[CASO 1: TEMPERATURAS, Dubuque – EE.UU. 5](#_Toc171815155)

[1) IDENTIFICACION 5](#_Toc171815156)

[INTERPRETACION 6](#_Toc171815157)

[INTERPRETACION 6](#_Toc171815158)

[ANALISIS DE COMPONENTES 7](#_Toc171815159)

[ESTACIONALIDAD 7](#_Toc171815160)

[INTERPRETACION 8](#_Toc171815161)

[CICLICIDAD 8](#_Toc171815162)

[INTERPRETACION 9](#_Toc171815163)

[ESTACIONARIEDAD EN VARIANZA 9](#_Toc171815164)

[INTERPRETACION 9](#_Toc171815165)

[ESTACIONARIDAD EN MEDIA 9](#_Toc171815166)

[INTEPRETACION 10](#_Toc171815167)

[PRUEBA DE DICKEY-FULLER AUMENTADO (ADF) PARA LA ESTRUCTURA REGULAR 10](#_Toc171815168)

[INTERPRETACION 11](#_Toc171815169)

[LA PRUEBA ADF PARA LA SERIE YT PARA LA ESTRUCTURA ESTACIONAL. 11](#_Toc171815170)

[INTERPRETACION 12](#_Toc171815171)

[INTERPRETACION 13](#_Toc171815172)

[2) ESTIMACION 14](#_Toc171815173)

[SELECCION DE ORDENES 14](#_Toc171815174)

[MODELOS PROPUESTOS 14](#_Toc171815175)

[SIGNIFICANCIA DE LOS COEFICIENTES 14](#_Toc171815176)

[MODELO 1 14](#_Toc171815177)

[INTEPRETACION 15](#_Toc171815178)

[MODELO 2 15](#_Toc171815179)

[INTEPRETACION 15](#_Toc171815180)

[MODELO 3 15](#_Toc171815181)

[INTEPRETACION 16](#_Toc171815182)

[MODELO 4 16](#_Toc171815183)

[INTEPRETACION 16](#_Toc171815184)

[MODELO 5 16](#_Toc171815185)

[INTEPRETACION 17](#_Toc171815186)

[MODELOS SELECCIONADOS 17](#_Toc171815187)

[3) VALIDACION 17](#_Toc171815188)

[ANALISIS DE LA MULTICOLINEALIDAD 17](#_Toc171815189)

[INTERPRETACION 17](#_Toc171815190)

[ANALISIS DE ESTABILIDAD 18](#_Toc171815191)

[MODELO 1 18](#_Toc171815192)

[MODELO 2 18](#_Toc171815193)

[MODELO 3 18](#_Toc171815194)

[INTERPRETACION 19](#_Toc171815195)

[PRUEBA DE LJUNG-BOX 19](#_Toc171815196)

[MODELO 1 19](#_Toc171815197)

[MODELO 2 19](#_Toc171815198)

[MODELO 3 20](#_Toc171815199)

[INTEPRETACION 21](#_Toc171815200)

[NORMALIDAD DE LOS RESIDUALES 21](#_Toc171815201)

[MODELO 1 21](#_Toc171815202)

[MODELO 2 22](#_Toc171815203)

[MODELO 3 23](#_Toc171815204)

[INTEPRETACION 23](#_Toc171815205)

[SELECCION DEL MEJOR MODELO 23](#_Toc171815206)

[INTERPRETACION 24](#_Toc171815207)

[4) PRONOSTICO 24](#_Toc171815208)

[PRONOSTICO PARA LA SERIE 24](#_Toc171815209)

[MODELO 1 24](#_Toc171815210)

[MODELO 2 25](#_Toc171815211)

[MODELO 3 27](#_Toc171815212)

[SERIE ORIGINAL (YT) Y PRONOSTICADA. 28](#_Toc171815213)

[MODELO 1 28](#_Toc171815214)

[MODELO 2 28](#_Toc171815215)

[MODELO 3 29](#_Toc171815216)

[GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES 29](#_Toc171815217)

[MODELO 1 29](#_Toc171815218)

[MODELO 2 30](#_Toc171815219)

[MODELO 3 31](#_Toc171815220)

[INTERPRETACION 32](#_Toc171815221)

[CONCLUSION 32](#_Toc171815222)

[EVALUACIÓN DE MODELOS SARIMA 32](#_Toc171815223)

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

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)

