**Tabla de contenido**

[CASO 1: TEMPERATURAS, Dubuque – EE.UU. 3](#_Toc171981327)

[1 Identificación 3](#_Toc171981328)

[1. 1 Análisis de la tendencia y la estacionalidad 5](#_Toc171981329)

[1.1.1 Estacionalidad 5](#_Toc171981330)

[1.1.2 Análisis de tendencia 5](#_Toc171981331)

[1.2 Análisis de estacionariedad 6](#_Toc171981332)

[1.2.1 Estacionariedad en varianza 6](#_Toc171981333)

[1.2.2 Estacionariedad en media 7](#_Toc171981334)

[1.3 Identificación del modelo estacionario 11](#_Toc171981335)

[1.3.1 Identificación de las órdenes p y q 11](#_Toc171981336)

[2 Estimación 11](#_Toc171981337)

[3 Validación 12](#_Toc171981338)

[3.1 Análisis de los coeficientes estimados 12](#_Toc171981339)

[3.1.1 Significación de los coeficientes 12](#_Toc171981340)

[3.1.2 Examen de la matriz de correlaciones entre los coeficientes 12](#_Toc171981341)

[3.1.3 Condición de convergencia e invertibilidad 13](#_Toc171981342)

[3.2 Análisis de los residuos 14](#_Toc171981343)

[3.2.1 PRUEBA DE LJUNG-BOX 14](#_Toc171981344)

[3.2.2 Homocedasticidad o varianza constante 16](#_Toc171981345)

[3.2.3 Contraste de normalidad 17](#_Toc171981346)

[4 Pronostico 19](#_Toc171981347)

[4.1Pronosticos de cada modelo 19](#_Toc171981348)

[Modelo 1: 19](#_Toc171981349)

[Modelo 2: 20](#_Toc171981350)

[Modelo 3: 21](#_Toc171981351)

[SERIE ORIGINAL (YT) Y PRONOSTICADA. 22](#_Toc171981352)

[Gráfica del ajuste y pronóstico con valores reales 24](#_Toc171981353)

[Métricas basadas en el error 26](#_Toc171981354)

[AIC Y BIC 26](#_Toc171981355)

[Conclusión: Elección del Modelo SARIMA 26](#_Toc171981356)

[Modelos Evaluados: 26](#_Toc171981357)

[Elección del Mejor Modelo: 27](#_Toc171981358)

[CASO 2: VENTAS DE LIBRERÍA UNIVERSITARIA 28](#_Toc171981359)

[1 Identificación 29](#_Toc171981360)

[1. 1 Análisis de la tendencia y la estacionalidad 29](#_Toc171981361)

[1.1.1 Estacionalidad 29](#_Toc171981362)

[1.1.2 Análisis de tendencia 30](#_Toc171981363)

[1.2 Análisis de estacionariedad 31](#_Toc171981364)

[1.2.1 Estacionariedad en varianza 31](#_Toc171981365)

[1.2.2 Estacionariedad en media 32](#_Toc171981366)

[1.3 Identificación del modelo estacionario 36](#_Toc171981367)

[1.3.1 Identificación de las órdenes p y q 36](#_Toc171981368)

[2 Estimación 37](#_Toc171981369)

[3 Validación 38](#_Toc171981370)

[3.1 Análisis de los coeficientes estimados 38](#_Toc171981371)

[3.1.1 Significación de los coeficientes 38](#_Toc171981372)

[3.1.2 Examen de la matriz de correlaciones entre los coeficientes 39](#_Toc171981373)

[3.1.3 Condición de convergencia e invertibilidad 39](#_Toc171981374)

[3.2 Análisis de los residuos 40](#_Toc171981375)

[3.2.1 Media es igual a cero 40](#_Toc171981376)

[3.2.2 Homocedasticidad o varianza constante 43](#_Toc171981377)

[3.2.4 Contraste de normalidad 44](#_Toc171981378)

[4 Pronostico 47](#_Toc171981379)

[4.1Pronosticos de cada modelo 47](#_Toc171981380)

[Modelo 1: 47](#_Toc171981381)

[Modelo 2: 48](#_Toc171981382)

[Modelo 3: 49](#_Toc171981383)

[Métricas del modelo 51](#_Toc171981384)

[Conclusión 52](#_Toc171981385)

**Actividad 6**

# Librerias necesaria

library(forecast)# mdelos ARIMA

library(tseries) # series de tiempo

library(TSA)

library(urca) # prueba de raises unitarias

library(ggplot2) # graficos

library(dplyr)

library(lmtest) #Inferencia para los coeficientes

library(MASS) #Tranformacion de box.cox

library(nortest) #pruebas de normalidad

library(mFilter) #filtro de jodric prescot (p-h)

library(strucchange) # analisis de estabilidad http://127.0.0.1:29089/graphics/plot\_zoom\_png?width=941&height=778

library(fitdistrplus)

library(readxl)

library(TSstudio)

# CASO 1: TEMPERATURAS, Dubuque – EE.UU.

## 1 Identificación

data <- read\_excel("F:\\777--Programacion repos\\Una\\r\\data\\actividad-07.xlsx",sheet = "s01")

View(data)

# 1) IDENTIFICACION

# Graficar la serie

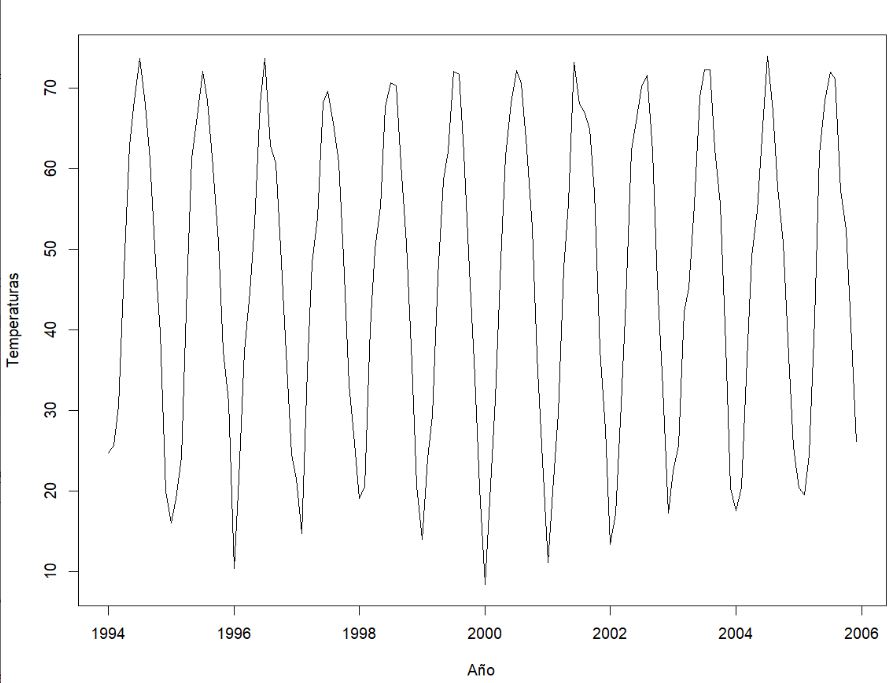
Yt <- ts(data$Temperatura, start = c(1994, 1), frequency = 12)

plot(Yt, xlab = "Año", ylab ="Temperaturas" )

# Descomposicion de la serie

Yt\_desc <- decompose(Yt, type = "additive")

plot(Yt\_desc, xlab = "Año/Meses")

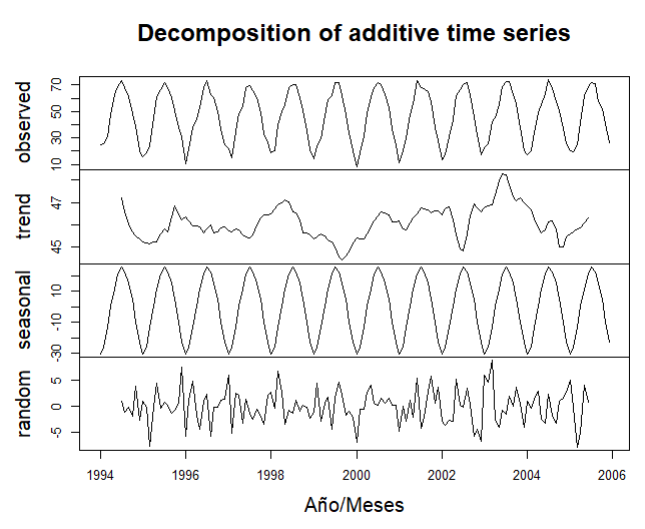


Podemos observar que la serie presenta un claro componente estacional, con aparente estacionariedad en la varianza y la media. En cuanto al componente cíclico, hay algunos indicios, pero no son definitivos. Finalmente, se destaca la ausencia total de una tendencia.

DESCOMPOSICION

Yt\_desc <- decompose(Yt, type = "additive")

plot(Yt\_desc, xlab = "Año/Meses")



Después de descomponer la serie, podemos confirmar algunos de los comportamientos descritos inicialmente.

### 1. 1 Análisis de la tendencia y la estacionalidad

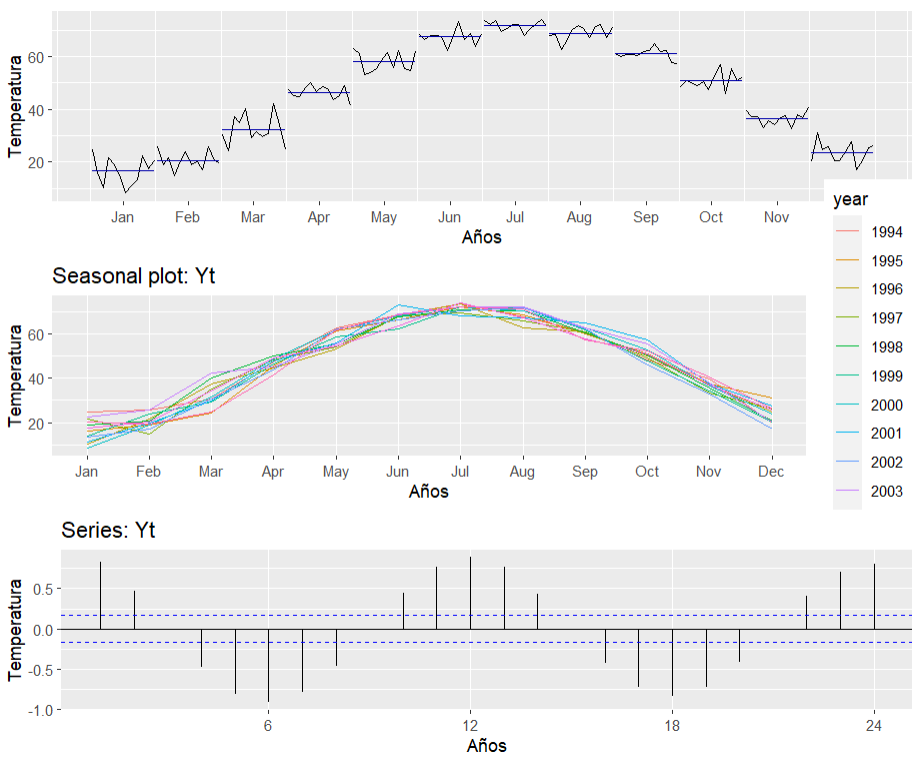
#### 1.1.1 Estacionalidad

plot1 <- ggsubseriesplot(Yt, xlab = "Años", ylab = "Temperatura")

plot2 <- ggseasonplot(Yt, xlab = "Años", ylab = "Temperatura")

plot3 <- ggAcf(Yt, xlab = "Años", ylab = "Temperatura")

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



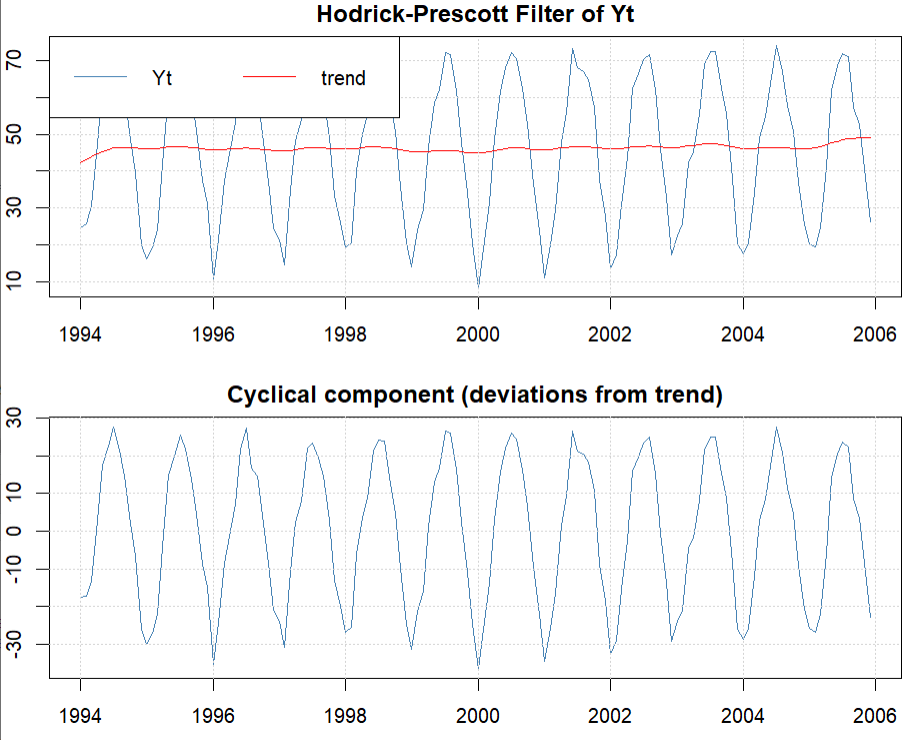
En el gráfico de líneas apiladas y en el de líneas separadas, se observa un claro comportamiento estacional. Esto se corrobora en el correlograma, cuyo comportamiento refleja sinusoidal característico de una serie estacional.

#### 1.1.2 Análisis de tendencia

lambda\_hp <- 1000

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

plot(data\_hp)

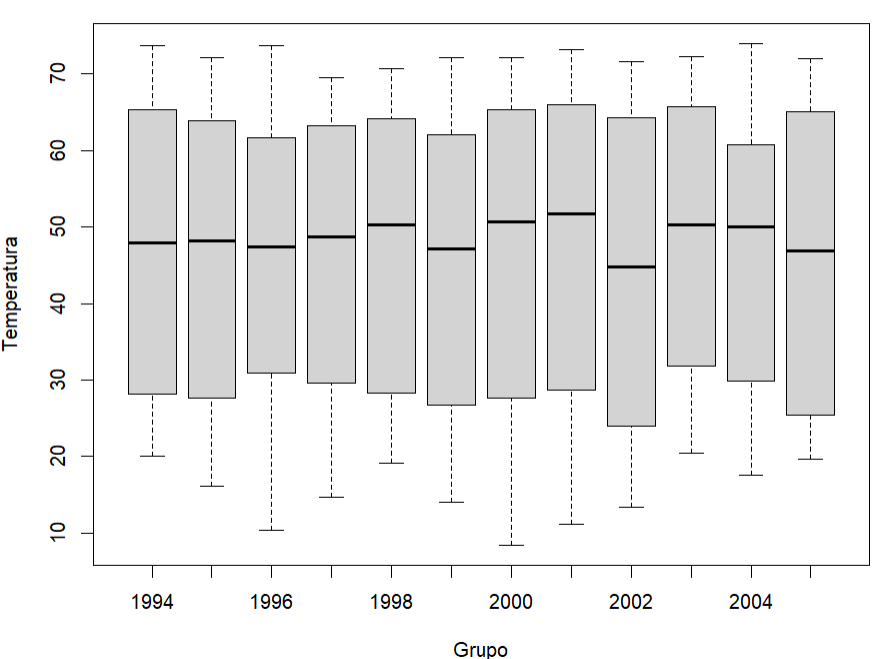


Claramente no hay tendencia y si aparenta haber un componente cicliclo que se repite cada 2 años(48 meses).

### 1.2 Análisis de estacionariedad

#### 1.2.1 Estacionariedad en varianza

boxplot(datos$Temperatura ~ datos$Año, xlab = "Grupo", ylab = "Temperatura" )



Podemos afirmar que la seri es estacionaria en varianza.

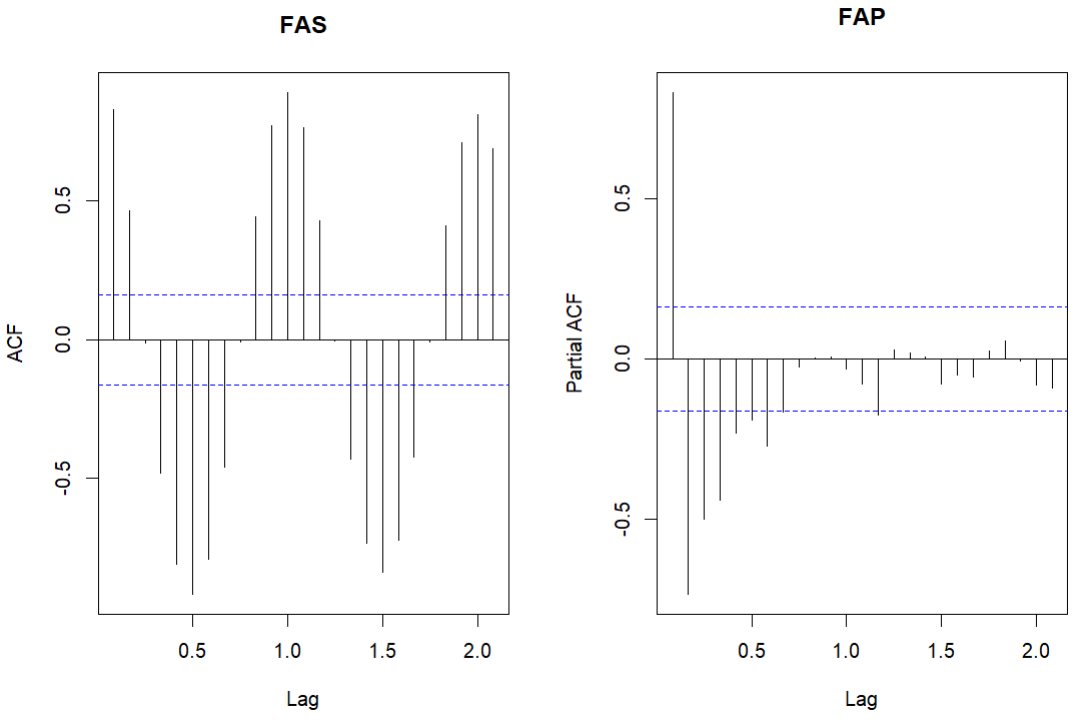
#### 1.2.2 Estacionariedad en media

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]



> FAP$acf[1]

[1] 0.8318019

Dado el gráfico FAS, observamos que decrece rápidamente a 0 en el tercer retardo. Junto con su comportamiento oscilatorio, esto indica que la serie es estacionaria. Esto se confirma con el valor del primer coeficiente resultante de las FAP, que es menor a 0.9, confirmando que la SERIE ES ESTACIONARIA EN MEDIA, pero de igual forma haremos una prueba confirmatoria.

