**Actividad 4**

**Estudiante: Maye Mamani Victor Raul**

Tabla de contenido

[CASO 1: VENTAS DE UN RESTAURANTE 3](#_Toc169690348)

[1. IDENTIFICACIÓN 3](#_Toc169690349)

[Análisis de la tendencia y la estacionalidad 3](#_Toc169690350)

[Conclusión: estacionalidad y tendencia 5](#_Toc169690351)

[Análisis de Estacionariedad 5](#_Toc169690352)

[Estacionariedad en Varianza 5](#_Toc169690353)

[Estacionariedad en Media 6](#_Toc169690354)

[Conclusión: Estacionariedad 7](#_Toc169690355)

[Identificación del modelo estacionario 7](#_Toc169690356)

[a) Identificación de las órdenes p y q 7](#_Toc169690357)

[b) Inclusión del termino independiente (δ) o intercepto 7](#_Toc169690358)

[Conclusión: Identificación de los modelos 8](#_Toc169690359)

[2. ESTIMACION 8](#_Toc169690360)

[3. VALIDACIÓN 9](#_Toc169690361)

[3.1. Análisis de los coeficientes estimados 9](#_Toc169690362)

[a. Significación de los coeficientes 9](#_Toc169690363)

[b. Examen de la matriz de correlaciones entre los coeficientes 9](#_Toc169690364)

[c. Condición de Convergencia e Invertibilidad 10](#_Toc169690365)

[d. Análisis de la estabilidad 10](#_Toc169690366)

[3.2. Análisis de los residuos 11](#_Toc169690367)

[a) Media es igual a cero 11](#_Toc169690368)

[b) Homocedasticidad o varianza constante 13](#_Toc169690369)

[c. Ausencia de correlación serial 14](#_Toc169690370)

[d) Contraste de normalidad 16](#_Toc169690371)

[CONCLUSION: ELECCION DEL MEJOR MODELO 18](#_Toc169690372)

[CASO 2: PRODUCCIÓN DE MAIZ 18](#_Toc169690373)

[1.IDENTIFICACION 19](#_Toc169690374)

[Análisis de la tendencia y estacionalidad 19](#_Toc169690375)

[Conclusión estacionalidad y tendencia: 19](#_Toc169690376)

[Análisis de la estacionariedad 19](#_Toc169690377)

[Estacionariedad en varianza 19](#_Toc169690378)

[Estacionariedad en media 21](#_Toc169690379)

[Conclusión de estacionariedad en media y varianza 23](#_Toc169690380)

[Diferenciación 23](#_Toc169690381)

[Conclusión de la diferenciación 24](#_Toc169690382)

[a. Identificación de las órdenes p y q 25](#_Toc169690383)

[b. Inclusión del término independiente o intercepto 25](#_Toc169690384)

[Conclusión de la inclusión del intercepto 26](#_Toc169690385)

[Conclusión: Identificación de los modelos 26](#_Toc169690386)

[2.ESTIMACION 26](#_Toc169690387)

[3. VALIDACION 27](#_Toc169690388)

[3.1. Análisis de los coeficientes estimados 27](#_Toc169690389)

[a. Significación de los coeficientes 27](#_Toc169690390)

[b. Examen de la matriz de correlaciones entre los coeficientes 27](#_Toc169690391)

[c. Condición de Convergencia e Invertibilidad 28](#_Toc169690392)

[d. Análisis de la estabilidad 29](#_Toc169690393)

[3.2 Análisis de residuos 30](#_Toc169690394)

[Homocedasticidad de los residuos 33](#_Toc169690395)

[Correlograma de los residuos 34](#_Toc169690396)

[Prueba de Ljung – Box 34](#_Toc169690397)

[CONCLUSION: 37](#_Toc169690398)

**LAS INTERPRETACIONES Y CONCLUSIONES DE CADA PASO SE ENCUENTRAN EN ESTE FORMATO**

**Ejemplo: Según el grafico anterior no hay tendencia**

# CASO 1: VENTAS DE UN RESTAURANTE

# 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)

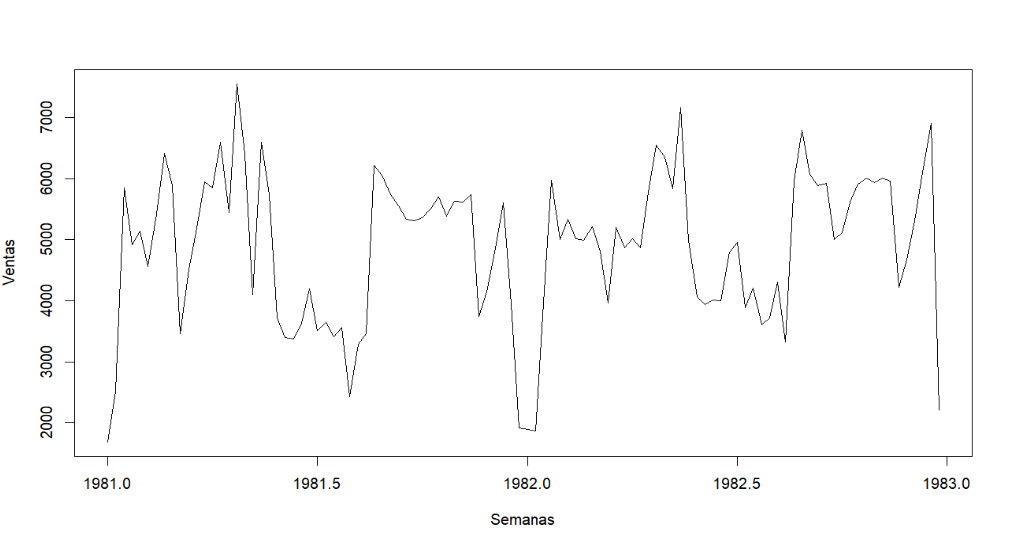
data <- read\_excel("F:\\777--Programacion repos\\Una\\r\\data\\actividad-04.xlsx",sheet = "datos1")

View(data)

# Gráfica de la serie

data\_ts <- ts(data$Ventas, start = 1981, frequency = 1)

plot(data\_ts, xlab="Semanas", ylab="Ventas")



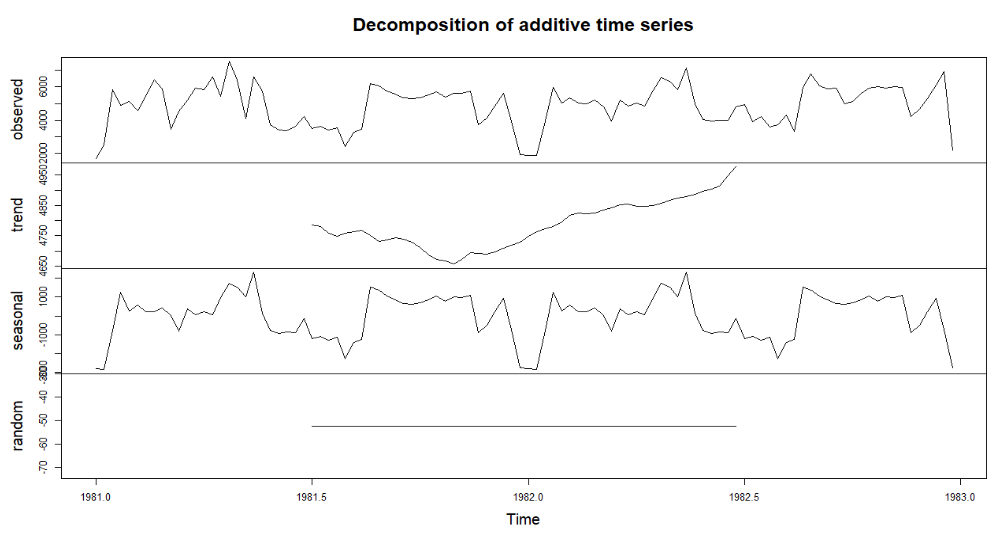
## 1. IDENTIFICACIÓN

### Análisis de la tendencia y la estacionalidad

#Descomposición de la serie

data\_des <- decompose(data\_ts, type = "additive")

plot(data\_des,type="l")



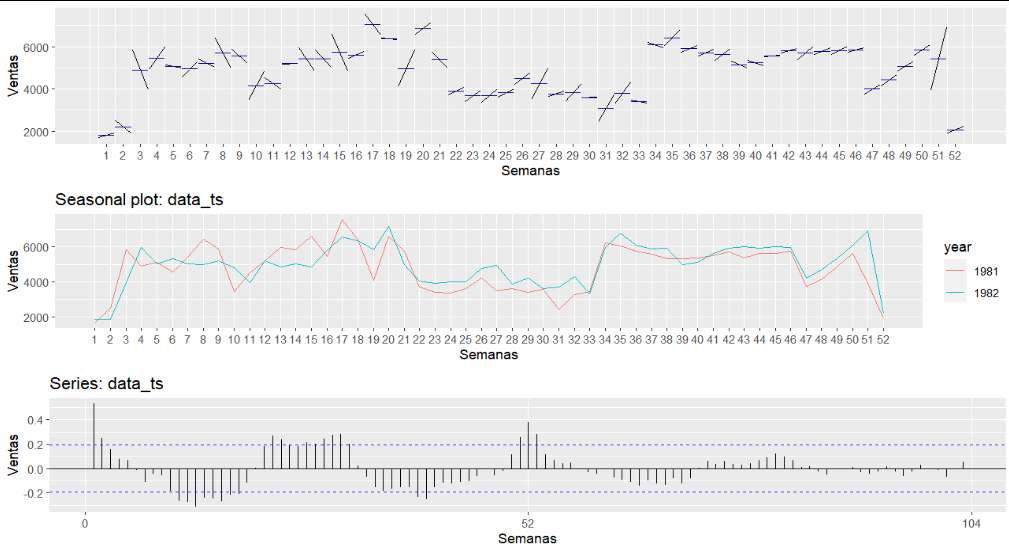
#Graficos de la serie para identificar estacionalidad

plot1 <- ggsubseriesplot(data\_ts, xlab = "Semanas", ylab = "Ventas")

plot2 <- ggseasonplot(data\_ts, xlab = "Semanas", ylab = "Ventas")

plot3 <- ggAcf(data\_ts, xlab = "Semanas", ylab = "Ventas")

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

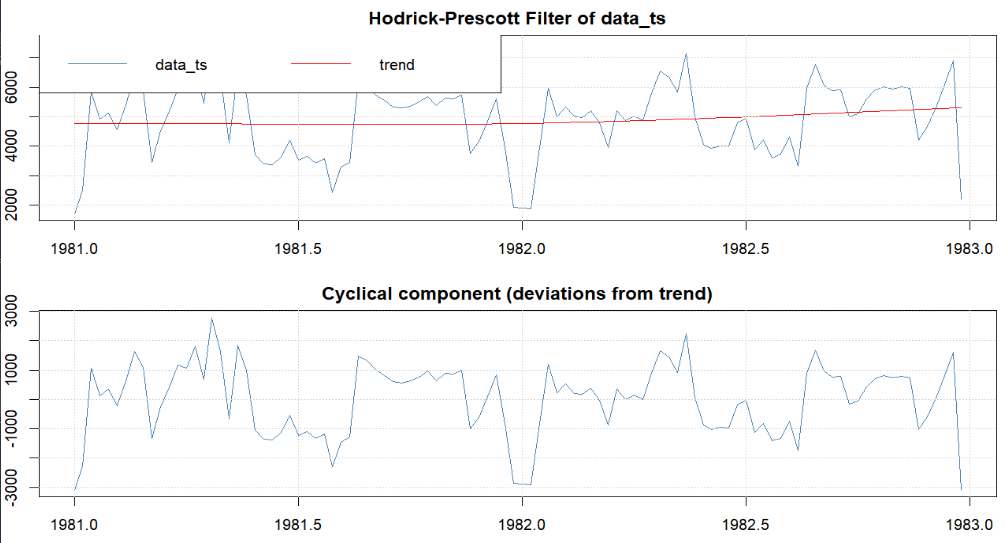


#Análisis de tendencia

lambda\_hp <- 129600

data\_hp <- hpfilter(data\_ts, type="lambda", freq=lambda\_hp)

plot(data\_hp)



#### Conclusión: estacionalidad y tendencia

**INTERPRETACION:**

**Según los gráficos anteriores NO parece haber una tendencia clara, pero sí hay estacionalidad y parece haber un patrón cíclico que se repite anualmente**