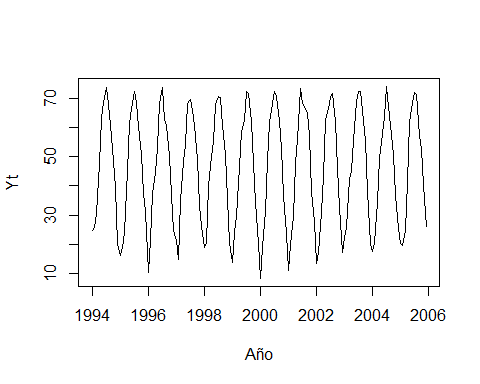
library(strucchange)#para analisis de estabilidad - Chow

library(survival)  
library(fitdistrplus)

library(readxl)  
  
  
datos <- read\_excel("D:/SERIES DE TIEMPO/ACTIVIDAD 7/1.xlsx")  
View(datos)

# IDENTIFICACION

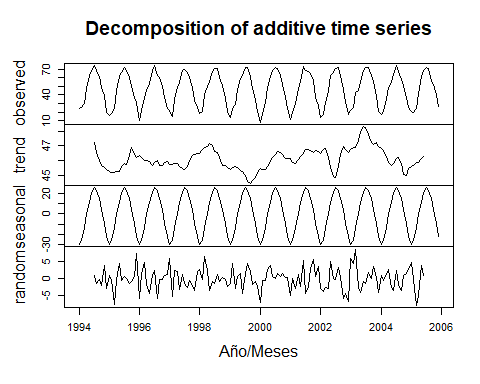
# Graficar la serie   
Yt <- ts(datos$Yt, start = c(1994, 1), frequency = 12)   
plot(Yt, xlab = "Año", ylab ="Yt" )



##### INTERPRETACION

Como podemos ver la serie tiene un claro componente estacional, y aparente estacionariedad en varianza y media, en cuanto al componente ciclico existen algunos indicios del mismo sin embargo nada claro. Finalmente podemos observar la ausencia total en cuanto a la tendencia.

# Descomposicion de la serie   
Yt\_desc <- decompose(Yt, type = "additive")  
plot(Yt\_desc, xlab = "Año/Meses")



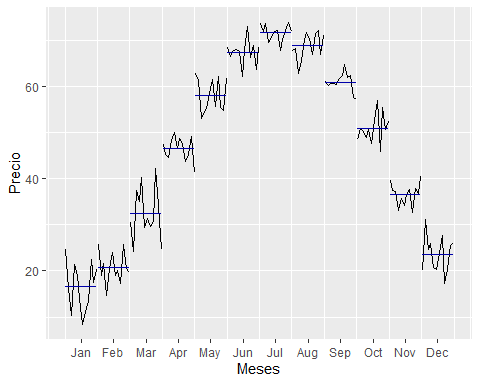
##### INTERPRETACION

Tras la descomposicion de la serie podemos corroborar algunos de los comportamiento descritos en un inicio.

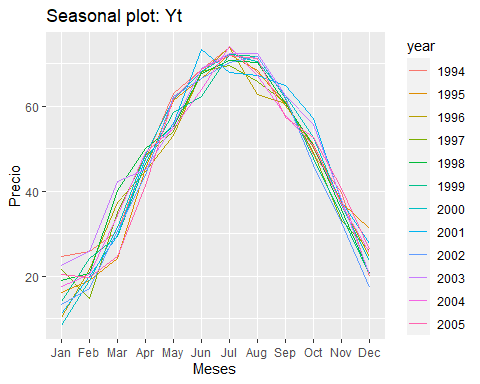
## ANALISIS DE COMPONENTES

### ESTACIONALIDAD

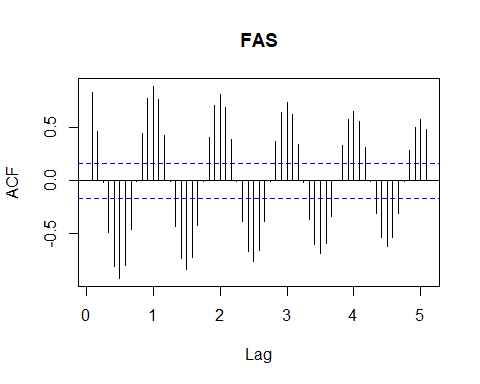
ggsubseriesplot(Yt, xlab = "Meses", ylab = "Precio" )



ggseasonplot(Yt, xlab = "Meses",ylab = "Precio" )