ESTRUCTURA REGULAR

Y\_ru = ur.df(Yt, lags = 1)

summary(Y\_ru)

###############################################

# Augmented Dickey-Fuller Test Unit Root Test #

###############################################

Test regression none

Call:

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

Residuals:

Min 1Q Median 3Q Max

-15.2431 -3.9445 0.4172 6.3175 25.4286

Coefficients:

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

z.lag.1 -0.04036 0.01381 -2.922 0.00406 \*\*

z.diff.lag 0.66696 0.06374 10.464 < 2e-16 \*\*\*

---

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

Residual standard error: 8.231 on 140 degrees of freedom

Multiple R-squared: 0.4455, Adjusted R-squared: 0.4376

F-statistic: 56.24 on 2 and 140 DF, p-value: < 2.2e-16

Value of test-statistic is: -2.9215

Critical values for test statistics:

1pct 5pct 10pct

tau1 -2.58 -1.95 -1.62

***Observamos que el T calculado (-2.9215) es MENOR que el T critico (-1.95*** ***entonces rechazamos la hipótesis nula de no estacionariedad y se concluye que, en efecto, la serie temporal*** ES ESTACIONARIA EN SU ESTRUCTURA REGULAR.

***ESTRUCTURA ESTACIONAL***

Y\_ru2 = ur.df(Yt, lags = 12)

summary(Y\_ru2)

###############################################

# Augmented Dickey-Fuller Test Unit Root Test #

###############################################

Test regression none

Call:

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

Residuals:

Min 1Q Median 3Q Max

-7.818 -2.202 0.105 2.034 10.954

Coefficients:

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

z.lag.1 -0.001165 0.007369 -0.158 0.87463

z.diff.lag1 -0.705466 0.088952 -7.931 1.38e-12 \*\*\*

z.diff.lag2 -0.745286 0.096910 -7.690 4.86e-12 \*\*\*

z.diff.lag3 -0.710238 0.100233 -7.086 1.08e-10 \*\*\*

z.diff.lag4 -0.819476 0.091704 -8.936 6.51e-15 \*\*\*

z.diff.lag5 -0.873259 0.089672 -9.738 < 2e-16 \*\*\*

z.diff.lag6 -0.751633 0.094939 -7.917 1.49e-12 \*\*\*

z.diff.lag7 -0.785548 0.094532 -8.310 1.86e-13 \*\*\*

z.diff.lag8 -0.844798 0.088407 -9.556 2.26e-16 \*\*\*

z.diff.lag9 -0.846600 0.092251 -9.177 1.77e-15 \*\*\*

z.diff.lag10 -0.716051 0.101155 -7.079 1.12e-10 \*\*\*

z.diff.lag11 -0.525273 0.098512 -5.332 4.74e-07 \*\*\*

z.diff.lag12 -0.246496 0.088293 -2.792 0.00612 \*\*

---

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

Residual standard error: 3.884 on 118 degrees of freedom

Multiple R-squared: 0.8872, Adjusted R-squared: 0.8748

F-statistic: 71.42 on 13 and 118 DF, p-value: < 2.2e-16

Value of test-statistic is: -0.1581

Critical values for test statistics:

1pct 5pct 10pct

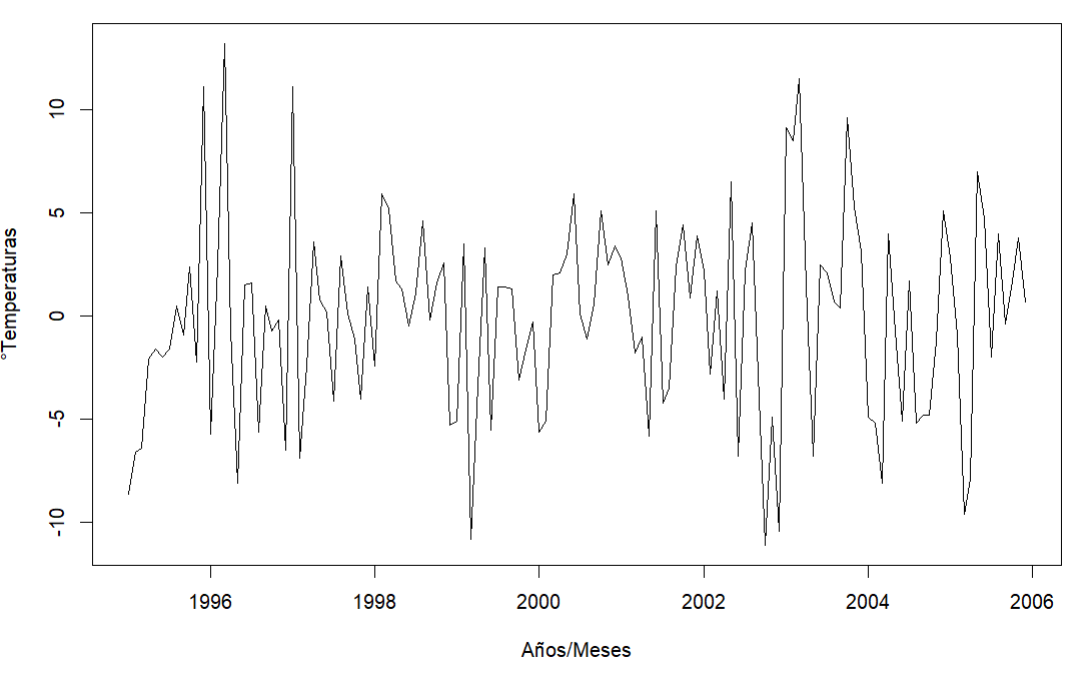
tau1 -2.58 -1.95 -1.62

***Observamos que el T calculado (-0.1581) es MAYOR que el T critico (-1.95 entonces aceptamos la hipótesis nula de no estacionariedad y se concluye que, en efecto, la serie temporal NO ES ESTACIONARIA EN SU ESTRUCTURA ESTACIONAL.***

Diferenciamos la serie en la parte estacional

D12.Yt <- diff(Yt, 12)

plot(D12.Yt, xlab="Años/Meses", ylab="°Temperaturas")



D12.Yt\_adf\_s <- ur.df(D12.Yt, lags = 12)

summary(D12.Yt\_adf\_s)

###############################################

# Augmented Dickey-Fuller Test Unit Root Test #

###############################################

Test regression none

Call:

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

Residuals:

Min 1Q Median 3Q Max

-11.9682 -2.6464 0.6982 2.6292 10.8921

Coefficients:

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

z.lag.1 -1.68474 0.37248 -4.523 1.6e-05 \*\*\*

z.diff.lag1 0.75804 0.34055 2.226 0.02814 \*

z.diff.lag2 0.66441 0.32021 2.075 0.04041 \*

z.diff.lag3 0.64190 0.30149 2.129 0.03556 \*

z.diff.lag4 0.54120 0.28226 1.917 0.05788 .

z.diff.lag5 0.46490 0.26301 1.768 0.08000 .

z.diff.lag6 0.56788 0.24280 2.339 0.02122 \*

z.diff.lag7 0.55071 0.22243 2.476 0.01487 \*

z.diff.lag8 0.49220 0.19736 2.494 0.01418 \*

z.diff.lag9 0.45317 0.17387 2.606 0.01047 \*

z.diff.lag10 0.43504 0.15177 2.866 0.00501 \*\*

z.diff.lag11 0.33757 0.12742 2.649 0.00930 \*\*

z.diff.lag12 -0.09392 0.09450 -0.994 0.32257

---

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

Residual standard error: 4.312 on 106 degrees of freedom

Multiple R-squared: 0.597, Adjusted R-squared: 0.5476

F-statistic: 12.08 on 13 and 106 DF, p-value: 1.107e-15

Value of test-statistic is: -4.523

Critical values for test statistics:

1pct 5pct 10pct

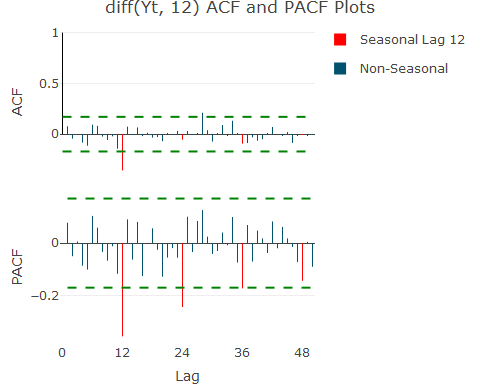
tau1 -2.58 -1.95 -1.62

Observamos que el T calculado (-4.523) es MENOR que el T critico (-1.95 entonces rechazamos la hipótesis nula de no estacionariedad y se concluye que, en efecto, la serie temporal ES ESTACIONARIA EN SU ESTRUCTURA ESTACIONAL

Volvemos a analizar el correlograma

### 1.3 Identificación del modelo estacionario

#### 1.3.1 Identificación de las órdenes p y q



* **SARIMA (0,0,0)x(2,1,0)12**
* **SARIMA (0,0,0)x(3,1,0)12**
* **SARIMA (0,0,0)x(0,1,1)12**

## 2 Estimación

mod1 = Arima(Yt, order = c(0,0,0), seasonal =list(order = c(2,1,0)))

coeftest(mod1)

z test of coefficients:

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

sar1 -0.528935 0.090505 -5.8443 5.088e-09 \*\*\*

sar2 -0.305888 0.094092 -3.2510 0.00115 \*\*

---

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

mod2 = Arima(Yt, order = c(0,0,0), seasonal =list(order = c(3,1,0)))

coeftest(mod2)

z test of coefficients:

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

sar1 -0.621312 0.089040 -6.9779 2.996e-12 \*\*\*

sar2 -0.483895 0.099508 -4.8629 1.157e-06 \*\*\*

sar3 -0.366663 0.096428 -3.8024 0.0001433 \*\*\*

---

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

mod3 = Arima(Yt, order = c(0,0,0), seasonal =list(order = c(0,1,1)))

coeftest(mod3)

z test of coefficients:

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

sma1 -1.00000 0.10413 -9.6033 < 2.2e-16 \*\*\*

---

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

## 3 Validación

### 3.1 Análisis de los coeficientes estimados

#### 3.1.1 Significación de los coeficientes

Para el modelo 1

Para el modelo 2

Para el modelo 3

#### 3.1.2 Examen de la matriz de correlaciones entre los coeficientes

vcov(mod1)

vcov(mod2)

vcov(mod3)

vcov(mod1)

sar1 sar2

sar1 0.008191179 0.003648156

sar2 0.003648156 0.008853282

vcov(mod2)

sar1 sar2 sar3

sar1 0.007928069 0.005058438 0.002110606

sar2 0.005058438 0.009901841 0.004191203

sar3 0.002110606 0.004191203 0.009298398

vcov(mod3)

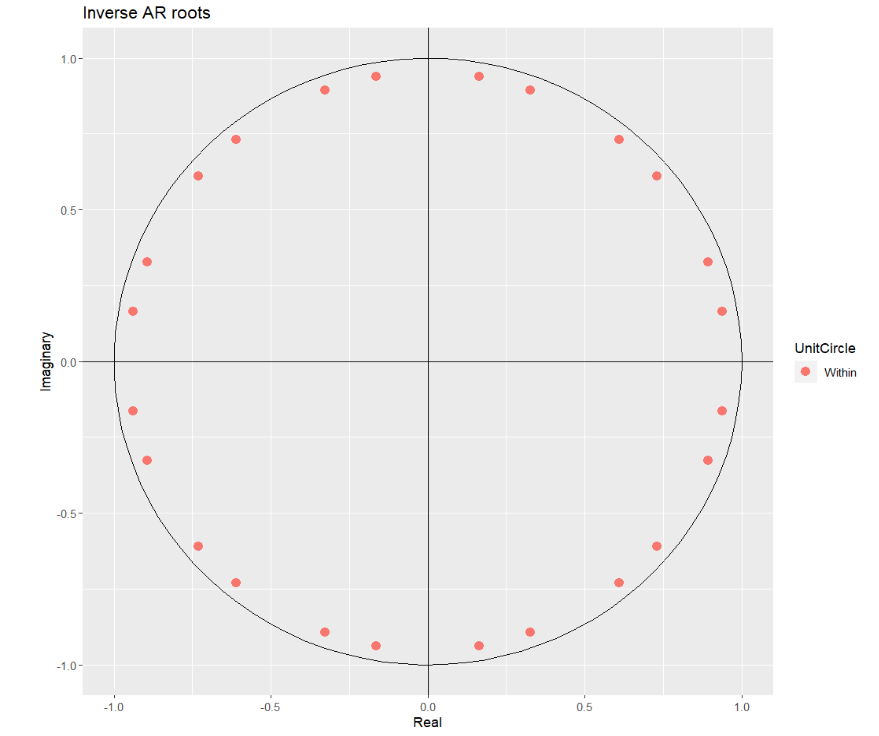
sma1

sma1 0.01084332

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

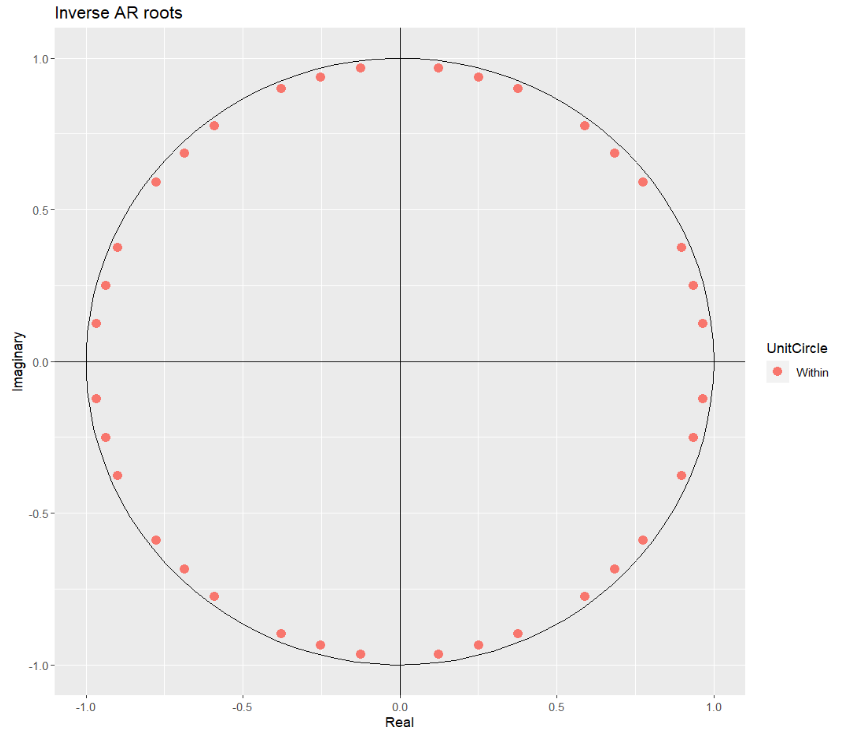
#### 3.1.3 Condición de convergencia e invertibilidad

autoplot(mod1)



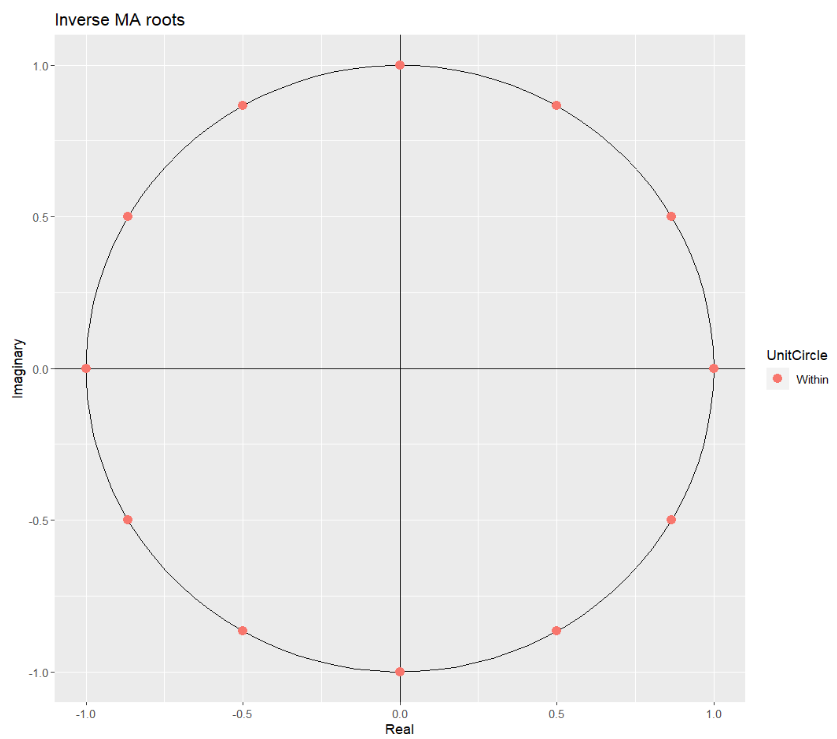
En la figura de raíces inversas de AR, se observa en el gráfico las raíces características se encuentran dentro del círculo, es decir que cumplen con la condición de estacionariedad para la parte autorregresiva.

autoplot(mod2)



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

autoplot(mod3)



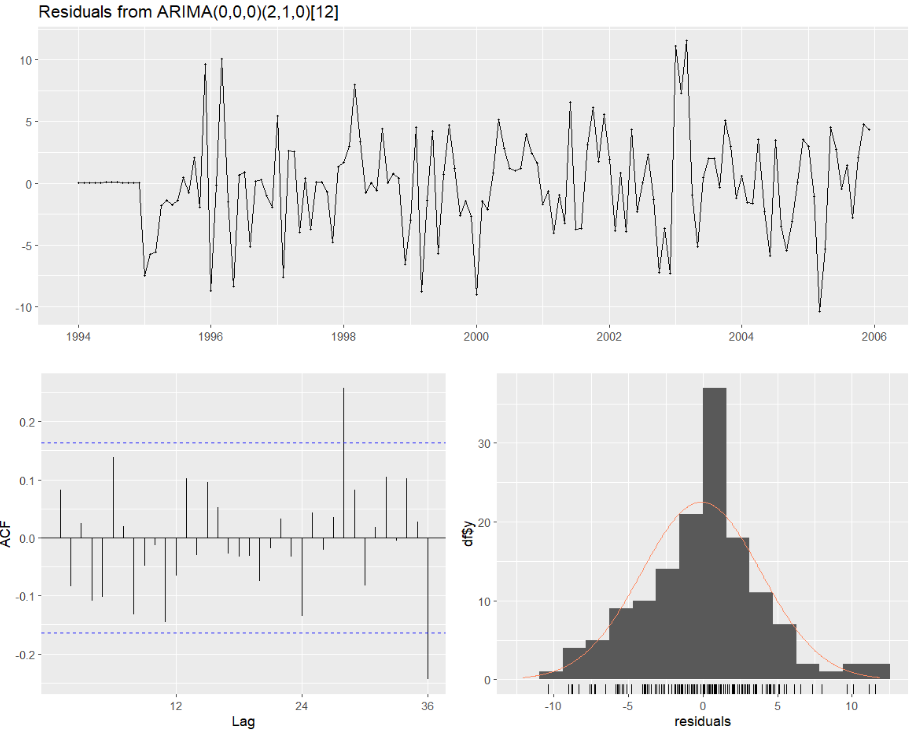
Al estar los valores dentro de la circunferencia unitaria es un indicativo de que el modelo se ajusta correctamente. Tanto en su parte AR, como en su parte MA.