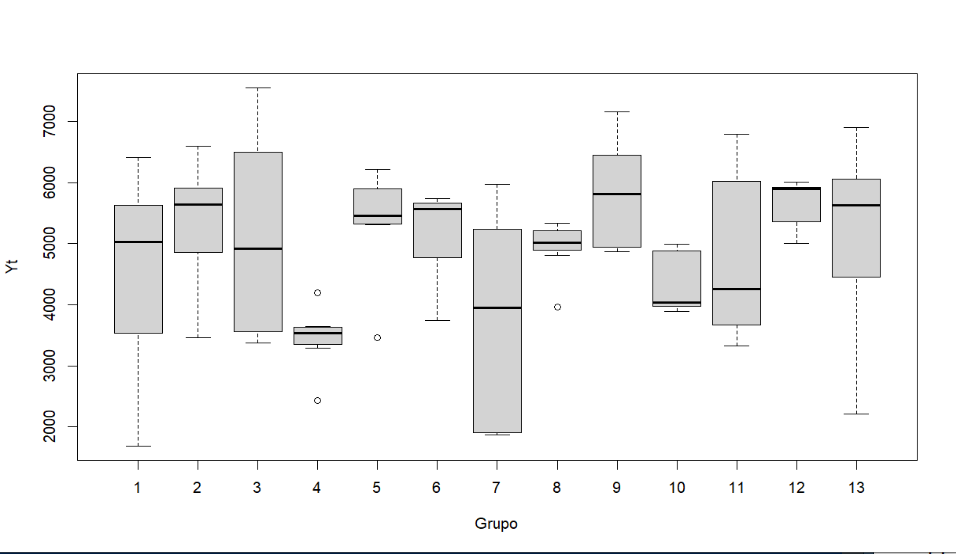
### Análisis de Estacionariedad

#### Estacionariedad en Varianza

#Estacionaridad en varianza

Grupo <- rep(1:13, each = 8)

boxplot(data$Ventas ~ Grupo, xlab = "Grupo", ylab="Yt")

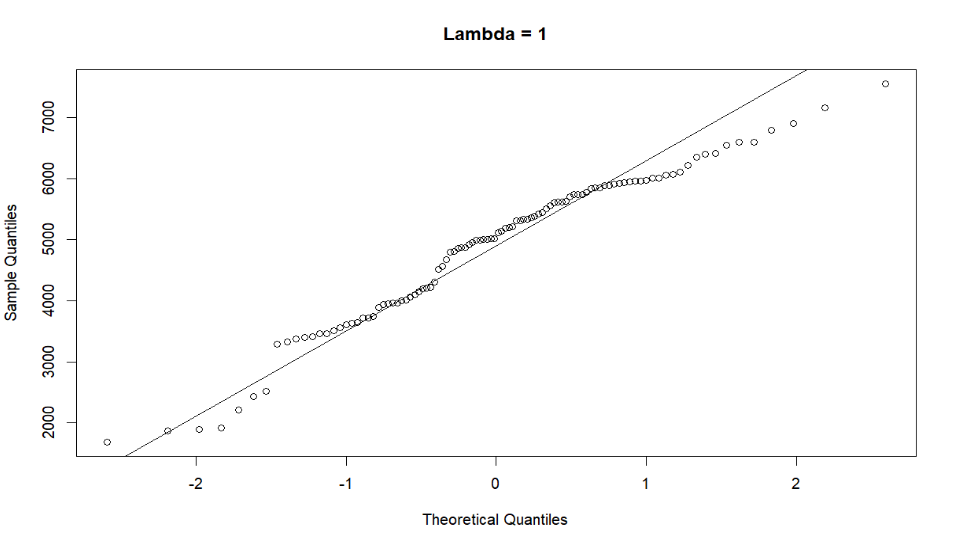


**Vemos que la variación no parece ser constante, por lo que probablemente necesitemos transformar la serie.**

**Si intentamos estimar un lambda adecuado nos dará un lambda de 1, ósea que será la serie original**

qqnorm(data\_ts, main = "Lambda = 1") # Yt: Original

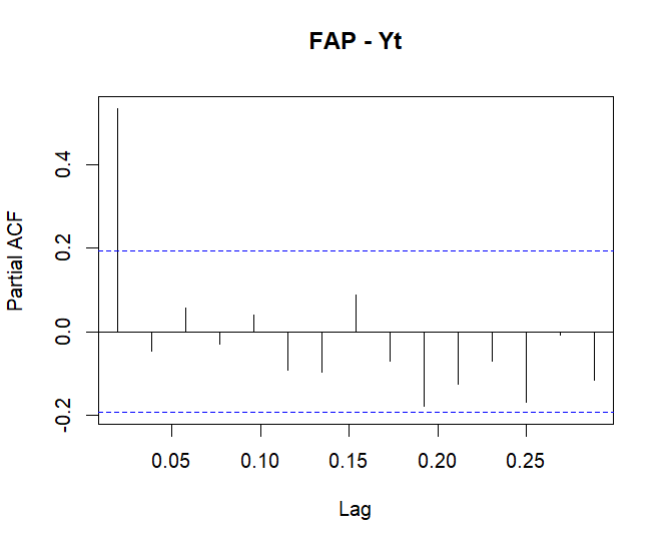
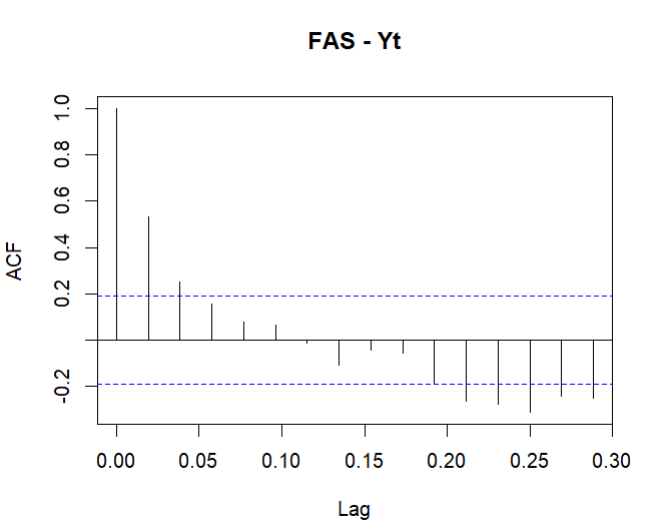
qqline(data\_ts)

****

#### Estacionariedad en Media

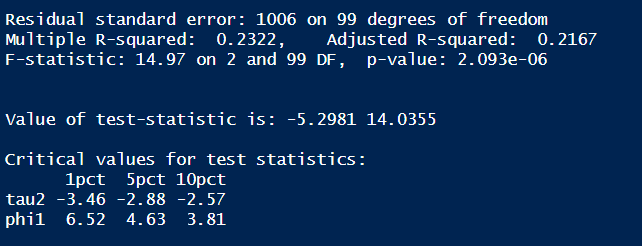
FAS <- acf(data\_ts, lag.max = 15, main="FAS - Yt", level = 0.95)

FAP <- pacf(data\_ts, lag.max = 15, main="FAP - Yt", level = 0.95)



**INTERPRETACION:**

**Los coeficientes de autocorrelación simple (FAS) decrecen hacia cero rápidamente de forma oscilatoria, además, el primer coeficiente de autocorrelación parcial (FAP) es menor a 0.9; por tanto, esta serie de ventas presenta altos indicios de ser estacionaria.**

****

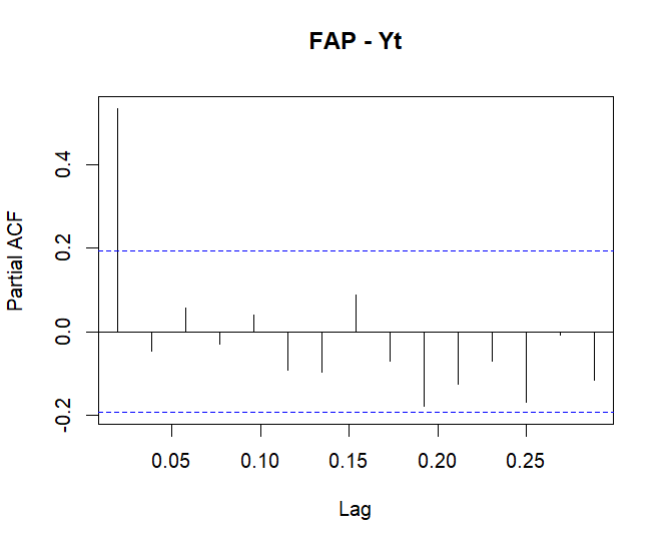
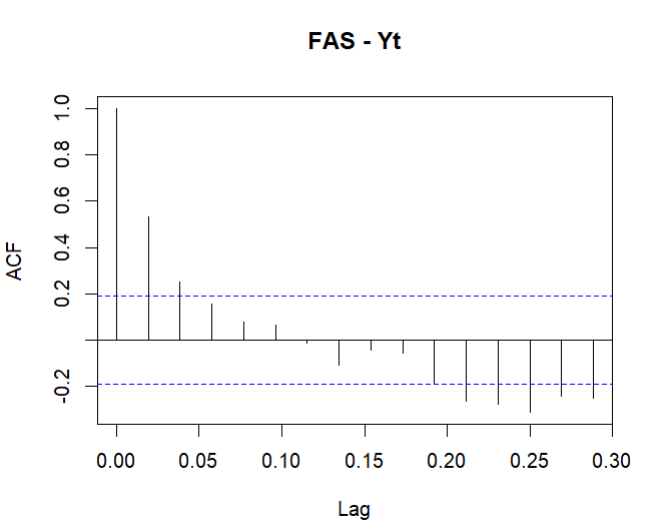
#### Conclusión: Estacionariedad

**INTERPRETACION:**

**De los resultados, el valor de T calculado =-5.29 8 1, T critico =-2.888 8 por lo tanto, aceptamos la hipótesis nula (Ho) de la existencia de raíz unitaria, es decir que la serie es estacionaria en media y varianza**

### Identificación del modelo estacionario

#### a) Identificación de las órdenes p y q



**La FAS estimada muestra una estructura infinita decreciendo en forma de onda seno-coseno amortiguada que sugiere un modelo ARMA(p, q) con parte autorregresiva de orden = 1.**

**Los tres primeros coeficientes del FAS son significativamente distintos de cero y el resto no. Según esta interpretación el modelo adecuado sería un AR(1) y MA(3).**

#### b) Inclusión del termino independiente (δ) o intercepto

#incluir el intercepto

Z <- mean(data\_ts)

Co <- var(data\_ts)

Tn <- length(data\_ts)

Ta <- Tn - 1

Sigma <- Co/Ta

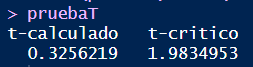
t <- Z/Sigma

tt <- qt(1-0.05/2,Ta-1)

pruebaT <- c(t, tt)

names(pruebaT) <- c("t-calculado","t-critico")

pruebaT



**INTERPRETACION:**

**Vemos que Tcalculado < Tcritico.Entonces se acepta la hipótesis nula, por lo cual no incluiremos el intercepto**

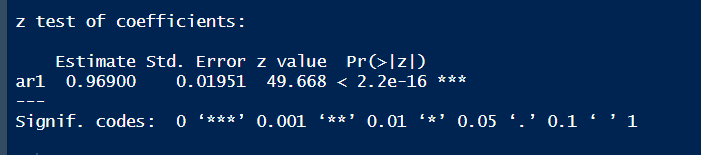
### Conclusión: Identificación de los modelos

**Resumiendo, se proponen los siguientes modelos para la serie estacionaria:**

## 2. ESTIMACION

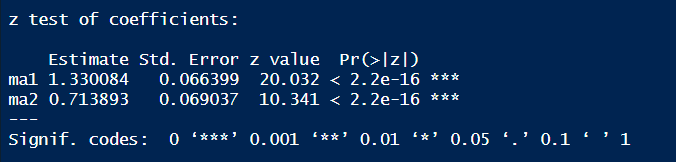
mod1 <- Arima(data\_ts, order = c(1, 0, 0), include.constant = FALSE)

coeftest(mod1)



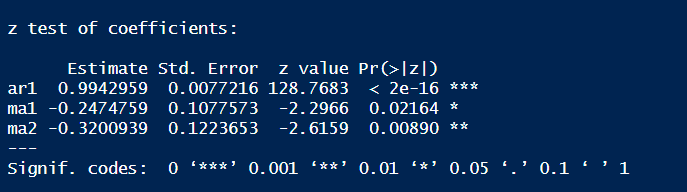
mod2 <- Arima(data\_ts, order = c(0, 0, 2), include.constant = FALSE)

coeftest(mod2)



mod3 <- Arima(data\_ts, order = c(1, 0, 2), include.constant = FALSE)

coeftest(mod3)



## 3. VALIDACIÓN

### 3.1. Análisis de los coeficientes estimados

#### a. Significación de los coeficientes

**# De las imágenes anteriores analizamos su significancia**

**Modelo 1: AR(1)**

**Modelo 2: MA(2)**

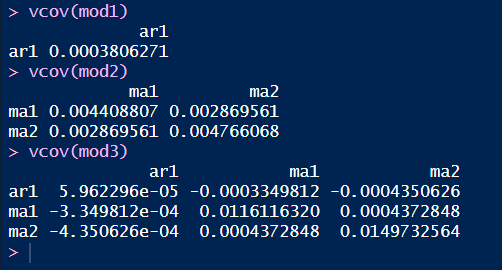
**Modelo 3: ARMA(1,2)**

#### b. Examen de la matriz de correlaciones entre los coeficientes

vcov(mod1)

vcov(mod2)

vcov(mod3)