FAS <- acf(Yt, lag.max = 61, main="FAS", level = 0.95)

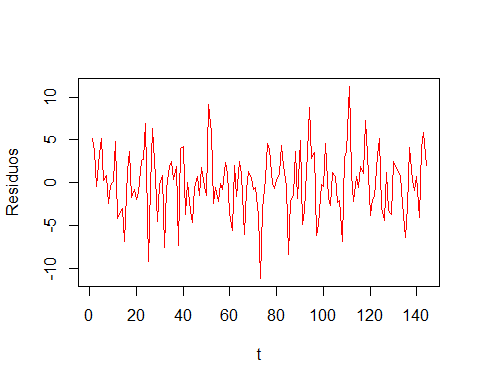


##### INTERPRETACION

Tanto en el grafico de lineas apiladas y el de lineas separadas se observan claron indicios de estacionalidad en la serie temporal, siendo esto corroborado en el grafico del correlograma cuyo comportamiento refleja el de una serie estacionaria teniendo un descenso exponencial a 0 ademas de una comportamiento sinusoidal propio de una serie estacionaria.

### CICLICIDAD

# Generación de t  
datos$t <- seq(1:NROW(datos))  
# Construir series  
cosP <- cos(2\*pi/144\*12\*datos$t)  
senP <- sin(2\*pi/144\*12\*datos$t)  
# Ajuste del modelo  
ciclo <- lm(Yt ~ cosP + senP)  
plot(ciclo$residuals, type = "l", xlab="t", ylab="Residuos", col = "red")

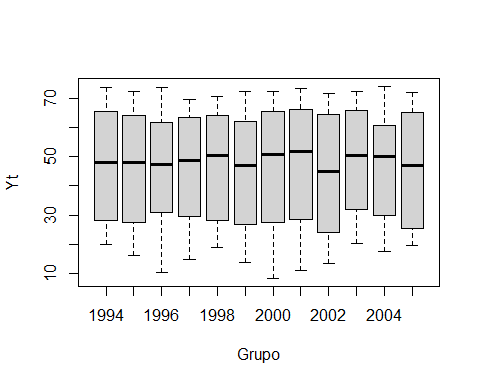


##### INTERPRETACION

Dado que la serie tiene una duracion mayor a un año es necesario hacer un analisis de la ciclidad de la misma donde despues de construir la serie respecto a sus funciones seno y coseno podemos observar ciertos indicios de ciclicidad cada 5 aproximadamente empezando desde el segundo año respecto a la serie temporal.

### ESTACIONARIEDAD EN VARIANZA

boxplot(datos$Yt ~ datos$año, xlab = "Grupo", ylab = "Yt" )



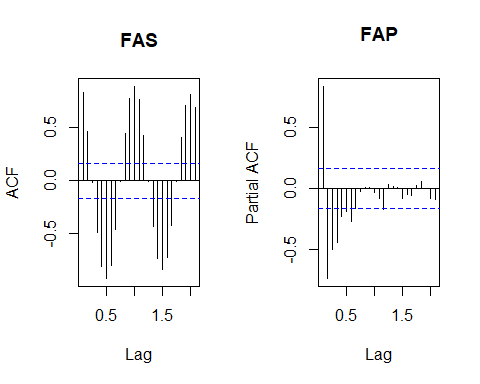
##### INTERPRETACION

Dado el grafico podemos inferir que este tiene un comportamiento propio de una serie que es estacionaria en varianza.

### ESTACIONARIDAD EN MEDIA

par(mfrow=c(1,2))  
FAS <- acf(Yt,lag.max = 25, main = "FAS", level = 0.95)

FAP <- pacf(Yt, lag.max = 25, main = "FAP", level = 0.95 )



FAP$acf[1]

## [1] 0.8318019

##### INTEPRETACION

Dado el grafico FAS observamos que este decrece rapidamente a 0 en el tercer retardo añadido al comportamiento oscilatorio del mismo tenemos indicios de que la serie es estacionaria, esto confirmado dado le valor del primer coeficiente resultante de las FAP cuyo valor valor es menor a 0.9 lo cual nos confirma que efectivamente la serie es estacionaria en media.

