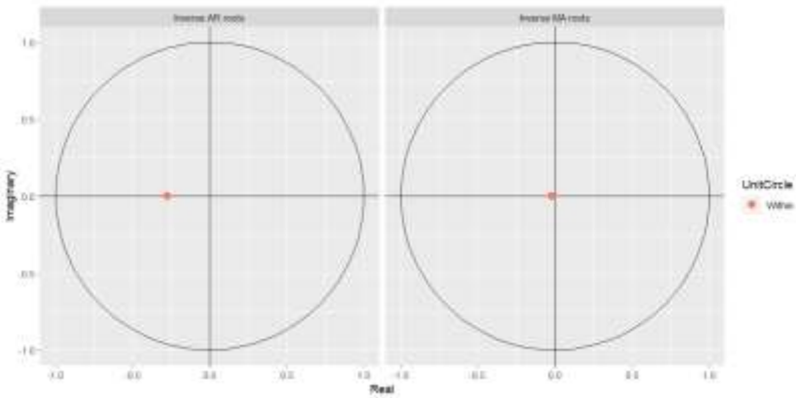
En la figura de raíces inversas de AR, se observa en el gráfico las raíces características se encuentran dentro del círculo, es decir que cumplen con la condición de estacionariedad para la parte autorregresiva.

```
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 = 88.171, p-value < 2.2e-16
```

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

```
supF test

data:  Chow_mod2
sup.F = 87.373, p-value < 2.2e-16
```

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

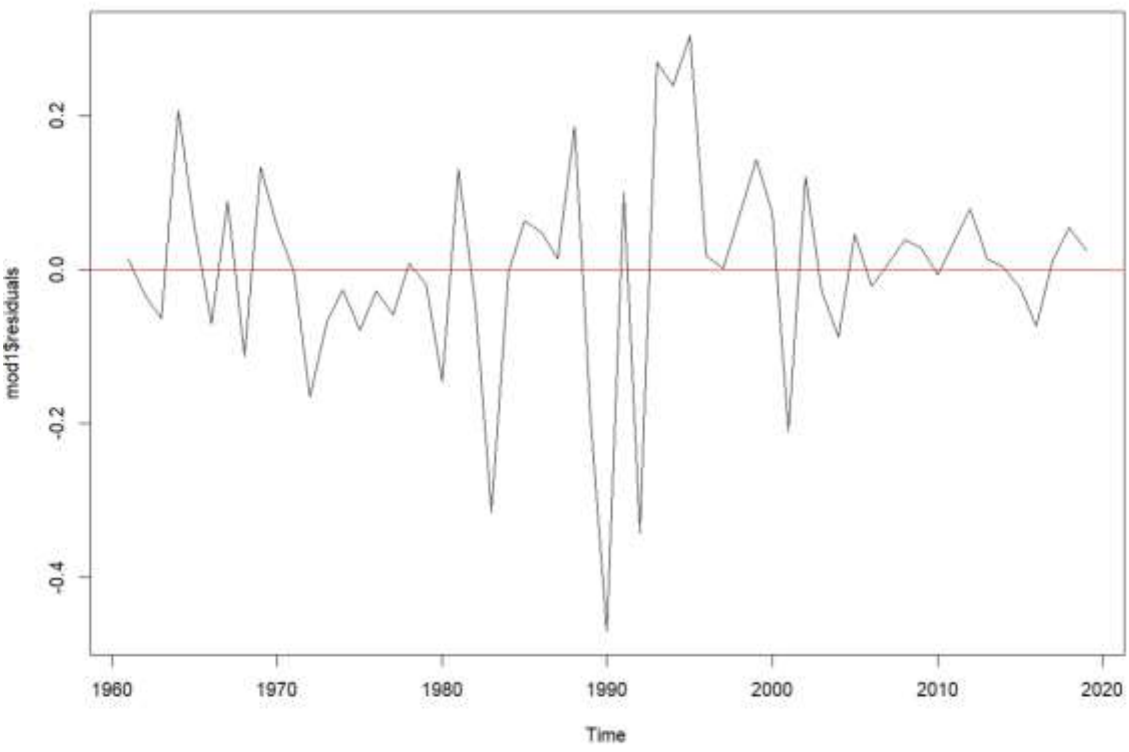
```
supF test

data:  Chow_mod3
sup.F = 88.253, 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)
```

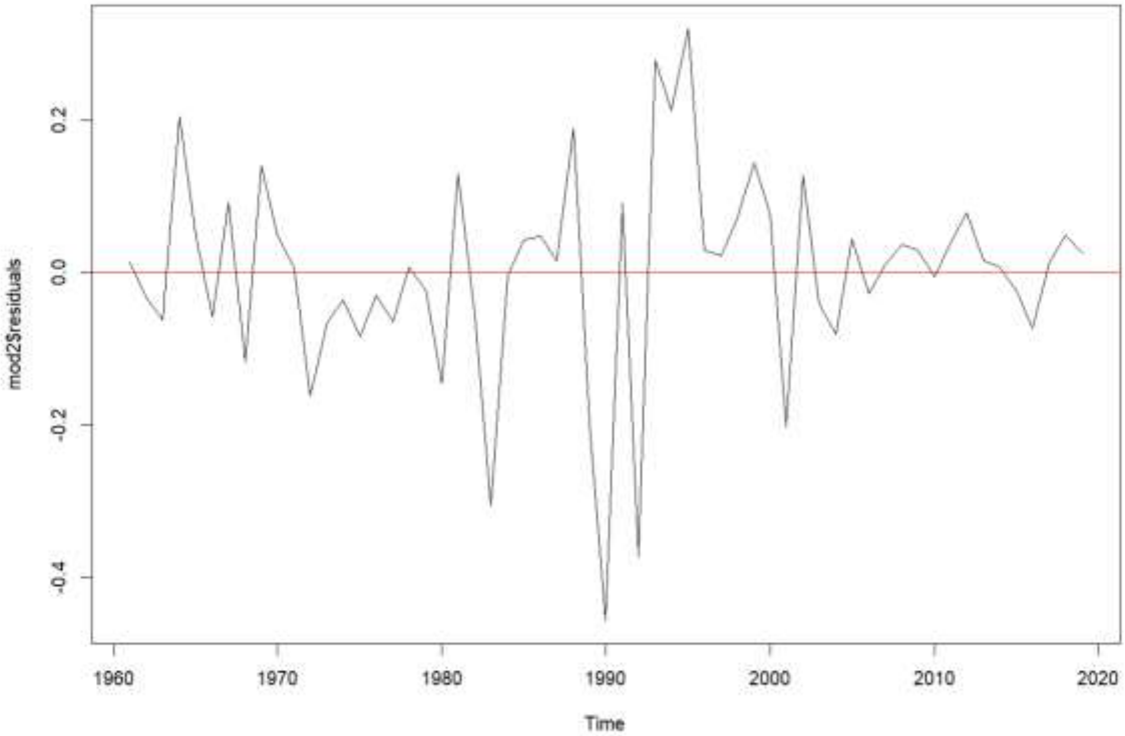



```
        One Sample t-test

data:   mod1$residuals
```

```
t = 0.0060581, df = 58, p-value = 0.9952
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.03506863  0.03528154
sample estimates:
   mean of x
0.0001064562
```