**INTERPRETACION:**

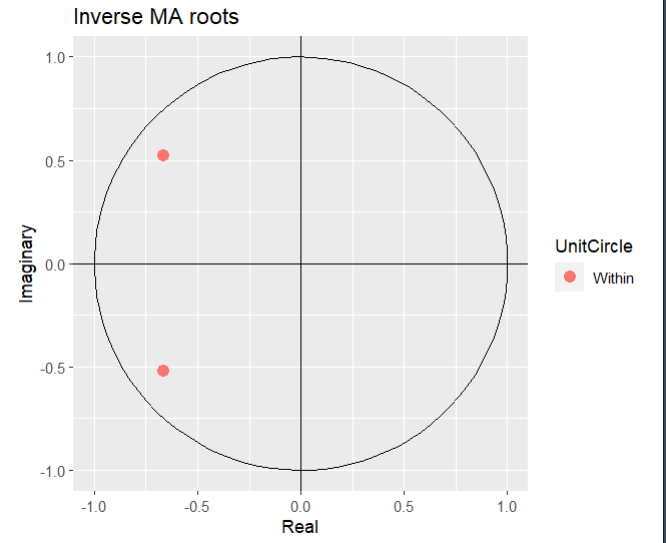
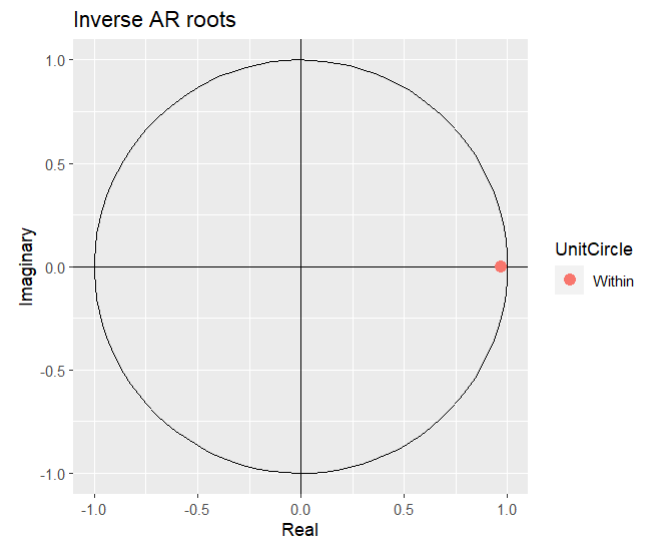
**Se observa claramente que ningún coeficiente esta próximo ni cercano a 0.9, por tanto, podemos indicar que no hay problema de multicolinealidad en los modelos propuestos.**

#### c. Condición de Convergencia e Invertibilidad

autoplot(mod1)

autoplot(mod2)

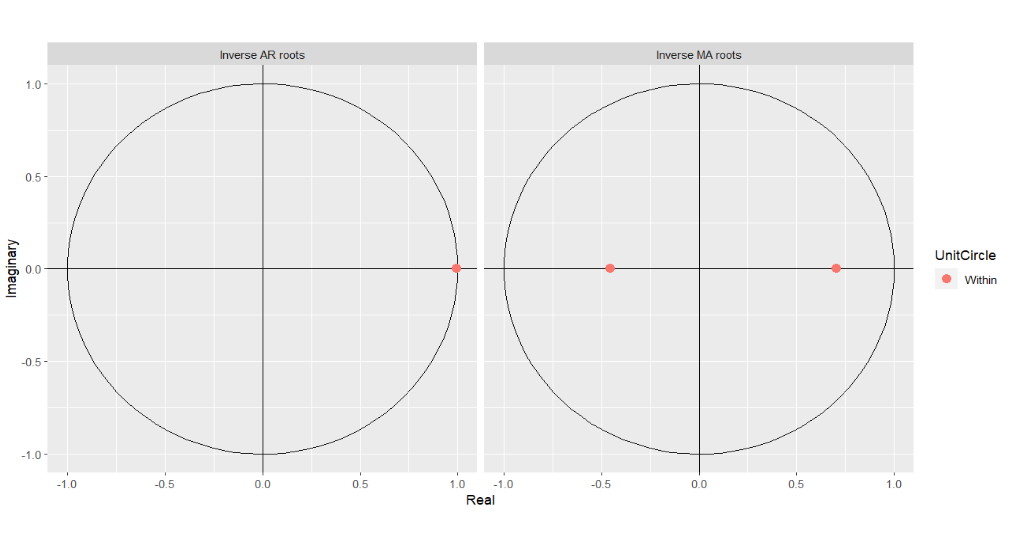
autoplot(mod3)



**INTERPRETACION:**

**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.**

**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.**



**INTERPRETACION:**

**En la última figura 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.**

#### d. Análisis de la estabilidad

Chow\_mod1 <- Fstats(mod1$fitted ~ 1, from = 0.43)

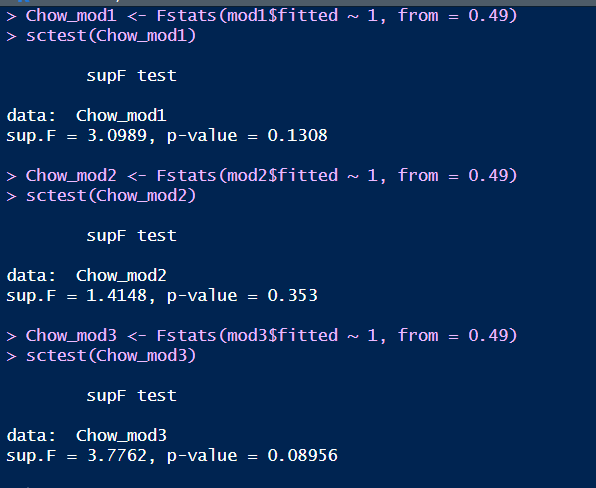
sctest(Chow\_mod1)

Chow\_mod2 <- Fstats(mod2$fitted ~ 1, from = 0.43)

sctest(Chow\_mod2)

Chow\_mod3 <- Fstats(mod3$fitted ~ 1, from = 0.43)

sctest(Chow\_mod3)



**INTERPRETACION:**

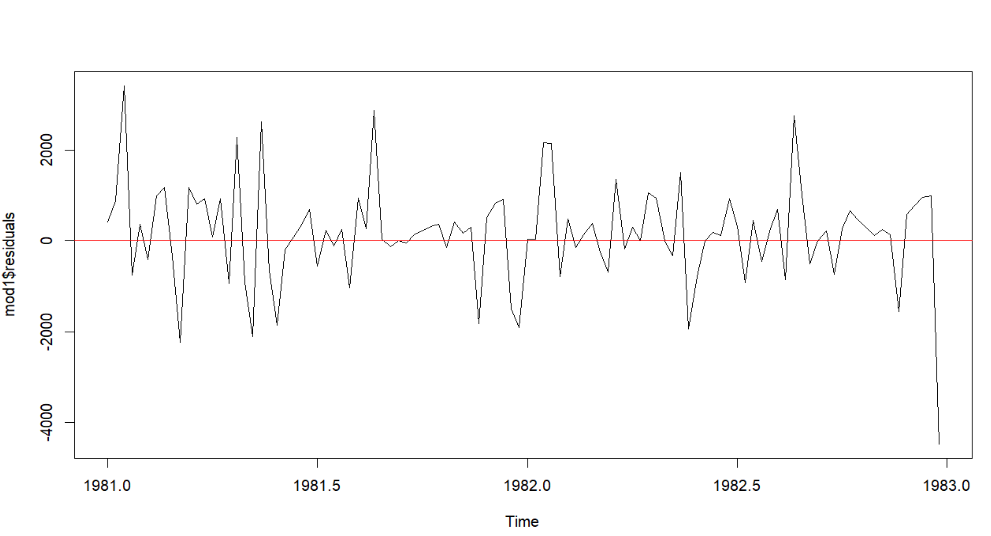
**En las tres pruebas se acepta la hipótesis nula (p > α = 0.05), es decir, existe estabilidad de coeficientes.**

### 3.2. Análisis de los residuos

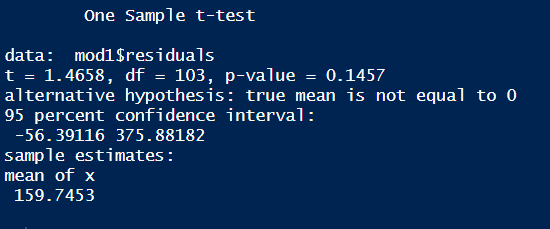
#### a) Media es igual a cero

plot(mod1$residuals)

abline(h = 0, col = "red")



t.test(mod1$residuals, mu = 0)



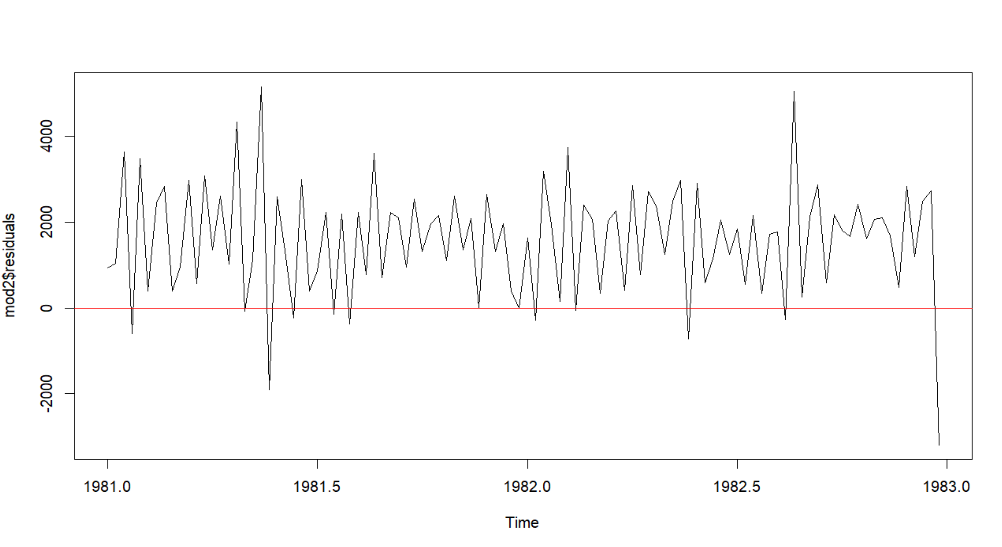
**INTERPRETACION:**

**Segun el grafico vemos que un buen numero de residuales están en torno a la media igual a cero.**

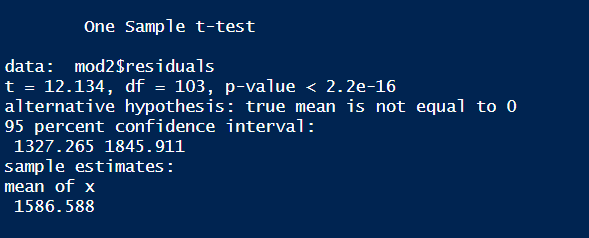
**Y confirmando lo anterior p=0.1457 > 0.05, se acepta Ho, es decir la media es igual a cero**

plot(mod2$residuals)

abline(h = 0, col = "red")



t.test(mod2$residuals, mu = 0)



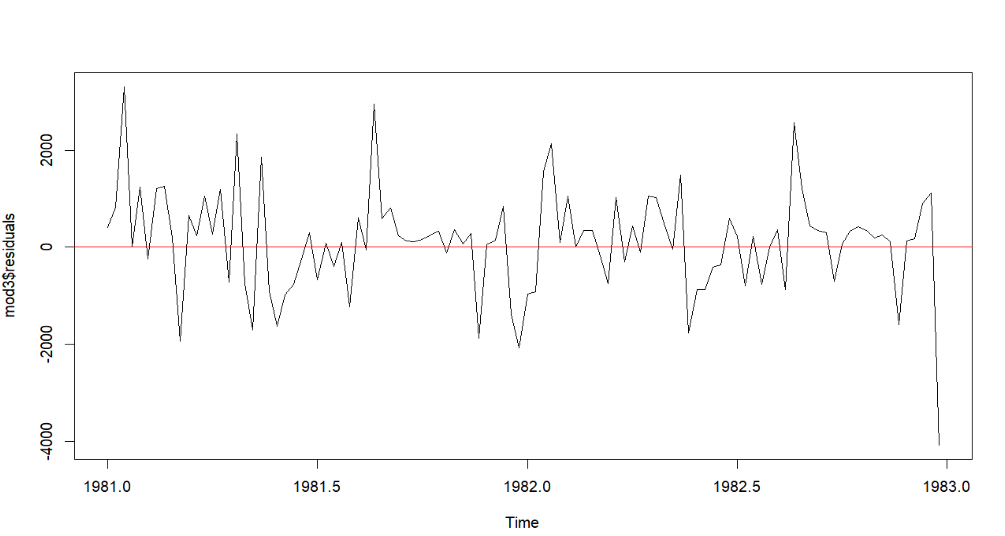
**INTERPRETACION:**

**Segun el grafico vemos que los residuales NO están en torno a la media igual a cero.**

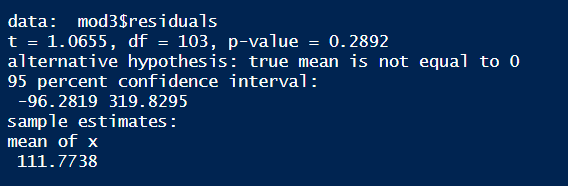
**Y confirmando lo anterior p=0.00000 < 0.05, se rechaza Ho, es decir la media no es igual a cero.**

plot(mod3$residuals)

abline(h = 0, col = "red")



t.test(mod3$residuals, mu = 0)