### 3.2 Análisis de los residuos

#### 3.2.1 PRUEBA DE LJUNG-BOX

**Modelo 1**

checkresiduals(mod1)



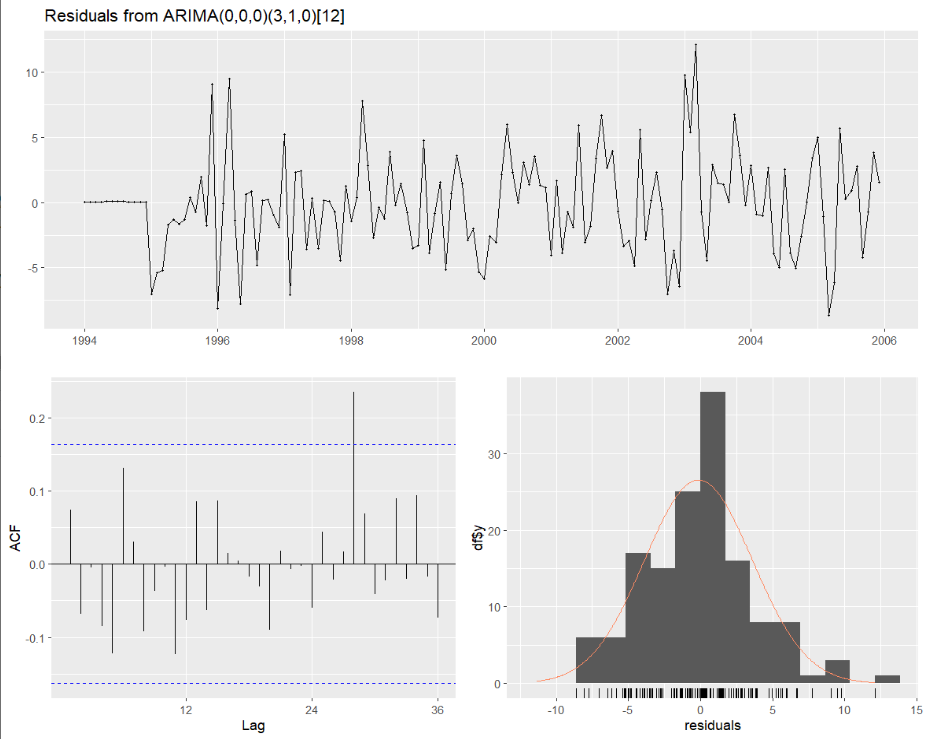
Ljung-Box test

data: Residuals from ARIMA(0,0,0)(2,1,0)[12]

Q\* = 24.411, df = 22, p-value = 0.3261

Model df: 2. Total lags used: 24

Modelo 2



Ljung-Box test

data: Residuals from ARIMA(0,0,0)(3,1,0)[12]

Q\* = 17.906, df = 21, p-value = 0.6549

Model df: 3. Total lags used: 24



Ljung-Box test

data: Residuals from ARIMA(0,0,0)(0,1,1)[12]

Q\* = 21.368, df = 23, p-value = 0.5586

Model df: 1. Total lags used: 24

Los coeficientes p asociados a las pruebas de Ljung-Box aplicadas a los modelos indican que no hay evidencia suficiente para rechazar la hipótesis nula, que afirma que no hay autocorrelación en los residuos. Esto sugiere que los residuos del modelo son consistentes con el ruido blanco.

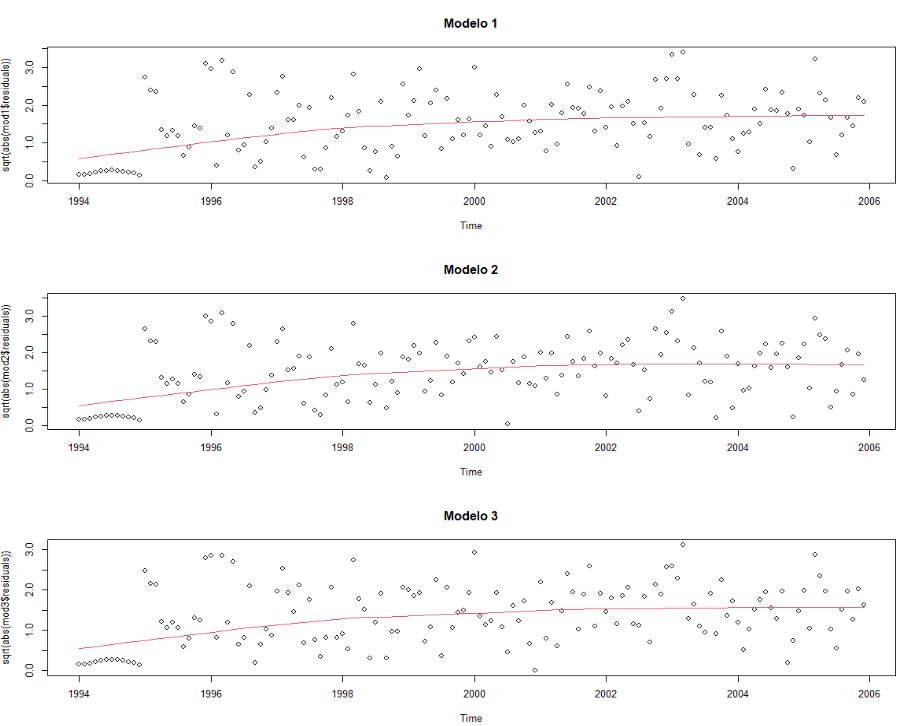
#### 3.2.2 Homocedasticidad o varianza constante

par(mfrow = c(3,1))

scatter.smooth(sqrt(abs(mod1$residuals)), lpars=list(col=2), main = "Modelo 1")

scatter.smooth(sqrt(abs(mod2$residuals)), lpars=list(col=2), main = "Modelo 2")

scatter.smooth(sqrt(abs(mod3$residuals)), lpars=list(col=2), main = "Modelo 3")



Se observa que los datos parecen 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 = 12.671, df = 1, p-value = 0.0003713

obs=get(mod2$series)

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

studentized Breusch-Pagan test

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

BP = 10.96, df = 1, p-value = 0.0009308

obs=get(mod3$series)

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

studentized Breusch-Pagan test

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

BP = 11.298, df = 1, p-value = 0.0007759

El valor de probabilidad (p-valor) asociado al estadístico BP asume los valores de probabilidad de 0.0003, 0.0009 y 0.0007 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.

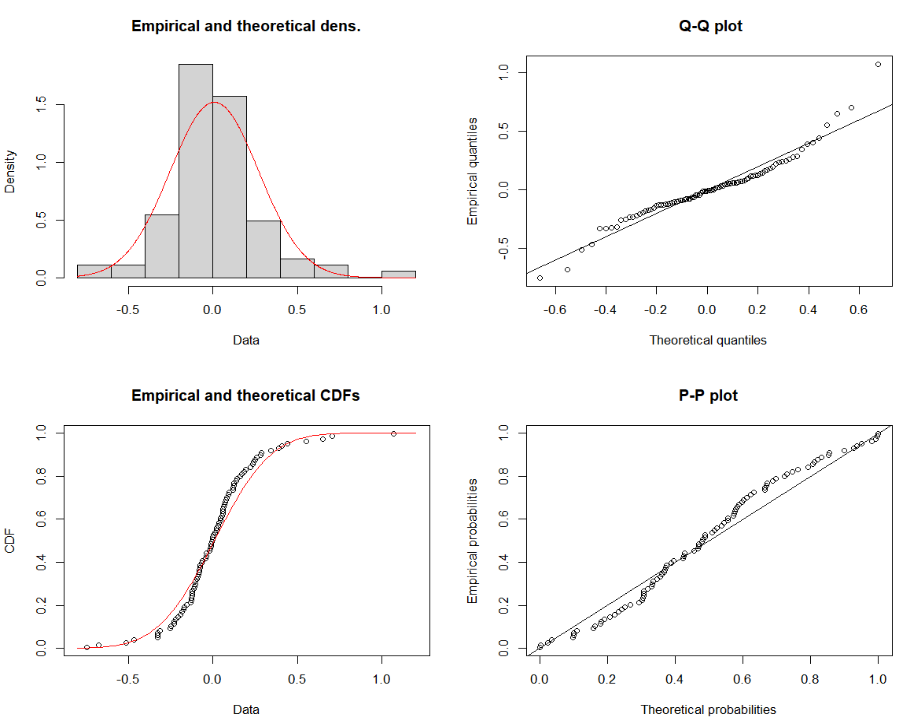
#### 3.2.3 Contraste de normalidad

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

plot(ajuste\_m1)

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

JB\_m1



Jarque Bera Test

data: resid\_m1

X-squared = 41.778, df = 2, p-value = 8.471e-10

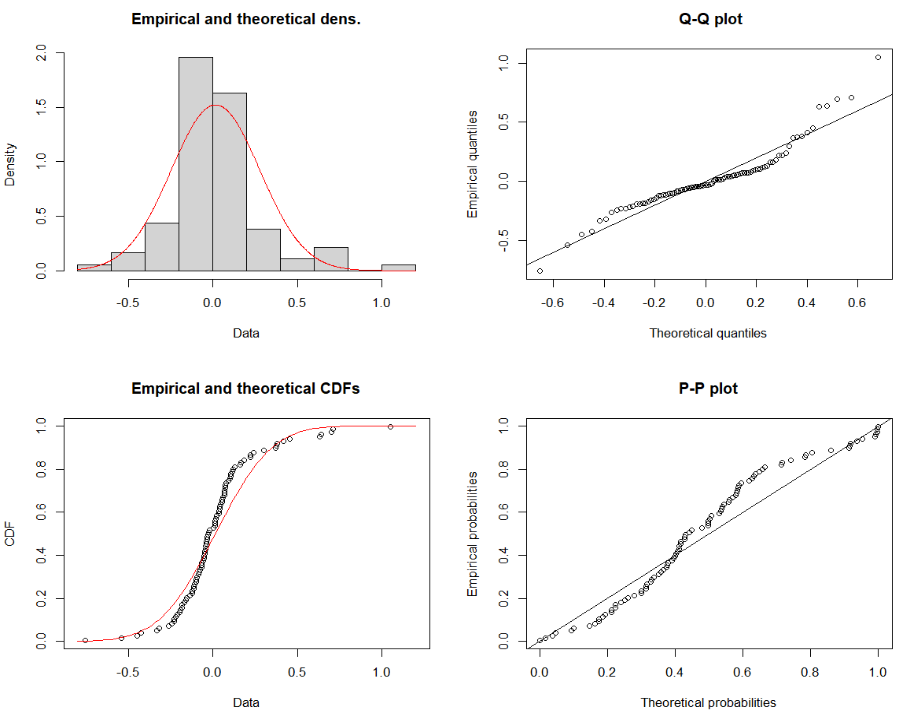
En las figuras se observa que los residuales del modelo 1 no presentan indicios de ajuste hacia la distribución normal. Y en la prueba JB, como p = 0.0000 < 0.05, se rechaza la Ho, es decir, los residuos NO se aproximan a una distribución normal.

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

plot(ajuste\_m2)

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

JB\_m2



Jarque Bera Test

data: resid\_m2

X-squared = 47.925, df = 2, p-value = 3.919e-11

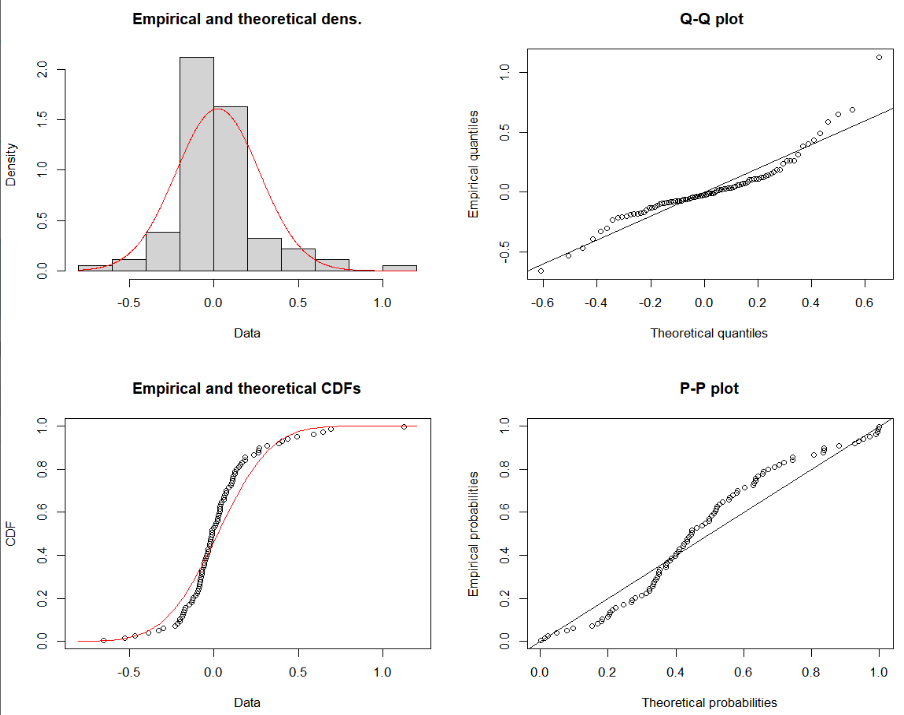
En las figuras se observa que los residuales del modelo 2 no presentan indicios de ajuste hacia la distribución normal.Y en la prueba JB, como p = 0.00000 < 0.05, se rechaza la Ho, es decir, los residuos NO se aproximan a una distribución normal.

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

plot(ajuste\_m3)

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

JB\_m3



Jarque Bera Test

data: resid\_m3

X-squared = 87.238, df = 2, p-value < 2.2e-16

En las figuras se observa que los residuales del modelo 3 NO 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

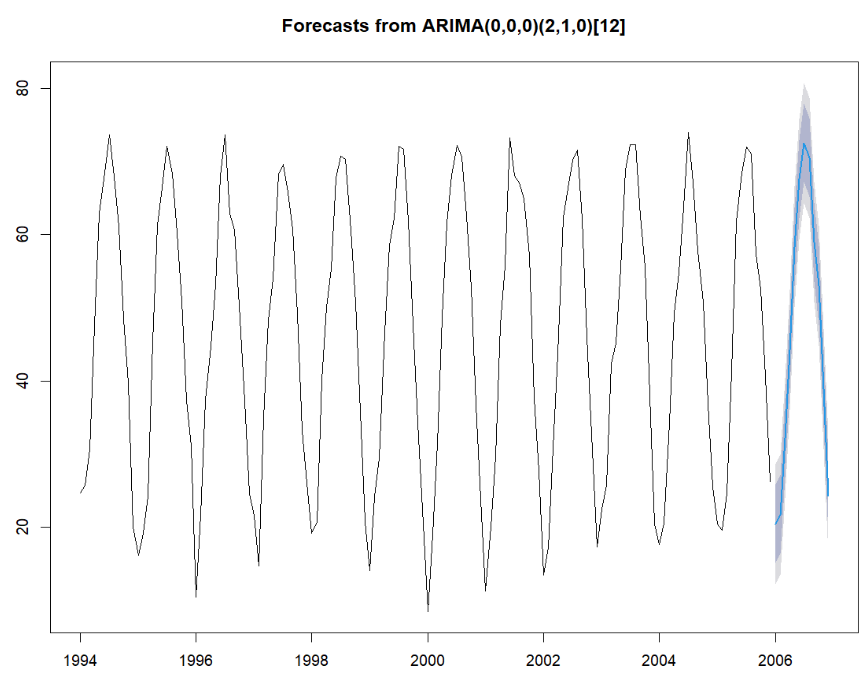
### 4.1Pronosticos de cada modelo

#### Modelo 1:

pron <- forecast(mod1, h = 12)

plot(pron)

summary(pron)



Forecast method: ARIMA(0,0,0)(2,1,0)[12]

Model Information:

Series: Yt

ARIMA(0,0,0)(2,1,0)[12]

Coefficients:

sar1 sar2

-0.5289 -0.3059

s.e. 0.0905 0.0941

sigma^2 = 17.6: log likelihood = -377.83

AIC=761.66 AICc=761.84 BIC=770.3

Error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set -0.1083612 3.986116 2.947049 -2.127589 9.655297 0.7865153 0.08213804

Forecasts:

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

Jan 2006 20.41783 15.04138 25.79428 12.19526 28.64040

Feb 2006 21.66666 16.29021 27.04311 13.44409 29.88923

Mar 2006 32.15547 26.77903 37.53192 23.93290 40.37805

Apr 2006 44.25504 38.87859 49.63149 36.03247 52.47761

May 2006 58.31157 52.93512 63.68802 50.08900 66.53415

Jun 2006 67.57403 62.19758 72.95048 59.35146 75.79661

Jul 2006 72.53786 67.16141 77.91431 64.31529 80.76043

Aug 2006 70.57488 65.19843 75.95133 62.35231 78.79745

Sep 2006 58.97984 53.60339 64.35629 50.75727 67.20241

Oct 2006 53.06907 47.69262 58.44552 44.84650 61.29164

Nov 2006 38.95711 33.58066 44.33356 30.73454 47.17968

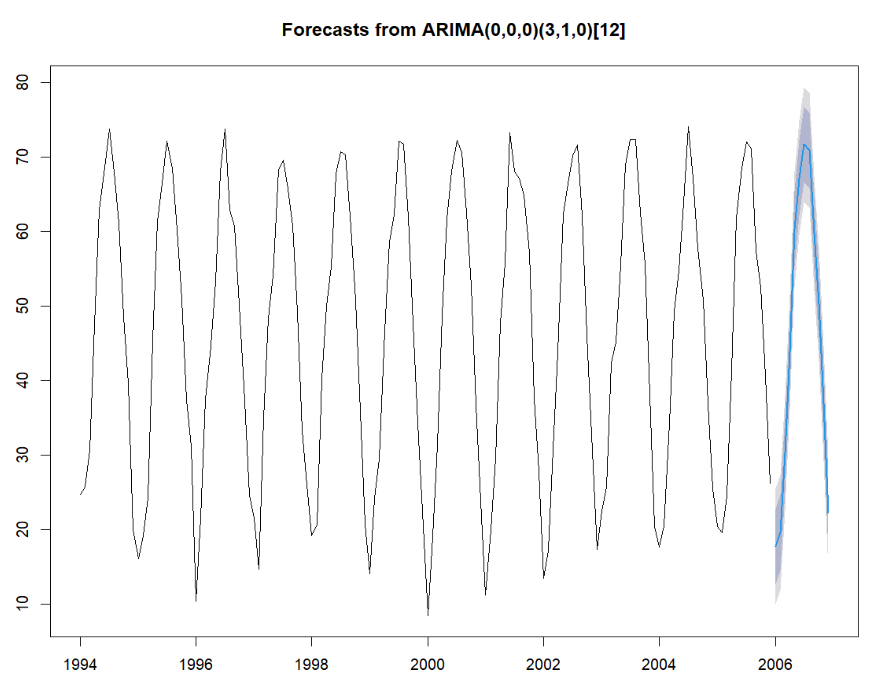
Dec 2006 24.26971 18.89326 29.64616 16.04714 32.49228

#### Modelo 2:

pron2 <- forecast(mod2, h = 12)

plot(pron2)

summary(pron2)



Forecast method: ARIMA(0,0,0)(3,1,0)[12]

Model Information:

Series: Yt

ARIMA(0,0,0)(3,1,0)[12]