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 = 0.9952 > 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)
```

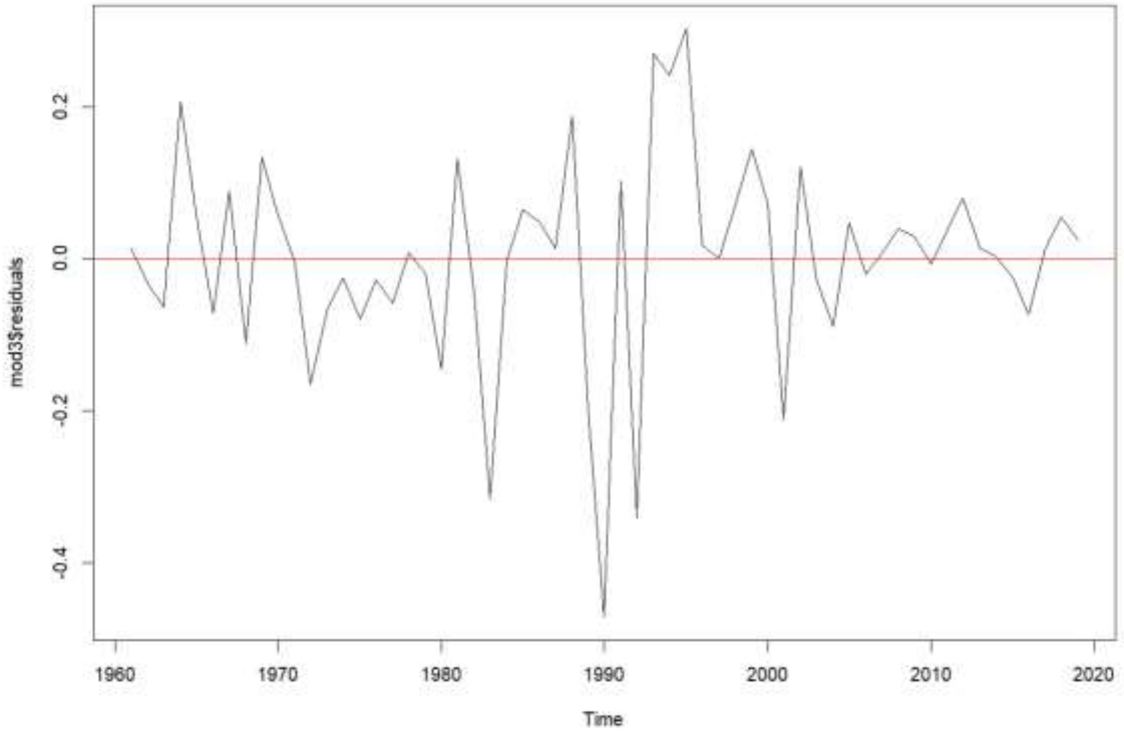



```
One Sample t-test

data:  mod2$residuals
t = 0.00428, df = 58, p-value = 0.9966
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.03517293  0.03532366
sample estimates:
   mean of x
7.536586e-05
```

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 = 0.9966 > α = 0.05, se acepta Ho, es decir la media es igual a cero.

```
plot(mod3$residuals)
```

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



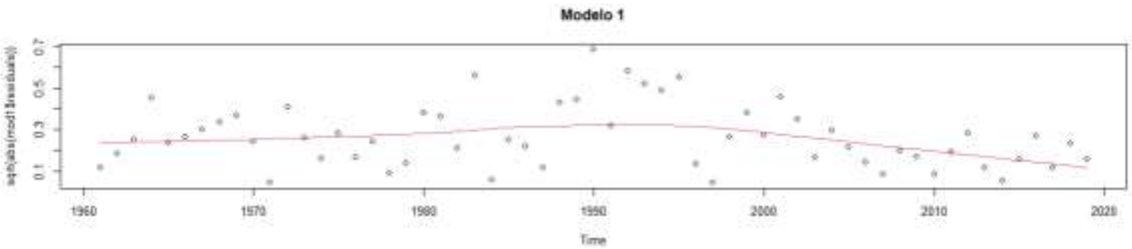

```
One Sample t-test

data:  mod3$residuals
t = 0.0068992, df = 58, p-value = 0.9945
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.03505327  0.03529574
sample estimates:
   mean of x
0.0001212334
```

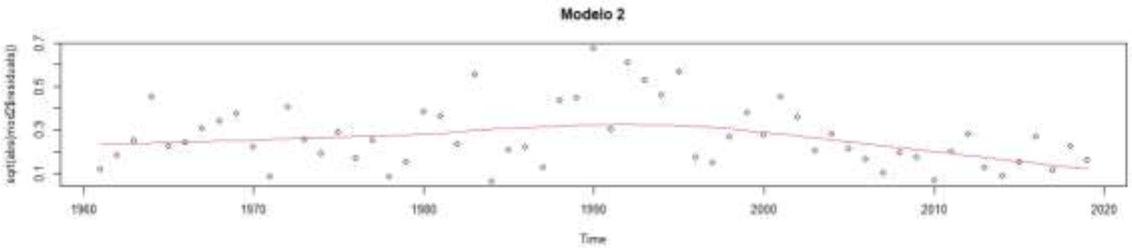
Parece indicar que un buen número de residuales están en torno a la media igual a cero. Para confirmar lo mencionado se realiza la prueba t: Como p = 0.9931 > α = 0.05, se acepta Ho, es decir la media es igual a cero.

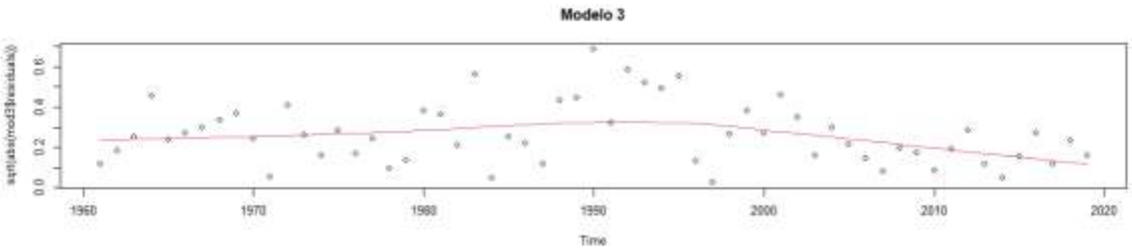
3.2.2 Homocedasticidad o varianza constante

```r
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 = 3.7168, df = 1, p-value = 0.05387
```

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

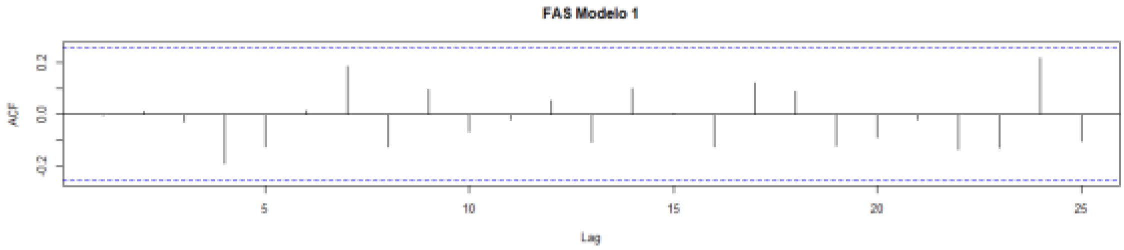
```
studentized Breusch-Pagan test