**INTERPRETACION:**

**Segun el grafico vemos que un buen número de residuales están en torno a la media igual a cero.**

**Y confirmando lo anterior p=0.2892 > 0.05, se acepta Ho, es decir la media es igual a cero**

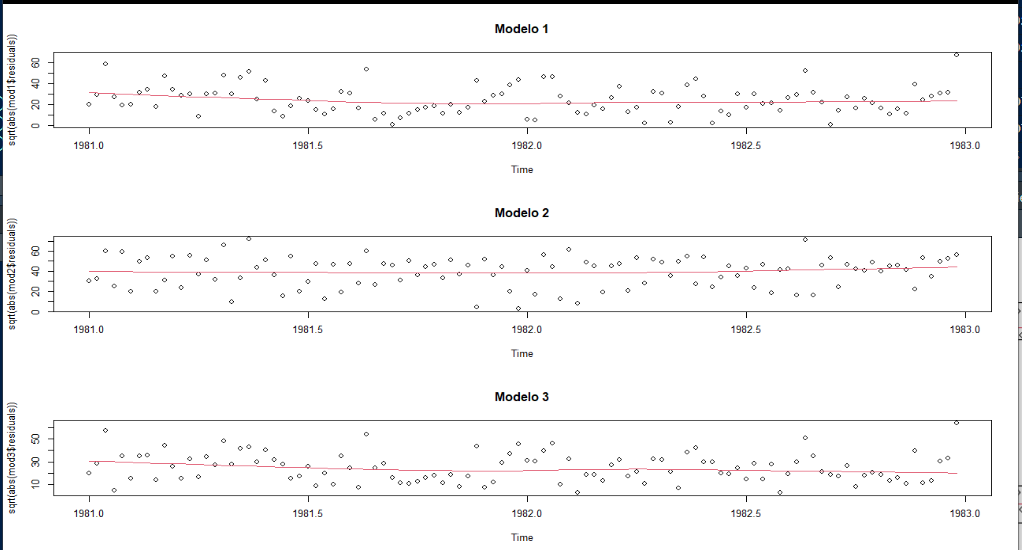
#### b) Homocedasticidad o varianza constante

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")



**INTERPRETACION:**

**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)))

obs=get(mod2$series)

bptest(resid(mod2)~I(obs-resid(mod2)))

obs=get(mod3$series)

bptest(resid(mod3)~I(obs-resid(mod3)))

studentized Breusch-Pagan test

data: resid(mod1) ~ I(obs - resid(mod1))

BP = 0.010868, df = 1, p-value = 0.917

studentized Breusch-Pagan test

data: resid(mod2) ~ I(obs - resid(mod2))

BP = 0.82445, df = 1, p-value = 0.3639

studentized Breusch-Pagan test

data: resid(mod3) ~ I(obs - resid(mod3))

BP = 0.22171, df = 1, p-value = 0.6377

**INTERPRETACION:**

**El valor de probabilidad (p-valor) asociado al estadístico BP asume un valor de 0.917 para el modelo 1, 0.3639 para el modelo 2 y 0.6377 para el modelo 3. Entonces el p-value de los 3 modelos es mayor a 0.05, con esto podemos decir que los residuales de los 3 modelos son constantes**

#### c. 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,

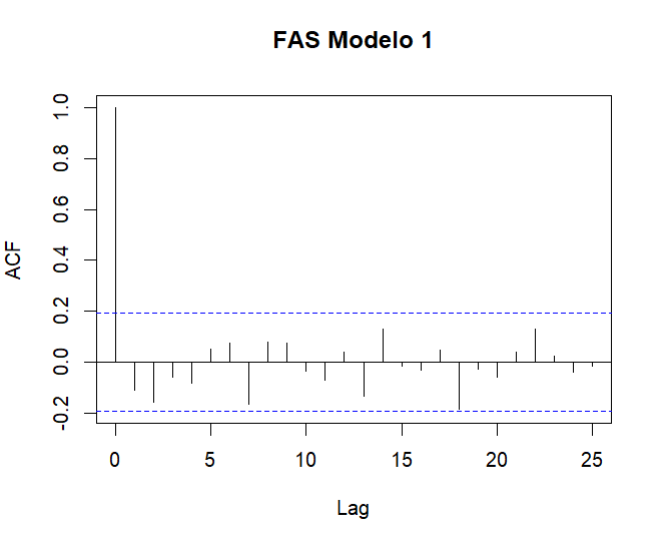
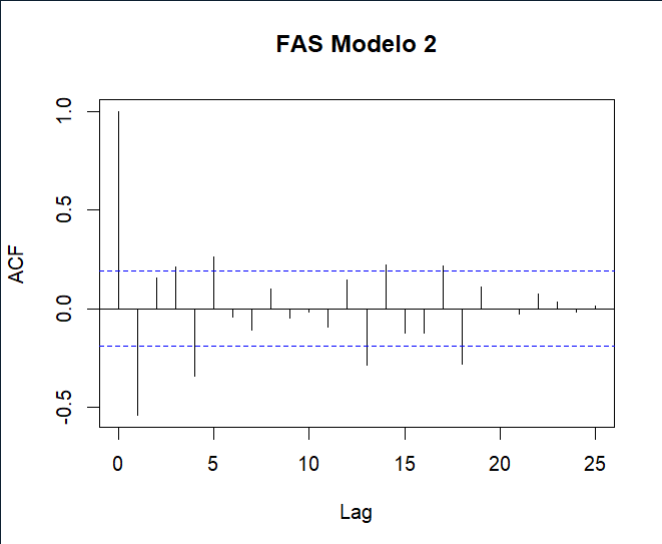
                main="FAS Modelo 1", level = 0.95)

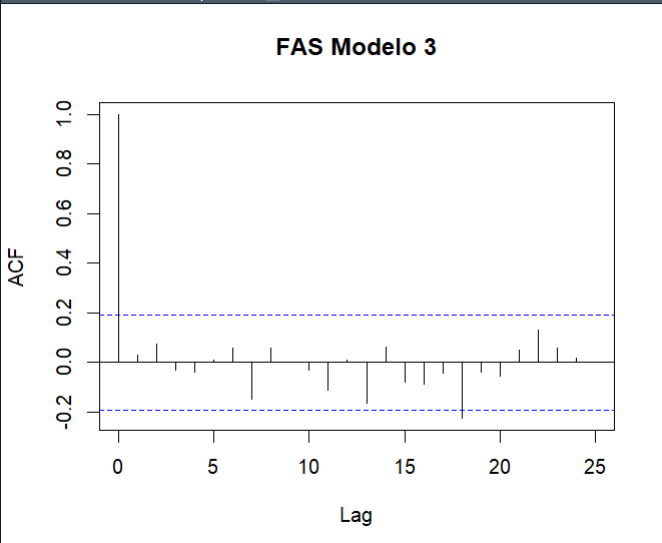
FAS\_e.m2 <- acf(resid\_m2, lag.max = 25,

                main="FAS Modelo 2", level = 0.95)

FAS\_e.m3 <- acf(resid\_m3, lag.max = 25,

                main="FAS Modelo 3", level = 0.95)





**INTERPRETACION:**

**Se observa que solo un coeficiente del FAS para los modelos 1 y 3 son significativos. Por el contrario el modelo 2 presenta muchos coeficientes que superan los limites de confianza indicando problemas de autocorrelación entre residuales**

**Para confirmar lo anterior usaremos la Prueba de Ljung - Box**

Box.test(resid\_m1,type = "Ljung-Box")

Box-Ljung test

data: resid\_m1

X-squared = 1.3104, df = 1, p-value = 0.2523

Box.test(resid\_m2,type = "Ljung-Box")

Box-Ljung test

data: resid\_m2

X-squared = 31.421, df = 1, p-value = 2.077e-08

Box.test(resid\_m3,type = "Ljung-Box")

Box-Ljung test

data: resid\_m3

X-squared = 0.080731, df = 1, p-value = 0.7763

**INTERPRETACION:**

**Confirmamos lo mencionado en los correlogramas, en los modelos 1 y 3 se acepta la hipótesis nula(p-valor > 0.05) de que los coeficientes de autocorrelación son cero; es decir, los residuos son independientes o están incorrelacionados. Y en el modelo 2 se rechaza la hipótesis nula (pvalor < 0.05) indicándonos que los residuos están correlacionados.**

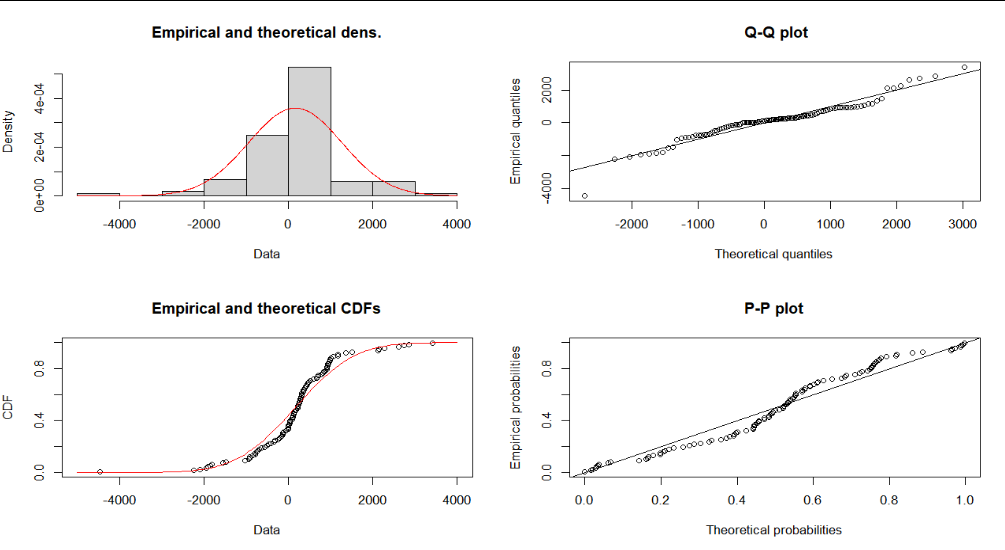
#### d) Contraste de normalidad

ajuste\_m1<-fitdist(data = resid\_m1, distr="norm")

plot(ajuste\_m1)

JB\_m1 <- jarque.bera.test(resid\_m1)

JB\_m1



Jarque Bera Test

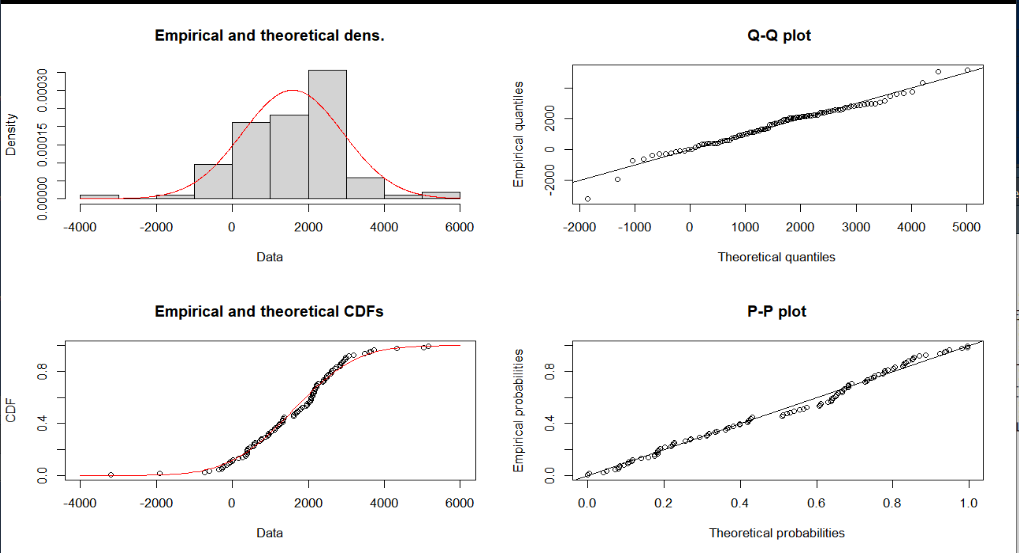
data: resid\_m1

X-squared = 40.1, df = 2, p-value = 1.96e-09

**INTERPRETACION:**

**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 = 0.000000 < 0.05, se rechaza Ho, es decir, los residuos NO se aproximan a una distribucion normal.**