Coefficients:

sar1 sar2 sar3

-0.6213 -0.4839 -0.3667

s.e. 0.0890 0.0995 0.0964

sigma^2 = 15.55: log likelihood = -371.61

AIC=751.23 AICc=751.54 BIC=762.76

Error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set -0.1418842 3.732405 2.797088 -2.149689 8.981968 0.7464933 0.07438451

Forecasts:

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

Jan 2006 17.69478 12.64106 22.74850 9.96578 25.42378

Feb 2006 19.55880 14.50508 24.61252 11.82980 27.28780

Mar 2006 30.26752 25.21380 35.32124 22.53853 37.99652

Apr 2006 43.72280 38.66907 48.77652 35.99380 51.45179

May 2006 60.28285 55.22913 65.33657 52.55385 68.01184

Jun 2006 67.13104 62.07732 72.18476 59.40204 74.86004

Jul 2006 71.65001 66.59629 76.70373 63.92101 79.37901

Aug 2006 70.87434 65.82062 75.92806 63.14534 78.60334

Sep 2006 59.72455 54.67083 64.77827 51.99556 67.45355

Oct 2006 50.24650 45.19278 55.30022 42.51750 57.97550

Nov 2006 36.87637 31.82265 41.93009 29.14738 44.60537

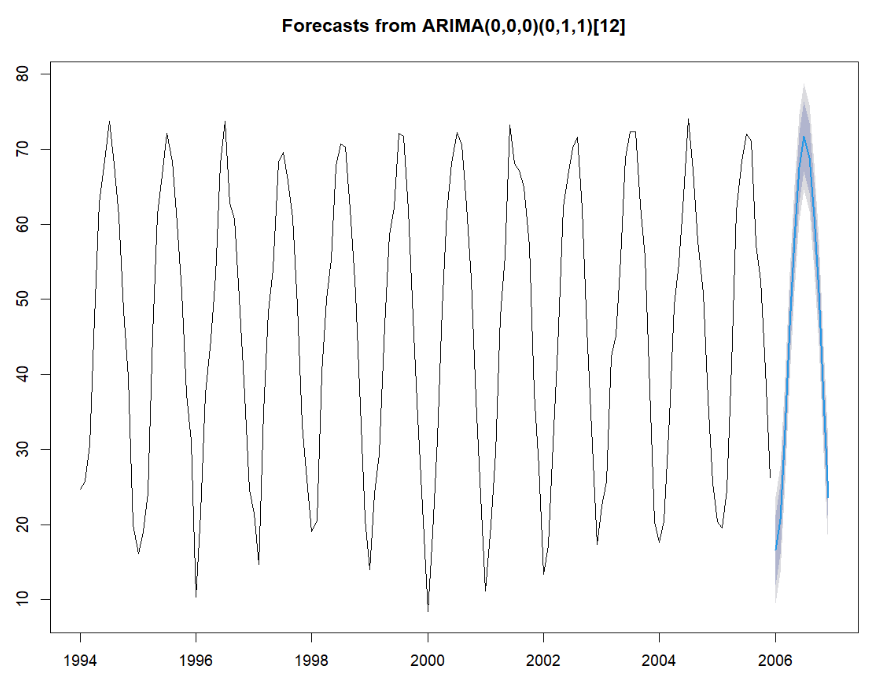
Dec 2006 22.16056 17.10684 27.21428 14.43157 29.88956

#### Modelo 3:

pron3 <- forecast(mod3, h = 12)

plot(pron3)

summary(pron3)



Forecast method: ARIMA(0,0,0)(0,1,1)[12]

Model Information:

Series: Yt

ARIMA(0,0,0)(0,1,1)[12]

Coefficients:

sma1

-1.0000

s.e. 0.1041

sigma^2 = 12.08: log likelihood = -366.13

AIC=736.25 AICc=736.35 BIC=742.02

Error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set -0.1628102 3.314491 2.46352 -2.580408 8.282389 0.6574699 0.07191584

Forecasts:

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

Jan 2006 16.60833 11.97302 21.24365 9.519228 23.69744

Feb 2006 20.65000 16.01468 25.28531 13.560894 27.73910

Mar 2006 32.47500 27.83968 37.11031 25.385893 39.56410

Apr 2006 46.52500 41.88968 51.16031 39.435892 53.61410

May 2006 58.09166 53.45635 62.72698 51.002558 65.18077

Jun 2006 67.49999 62.86468 72.13531 60.410890 74.58910

Jul 2006 71.71666 67.08134 76.35198 64.627557 78.80576

Aug 2006 68.88333 64.24801 73.51864 61.794223 75.97243

Sep 2006 61.02499 56.38968 65.66031 53.935891 68.11410

Oct 2006 50.97500 46.33968 55.61031 43.885892 58.06410

Nov 2006 36.65000 32.01468 41.28531 29.560893 43.73910

Dec 2006 23.64166 19.00635 28.27698 16.552561 30.73077

### SERIE ORIGINAL (YT) Y PRONOSTICADA.

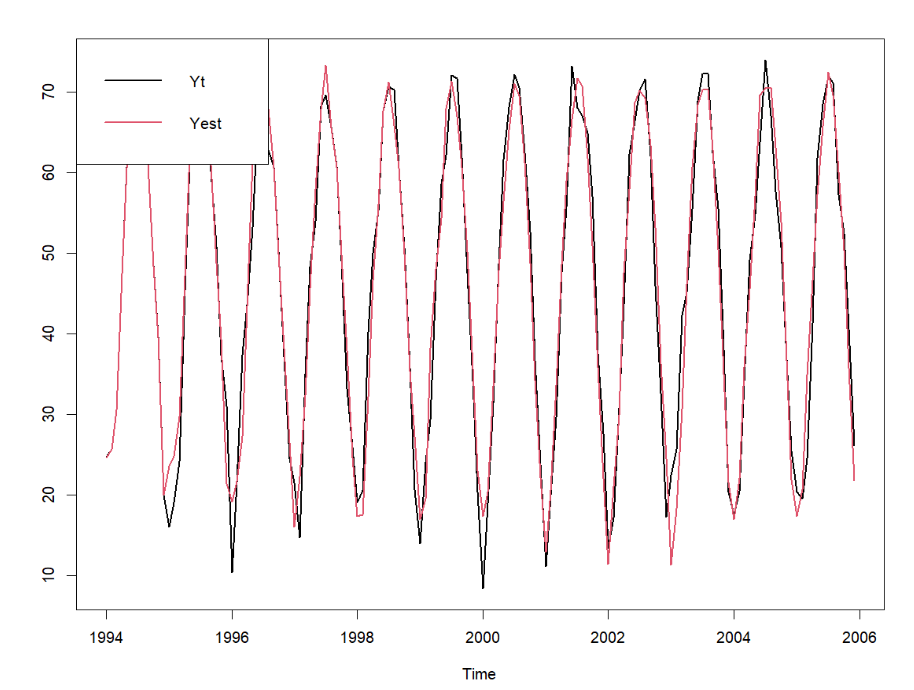
**Modelo 1**

Yt\_S <- mod1$fitted

grafico\_comparativo <- cbind(Yt, Yt\_S)

ts.plot(grafico\_comparativo, col = c(1:2), lwd = 2)

legend("topleft", c("Yt", "Yest"), lty=c(1,1), col=c(1:2), lwd = 2)

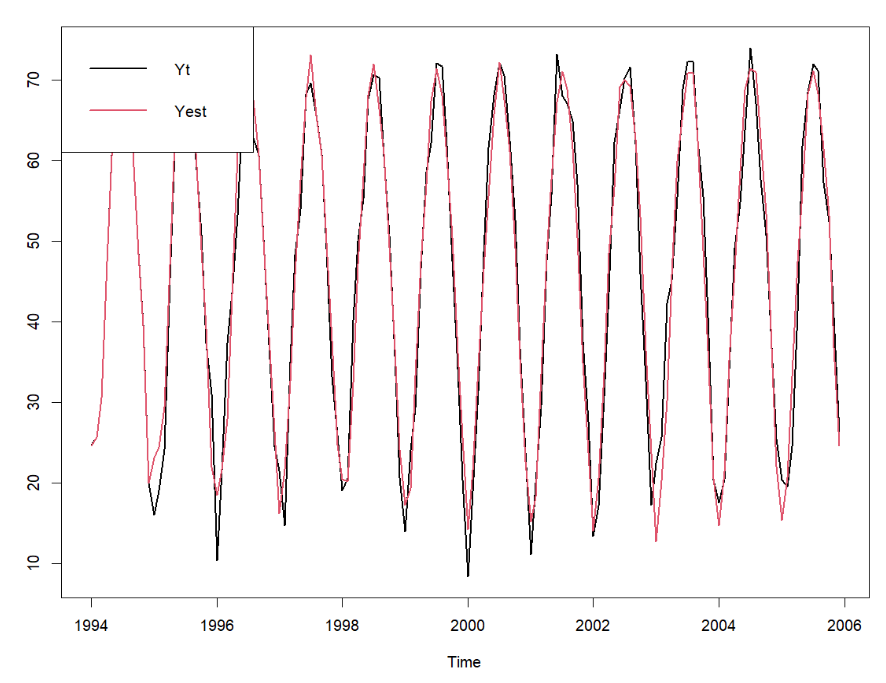


Yt\_S2 <- mod2$fitted

grafico\_comparativo2 <- cbind(Yt, Yt\_S2)

ts.plot(grafico\_comparativo2, col = c(1:2), lwd = 2)

legend("topleft", c("Yt", "Yest"), lty=c(1,1), col=c(1:2), lwd = 2)

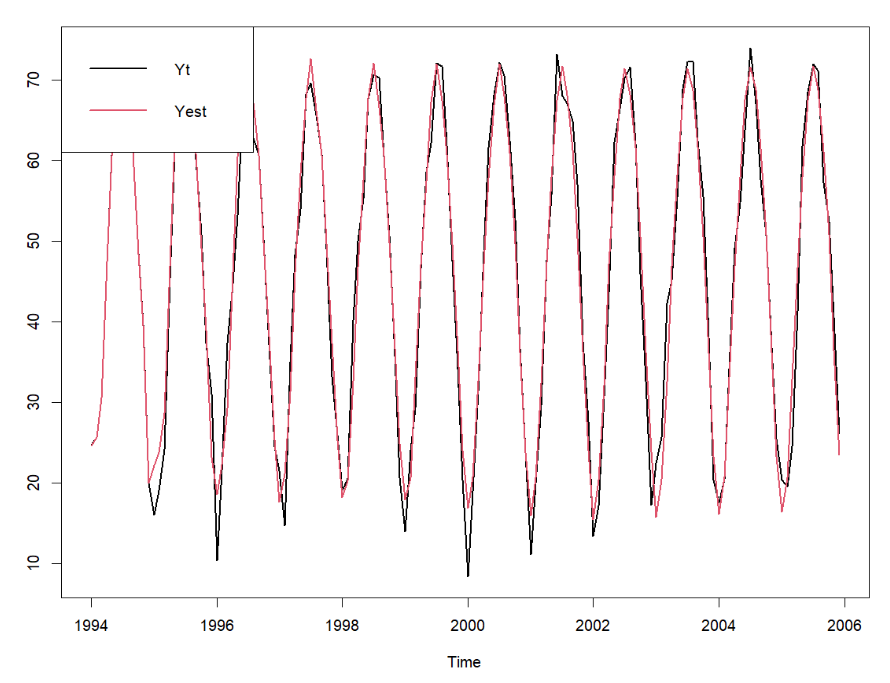


Yt\_S3 <- mod3$fitted

grafico\_comparativo3 <- cbind(Yt, Yt\_S3)

ts.plot(grafico\_comparativo3, col = c(1:2), lwd = 2)

legend("topleft", c("Yt", "Yest"), lty=c(1,1), col=c(1:2), lwd = 2)



### Gráfica del ajuste y pronóstico con valores reales

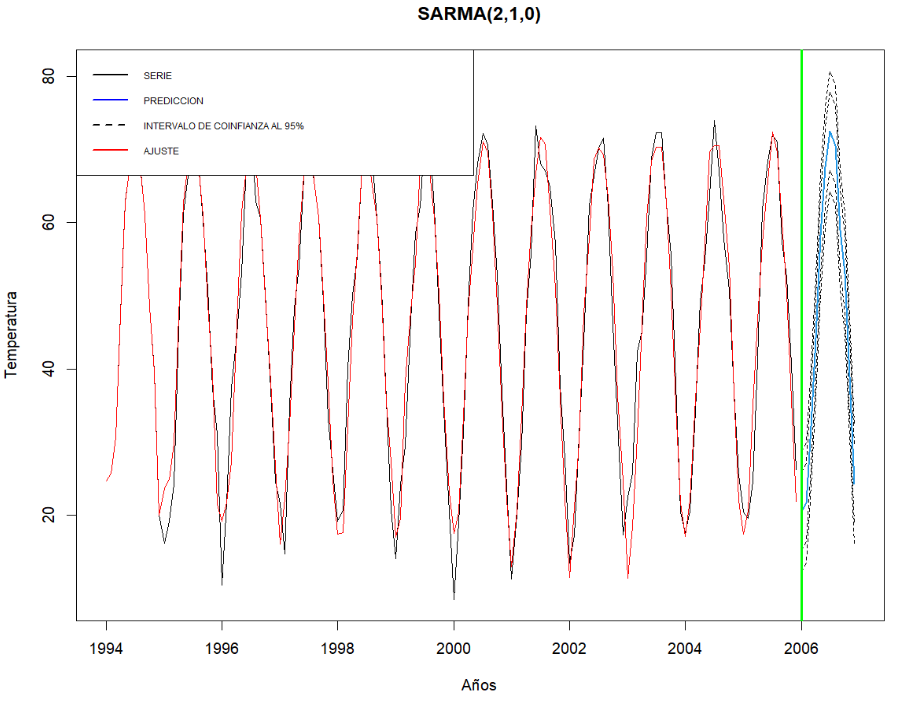
**MODELO 1**

plot(pron, shaded = FALSE, xlab = "Años", ylab = "Temperatura",main = "SARMA(2,1,0)")

lines(pron$fitted, col = "red")

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

abline(v=2006, lwd = 3, col="green")



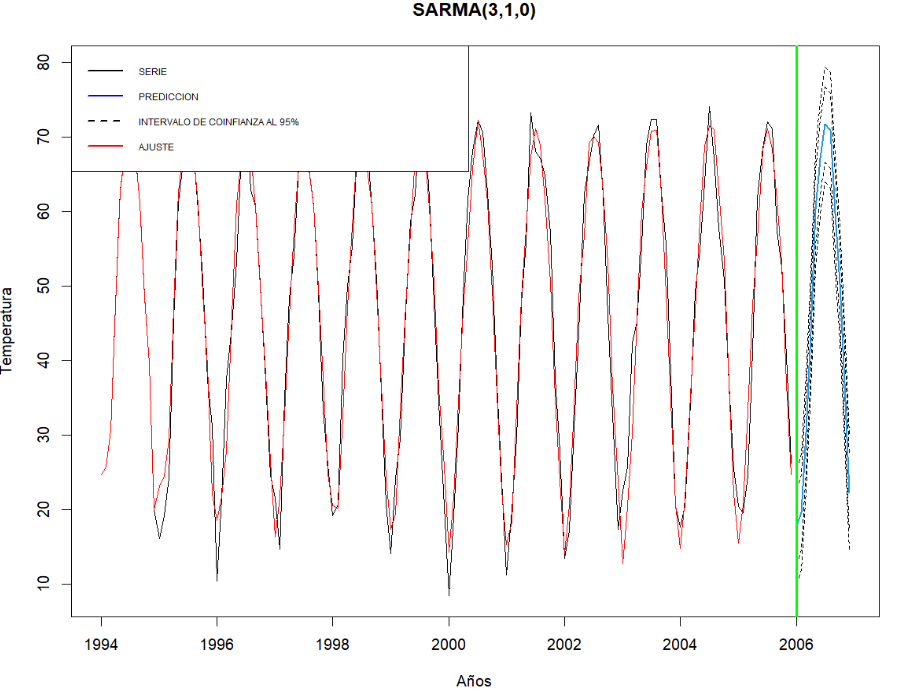
**MODELO 2**

plot(pron2, shaded = FALSE, xlab = "Años", ylab = "Temperatura",main = "SARMA(3,1,0)")

lines(pron2$fitted, col = "red")

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

abline(v=2006, lwd = 3, col="green")

****

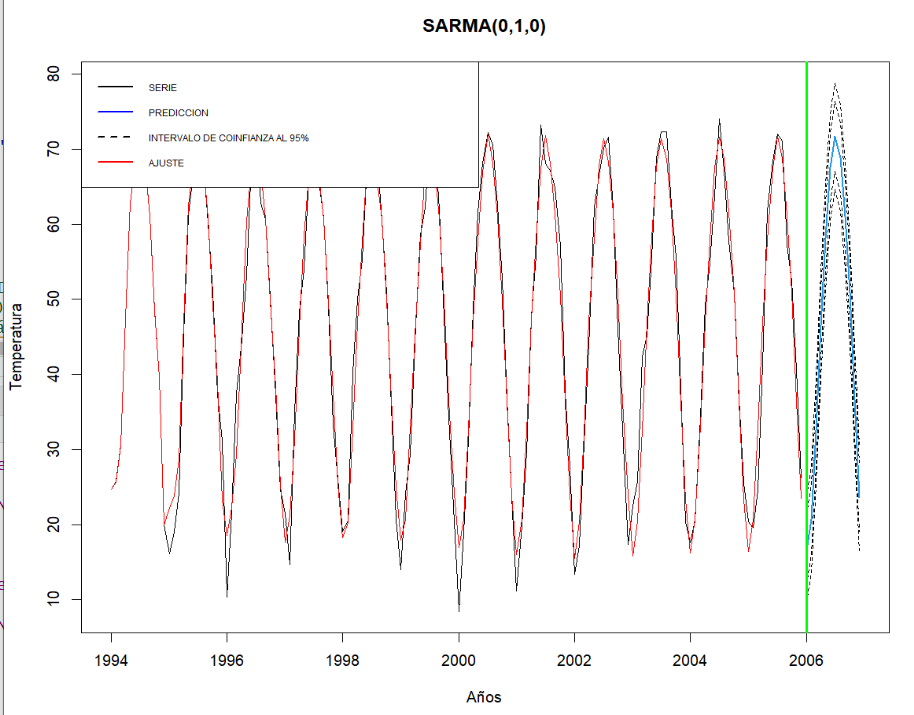
**MODELO 3**

plot(pron3, shaded = FALSE, xlab = "Años", ylab = "Temperatura",main = "SARMA(0,1,0)")

lines(pron3$fitted, col = "red")

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

abline(v=2006, lwd = 3, col="green")

****

### Métricas basadas en el error

accuracy(Pron1)

ME RMSE MAE MPE MAPE MASE ACF1

Training set -0.1083612 3.986116 2.947049 -2.127589 9.655297 0.7865153 0.08213804

accuracy(Pron2)

ME RMSE MAE MPE MAPE MASE ACF1

Training set -0.1418842 3.732405 2.797088 -2.149689 8.981968 0.7464933 0.07438451

accuracy(Pron3)

ME RMSE MAE MPE MAPE MASE ACF1

Training set -0.1628102 3.314491 2.46352 -2.580408 8.282389 0.6574699 0.07191584

Basado en las métricas MAE y RMSE, el modelo 3 es el que mejor rendimiento tiene con el menor MAE y RMSE

### AIC Y BIC

AIC(mod1); BIC(mod1)

[1] 761.6553

[1] 770.3037

AIC(mod2); BIC(mod2)

[1] 751.2253

[1] 762.7565

AIC(mod3); BIC(mod3)

[1] 736.2537

[1] 742.0193

## Conclusión: Elección del Modelo SARIMA

#### Modelos Evaluados:

1. **SARIMA(1,1,0)(1,0,0)[12]**
   * **RMSE**: 3.986116
   * **MAE**: 2.947049
   * **AIC**: 761.6553
   * **BIC**: 770.3037
2. **SARIMA(0,1,5)(1,0,0)[12]**
   * **RMSE**: 3.732405
   * **MAE**: 2.797088
   * **AIC**: 751.2253
   * **BIC**: 762.7565
3. **SARIMA(1,1,2)(1,0,0)[12]**
   * **RMSE**: 3.314491
   * **MAE**: 2.46352
   * **AIC**: 736.2537
   * **BIC**: 742.0193

#### Elección del Mejor Modelo:

Basado en los resultados de las métricas y los criterios de información (AIC y BIC):

* El modelo **SARIMA(1,1,2)(1,0,0)[12]** tiene el menor RMSE (3.314491) y MAE (2.46352) entre los modelos evaluados.
* Además, presenta el menor AIC (736.2537) y BIC (742.0193), lo cual indica un mejor ajuste del modelo a los datos en comparación con los otros modelos evaluados.