#### *PRUEBA DE DICKEY-FULLER AUMENTADO (ADF) PARA LA 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

##### INTERPRETACION

Como -2.9215 < -1.95 entonces rechazamos la hipótesis nula de no estacionariedad y se concluye que, en efecto, la serie temporal es estacionaria.

#### *LA PRUEBA ADF PARA LA SERIE YT PARA LA ESTRUCTURA ESTACIONAL.*

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

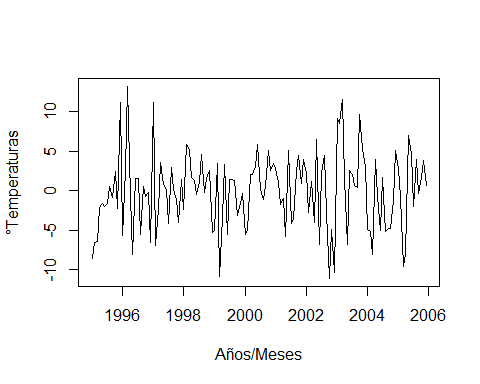
##   
## ###############################################   
## # Augmented Dickey-Fuller Test Unit Root Test #   
## ###############################################   
##   
## Test regression none   
##   
##   
## Call:  
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
##   
## Residuals:  
## 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

##### INTERPRETACION

Como -0.1581 > -1.95 entonces aceptamos la hipótesis nula de no estacionariedad y se concluye que, la serie temporal no es estacionaria en su estructura estacional.

###### Dado que la serie no es estacionaria en su parte estacionaria procedemos a la diferenciacia estacional de la serie. DIFERENCIACION 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

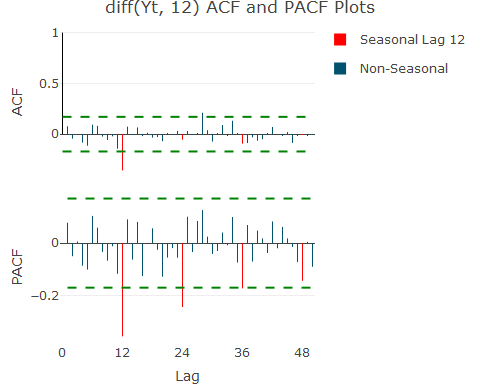
##### INTERPRETACION

Como -4.523 < -1.95 entonces rechazamos la hipótesis nula de no estacionariedad y se concluye que, la serie temporal es estacionaria en su estructura estacional.

## ESTIMACION

### SELECCION DE ORDENES

ts\_cor(diff(Yt,12), lag.max = 50)



### MODELOS PROPUESTOS

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

### SIGNIFICANCIA DE LOS COEFICIENTES

#### MODELO 1

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

##   
## z test of coefficients:  
##   
## Estimate Std. Error z value Pr(>|z|)   
## ma1 0.111808 0.096736 1.1558 0.2478   
## sar1 0.029919 0.097903 0.3056 0.7599   
## sar2 -0.149410 0.099714 -1.4984 0.1340   
## sma1 -0.999997 0.123433 -8.1015 5.427e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##### INTEPRETACION

Dado el modelo SARIMA (0,0,1)x(2,1,1)12 observamos que el unico coeficiente signicativo es es el SMA(1) . Sin embargo el modelo en general no es significativo dado los valores de los demas coeficientes.

#### MODELO 2

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

##   
## z test of coefficients:  
##   
## Estimate Std. Error z value Pr(>|z|)   
## ma1 0.105256 0.096834 1.0870 0.27705   
## sar1 -0.033879 0.099010 -0.3422 0.73222   
## sar2 -0.179178 0.096085 -1.8648 0.06221 .   
## sar3 -0.196649 0.107782 -1.8245 0.06808 .   
## sma1 -0.999996 0.204416 -4.8920 9.984e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##### INTEPRETACION

Dado el modelo SARIMA (0,0,1)x(3,1,1)12 observamos que el unico coeficiente signicativo es nuevamente el SMA(1) y poco significativos los coeficientes asociados a SAR(2) y SAR(3) sin embargo no son significativos para el modelo ya que en general el modelo no se ajusta bien a la serie temporal.

#### MODELO 3

mod3 = Arima(Yt, order = c(0,0,0), seasonal =list(order = c(2,1,0)))  
coeftest(mod3)

##   
## 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

##### INTEPRETACION

Dado el modelo SARIMA (0,0,0)x(2,1,0)12 podemos inferenciar que tanto los coeficientes asociados al SAR(1) y SAR(2) son signifcativos lo que nos indicios de que el modelo logra ajustarse de manera adecuada a la serie temporal.