****

Jarque Bera Test

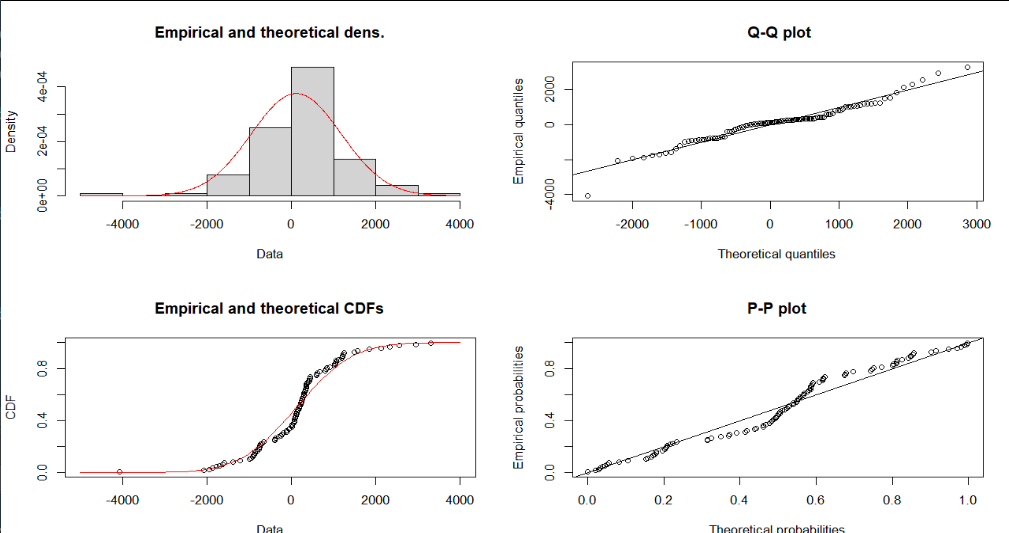
data: resid\_m2

X-squared = 6.7366, df = 2, p-value = 0.03445

**INTERPRETACION:**

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

**En la prueba JB, como p = 0.03445 < 0.05, se rechaza Ho, es decir, los residuos NO se aproximan a una distribucion normal.**

****

Jarque Bera Test

data: resid\_m3

X-squared = 22.824, df = 2, p-value = 1.106e-05

**INTERPRETACION:**

**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 = 0.000001 < 0.05, se rechaza Ho, es decir, los residuos NO se aproximan a una distribución normal.**

## CONCLUSION: ELECCION DEL MEJOR MODELO

**CONCLUIMOS QUE EL MEJOR MODELO ES EL AR(1) POR LO SIGUIENTE:**

* **SUS COEFICIENTES SON SIGNIFICATIVOS**
* **LA MEDIA DE LOS RESIDUALES TIENDE A CERO**
* **CUMPLE CON LA CONDICION DE ESTACIONALIDAD**
* **NO TIENE MULTICOLINEALIDAD**
* **LOS RESIDUALES SON CONSTANTES**

# CASO 2: PRODUCCIÓN DE MAIZ

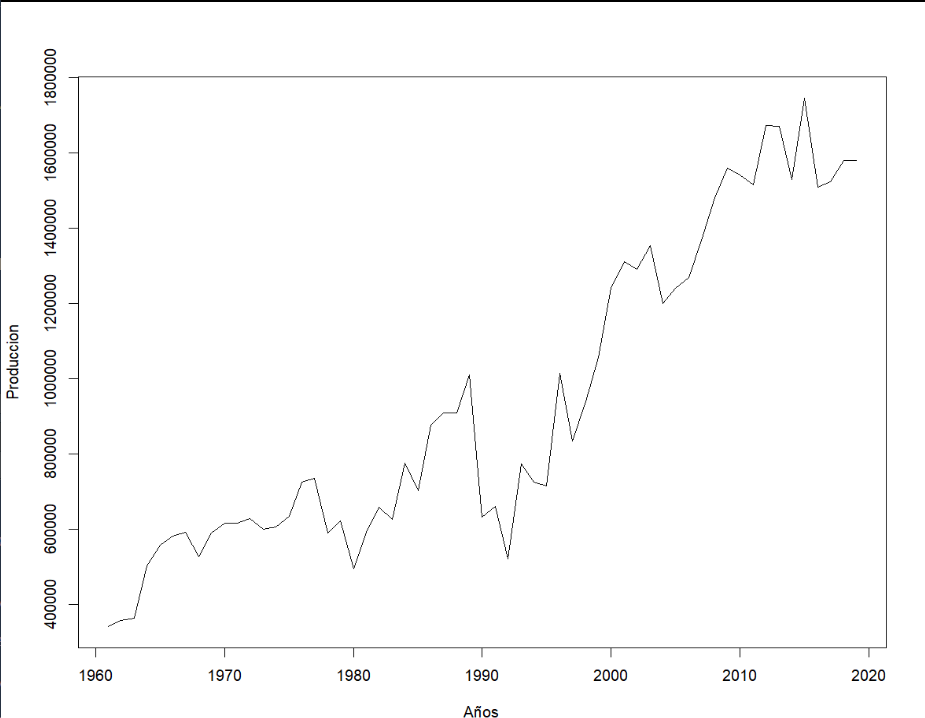
data <- read\_excel("F:\\777--Programacion repos\\Una\\r\\data\\actividad-04.xlsx",sheet = "datos2")

View(data)

# Gráfica de la serie

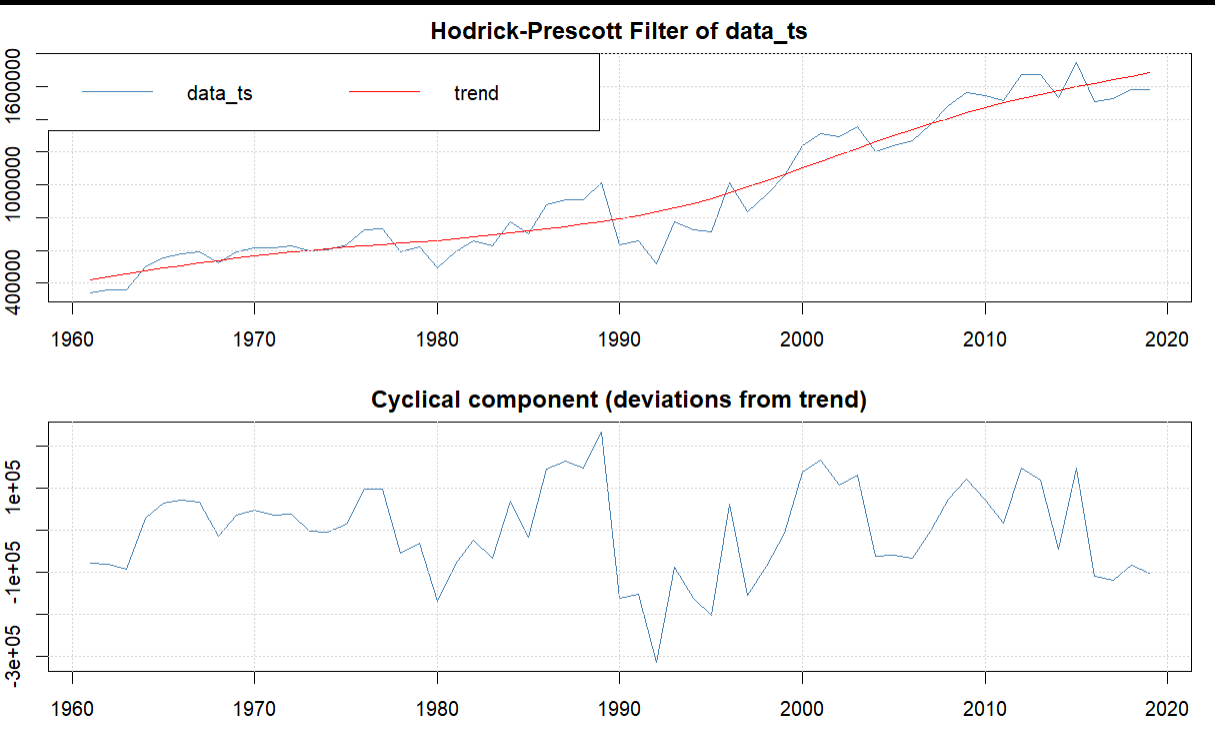
data\_ts <- ts(data$Produccion, start = c(1961,1), frequency = 1)

plot(data\_ts, xlab="Años", ylab="Produccion")



## 1.IDENTIFICACION

### Análisis de la tendencia y estacionalidad

****

#### Conclusión estacionalidad y tendencia:

**INTERPRETACION:**

**Se observa que la serie si presenta una tendencia creciente, y no tiene estacionalidad dado que la serie es anual y tiene periodicidad menor a un año**

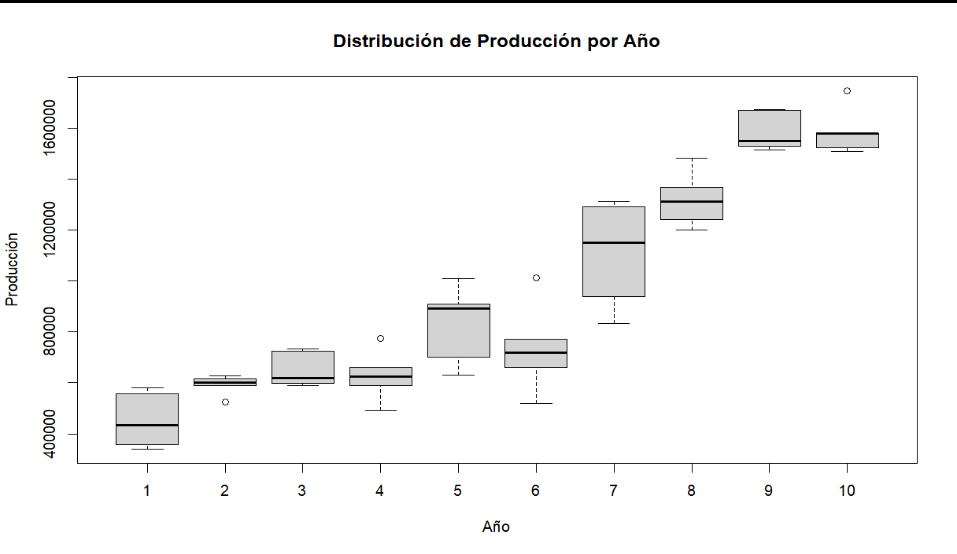
### Análisis de la estacionariedad

#### Estacionariedad en varianza

Grupo <- rep(1:9, each = 6)

Grupo <- c(Grupo, rep(x = 10,5))

boxplot(data$Produccion ~ Grupo, xlab = "Año", ylab = "Producción", main = "Distribución de Producción por Año")

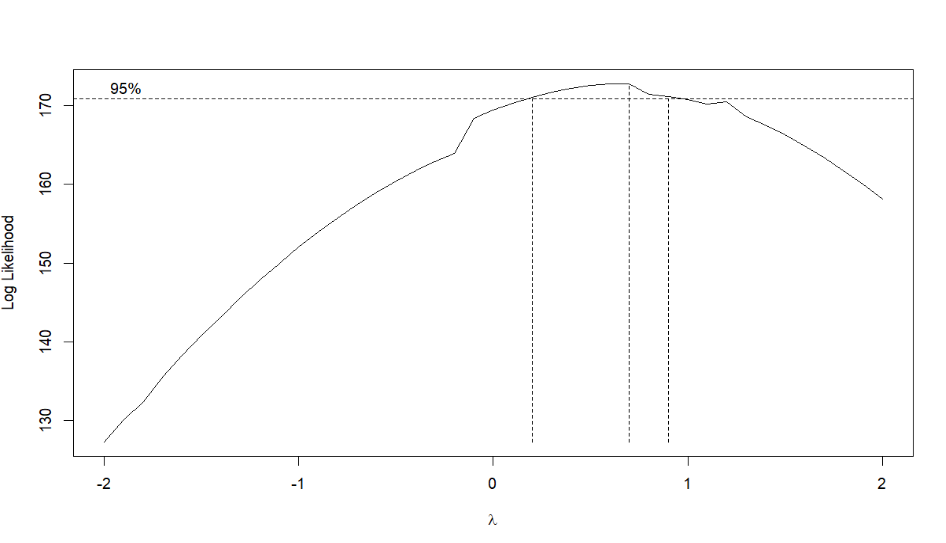


**INTERPRETACION:**

**Dado el grafico podemos identificar que la serie no parece tener una varianza constante, en otras palabras no aparenta ser estacionaria en varianza.**

**VALOR OPTIMO PARA LAMDA**

b = BoxCox.ar(data\_ts)



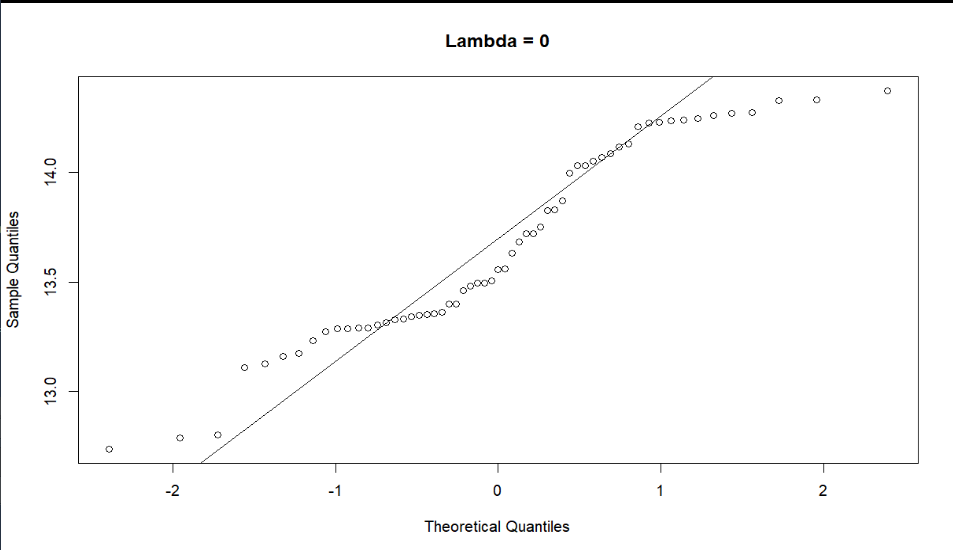
lambda <- b$mle

round(lambda,0)