data:  resid(mod2) ~ I(obs - resid(mod2))
BP = 4.0436, df = 1, p-value = 0.04434
```

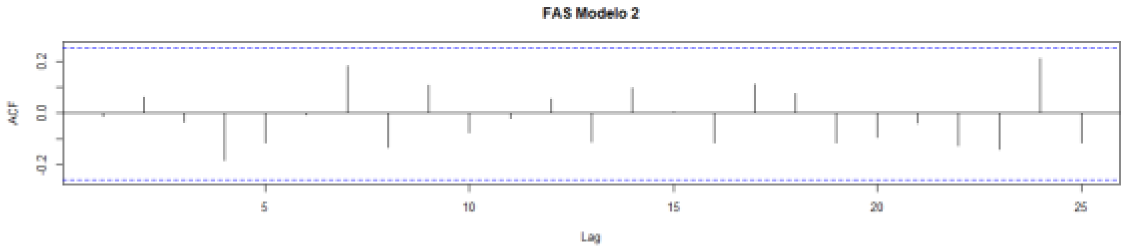
```
obs=get(mod3$series)
bptest(resid(mod3)~I(obs-resid(mod3)))
```

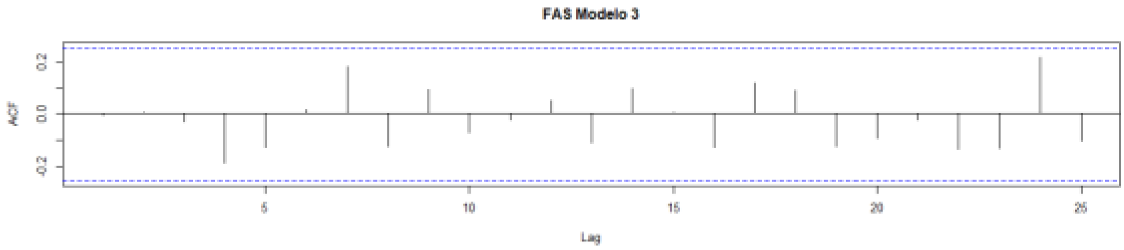
```
studentized Breusch-Pagan test

data:  resid(mod3) ~ I(obs - resid(mod3))
BP = 3.6964, df = 1, p-value = 0.05453
```

El valor de probabilidad (p-valor) asociado al estadístico BP asume los valores de probabilidad de 0.0538, 0.04434 y 0.0545 para los modelos 1,2 y 3 respectivamente, que son menores a α=0.05, por lo cual podemos afirmar que los residuales de estos modelos NO 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 = 0.00054983, df = 1, p-value = 0.9813
```

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

data:  resid_m2
X-squared = 0.013791, df = 1, p-value = 0.9065
```

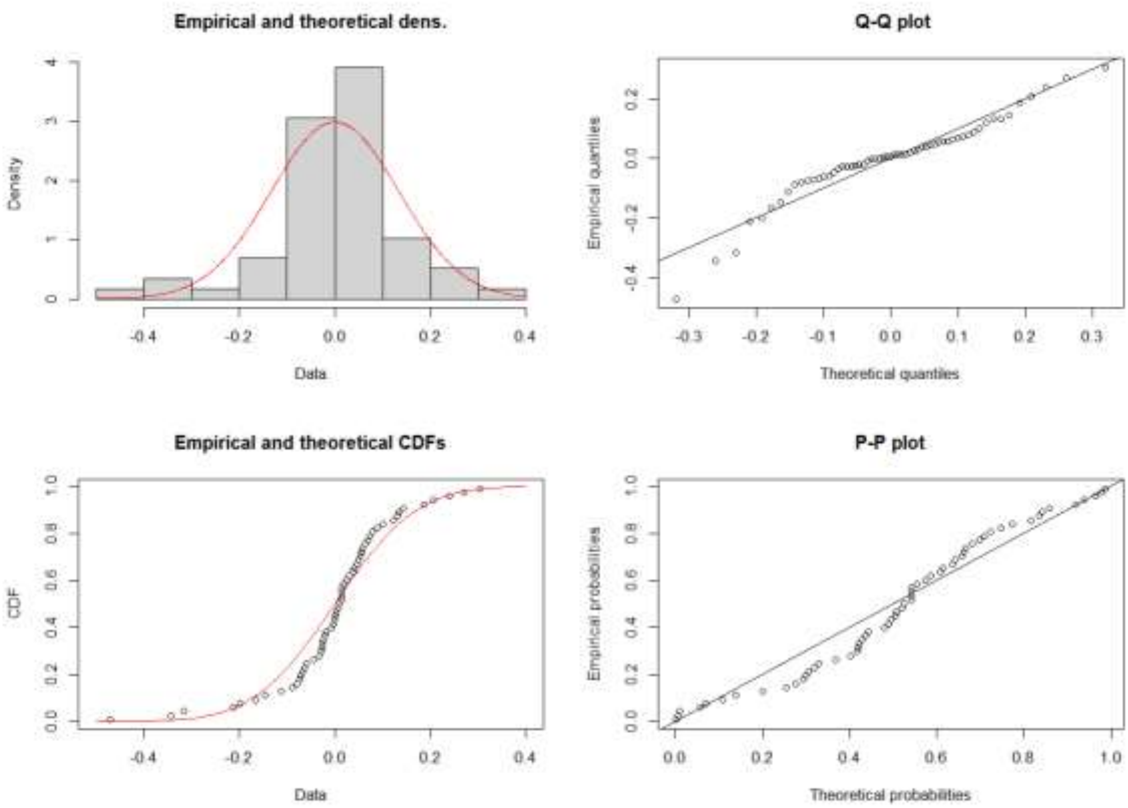
```
Box.test(resid_m3,type = "Ljung-Box")

Box-Ljung test
```

```
data:  resid_m3
X-squared = 0.0011834, df = 1, p-value = 0.9726
```

Con los valores 0.9813, 0.9065 y 0.9065 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
```

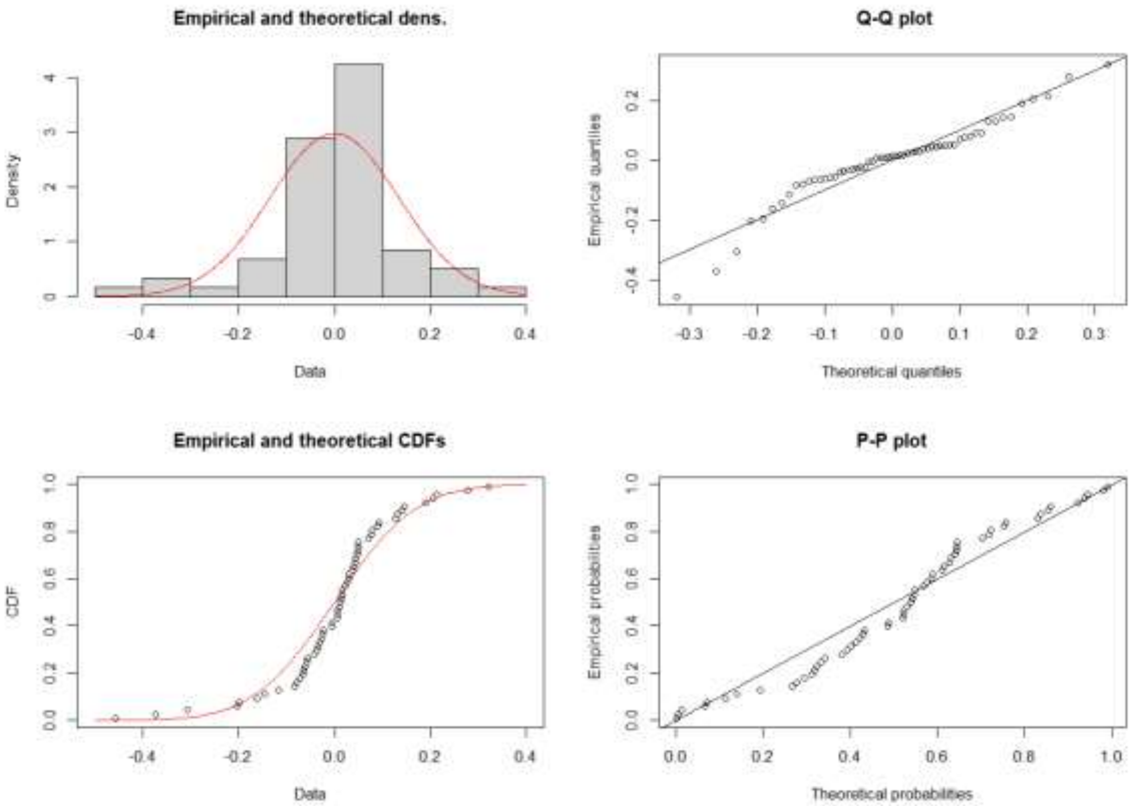



```
Jarque Bera Test

data:  resid_m1
X-squared = 19.141, df = 2, p-value = 6.976e-05
```