#### MODELO 4

mod4 = Arima(Yt, order = c(0,0,0), seasonal =list(order = c(3,1,0)))  
coeftest(mod4)

##   
## 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

##### INTEPRETACION

Dado el modelo SARIMA (0,0,0)x(3,1,0)12 observamos que los coeficientes asociados a SAR(1), SAR(2) y SAR(3) son significativos para el modelo lo que nos indica que el modelo en general tiene un buen ajuste dada la serie temporal.

#### MODELO 5

mod5 = Arima(Yt, order = c(0,0,0), seasonal =list(order = c(0,1,1)))  
coeftest(mod5)

##   
## 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

##### INTEPRETACION

Dado el modelo SARIMA (0,0,0)x(2,1,0)12 podemos inferenciar que tanto los coeficientes asociados al SAR(1) y SAR(2) son signifcativos lo que nos indicios de que el modelo logra ajustarse de manera adecuada a la serie temporal.

### MODELOS SELECCIONADOS

Dados los resultados de la estimacion de los coeficientes de los modelos propuestos seleccionamos los siguientes ya que estos presentan un ajuste considerable respecto la serie temporal.

* 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

### VALIDACION

### ANALISIS DE LA MULTICOLINEALIDAD

vcov(mod3)

## sar1 sar2  
## sar1 0.008191179 0.003648156  
## sar2 0.003648156 0.008853282

vcov(mod4)

## 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(mod5)

## sma1  
## sma1 0.01084332

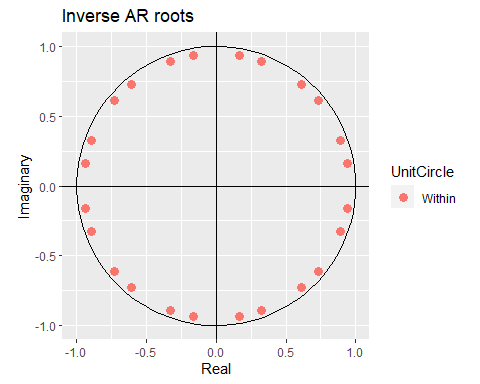
##### INTERPRETACION

Dados los coeficientes asociados a un nuestras variables consideradas en cada modelos podemos observar que ninguna tiene un valor considerable respecto a la multicolinealidad lo que indica una independencia considerable respecto a las demas variables predictoras y un posterior buen modelo predictor.

### ANALISIS DE ESTABILIDAD

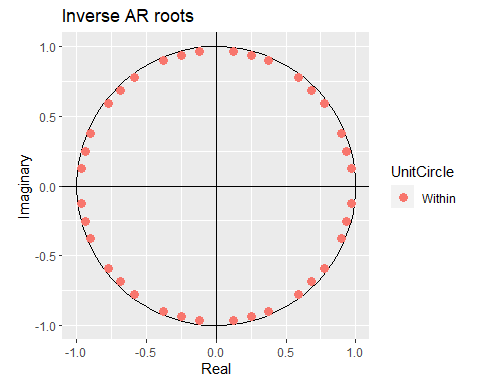
#### MODELO 1

autoplot(mod3)



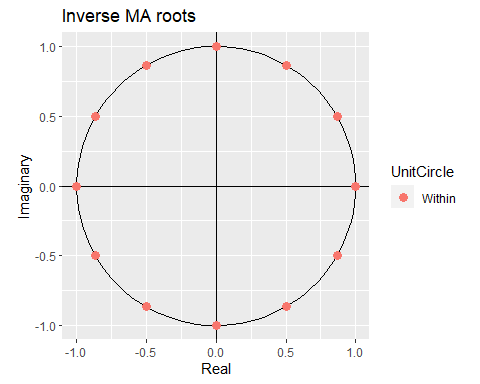
#### MODELO 2

autoplot(mod4)



#### MODELO 3

autoplot(mod5)



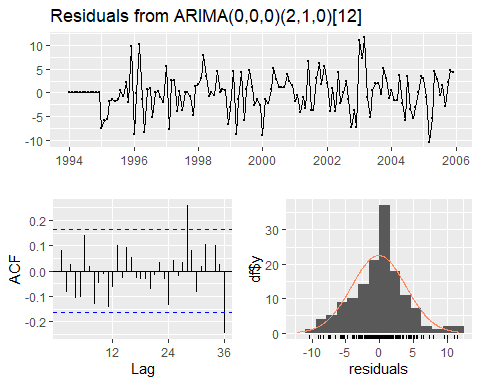
##### INTERPRETACION

Dado que nuestros modelos no presentan coeficientes en la parte ARIMA(0,0,0) se inferencia que los mismos son invertibles y estacionarios.