**Observamos que el valor optimo para lamda es 0.7 que se acerca mas a 1, por lo que lambda seria 1. Entonces no hacemos una transformación, ya que sigue siendo la serie original.**

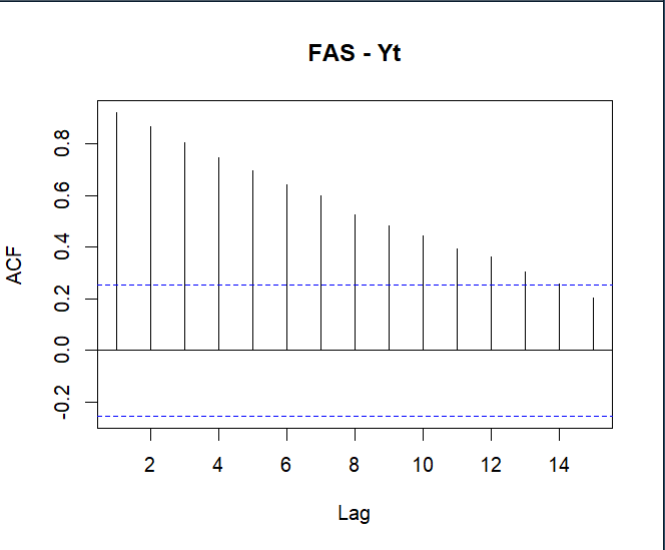
qqnorm(data\_ts, main = "Lambda = 1") # Yt: Original

qqline(data\_ts)

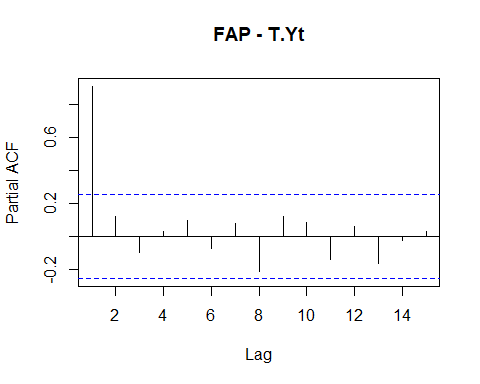
****

#### Estacionariedad en media

FAS <- acf(data\_ts, lag.max = 15, main="FAS - Yt", level = 0.95)



FAP <- pacf(data\_ts, lag.max = 15, main="FAP - Yt", level = 0.95)



FAP[1]

##   
## Partial autocorrelations of series 'T.Yt', by lag  
##   
## 1   
## 0.921

**INTERPRETACION:**

**Observamos que el comportamiento de las FAS es decreciente lentamente lo que nos indica que no hay estacionariedad de la serie. Y vemos que el primer coeficiente del FAP es significativo siendo 0.921 mayor a 0.9. Por lo que la serie no es estacionaria.**

# 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

-360220 -53502 8013 77557 200586

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 7.821e+04 3.917e+04 1.997 0.0510 .

z.lag.1 -2.019e-01 9.739e-02 -2.074 0.0430 \*

tt 4.604e+03 2.333e+03 1.974 0.0536 .

z.diff.lag -2.523e-01 1.328e-01 -1.899 0.0630 .

---

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

Residual standard error: 108100 on 53 degrees of freedom

Multiple R-squared: 0.1903, Adjusted R-squared: 0.1445

F-statistic: 4.153 on 3 and 53 DF, p-value: 0.01024

Value of test-statistic is: -2.0737 2.7715 2.1589

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

**INTERPRETACION:**

**Con un T calculado de -2.0737 > T critico de -3.45. Concluimos que se rechaza la hipótesis nula de la existencia de raíz unitaria, es decir la serie NO es estacionaria en media y varianza**

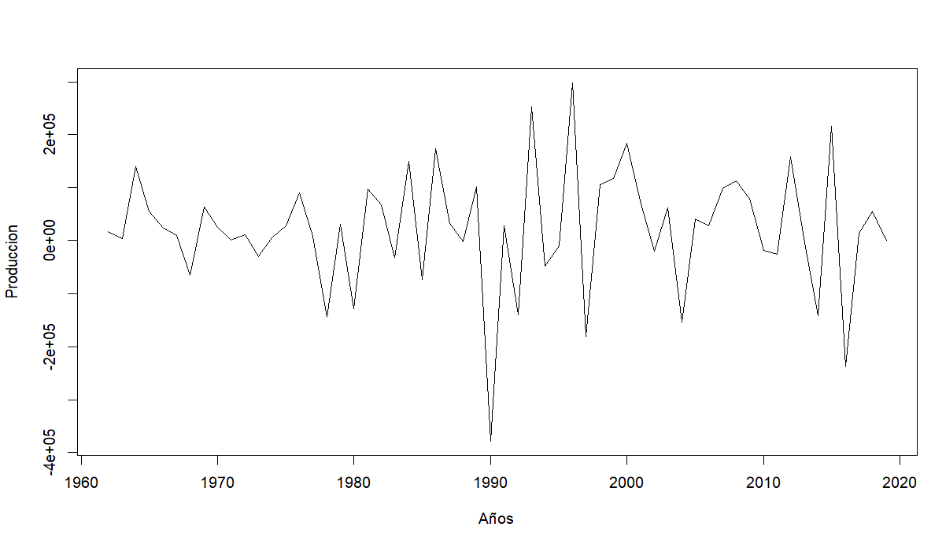
#### Conclusión de estacionariedad en media y varianza

**Dado lo anterior concluimos que según los métodos graficos y la prueba de raíz unitaria de dickey Fuller que NO HAY ESTACIONARIEDAD EN VARIANZA Y MEDIA**

### Diferenciación

data\_diff <- diff(data\_ts)

plot(data\_diff, xlab = "Años", ylab="Produccion")



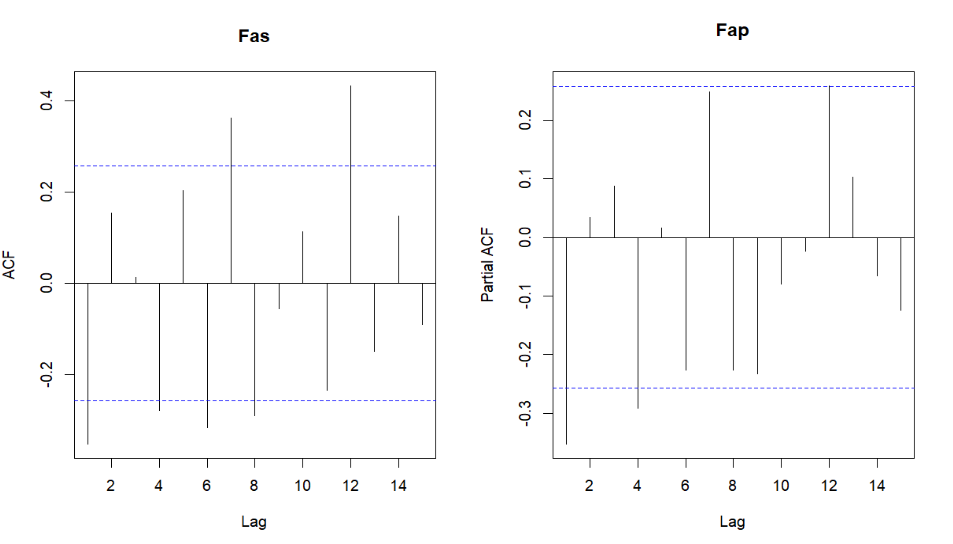
**Después de realizar la primera diferencia la serie aparenta ser estacionaria**

#Correlograma de la 1ra diferencia

par(mfrow =c(1,2))

FAS <- acf(data\_diff, lag.max=15, main="Fas", level = 0.95)

FAP <- pacf(data\_diff, lag.max=15, main="Fap", level =0.95)



*INTERPRETACION:*

**Según el FAS se observa que disminuye rápidamente a 0, y en el FAP vemos que el primer coeficiente es menor a 0.9. Esto nos da a entender que la serie es ESTACIONARIA**

**#Dickey fuller**

data\_adf <- ur.df(data\_diff, type="drift", lags = 1)

summary(data\_adf)

Value of test-statistic is: -5.7842 16.7287

Critical values for test statistics:

1pct 5pct 10pct

tau2 -3.51 -2.89 -2.58

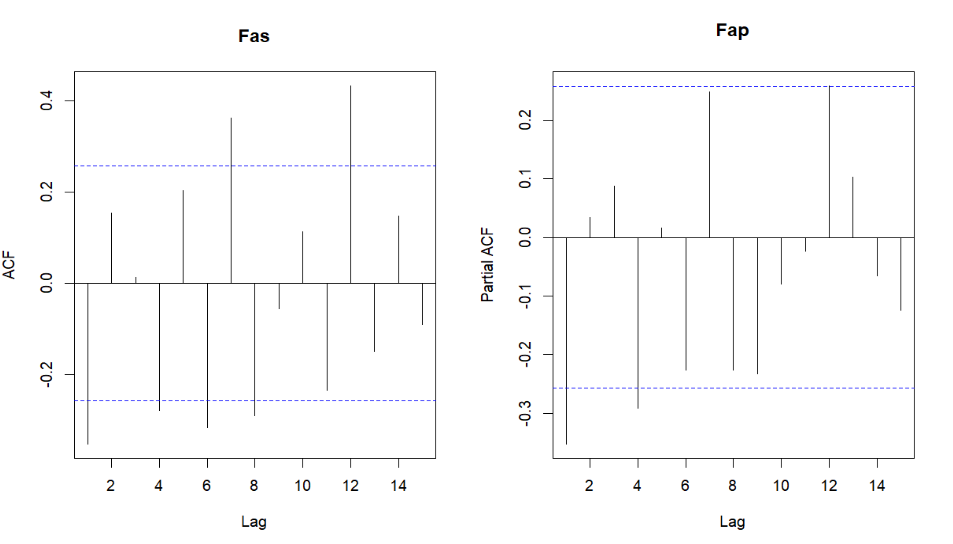
phi1 6.70 4.71 3.86

**Vemos que T calculado -5.7842 < T critico -2.89. Por lo que haceptamos Ho, ósea que si hay raíz unitaria. Lo que nos indica que si hay ESTACIONARIEDAD EN MEDIA Y VARIANZA.**

#### Conclusión de la diferenciación

**Dado el análisis anterior usando las autocorrelaciones y la prueba de raíz unitaria se observa que la serie es estacionaria después de aplicar la primera diferencia.**

### a. Identificación de las órdenes p y q



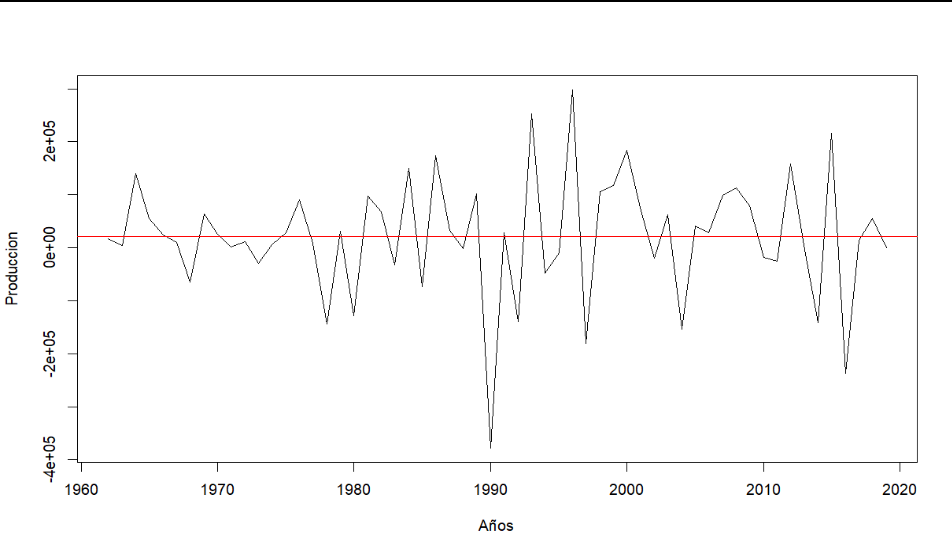
**INTERPRETACION:**

**Según las FAS y FAP podemos ver que el comportamiento de estas podrían sugerirnos un modelo ARMA(1,2) ya que no muestran un patrón fijo y van seguidos de una mezcla oscilaciones sinusoidales**

### b. Inclusión del término independiente o intercepto

plot(data\_diff, xlab="Años", ylab="Produccion")

abline(h = mean(data\_diff), col = "red")



Z <- mean(data\_diff)

Co <- var(data\_diff)

Tn <- length(data\_ts)

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

9.242922e-05 2.002465e+00

#### Conclusión de la inclusión del intercepto

**El primer grafico nos muestra que la serie estacionaria parece oscilar en torno a una media de 0. Para verificar esto se hizo la prueba estadistica donde el T calculado < T crítico, entonces aceptamos la hipótesis nula. Finalmente concluimos que el intercepto no debe ser incluido en el modelo.**