Por lo tanto, el SARIMA(1,1,2)(1,0,0)[12] es el modelo recomendado basado en las métricas de evaluación (RMSE, MAE) y los criterios de información (AIC, BIC). Este modelo parece ofrecer la mejor combinación de precisión en la predicción y ajuste al conjunto de datos observados.

# CASO 2: VENTAS DE LIBRERÍA UNIVERSITARIA

library(forecast)#modelo ARIMA

library(tseries)#Para series de tiempo

library(TSstudio)#correlograma parte regular y estacional

library(TSA)#modelos ARMA

library(ggplot2)#para hacer graficos

library(urca)#para hacer test de raiz unitaria (detectar hay)

library(dplyr)#para la manipulacionde datos

library(lmtest)#inferencia para los coeficioentes

library(MASS)#transformacion de Box-Cox

library(nortest)#pruebas de normalidad

library(mFilter)#para hoodrick - prescot

library(zoo)

library(TTR)

library(sandwich)

library(strucchange)#para analisis de estabilidad - Chow

library(survival)

library(fitdistrplus)

datos <- read\_excel("F:\\777--Programacion repos\\Una\\r\\data\\actividad-07.xlsx",sheet = "02")

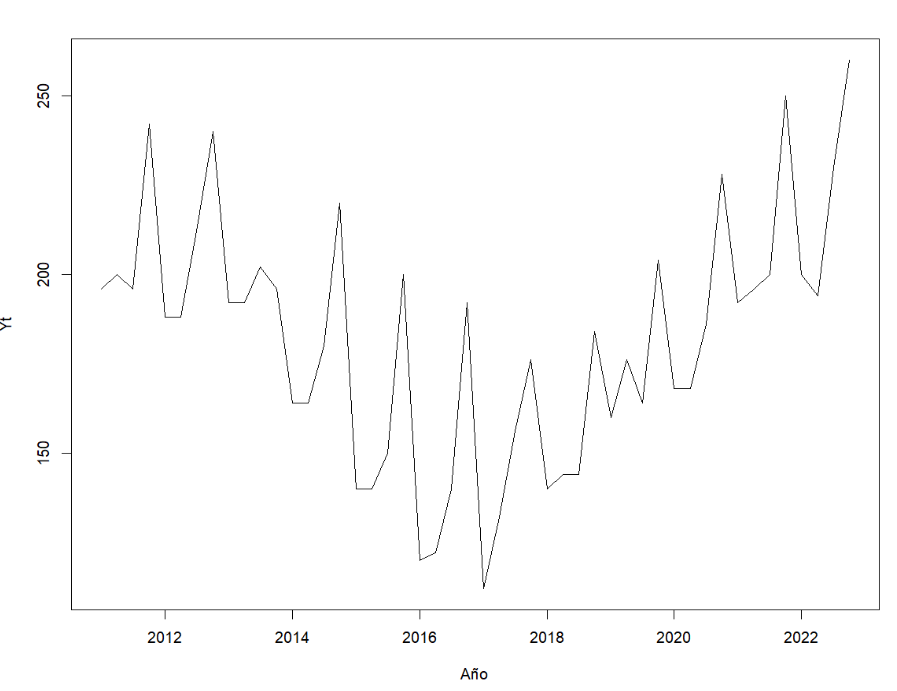
View(datos)

# 1) IDENTIFICACION

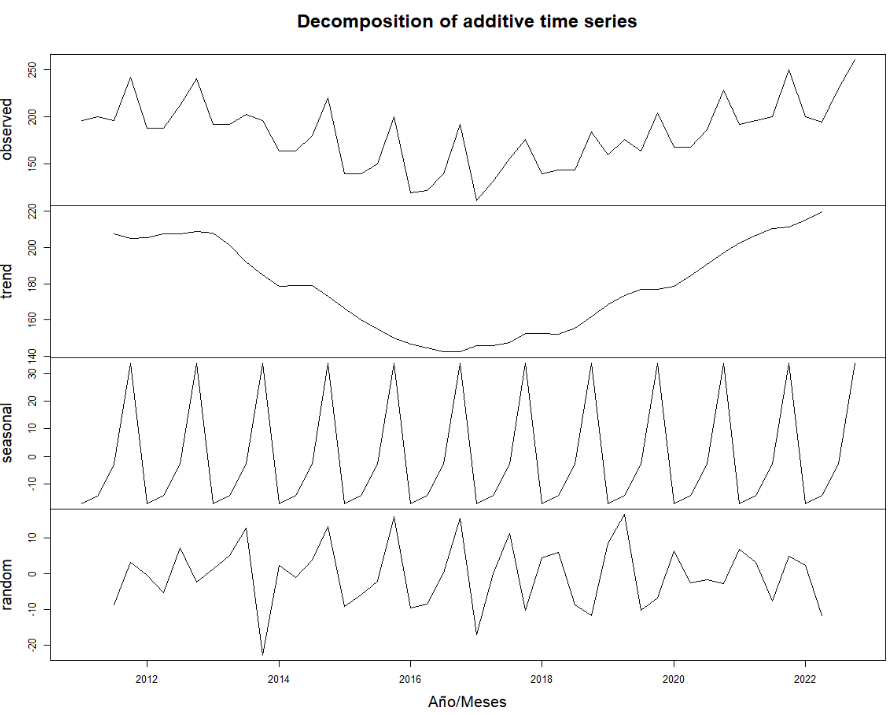
# Graficar la serie

Yt <- ts(datos$Ventas, start = c(2011, 1), frequency = 4)

plot(Yt, xlab = "Año", ylab ="Yt" )



Como se puede apreciar la serie tiene una tendencia el cual decrece asta el año 2017 y a partir de hay la serie tiene una tendencia creciente pero en cuanto a la estacionariedad en varianza tiene indicios que si hay estacionariedad en varianza.



Tras la descomposición de la serie se puede observar que si existe tendencia tal como se describió al inicio, y entre otros factores como la parte aleatoria.

## 1 Identificación

### 1. 1 Análisis de la tendencia y la estacionalidad

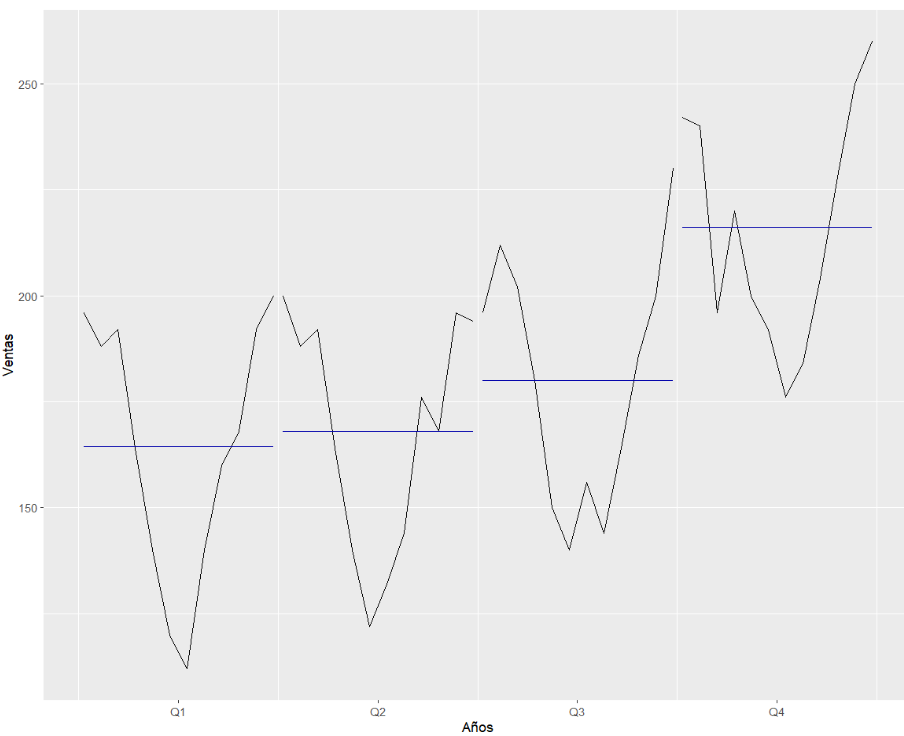
#### 1.1.1 Estacionalidad

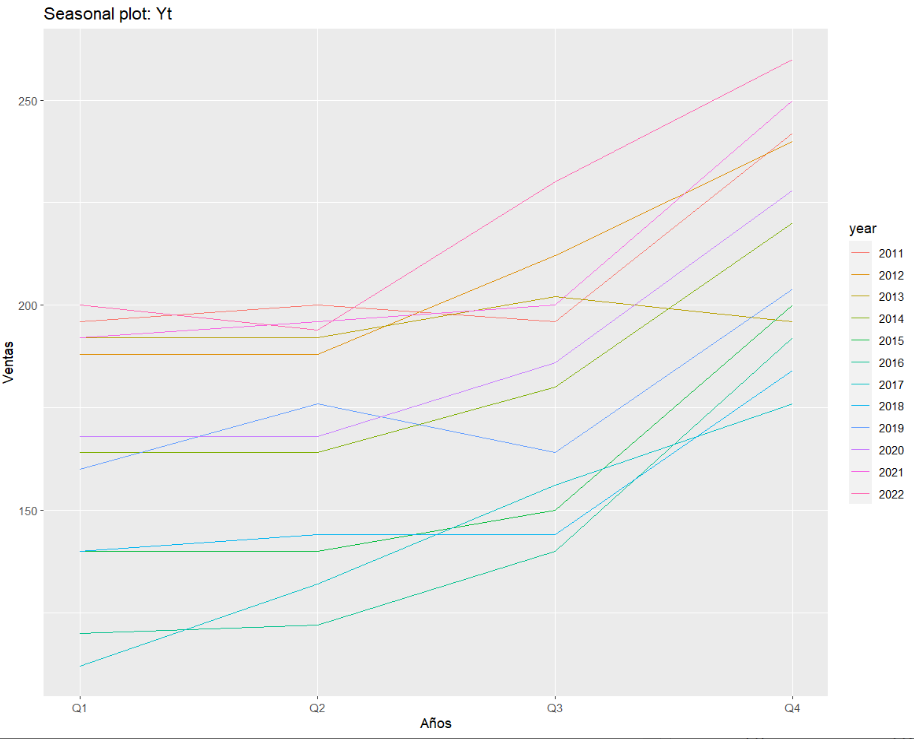
# Estacionalidad

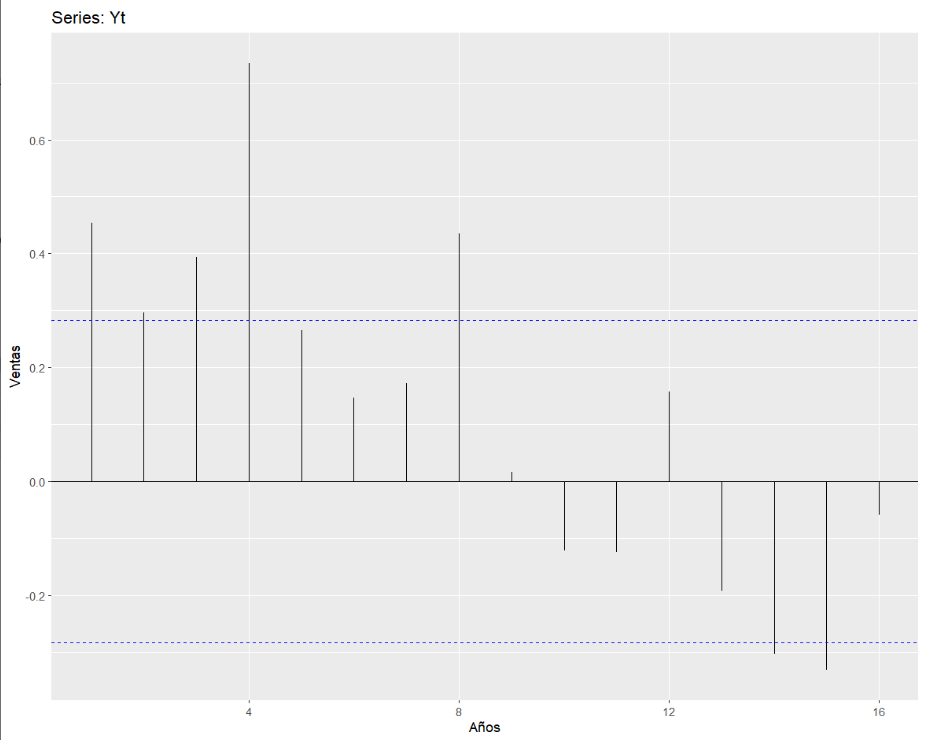
ggsubseriesplot(Yt, xlab = "Años", ylab = "Ventas" )

ggseasonplot(Yt, xlab = "Años",ylab = "Ventas" )

ggAcf(Yt, xlab = "Años", ylab = "Ventas")







Como se puede apreciar en las líneas apiladas y las líneas separadas se observa que no presentan indicios de estacionalidad en la serie temporal el cual se confirma con el correlograma que tiene uncomportamiento en el que decrece lentamente lo cual es una clara señal que no existe la estacionariedad en la serie.

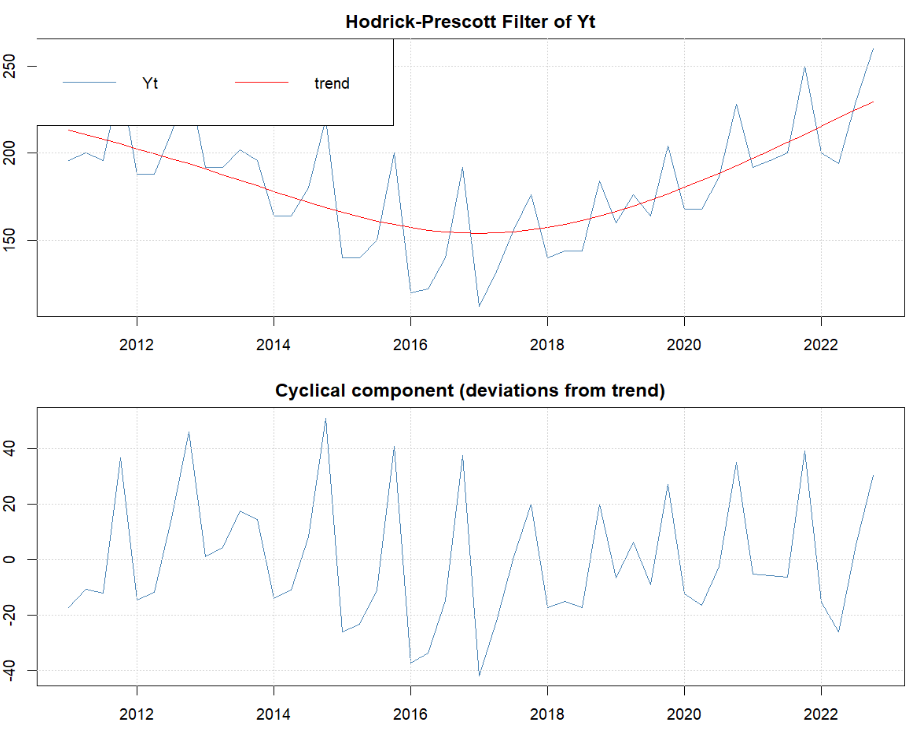
#### 1.1.2 Análisis de tendencia

#Análisis de tendencia

lambda\_hp <- 1600

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

plot(data\_hp)

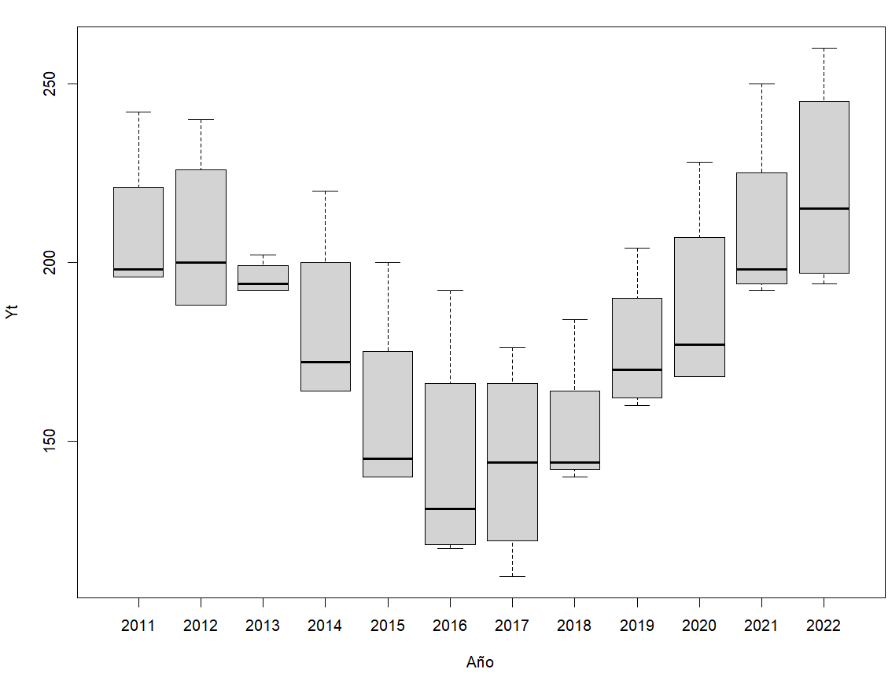


Se aprecia que la tendencia inicialmente decrece hasta el año 2017 y de ahí en adelante crece.