### PRUEBA DE LJUNG-BOX

#### MODELO 1

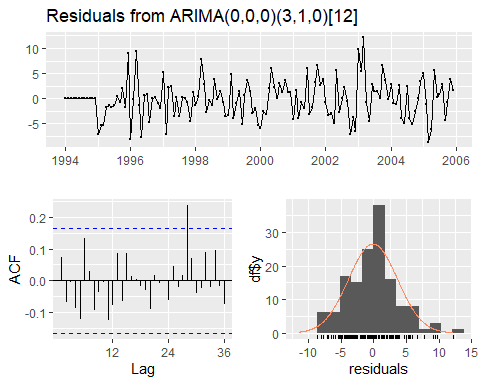
checkresiduals(mod3)



##   
## 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

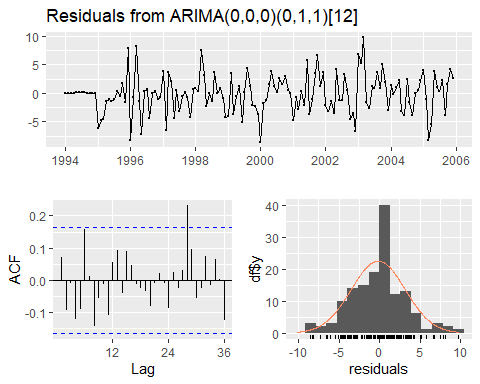
checkresiduals(mod4)



##   
## 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

#### MODELO 3

checkresiduals(mod5)



##   
## 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

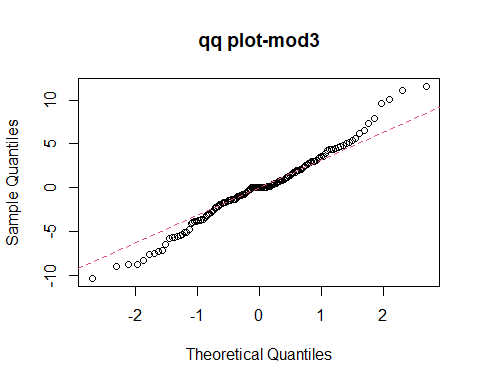
##### INTEPRETACION

Dados los coeficientes p asociados a nuestras pruebas de LJUNG BOX aplicadas a nuestros modelos indica 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.

### NORMALIDAD DE LOS RESIDUALES

#### MODELO 1

qqnorm(mod3$residuals, main = "qq plot-mod3")  
qqline(mod3$residuals, col= 2, lwd = 1, lty = 2)

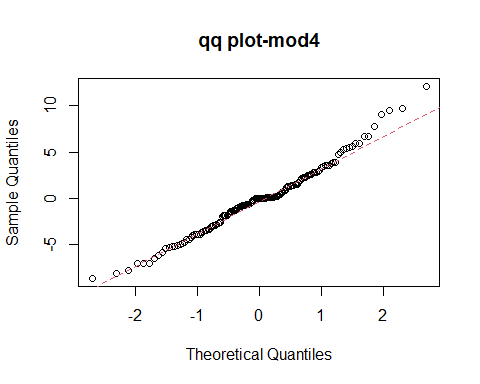


jarque.bera.test(mod1$residuals)

##   
## Jarque Bera Test  
##   
## data: mod1$residuals  
## X-squared = 1.0696, df = 2, p-value = 0.5858

#### *MODELO 2*

qqnorm(mod4$residuals, main = "qq plot-mod4")  
qqline(mod4$residuals, col= 2, lwd = 1, lty = 2)

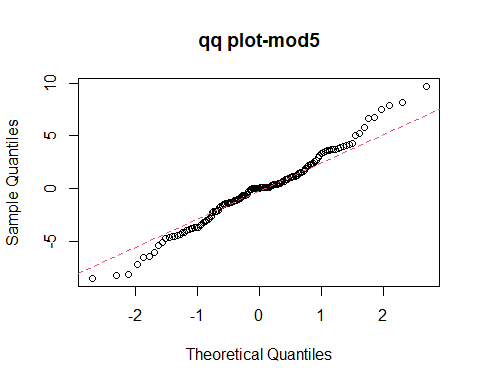


jarque.bera.test(mod4$residuals)