#### Conclusión: Identificación de los modelos

**Resumiendo, se proponen los siguientes modelos para la serie estacionaria:**

## 2.ESTIMACION

#ARIMA(1,1,0)

mod1 <- Arima(data\_ts, order = c(1,1,0))

coeftest(mod1)

z test of coefficients:

Estimate Std. Error z value Pr(>|z|)

ar1 -0.30209 0.12375 -2.441 0.01464 \*

---

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

**Dado la significancia de los coeficientes del modelo AR(1) podemos ver que es significativo para el modelo.**

#ARIMA(0,1,2)

mod2 <- Arima(data\_ts, order = c(0,1,2))

coeftest(mod2)

z test of coefficients:

Estimate Std. Error z value Pr(>|z|)

ma1 -0.24076 0.11944 -2.0158 0.04382 \*

ma2 0.27947 0.17537 1.5936 0.11102

---

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

#ARIMA(1,1,2)

mod3 <- Arima(data\_ts, order = c(1,1,2))

coeftest(mod3)

**Dado la significancia de los coeficientes del modelo MA(2) podemos ver que solo MA1 es significativo para el modelo.**

z test of coefficients:

Estimate Std. Error z value Pr(>|z|)

ar1 -0.13998 0.32033 -0.4370 0.6621

ma1 -0.12513 0.29434 -0.4251 0.6708

ma2 0.25508 0.18971 1.3446 0.1788

*INTERPRETACION:*

**Dado la significancia de los coeficientes del modelo ARMA(1,2) podemos ver que ningún coeficiente es significativo para el modelo.**

## 3. VALIDACION

### 3.1. Análisis de los coeficientes estimados

#### a. Significación de los coeficientes

**De los coeficientes anteriores analizamos su significancia**

**Modelo 1: AR(1)**

Modelo 2: MA(2)

Modelo 3: ARMA(1,2)

#### b. Examen de la matriz de correlaciones entre los coeficientes

vcov(mod1)

ar1

ar1 0.01531465

vcov(mod2)

ma1 ma2

ma1 0.014264983 0.002097739

ma2 0.002097739 0.030752962

vcov(mod3)

ar1 ma1 ma2

ar1 0.10261333 -0.08610588 0.02921369

ma1 -0.08610588 0.08663420 -0.02540416

ma2 0.02921369 -0.02540416 0.03598994

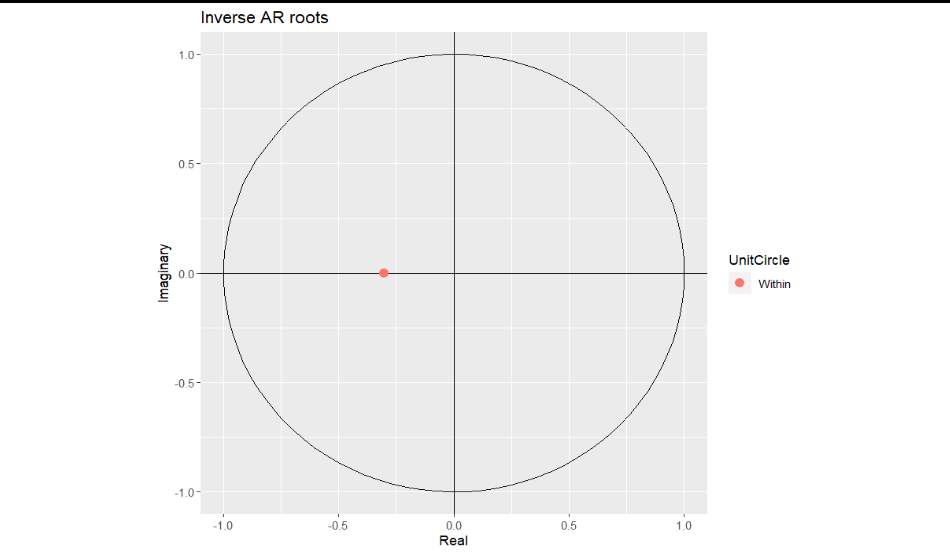
**Se observa claramente que ningún coeficiente esta próximo ni cercano a 0.9, por tanto, podemos indicar que no hay problema de multicolinealidad en los modelos propuestos.**

#### c. Condición de Convergencia e Invertibilidad

autoplot(mod1)

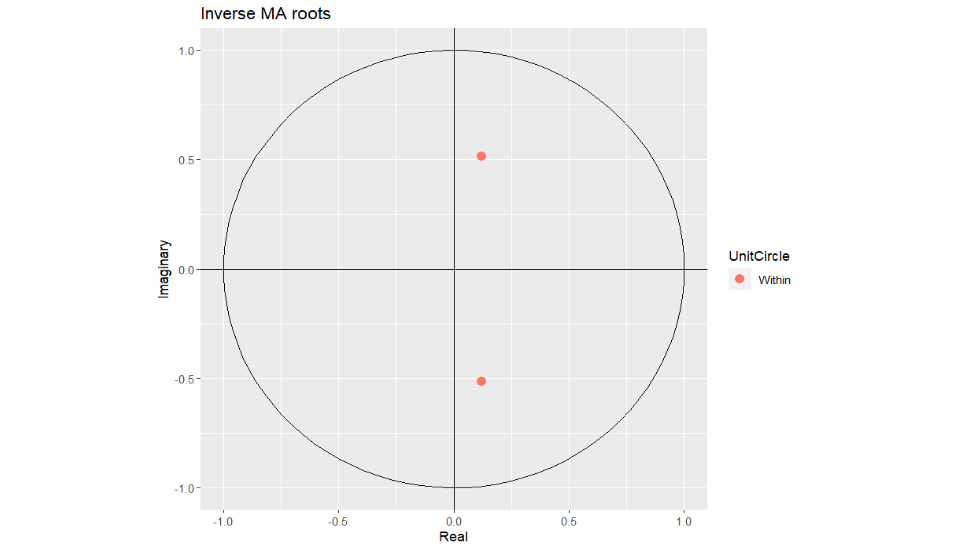
autoplot(mod2)

autoplot(mod3)



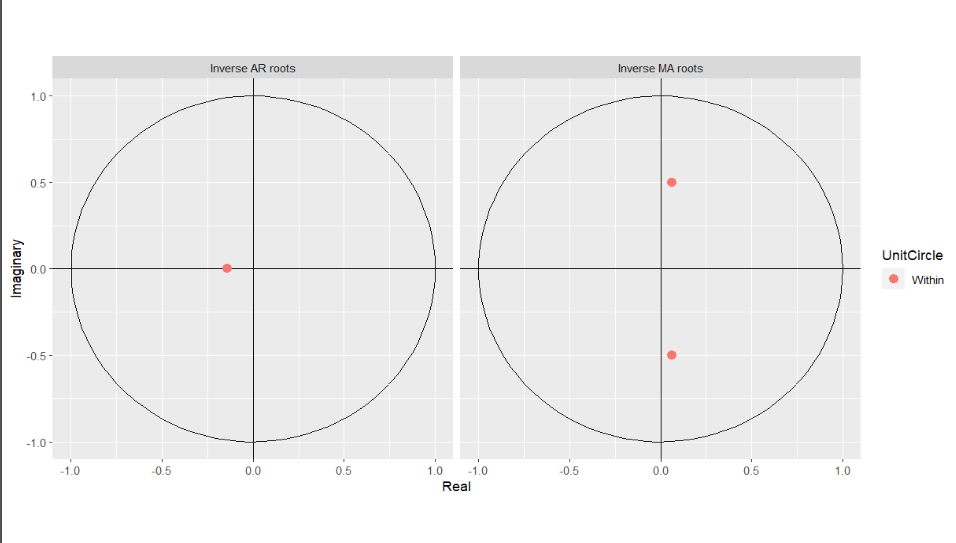
**INTERPRETACION:**

**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.**



**INTERPRETACION:**

**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.**



**INTERPRETACION:**

**En la última figura 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.**

#### d. Análisis de la estabilidad

Chow\_mod1 <- Fstats(mod1$fitted ~ 1, from = 0.40)

sctest(Chow\_mod1)

Chow\_mod2 <- Fstats(mod2$fitted ~ 1, from = 0.40)

sctest(Chow\_mod2)

Chow\_mod3 <- Fstats(mod3$fitted ~ 1, from = 0.40)

sctest(Chow\_mod3)

supF test

data: Chow\_mod1

sup.F = 2.2185, p-value = 0.3873

supF test

data: Chow\_mod2

sup.F = 3.8142, p-value = 0.1754

supF test

data: Chow\_mod3

sup.F = 1.7002, p-value = 0.5028

**En las tres pruebas se acepta la hipótesis nula (p > 0.05), es decir,SI existe estabilidad de coeficientes.**

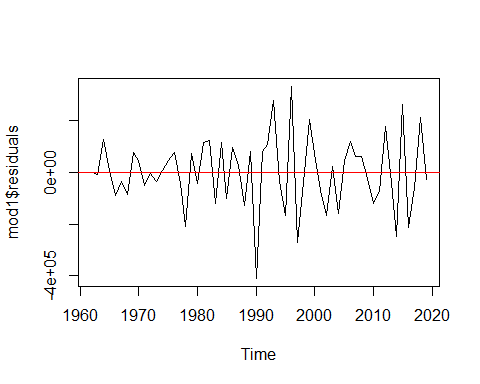
### 3.2 Análisis de residuos

**Ho: E(at) = 0   
Hi: E(at) != 0**  **#si el p value es mayor que la significancia aceptamos ho**

plot(mod1$residuals)

abline(h = 0, col = "red")

t.test(mod1$residuals, mu = 0)



One Sample t-test

data: mod1$residuals

t = 1.9513, df = 58**, p-value = 0.05586**

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

-707.3088 55405.9293

sample estimates:

mean of x

27349.31

**INTERPRETACION:**

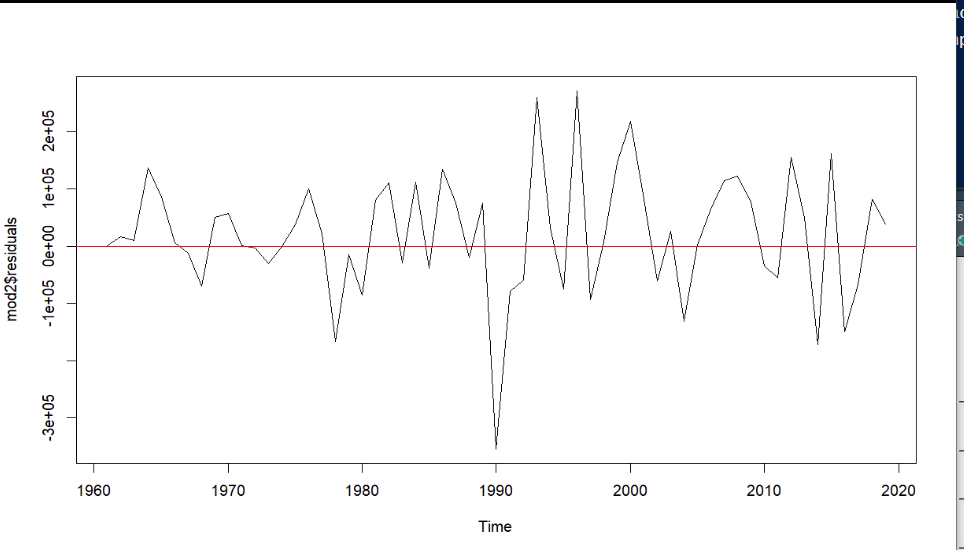
**Según el grafico vemos que un buen numero de residuales están en torno a la media igual a cero.**

**Y confirmando lo anterior p=0.005586 > 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 = 1.4724, df = 58, p-value = 0.1463

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

-7411.469 48648.560

sample estimates:

mean of x

20618.55

**INTERPRETACION:**

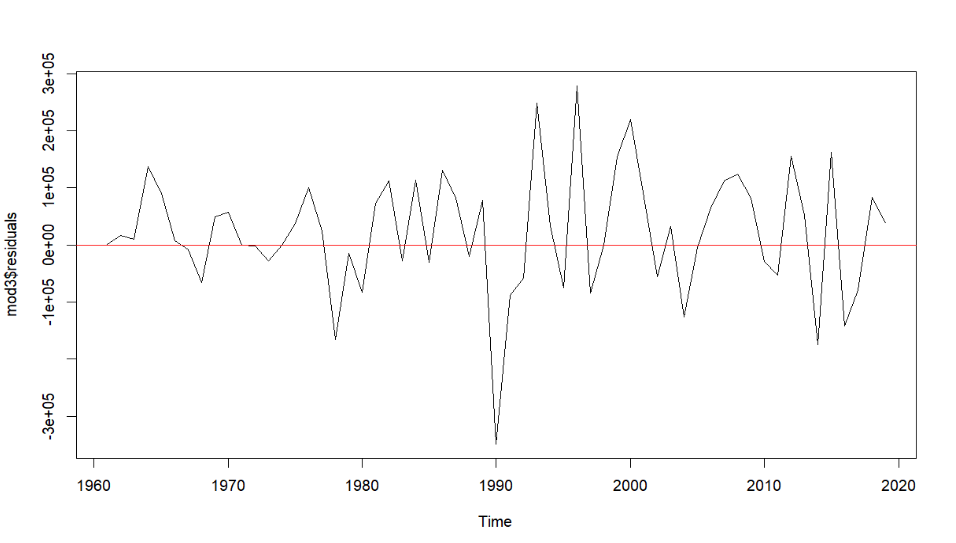
**Segun el grafico vemos que un buen numero de residuales están en torno a la media igual a cero.**