### 1.2 Análisis de estacionariedad

#### 1.2.1 Estacionariedad en varianza

boxplot(datos$Ventas ~ datos$Año, xlab = "Año", ylab="Yt")

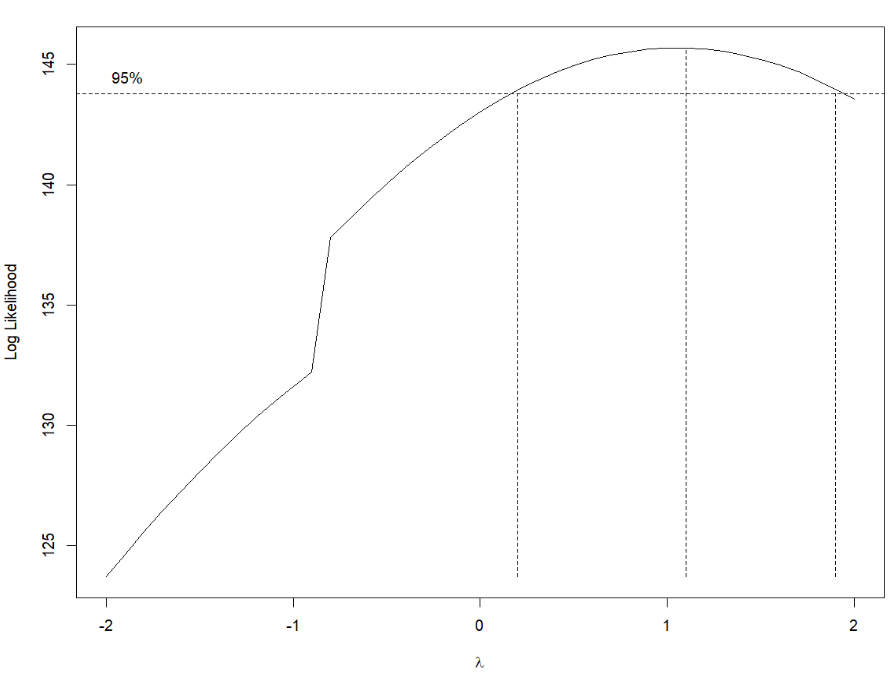


Dado el grafico podemos inferir que la siete tiene un comportamiento en el cual parece presentar la estacionariedad varianza si no fuera por la tercera caja que muestra tener una variación.

b <- BoxCox.ar(Yt)

lambda <- b$mle

round(lambda,2)



[1] 1.1

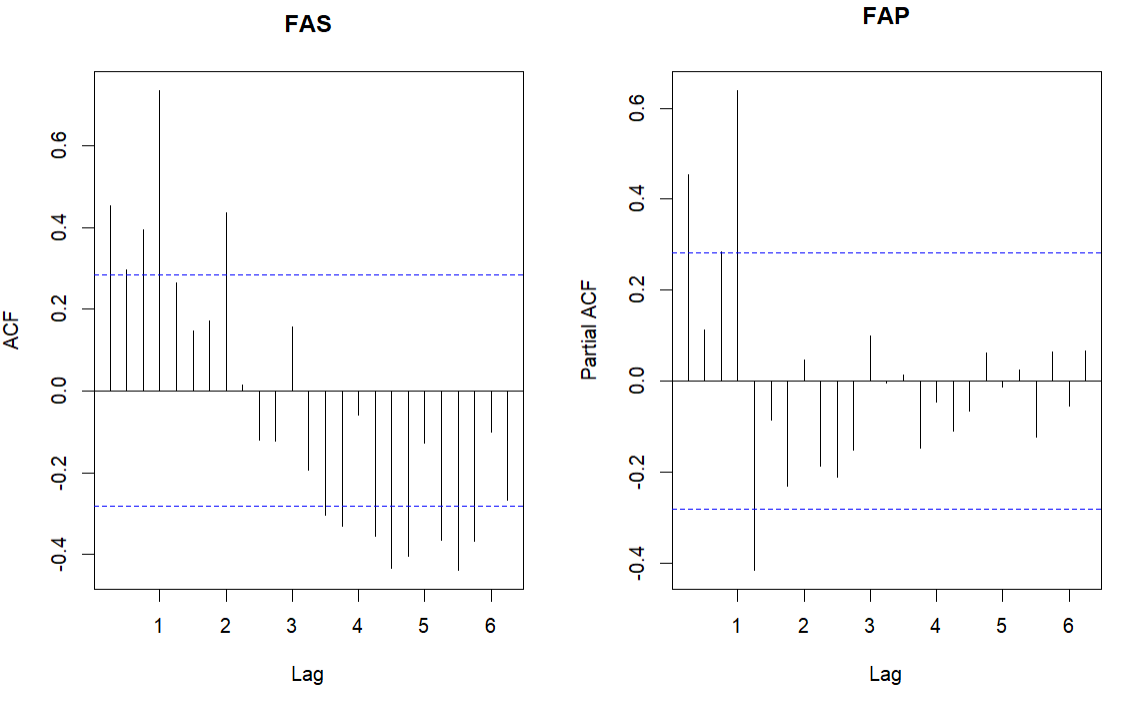
Segun el dato 1.1 y el grafico de boxcox se puede apreciar que la serie no necesita una transformación por lo que podemos estar afirmando que si es estacionaria en varianza.

#### 1.2.2 Estacionariedad en media

par(mfrow = c(1,2))

FAS <- acf(Yt,lag.max = 25, main = "FAS")

FAP <- pacf(Yt, lag.max = 25, main = "FAP")



Dado el grafico del FAS observamos que decrece lentamente y en el FAP el retardo 9 recién llega a 0 lo que es un gran indicio de que la serie no es estacionaria en media, por tanto, se debe hacer una diferenciación para estabilizarlo.

PARTE REGULAR

data\_adf <- ur.df(Yt,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

-68.867 -18.850 -4.859 21.205 57.345

Coefficients:

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

(Intercept) 75.2917 30.3705 2.479 0.0172 \*

z.lag.1 -0.4102 0.1670 -2.456 0.0182 \*

z.diff.lag -0.1473 0.1600 -0.920 0.3625

---

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

Residual standard error: 31.35 on 43 degrees of freedom

Multiple R-squared: 0.2356, Adjusted R-squared: 0.2

F-statistic: 6.625 on 2 and 43 DF, p-value: 0.003104

Value of test-statistic is: -2.456 3.0737

Critical values for test statistics:

1pct 5pct 10pct

tau2 -3.58 -2.93 -2.60

phi1 7.06 4.86 3.94

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

PARTE ESTACIONAL

data\_adf2 <- ur.df(Yt , lags = 4)

summary(data\_adf2)

###############################################

# Augmented Dickey-Fuller Test Unit Root Test #

###############################################

Test regression none

Call:

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

Residuals:

Min 1Q Median 3Q Max

-39.801 -8.880 -3.702 12.350 38.205

Coefficients:

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

z.lag.1 0.003761 0.014585 0.258 0.79791

z.diff.lag1 -0.452482 0.142315 -3.179 0.00293 \*\*

z.diff.lag2 -0.447225 0.149674 -2.988 0.00490 \*\*

z.diff.lag3 -0.353588 0.152878 -2.313 0.02624 \*

z.diff.lag4 0.492846 0.142303 3.463 0.00134 \*\*

---

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

Residual standard error: 17.13 on 38 degrees of freedom

Multiple R-squared: 0.7785, Adjusted R-squared: 0.7493

F-statistic: 26.71 on 5 and 38 DF, p-value: 1.759e-11

Value of test-statistic is: 0.2579

Critical values for test statistics:

1pct 5pct 10pct

tau1 -2.62 -1.95 -1.61

En la parte regular tenemos t-calculado = -0.036 lo cual es mayor que el t-critico = -1.95 por tanto se rechaza la hipotesis nula de la existencia de la raiz unitaria, es decir que la serie no es estacionaria en media en la estructura Regular.

En la parte ESTACIONAL tenemos t-caculado = 0.2579 lo cual es mayor que el t-critico = -1.95 por tanto se rechaza la hipotesis nula de la existencia de la raiz unitaria, es decir que la serie no es estacionaria en media en la estructura estacional por tanto se necesita realizar la diferenciacion en la parte regular y estacional.

Diferenciamos debido a que la serie no es estacionaria - PARTE REGULAR

par(mfrow=c(1,1))

D1.Yt <- diff(Yt);

plot(D1.Yt,xlab="tiempo", ylab= "Ventas" )

# PRUEBA DE dF para la parte regular

D1.Yt\_adf <- ur.df(D1.Yt, lags = 1)

summary(D1.Yt\_adf)

###############################################

# Augmented Dickey-Fuller Test Unit Root Test #

###############################################

Test regression none

Call:

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

Residuals:

Min 1Q Median 3Q Max

-52.521 -18.555 -6.299 20.322 62.184

Coefficients:

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

z.lag.1 -1.9387 0.2379 -8.148 2.93e-10 \*\*\*

z.diff.lag 0.4195 0.1438 2.917 0.0056 \*\*

---

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

Residual standard error: 30.62 on 43 degrees of freedom

Multiple R-squared: 0.7294, Adjusted R-squared: 0.7168

F-statistic: 57.95 on 2 and 43 DF, p-value: 6.242e-13

Value of test-statistic is: -8.1482

Critical values for test statistics:

1pct 5pct 10pct

tau1 -2.62 -1.95 -1.61

Con la prueba de dickey fuller aumentada observamos que el T calculado (-8.1482) es MENOR que el T critico (-1.95) por tanto, rechazamos la hipotesis nula de no estacionariedad y se concluye que la serie ES ESTACIONARIA EN MEDIA EN SU PARTE REGULAR

Diferenciacion PARTE ESTACIONAL

D4.D1.Yt <- diff(Yt,4)

#raiz unitaria parte estacional

D4.D1.Yt\_adf <- ur.df(D4.D1.Yt,lags = 4)

summary(D4.D1.Yt\_adf)

###############################################

# Augmented Dickey-Fuller Test Unit Root Test #

###############################################

Test regression none

Call:

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

Residuals:

Min 1Q Median 3Q Max

-44.808 -9.101 -0.004 9.938 35.864

Coefficients:

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

z.lag.1 -0.22327 0.18158 -1.230 0.227

z.diff.lag1 -0.32948 0.20987 -1.570 0.126

z.diff.lag2 -0.27542 0.21017 -1.310 0.199

z.diff.lag3 -0.06103 0.19482 -0.313 0.756

z.diff.lag4 -0.24050 0.17027 -1.412 0.167

Residual standard error: 17.02 on 34 degrees of freedom

Multiple R-squared: 0.3237, Adjusted R-squared: 0.2242

F-statistic: 3.254 on 5 and 34 DF, p-value: 0.01656

Value of test-statistic is: -1.2296

Critical values for test statistics:

1pct 5pct 10pct

tau1 -2.62 -1.95 -1.61

Viendo el gráfico vemos que claramente se ha estabilizado en media. Y confirmando esto con la prueba de dickey fuller aumentada observamos que el T calculado (-1.2296) es MAYOR que el T critico (-1.95) por tanto, aceptamos la hipotesis nula de no estacionariedad y se concluye que la serie NO ES ESTACIONARIA EN MEDIA EN SU PARTE ESTACIONAL.

D4D4.D1.Yt <- diff(D4.D1.Yt,4)

D4D4.D1.Yt\_adf <- ur.df(D4D4.D1.Yt,lags = 4)

summary(D4D4.D1.Yt\_adf)

###############################################

# Augmented Dickey-Fuller Test Unit Root Test #

###############################################

Test regression none

Call:

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

Residuals:

Min 1Q Median 3Q Max

-41.061 -9.650 -1.388 10.320 35.235

Coefficients:

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

z.lag.1 -2.9621 0.6247 -4.742 3.94e-05 \*\*\*

z.diff.lag1 1.3885 0.5330 2.605 0.0137 \*

z.diff.lag2 0.9410 0.4320 2.178 0.0366 \*

z.diff.lag3 0.6882 0.3003 2.292 0.0284 \*

z.diff.lag4 0.2574 0.1702 1.512 0.1400

---

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

Residual standard error: 17.07 on 33 degrees of freedom

Multiple R-squared: 0.7568, Adjusted R-squared: 0.7199

F-statistic: 20.53 on 5 and 33 DF, p-value: 2.817e-09

Value of test-statistic is: -4.742

Critical values for test statistics:

1pct 5pct 10pct

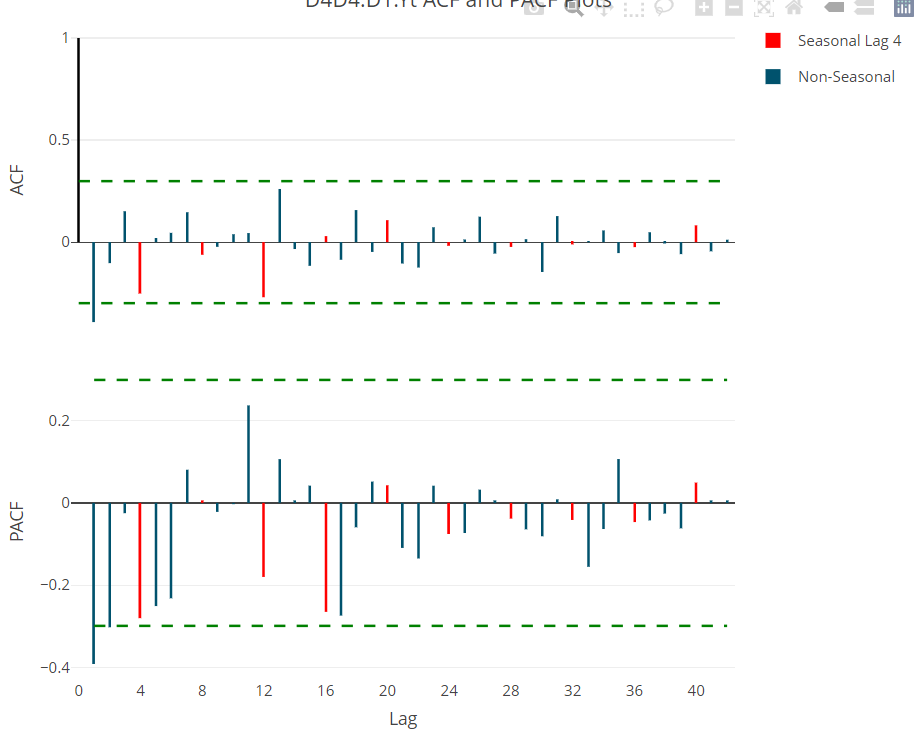
tau1 -2.62 -1.95 -1.61

Tenemos t-calculado = -4.742 lo cual es menor que el t-critico = -1.95 por tanto se acepta la hipotesis nula de la existencia de la raiz unitaria el cual se corrobora con el grafico.

### 1.3 Identificación del modelo estacionario

#### 1.3.1 Identificación de las órdenes p y q

ts\_cor(D4D4.D1.Yt,lag.max = 50)



Planteamos los modelos:

MODELO 1: SARIMA (0,1,1)(0,1,1)4

MODELO 2: SARIMA (1,1,0)(0,1,1)4

MODELO 3: SARIMA (0,1,0)(0,1,1)4

MODELO 4: SARIMA (1,1,1,)(0,1,0,)4

## 2 Estimación

modelo1 <- Arima(Yt, order = c(0,1,1), seasonal = list(order = c(1,2,1)))

coeftest(modelo1)

z test of coefficients:

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

ma1 -0.64678 0.12833 -5.0400 4.656e-07 \*\*\*

sar1 -0.22605 0.16566 -1.3645 0.1724

sma1 -0.99999 0.18101 -5.5246 3.302e-08 \*\*\*

---

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

modelo2 <- Arima(Yt, order = c(0,1,1), seasonal = list(order = c(0,1,1)))

coeftest(modelo2)

z test of coefficients:

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

ma1 -0.53259 0.15051 -3.5385 0.0004024 \*\*\*

sma1 -0.56678 0.27880 -2.0329 0.0420583 \*

---

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

modelo3 <- Arima(Yt, order = c(0,1,1), seasonal = list(order = c(0,1,0)))

coeftest(modelo3)

z test of coefficients:

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

ma1 -0.69300 0.11788 -5.8789 4.129e-09 \*\*\*

---

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

modelo4 <- Arima(Yt, order = c(0,1,0), seasonal = list(order = c(0,1,1)))

coeftest(modelo4)

z test of coefficients:

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

sma1 -1.00000 0.33424 -2.9919 0.002773 \*\*

---

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

## 3 Validación

### 3.1 Análisis de los coeficientes estimados

#### 3.1.1 Significación de los coeficientes

coeftest(modelo1)

z test of coefficients:

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

ma1 -0.64678 0.12833 -5.0400 4.656e-07 \*\*\*

sar1 -0.22605 0.16566 -1.3645 0.1724

sma1 -0.99999 0.18101 -5.5246 3.302e-08 \*\*\*

---

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

coeftest(modelo2)

z test of coefficients:

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

ma1 -0.53259 0.15051 -3.5385 0.0004024 \*\*\*

sma1 -0.56678 0.27880 -2.0329 0.0420583 \*

---

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

coeftest(modelo3)

z test of coefficients:

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

ma1 -0.69300 0.11788 -5.8789 4.129e-09 \*\*\*

---

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

coeftest(modelo4)

z test of coefficients:

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

sma1 -1.00000 0.33424 -2.9919 0.002773 \*\*

---

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

#### 3.1.2 Examen de la matriz de correlaciones entre los coeficientes

vcov(mod1)

sar1 sar2

sar1 0.008191179 0.003648156

sar2 0.003648156 0.008853282

vcov(mod2)

sar1 sar2 sar3

sar1 0.007928069 0.005058438 0.002110606

sar2 0.005058438 0.009901841 0.004191203

sar3 0.002110606 0.004191203 0.009298398

vcov(mod3)

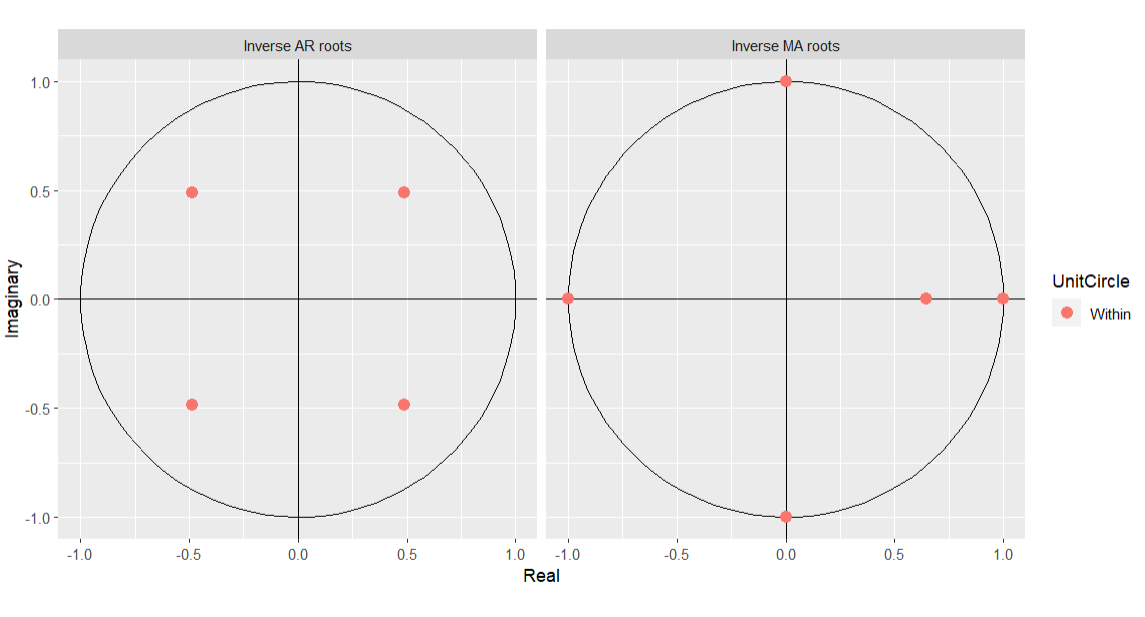
sma1

sma1 0.01084332

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.