##   
## Jarque Bera Test  
##   
## data: mod4$residuals  
## X-squared = 4.6937, df = 2, p-value = 0.09567

#### MODELO 3

qqnorm(mod5$residuals, main = "qq plot-mod5")  
qqline(mod5$residuals, col= 2, lwd = 1, lty = 2)



jarque.bera.test(mod5$residuals)

##   
## Jarque Bera Test  
##   
## data: mod5$residuals  
## X-squared = 1.4639, df = 2, p-value = 0.481

##### INTEPRETACION

Como podemos observar de manera exploratorio en nuestros graficos podemos observar claramente que los residuales se ajustan a la linea normal lo que indica que da indicios que la misma sigue una distribucion normal. Esto se puede confirmar con los resultados de nuestra prueba de Jarque Bera cuyo coeficiente asociado a p es menor a 0.05 en cada uno de los modelos lo que nos indica que absolutamente todos los modelos se ajustan a una distribucion normal.

### SELECCION DEL MEJOR MODELO

AIC(mod3); BIC(mod3)

## [1] 761.6553

## [1] 770.3037

AIC(mod4); BIC(mod4)

## [1] 751.2253

## [1] 762.7565

AIC(mod5); BIC(mod5)

## [1] 736.2537

## [1] 742.0193

##### INTERPRETACION

El modelo 5 (mod5) tiene el AIC más bajo (736.2537), lo que sugiere que es el modelo preferido según el criterio de AIC y también tiene el BIC más bajo (742.0193), lo que sugiere que es el modelo preferido según el criterio de BIC.

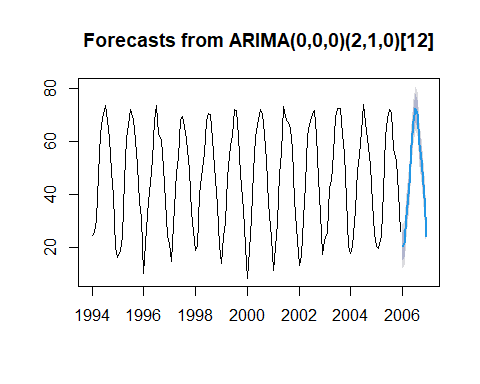
Es por esto que seleccionamos el modelo SARIMA(0,1,1)12 es el seleccionado. Sin embargo cabe mencionar que los demas modelos tambien tienen valores altamente competitivos con respecto a este, dicho esto procederemos a evaluar en el pronostico todos los modelos antes mencionados.

## **PRONOSTICO**

### PRONOSTICO PARA LA SERIE

#### MODELO 1

pron <- forecast(mod3, h = 12)  
plot(pron)

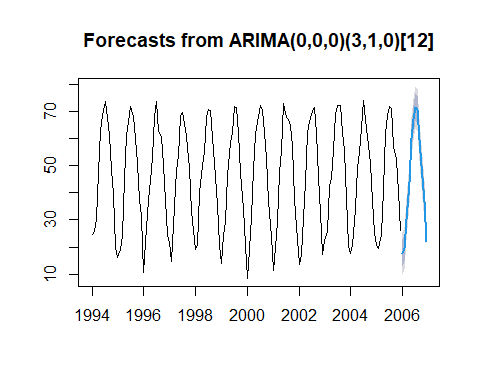


summary(pron)

##   
## 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  
## Training set -0.1083612 3.986116 2.947049 -2.127589 9.655297 0.7865153  
## ACF1  
## Training set 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(mod4, h = 12)  
plot(pron2)

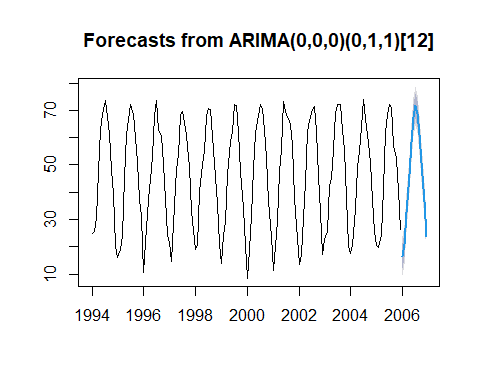


summary(pron2)