En las figuras se observa que los residuales del modelo 1 no presentan indicios de ajuste hacia la distribución normal. En la prueba JB, como p = 0.0000 < 0.05, se RECHAZA 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

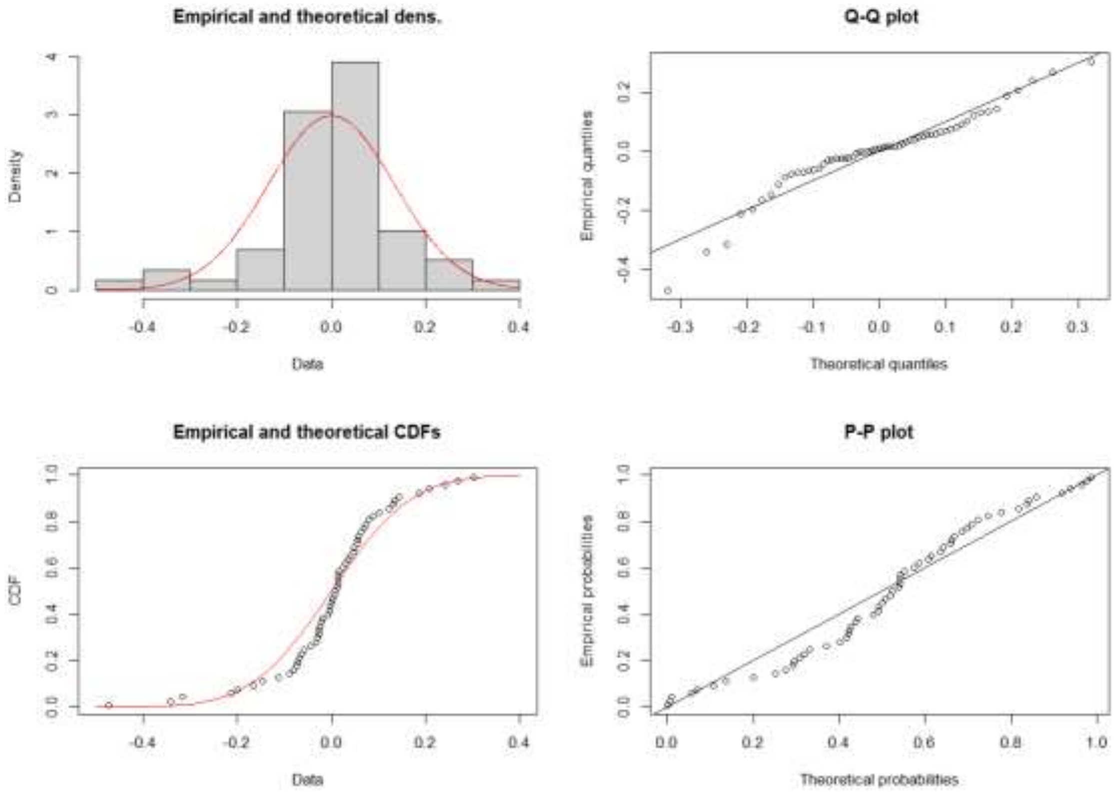



```
Jarque Bera Test

data:  resid_m2
X-squared = 18.168, df = 2, p-value = 0.0001135
```