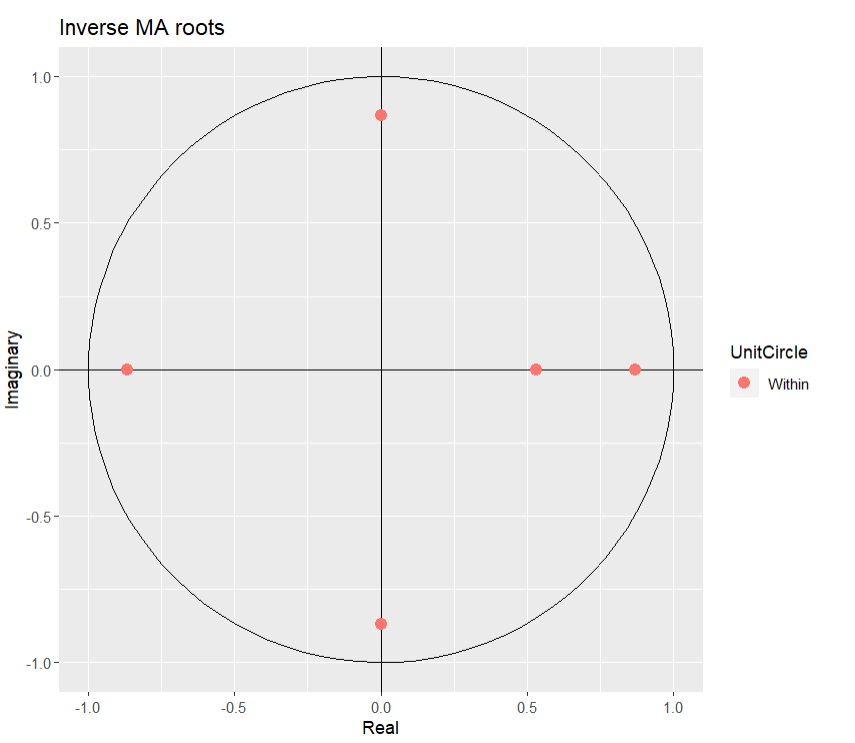
#### 3.1.3 Condición de convergencia e invertibilidad

autoplot(mod1)



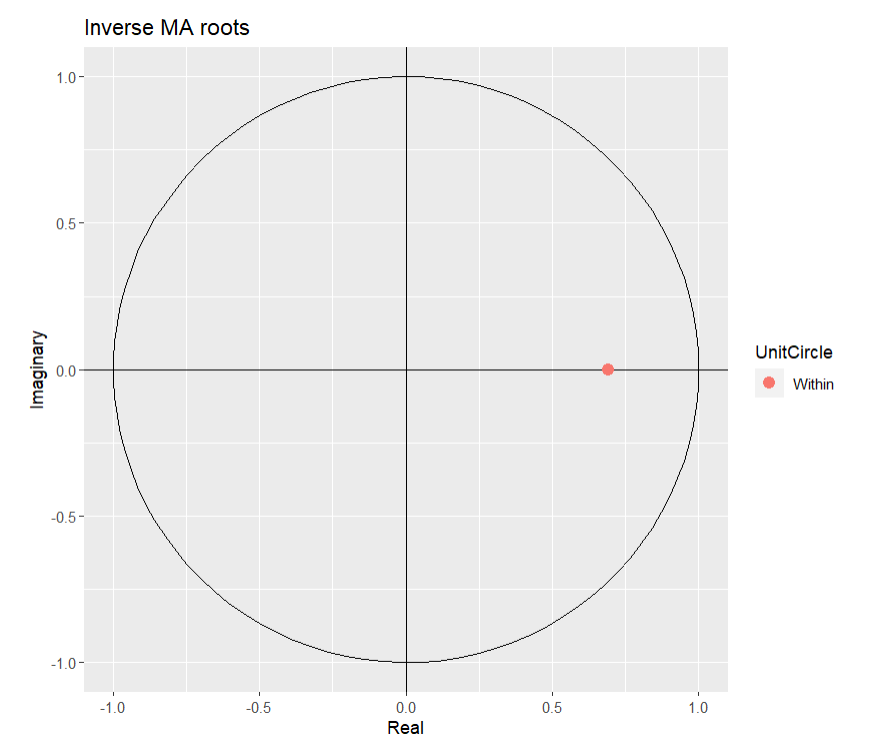
Dado para el primer modelo dado que los puntos estan dentro del circulo indica que son invertibles y estacinarios.

autoplot(mod2)



Para el segundo modelo ya que solo presenta la parte de ma() infiere que la parte autoregresiva ya presenta invertibilidad y estacionariedad. Pero segun el grafico se puede observar que que todo los puntos estan dentro del circulo señal clara de invertibilidad y estacionariedad.

autoplot(mod3)



Dado que todo los puntos están dentro del circulo lo cual no indica que son invertibles y estacionarios.

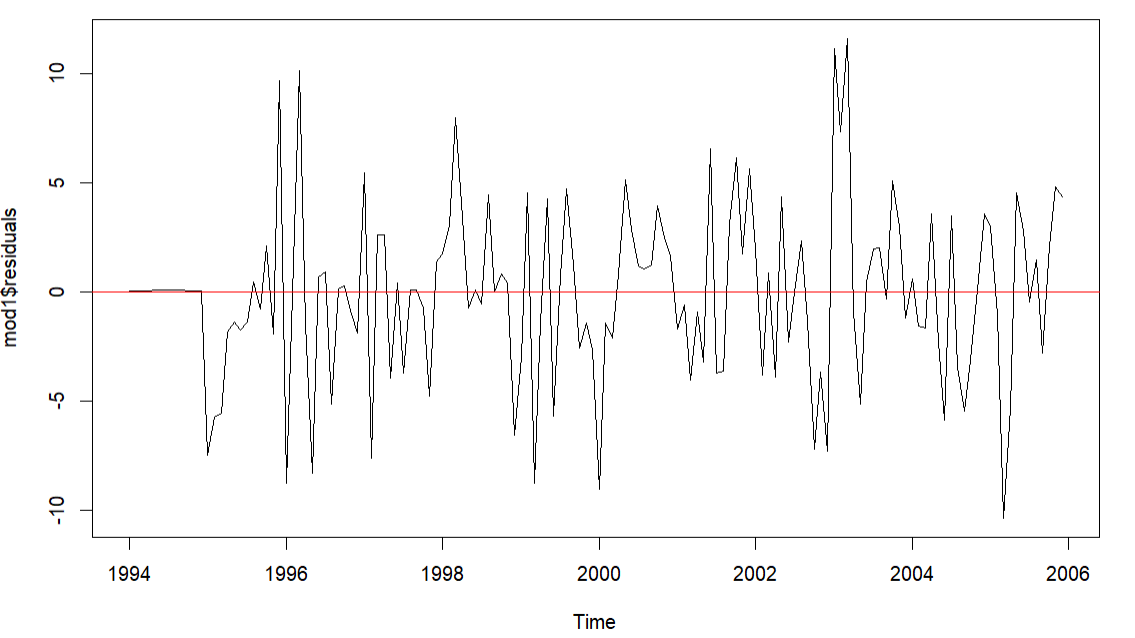
### 3.2 Análisis de los residuos

#### 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 = -0.3252, df = 143, p-value = 0.7455

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

-0.7670199 0.5502975

sample estimates:

mean of x

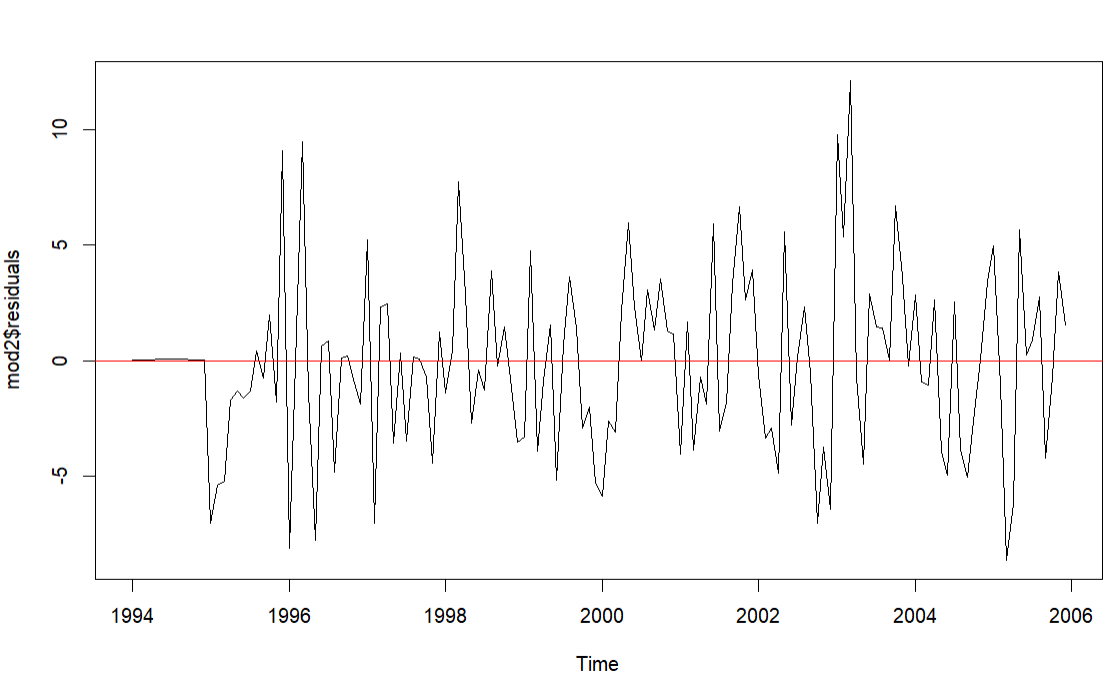
-0.1083612

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 = 0.7455 > 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 = -0.45491, df = 143, p-value = 0.6499

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

-0.7584022 0.4746337

sample estimates:

mean of x

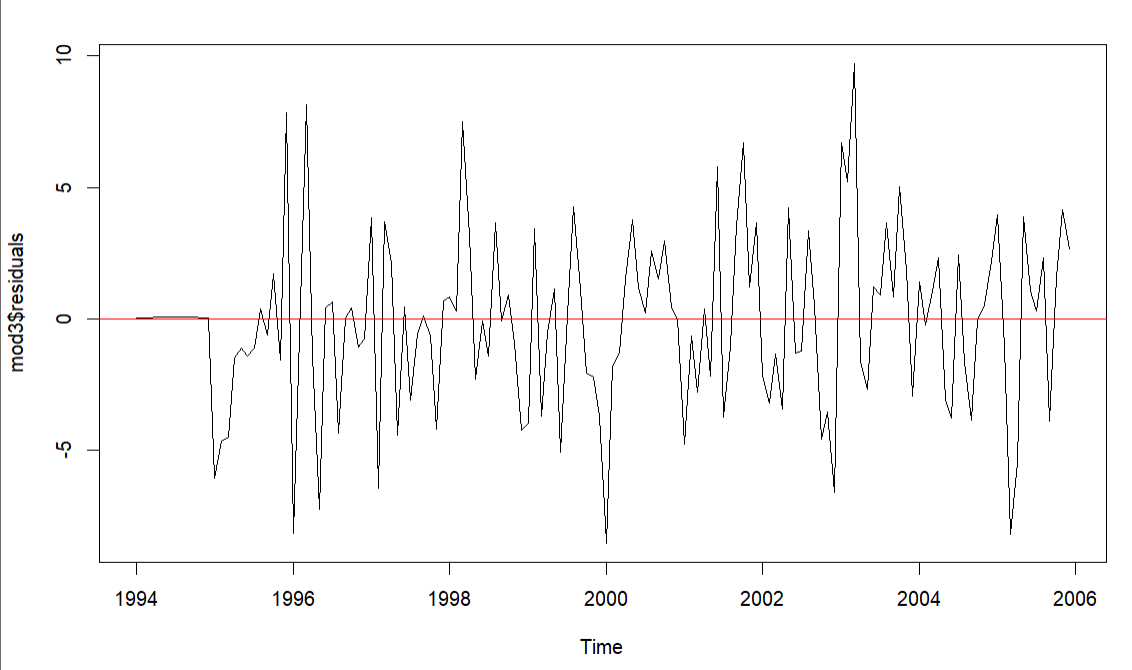
-0.1418842

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 = 0.6499 > α = 0.05, se acepta Ho, es decir la media es igual a cero.

plot(mod3$residuals)

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

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



One Sample t-test

data: mod3$residuals

t = -0.58811, df = 143, p-value = 0.5574

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

-0.7100319 0.3844116

sample estimates:

mean of x

-0.1628102

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 = 0.5574 > α = 0.05, se acepta Ho, es decir la media es igual a cero.

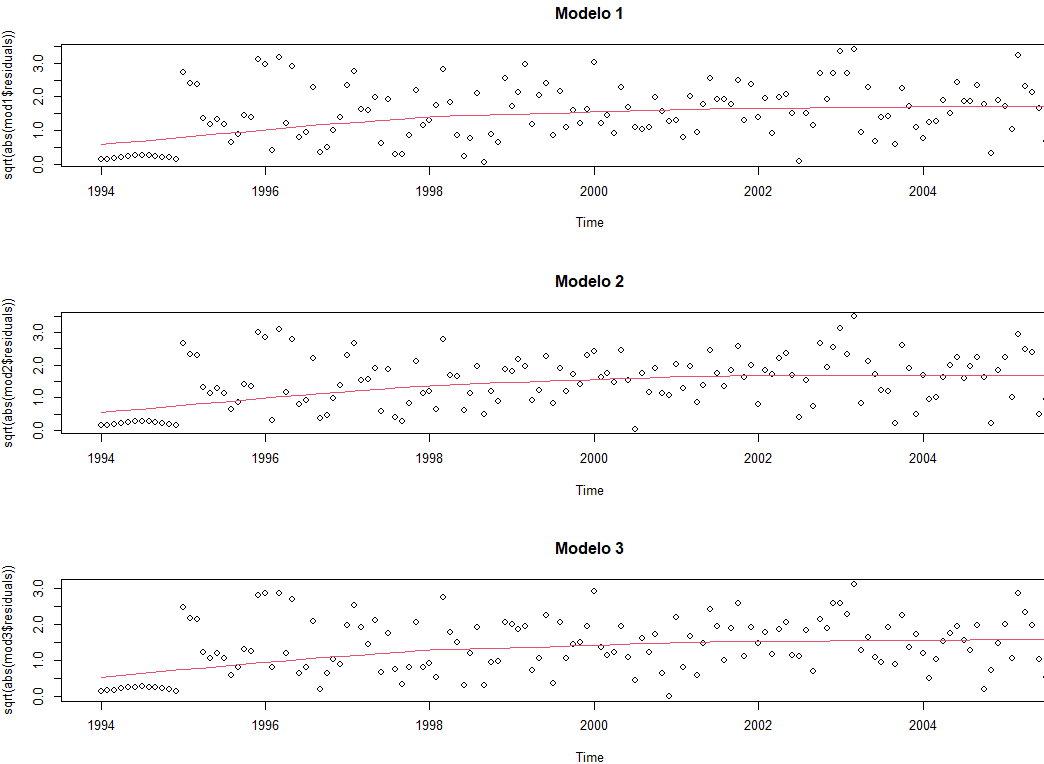
#### 3.2.2 Homocedasticidad o varianza constante

par(mfrow = c(3,1))

scatter.smooth(sqrt(abs(mod1$residuals)), lpars=list(col=2), main = "Modelo 1")

scatter.smooth(sqrt(abs(mod2$residuals)), lpars=list(col=2), main = "Modelo 2")

scatter.smooth(sqrt(abs(mod3$residuals)), lpars=list(col=2), main = "Modelo 3")



Se observa que los datos presentan una variabilidad considerable.

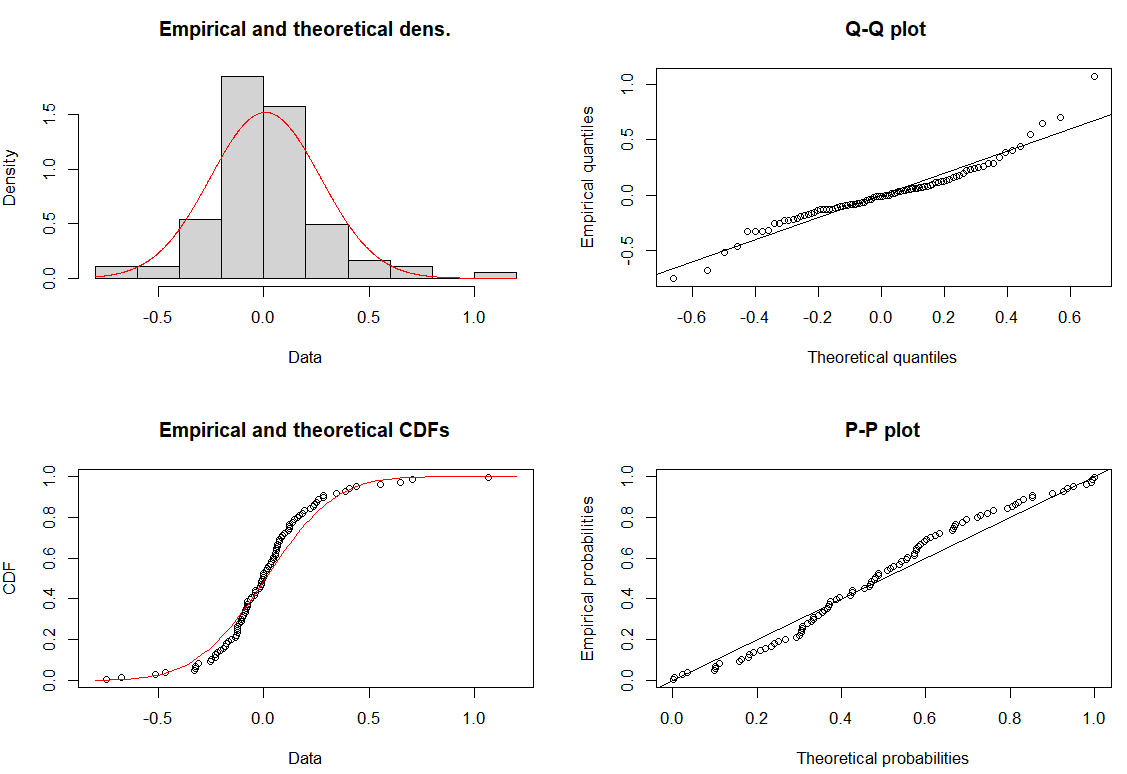
#### 3.2.4 Contraste de normalidad

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

plot(ajuste\_m1)

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

JB\_m1



Jarque Bera Test

data: resid\_m1

X-squared = 41.778, df = 2, p-value = 8.471e-10

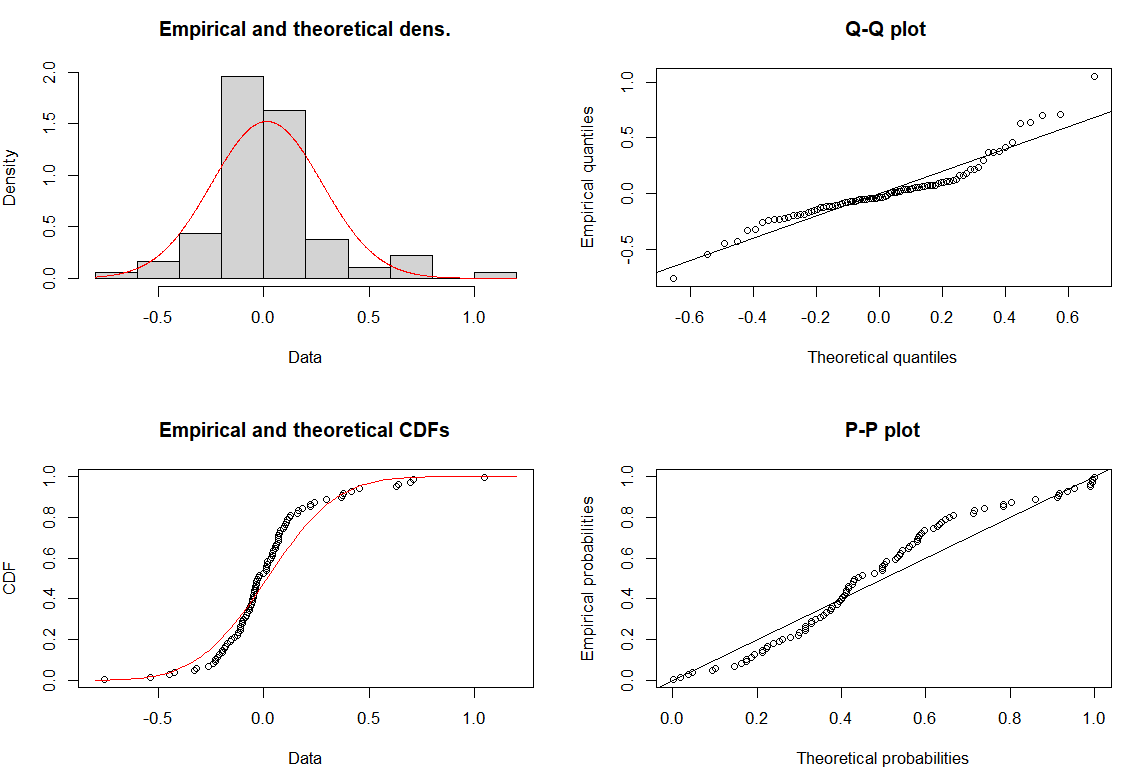
En las figuras se observa que los residuales del modelo 1 no presentan indicios de ajuste hacia la distribución normal. En la prueba JB, como p = 0.0000 < 0.05, se RECHAZA Ho, es decir, los residuos NO se aproximan a una distribución normal.