**Y confirmando lo anterior p=0.1463 > 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)



One Sample t-test

data: mod3$residuals

t = 1.5475, df = 58, p-value = 0.1272

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

-6336.359 49516.763

sample estimates:

mean of x

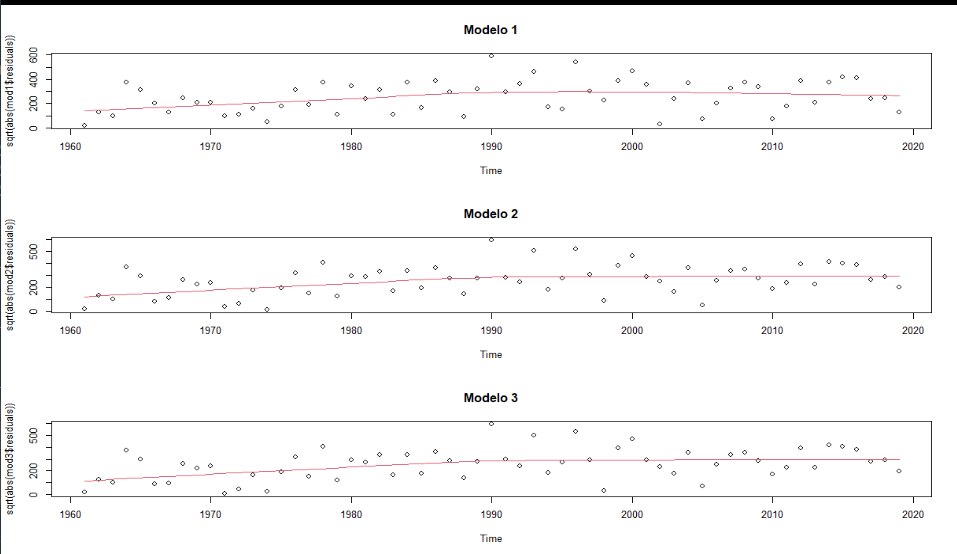
21590.2

**INTERPRETACION:**

**Segun el grafico vemos que un buen numero de residuales están en torno a la media igual a cero.**

**Y confirmando lo anterior p=0.1272 > 0.05, se acepta Ho, es decir la media es igual a cero**

#### Homocedasticidad de los residuos



***INTERPRETACION****:*

**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 = 0.57678, df = 1, p-value = 0.4476

obs=get(mod2$series)

bptest(resid(mod2)~I(obs-resid(mod2)))

studentized Breusch-Pagan test

data: resid(mod2) ~ I(obs - resid(mod2))

BP = 0.52042, df = 1, p-value = 0.4707

obs=get(mod3$series)

bptest(resid(mod3)~I(obs-resid(mod3)))

studentized Breusch-Pagan test

data: resid(mod3) ~ I(obs - resid(mod3))

BP = 0.57112, df = 1, p-value = 0.4498

**INTERPRETACION**:

**El valor de probabilidad (p-valor) asociado al estadístico BP asume un valor de 0.4476 para el modelo 1, 0.4707 para el modelo 2 y 0.4498 para el modelo 3. Entonces el p-value de los 3 modelos es mayor a 0.05, con esto podemos decir que los residuales de los 3 modelos son constantes**

#### Correlograma de los residuos

resid\_m1 <- as.vector(mod1$residuals)

resid\_m2 <- as.vector(mod2$residuals)

resid\_m3 <- as.vector(mod3$residuals)

par(mfrow = c(1,3))

FAS\_e.m1 <- acf(resid\_m1, lag.max = 25,

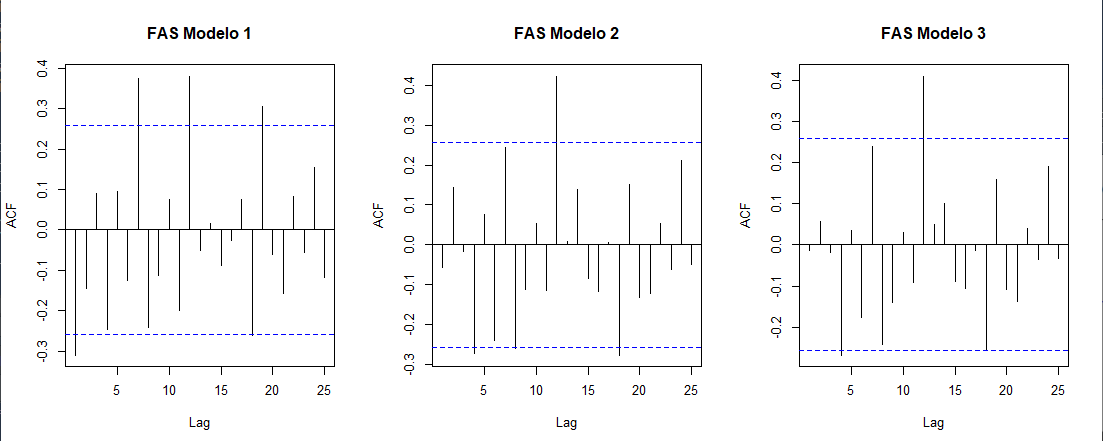
                main="FAS Modelo 1", level = 0.95)

FAS\_e.m2 <- acf(resid\_m2, lag.max = 25,

                main="FAS Modelo 2", level = 0.95)

FAS\_e.m3 <- acf(resid\_m3, lag.max = 25,

                main="FAS Modelo 3", level = 0.95)



**INTERPRETACION**:

**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, sobre todo los de los primeros retardos. Pero en el modelo 1 se ve varios coeficientes significativos indicando problemas de autocorrelación.**

**Prueba de Ljung – Box**

Box.test(resid\_m1,type = "Ljung-Box")

Box.test(resid\_m2,type = "Ljung-Box")

Box.test(resid\_m3,type = "Ljung-Box")

Box-Ljung test

data: resid\_m1

X-squared = 5.8325, df = 1, p-value = 0.01573

Box-Ljung test

data: resid\_m2

X-squared = 0.20182, df = 1, p-value = 0.6533

Box-Ljung test

data: resid\_m3

X-squared = 0.012917, df = 1, p-value = 0.9095

**INTERPRETACION**:

**Confirmamos lo mencionado en los correlogramas, en los modelos 1 y 3 se acepta la hipótesis nula(p-valor > 0.05) de que los coeficientes de autocorrelación son cero; es decir, los residuos son independientes o están incorrelacionados.**

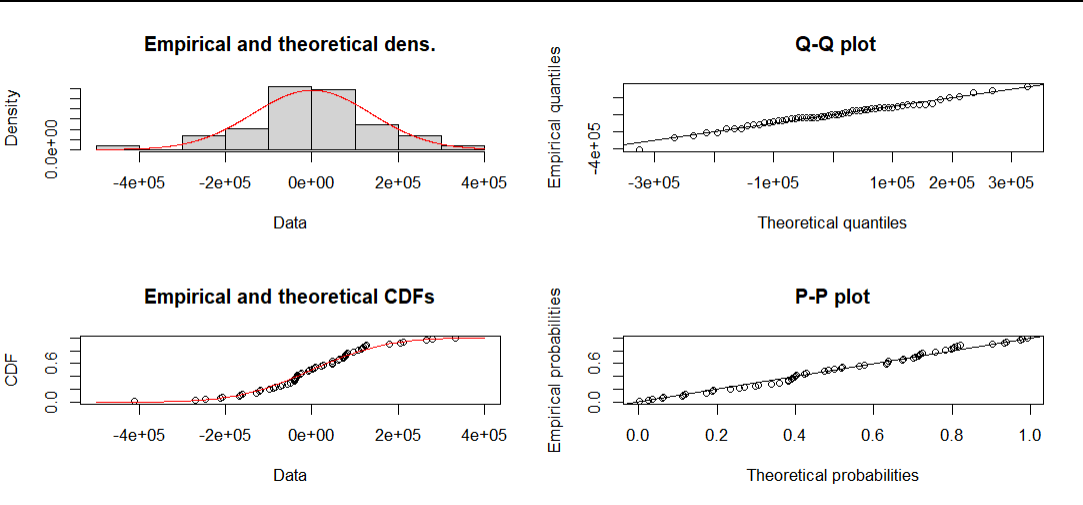
**Y en el modelo 1 se rechaza la hipótesis nula indicando que los residuos están correlacionados**  
d. Contraste de normalidad

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 = 1.2796, df = 2, p-value = 0.5274

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

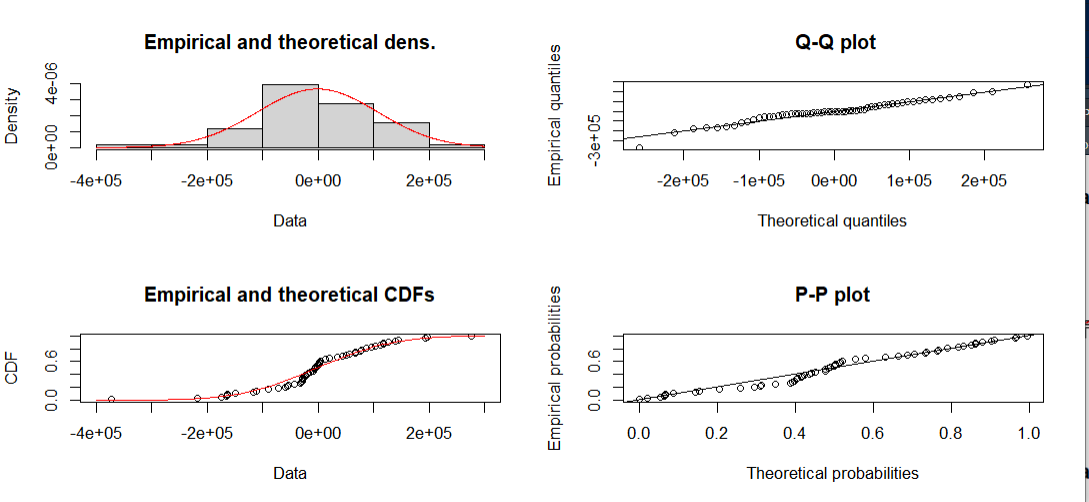
***En la prueba JB, como p = 0.5274 > 0.05, se acepta Ho, es decir, los residuos se aproximan a una distribución normal.***

ajuste\_m2<-fitdist(data = resid\_m2, distr="norm")

plot(ajuste\_m2)

JB\_m2 <- jarque.bera.test(resid\_m2)

JB\_m2



Jarque Bera Test

data: resid\_m2

X-squared = 7.3252, df = 2, p-value = 0.02567

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

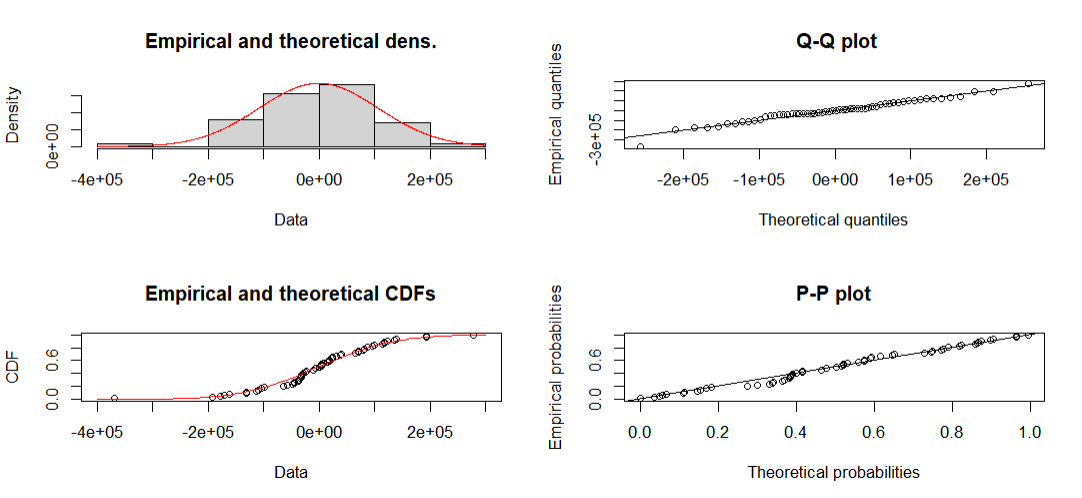
***En la prueba JB, como p = 0.0257 < 0.05, se rechaza Ho, es decir, los residuos NO se aproximan a una distribución normal.***

ajuste\_m3<-fitdist(data = resid\_m3, distr="norm")

plot(ajuste\_m3)

JB\_m3 <- jarque.bera.test(resid\_m3)

JB\_m3



***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 = 0.4199 < 0.05, se rechaza Ho, es decir, los residuos NO se aproximan a una distribución normal.***

## CONCLUSION:

**El modelo con mejores métricas fue el fue el AR(1) por las siguientes razones:**

* **Los residuos se aproximan a una distribución normal**
* **Los residuos son independientes**
* **Cumple con la condición de estacionalidad**
* **No tiene multicolinealidad**
* **Los coeficientes son estables**