En las figuras se observa que los residuales del modelo 2 no presentan indicios de ajuste hacia la distribución normal. En la prueba JB, como p = 0.00011 < 0.05, se RECHAZA 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 = 19.16, df = 2, p-value = 6.909e-05
```

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

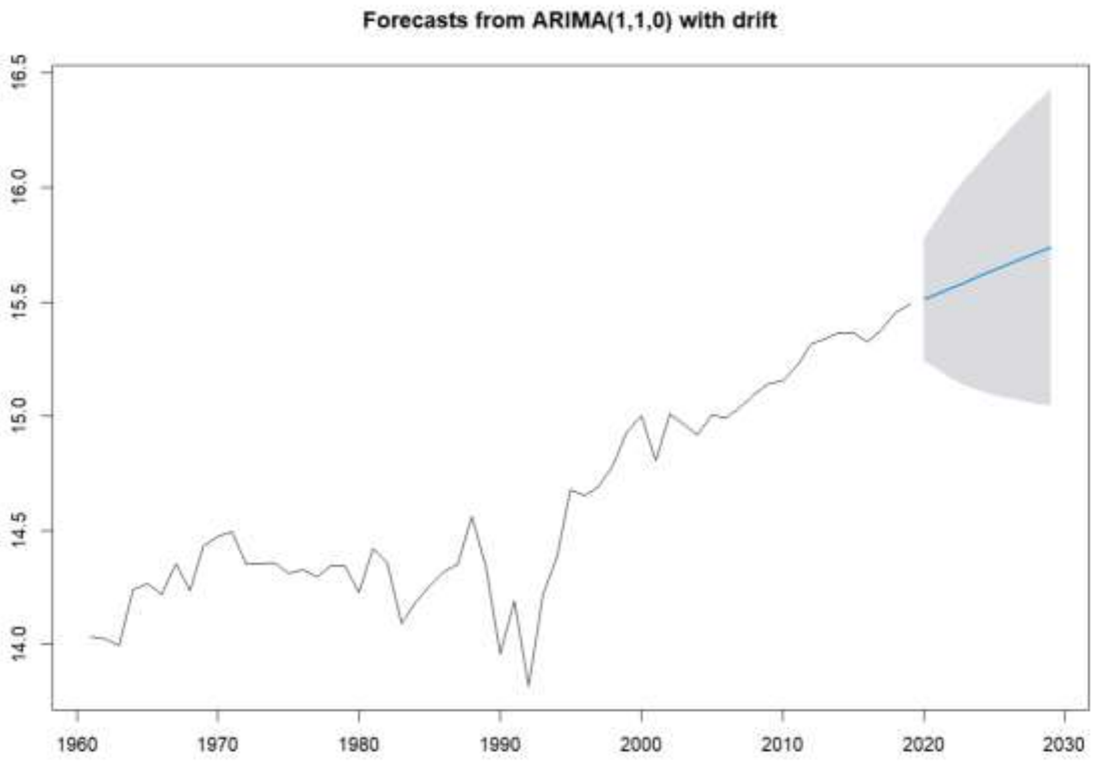
4 Pronostico

4.1Pronosticos de cada modelo

Modelo 1: ARIMA (1,1,0)

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

Alumno: Maye Mamani Victor Raul



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

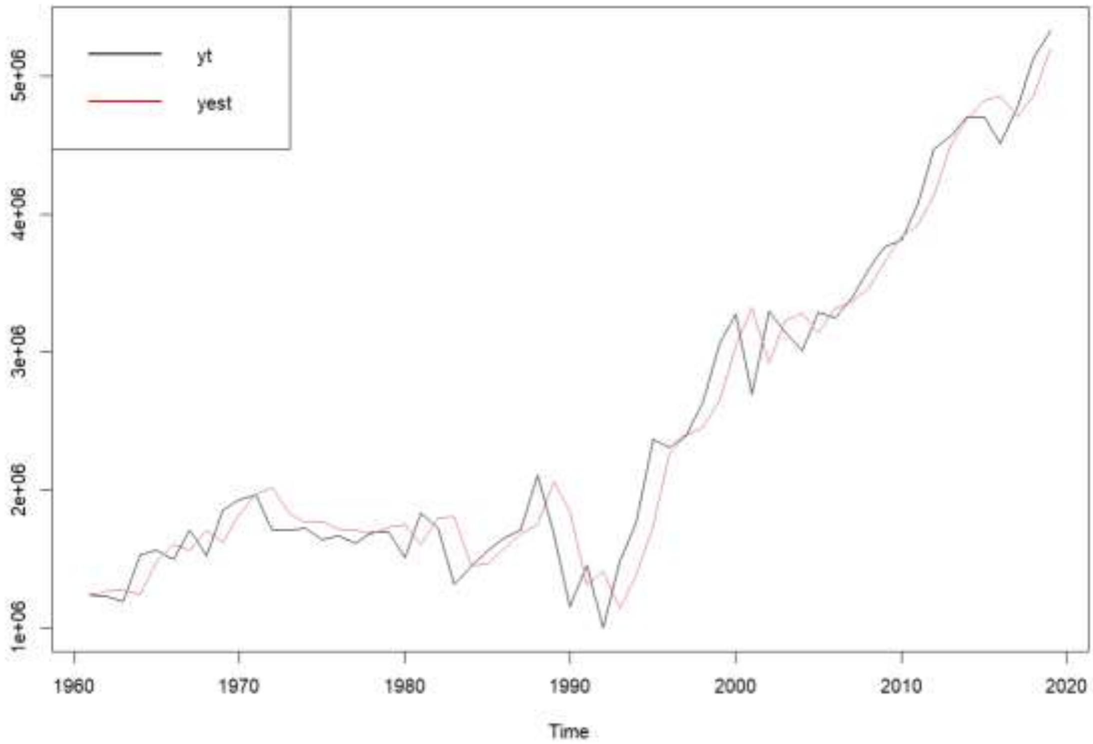
Los datos proyectados para los siguientes 10 años son:

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

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

Coefficients:
          ar1   drift
      -0.2584  0.0252
s.e.   0.1256  0.0141

sigma^2 = 0.01887:  log likelihood = 33.83
AIC=-61.65   AICc=-61.21   BIC=-55.47

Error measures:
                       ME      RMSE        MAE          MPE      MAPE
MASE         ACF1
Training set 0.0001064562 0.1338278 0.09156039 -0.006624538 0.6352687
0.9057662 -0.002976719

Forecasts:
     Point Forecast    Lo 95    Hi 95
2020       15.51088 15.24165 15.78011
2021       15.53692 15.20173 15.87210
2022       15.56186 15.16222 15.96151
2023       15.58709 15.13423 16.03995
2024       15.61225 15.11129 16.11321
2025       15.63742 15.09271 16.18213
2026       15.66259 15.07737 16.24781
2027       15.68776 15.06466 16.31086
2028       15.71293 15.05412 16.37174
2029       15.73810 15.04542 16.43078
```

Alumno: Maye Mamani Victor Raul

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:

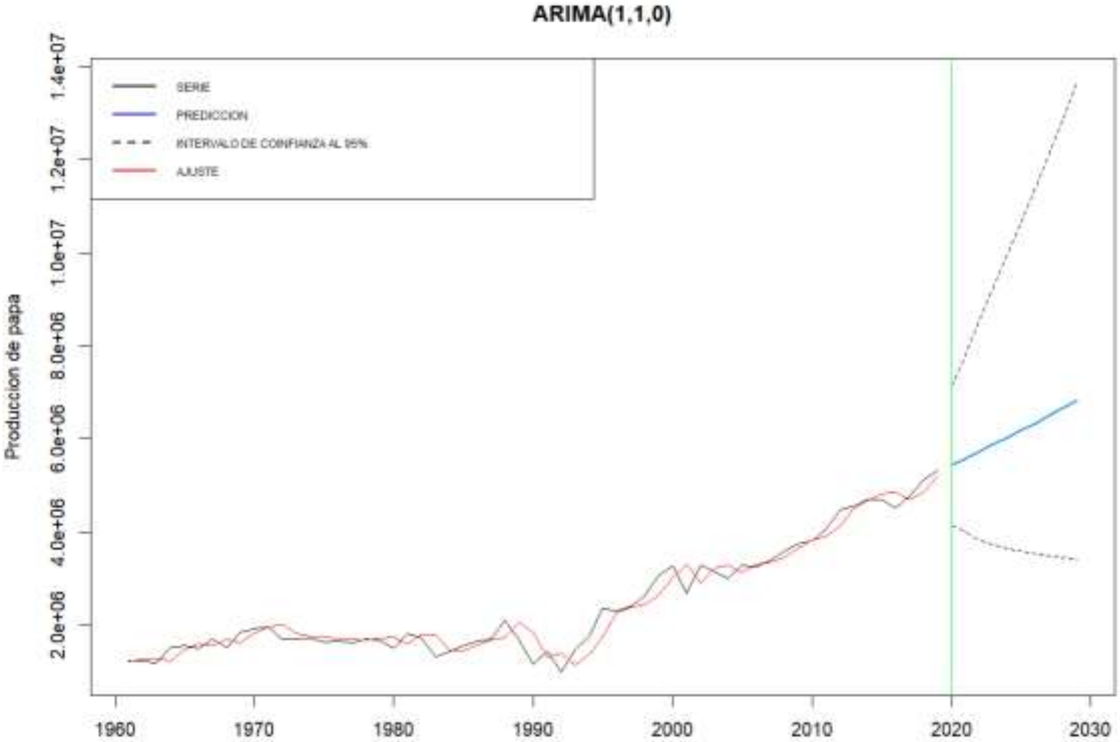
```
Pron1$mean <- exp(Pron1$mean)
Pron1$lower <- exp(Pron1$lower)
Pron1$upper <- exp(Pron1$upper)
Pron1$x <- exp(Pron1$x)
Pron1$fitted <- exp(Pron1$fitted)
Pron1$residuals <- exp(Pron1$residuals)
summary(Pron1)
```

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

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

Coefficients:
          ar1   drift
      -0.2584  0.0252
```

```
s.e.    0.1256  0.0141

sigma^2 = 0.01887:  log likelihood = 33.83
AIC=-61.65   AICc=-61.21   BIC=-55.47

Error measures:
                     ME     RMSE      MAE        MPE     MAPE      MASE
ACF1
Training set 14140.15 248384.4 184763.5 -0.9245591 9.410599 0.8924902
0.0205373