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

plot(ajuste\_m2)

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

JB\_m2



Jarque Bera Test

data: resid\_m2

X-squared = 47.925, df = 2, p-value = 3.919e-11

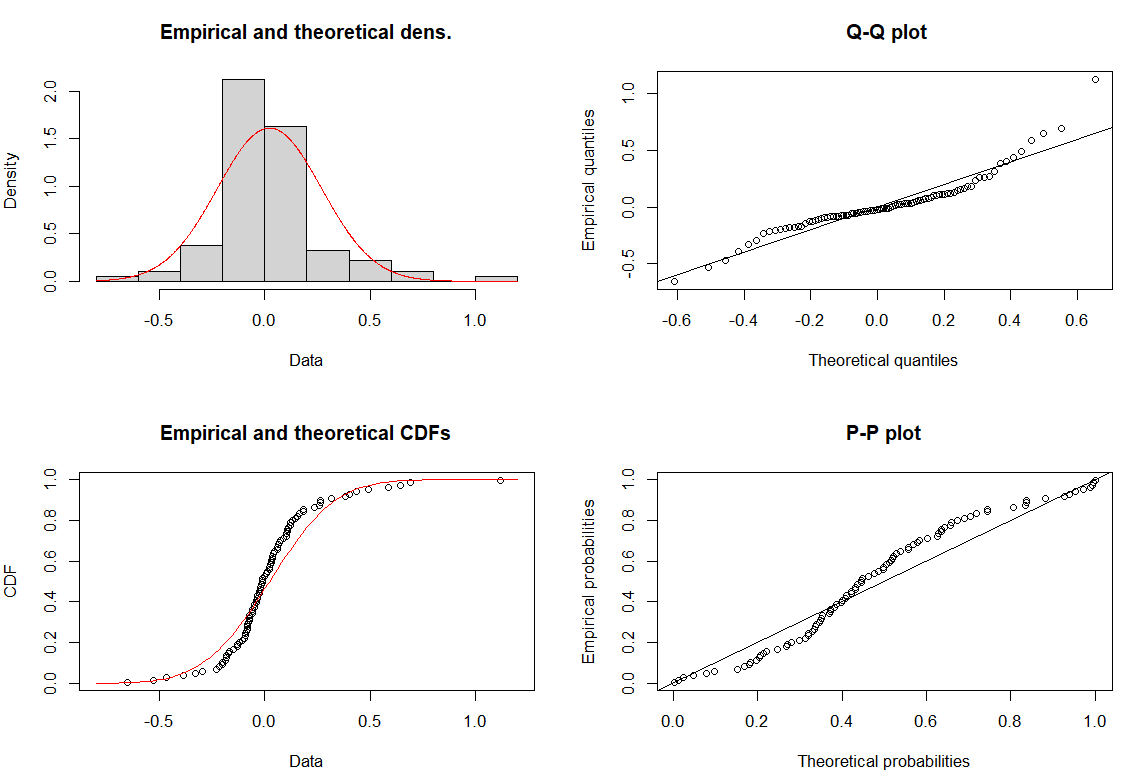
En las figuras se observa que los residuales del modelo 2 no presentan indicios de ajuste hacia la distribución normal. En la prueba JB, como p = 0.00011 < 0.05, se RECHAZA Ho, es decir, los residuos NO se aproximan a una distribución normal.

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

plot(ajuste\_m3)

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

JB\_m3



Jarque Bera Test

data: resid\_m3

X-squared = 87.238, df = 2, p-value < 2.2e-16

En las figuras se observa que los residuales del modelo 3 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

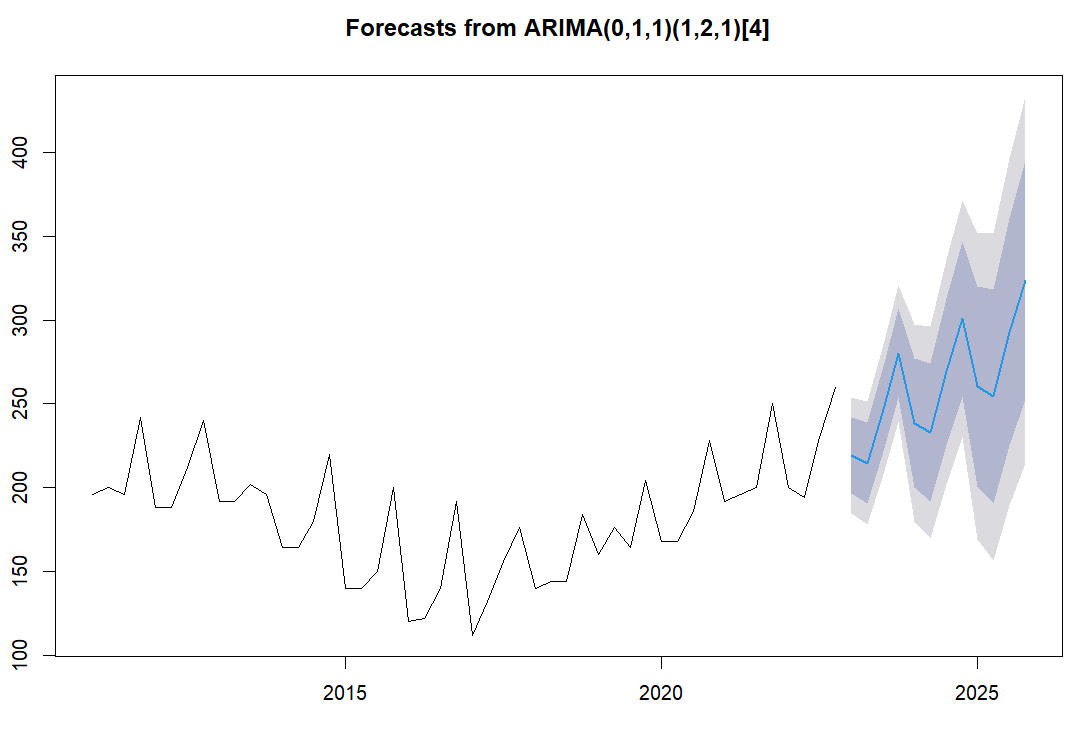
#### Modelo 1:

**Pronóstico para la serie**

Pron <- forecast(modelo1,h=12)

plot(Pron)

summary(Pron)

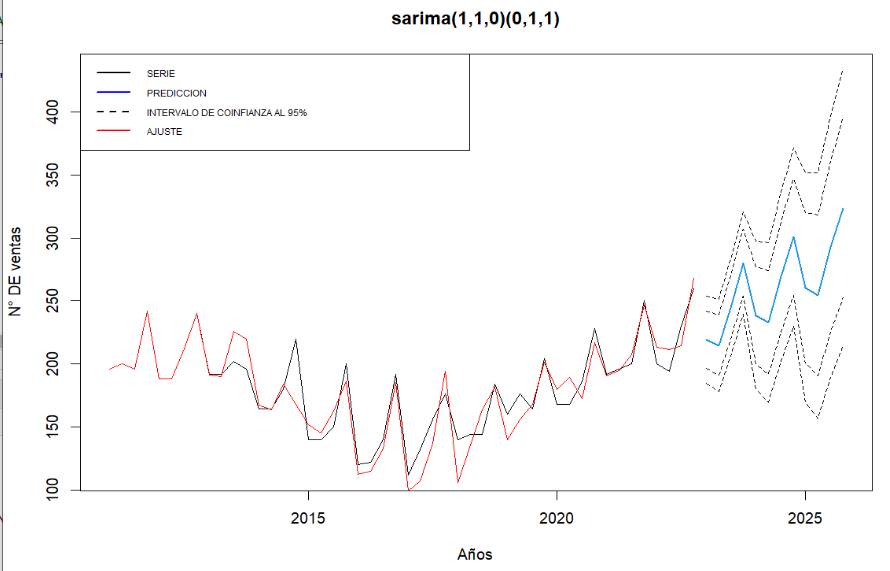


**Pronostico vs valores reales**

plot(Pron, shaded = FALSE, xlab = "Años", ylab = "N° DE ventas",main = "sarima(1,1,0)(0,1,1)")

lines(Pron$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)



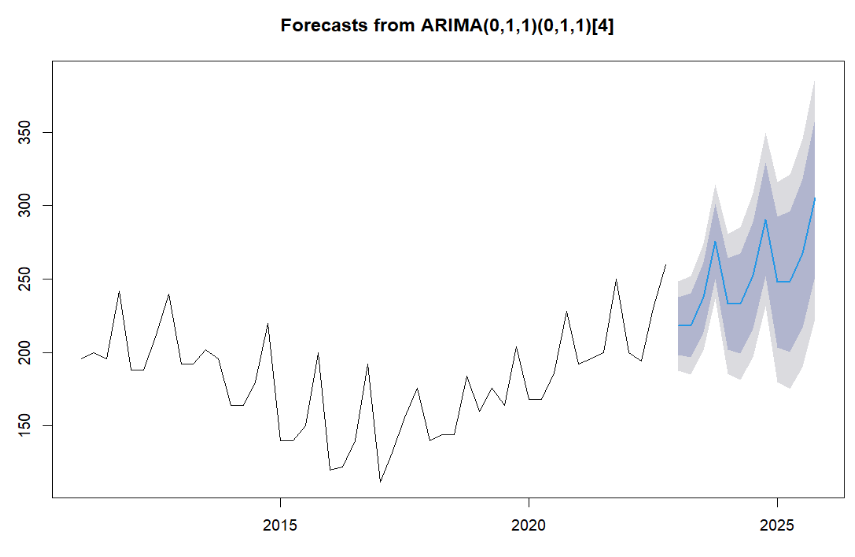
#### Modelo 2:

**Pronóstico para la serie**

Pron2 <- forecast(modelo2,h=12)

plot(Pron2)

summary(Pron2)

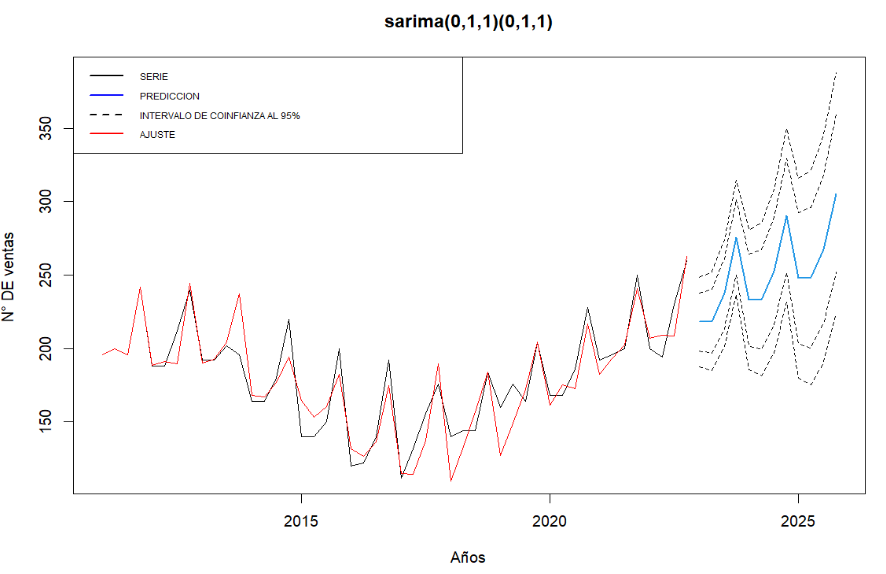


**Pronostico vs valores reales**

plot(Pron2, shaded = FALSE, xlab = "Años", ylab = "N° DE ventas ",main = "sarima(0,1,1)(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)



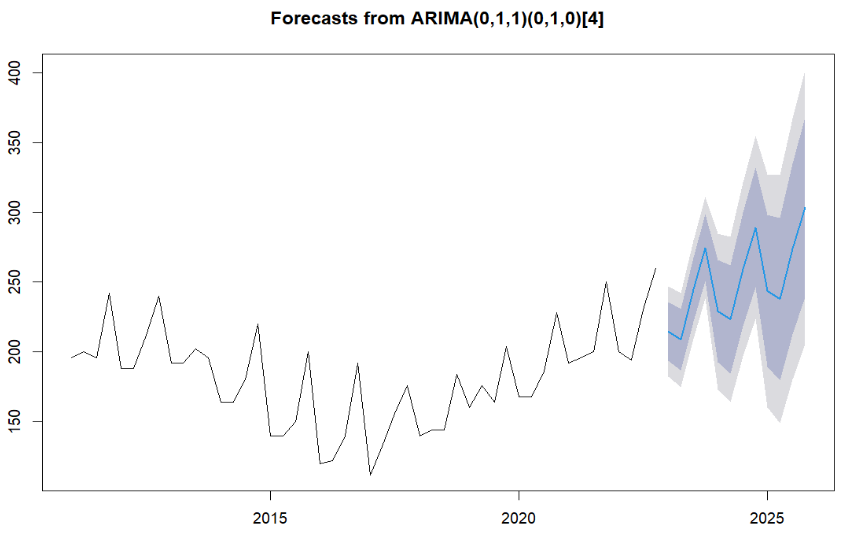
#### Modelo 3:

**Pronóstico para la serie**

Pron3 <- forecast(modelo3,h=12)

plot(Pron3)

summary(Pron3)

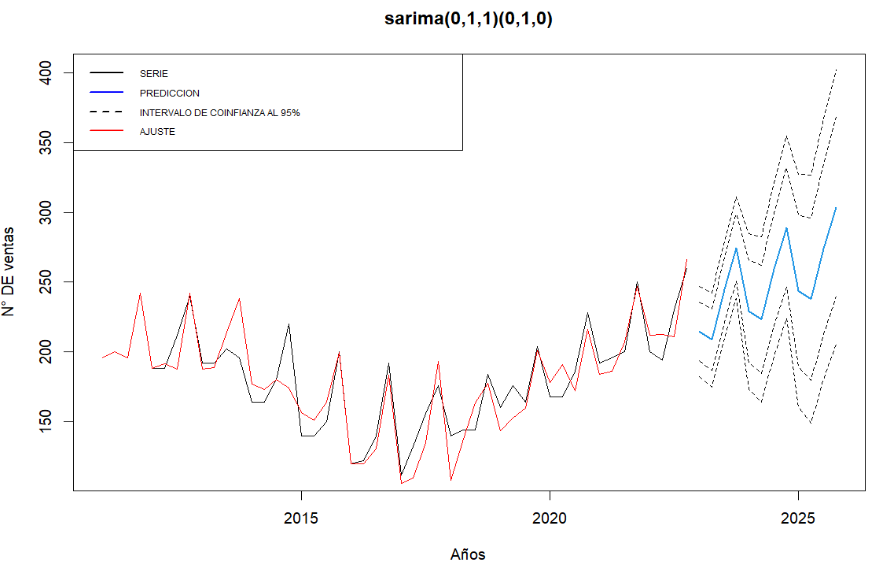
****

**Pronostico vs valores reales**

plot(Pron3, shaded = FALSE, xlab = "Años", ylab = "N° DE ventas",main = "sarima(0,1,1)(0,1,0)")

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)



### Métricas del modelo

accuracy(modelo1)

ME RMSE MAE MPE MAPE MASE ACF1

Training set 1.72136 14.63815 10.3043 1.177992 6.075034 0.5981389 0.0003568487

accuracy(modelo2)

ME RMSE MAE MPE MAPE MASE ACF1

Training set 2.217023 14.30075 10.39466 1.054772 6.066516 0.6033838 -0.03435107

accuracy(modelo3)

ME RMSE MAE MPE MAPE MASE ACF1

Training set 1.409756 15.37879 11.3276 0.7839319 6.51894 0.6575385 0.06208574

AIC(modelo1);BIC(modelo1)

[1] 348.2515

[1] 354.9057

AIC(modelo2);BIC(modelo2)

[1] 363.5213

[1] 368.8049

AIC(modelo3);BIC(modelo3)

[1] 366.4493

[1] 369.9717

## Conclusión

1. **Modelo 1:**
   * **RMSE:** 14.63815
   * **MAE:** 10.3043
   * **AIC:** 348.2515
   * **BIC:** 354.9057
2. **Modelo 2:**
   * **RMSE:** 14.30075
   * **MAE:** 10.39466
   * **AIC:** 363.5213
   * **BIC:** 368.8049
3. **Modelo 3:**
   * **RMSE:** 15.37879
   * **MAE:** 11.3276
   * **AIC:** 366.4493
   * **BIC:** 369.9717

Basado en estas métricas y criterios:

* El **Modelo 2** tiene el RMSE más bajo (14.30075) y un MAE comparativamente bajo (10.39466).
* Además, tiene el AIC más bajo (363.5213) y el BIC más bajo (368.8049), lo que indica un mejor ajuste del modelo en comparación con los otros dos.

Por lo tanto, según estas métricas y criterios, el **Modelo 2** sería la elección recomendada debido a su mejor desempeño general en la evaluación de la serie temporal.