Forecasts:
     Point Forecast   Lo 95    Hi 95
2020        5448652 4162585  7132061
2021        5592382 3999699  7819274
2022        5733648 3844744  8550562
2023        5880140 3738636  9248303
2024        6029935 3653833  9951224
2025        6183662 3586595 10661273
2026        6341277 3531975 11385075
2027        6502918 3487371 12126024
2028        6668678 3450819 12887161
2029        6838663 3420935 13670916
```

GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

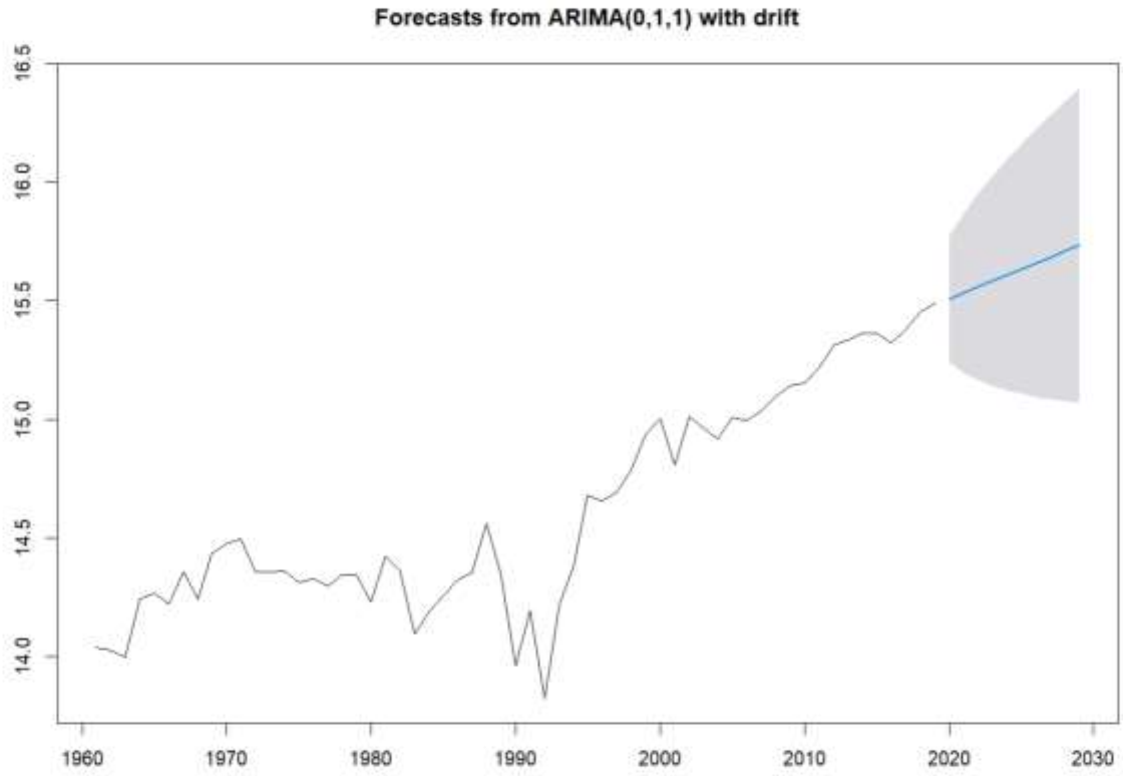
```
plot(Pron1, shaded = FALSE, xlab = "Años", ylab = "Produccion de
papa",main = "ARIMA(1,1,0)")
lines(Pron1$fitted, col = "red")
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE
COINFIANZA AL 95%", "AJUSTE"),col=c("black", "blue", "black", "red"),
lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=2020, lwd = 1, col="green")
```



Alumno: Maye Mamani Victor Raul

Modelo 2: ARIMA (0,1,1)

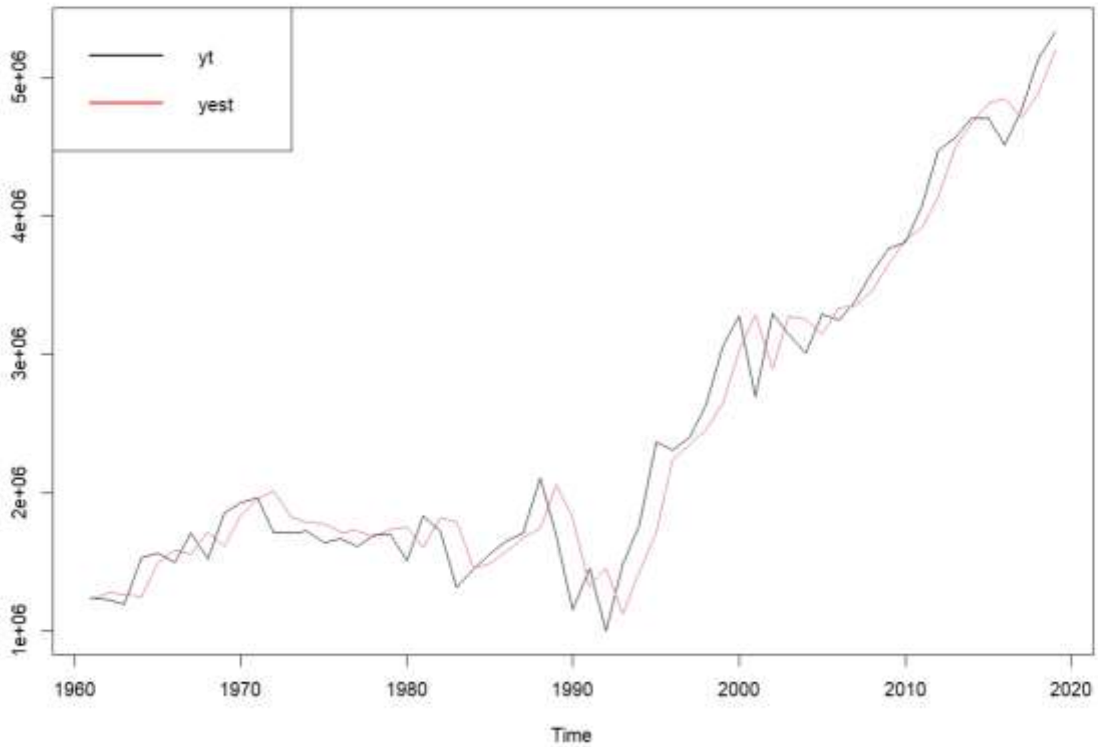
```
Pron2 <- forecast(mod2,level=c(95),h=10)
plot(Pron2)
summary(Pron2)
```



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)
```

Alumno: Maye Mamani Victor Raul



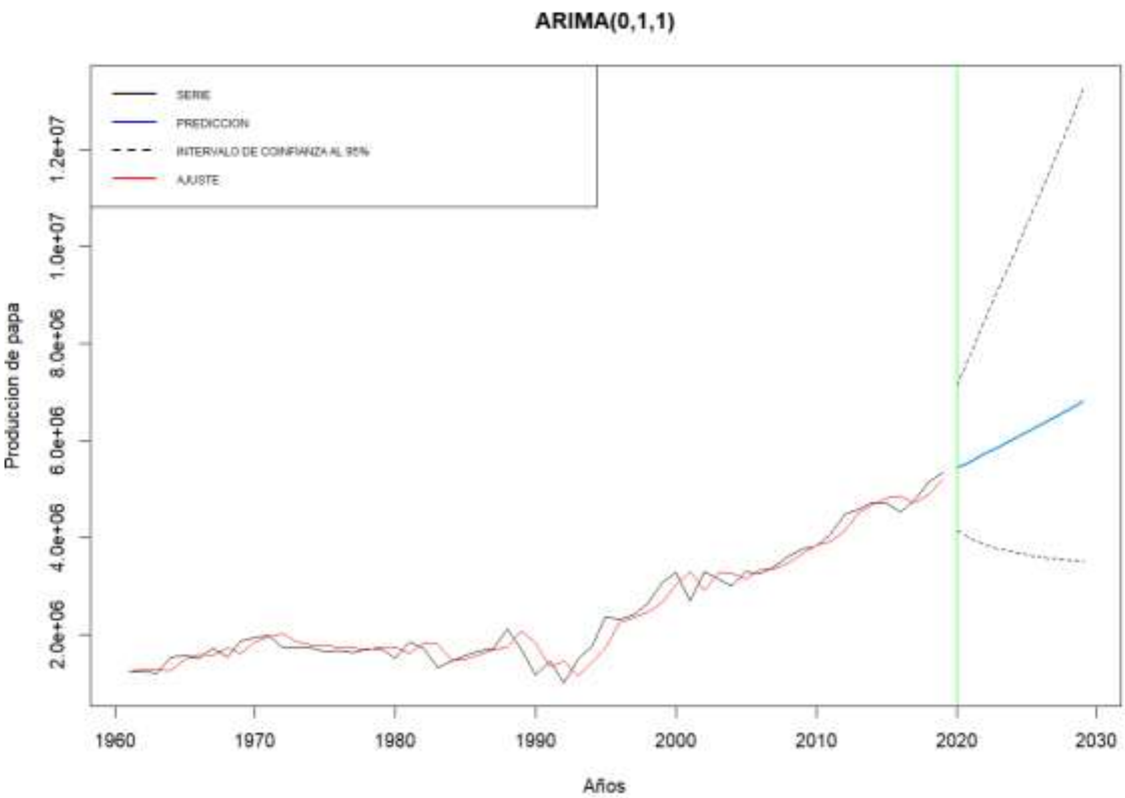

```r
Pron2$mean <- exp(Pron2$mean)
Pron2$lower <- exp(Pron2$lower)
Pron2$upper <- exp(Pron2$upper)
Pron2$x <- exp(Pron2$x)
Pron2$fitted <- exp(Pron2$fitted)
Pron2$residuals <- exp(Pron2$residuals)
summary(Pron2)
```

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

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

Coefficients:
          ma1   drift
      -0.2497  0.0251
s.e.   0.1257  0.0134

sigma^2 = 0.01895:  log likelihood = 33.71
AIC=-61.42   AICc=-60.97   BIC=-55.23

Error measures:
                   ME     RMSE      MAE        MPE     MAPE      MASE
ACF1
Training set 15276.54 247738.6 186812.5 -0.9300957 9.488785 0.9023876
0.0161299

Forecasts:
     Point Forecast   Lo 95   Hi 95
2020        5432074 4147596 7114346
2021        5570326 3975566 7804809
2022        5712097 3854420 8465099
2023        5857475 3763414 9116726
2024        6006554 3692664 9770368
```

```
2025         6159427 3636602 10432414
2026         6316190 3591751 11107191
2027         6476944 3555784 11797906
2028         6641788 3527063 12507107
2029         6810829 3504390 13236936
```

GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

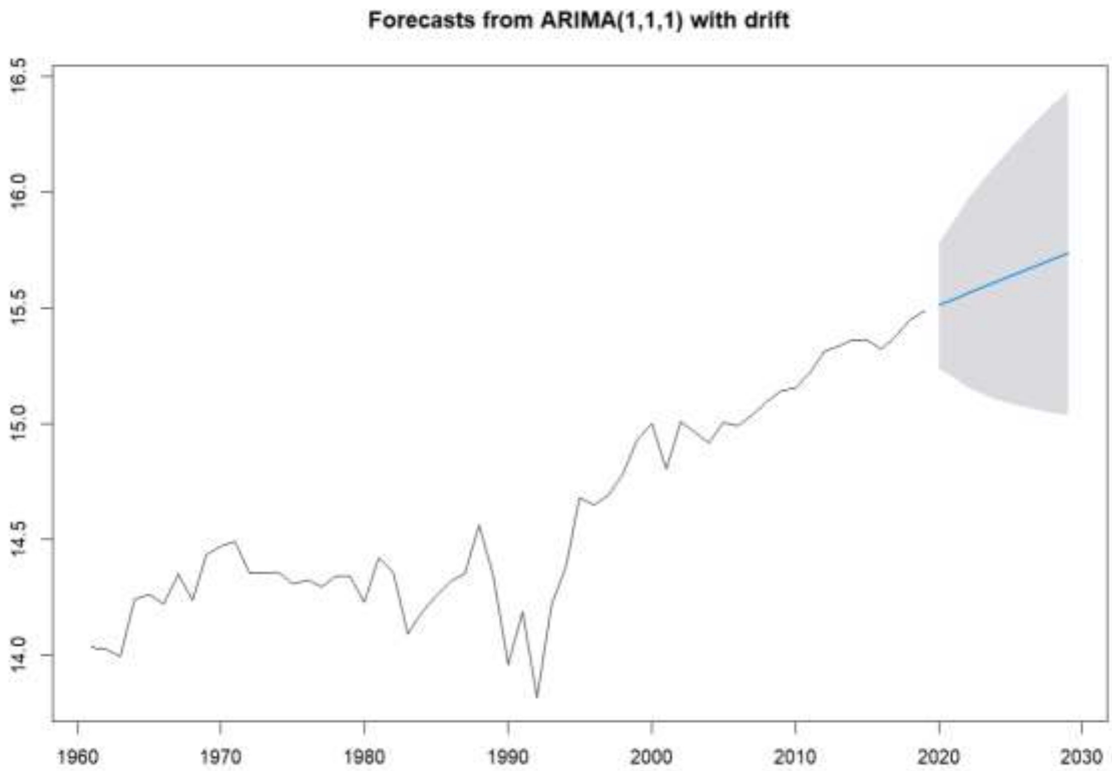
```
plot(Pron2, shaded = FALSE, xlab = "Años", ylab = "Produccion de
papa",main = "ARIMA(0,1,1)")
lines(Pron2$fitted, col = "red")
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE
COINFIANZA AL 95%", "AJUSTE"),col=c("black", "blue", "black", "red"),
lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=2020, lwd = 1, col="green")
```



Modelo 3: ARIMA (1,1,1)

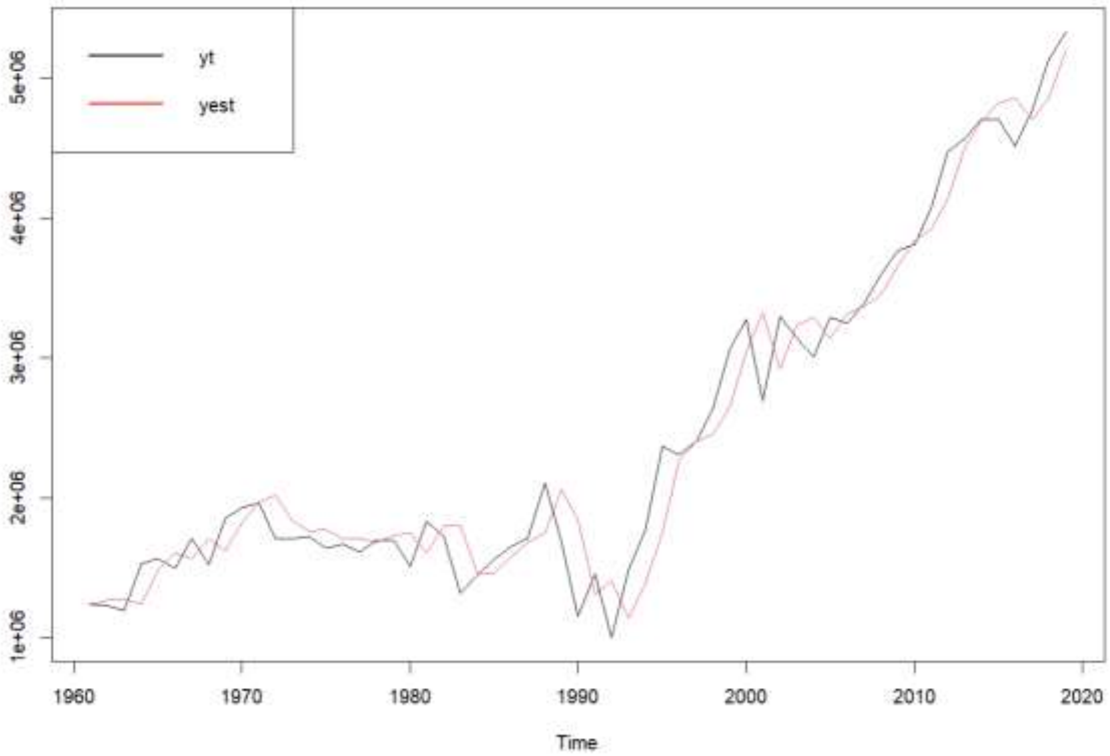
```
Pron3 <- forecast(modelo3,level=c(95),h=10)
plot(Pron3)
summary(Pron3)
```

Alumno: Maye Mamani Victor Raul



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

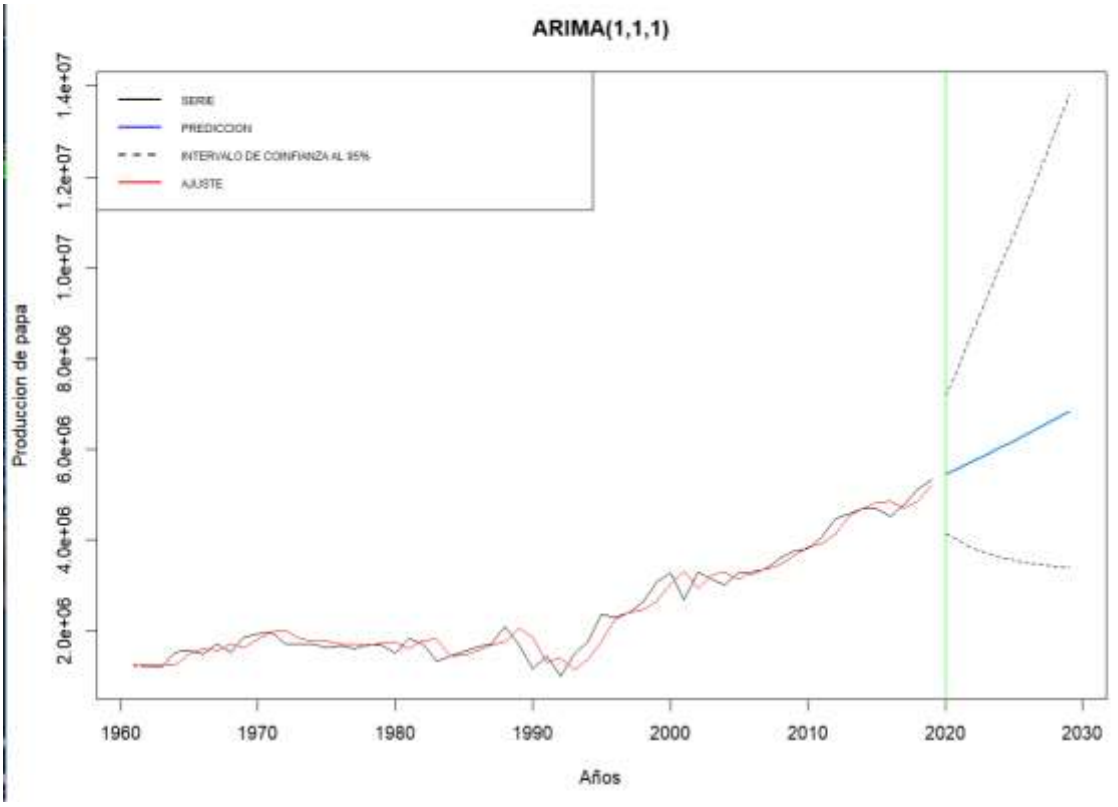
Alumno: Maye Mamani Victor Raul

```r
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(1,1,1) with drift

Model Information:
Series: data_tf
ARIMA(1,1,1) with drift

Coefficients:
          ar1     ma1   drift
      -0.2767  0.0197  0.0252
s.e.   0.4466  0.4630  0.0142

sigma^2 = 0.01921:  log likelihood = 33.83
AIC=-59.65   AICc=-58.9   BIC=-51.41

Error measures:
                   ME   RMSE      MAE       MPE     MAPE      MASE
ACF1
Training set 14067.32 248468 184674.9 -0.923185 9.407904 0.8920624 0.0
1885228

Forecasts:
     Point Forecast   Lo 95    Hi 95
2020        5449957 4153468  7151140
2021        5593637 3987640  7846440
2022        5734810 3827735  8592038
2023        5881330 3719202  9300394
2024        6031088 3631884 10015192
2025        6184802 3562618 10736986
2026        6342393 3506127 11473046
2027        6504012 3459822 12226688
2028        6669746 3421679 13001074
2029        6839703 3390293 13798673
```

```r
plot(Pron3, shaded = FALSE, xlab = "Años", ylab = "Produccion de
papa",main = "ARIMA(1,1,1)")
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=2020, 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 14140.15 248384.4 184763.5 -0.9245591 9.410599 0.8924902
0.0205373
```

```
accuracy(Pron2)
                    ME       RMSE        MAE          MPE       MAPE        MASE
ACF1
Training set 15276.54 247738.6 186812.5 -0.9300957 9.488785 0.9023876
0.0161299
```

```
accuracy(Pron3)
                    ME     RMSE        MAE         MPE       MAPE        MASE
ACF1
Training set 14067.32 248468 184674.9 -0.923185 9.407904 0.8920624 0.0
1885228
```

Basado en las métricas MAE y RMSE, el modelo ARIMA(1,1,0) es el mejor entre los tres, ya que tiene los valores más bajos(MAE = 184763.5 Y RMSE=247738.6) en ambas métricas. Esto indica que sus predicciones son más precisas y con menos variabilidad en los errores.

Conclusión

Modelos Evaluados:

1. **ARIMA(1,1,0)**:

  - **Significancia**: Coeficiente AR significativo.
  - **Errores**: Muestra mayores errores en comparación con otros modelos.
  - **Estabilidad**: No es estable.
2. **ARIMA(0,1,1)**:
  - **Significancia**: Coeficiente MA significativo.
  - **Errores**: Mejor ajuste en los residuales, con menores errores.
  - **Estabilidad**: Tampoco es completamente estable.
3. **ARIMA(1,1,1)**:
  - **Significancia**: Coeficientes no significativos.
  - **Errores**: Menor rendimiento en términos de ajuste y predicción.
  - **Estabilidad**: Similar falta de estabilidad.

Modelo Seleccionado: ARIMA(1,1,0) se selecciona como el más adecuado debido a la significancia del coeficiente AR y el mejor comportamiento en los errores residuales, a pesar de la falta de estabilidad observada en todos los modelos evaluados.