##   
## 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  
## Training set -0.1418842 3.732405 2.797088 -2.149689 8.981968 0.7464933  
## ACF1  
## Training set 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(mod5, h = 12)  
plot(pron3)



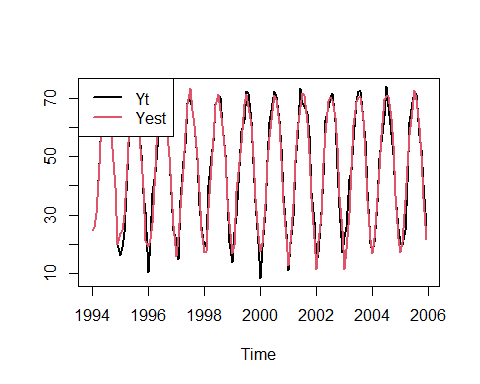
summary(pron3)

##   
## 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  
## Training set -0.1628102 3.314491 2.46352 -2.580408 8.282389 0.6574699  
## ACF1  
## Training set 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.

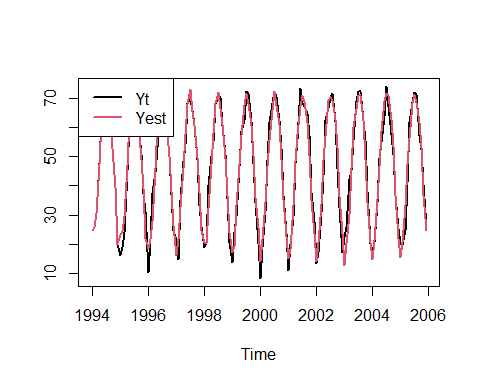
#### MODELO 1

Yt\_S <- mod3$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)



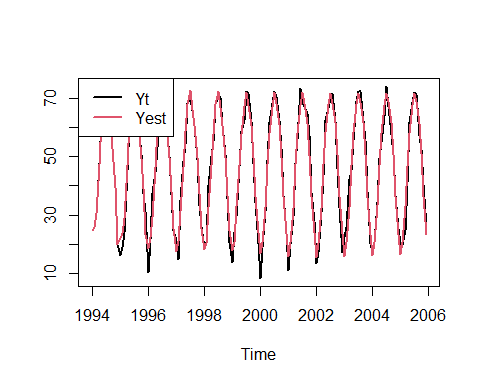
#### MODELO 2

Yt\_S2 <- mod4$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)



#### MODELO 3

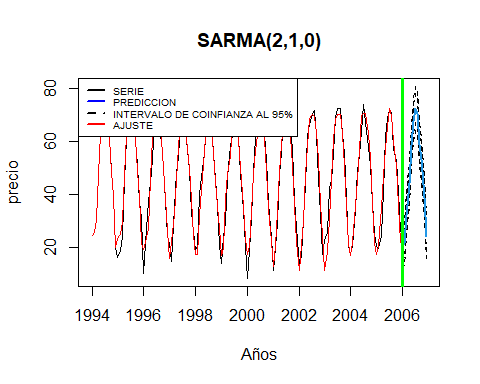
Yt\_S3 <- mod5$fitted  
grafico\_comparativo3 <- cbind(Yt, Yt\_S3)  
ts.plot(grafico\_comparativo3, col = c(1:2), lwd = 2)  
legend("topleft", c("Yt", "Yest"), lty=c(1,1), col=c(1:2), lwd = 2)



### GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

#### *MODELO 1*

plot(pron, shaded = FALSE, xlab = "Años", ylab = "precio",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")

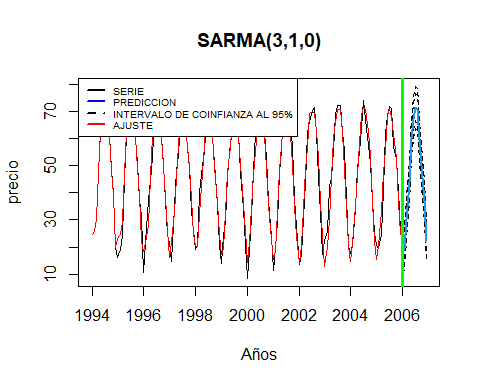


accuracy(pron)

## ME RMSE MAE MPE MAPE MASE  
## Training set -0.1083612 3.986116 2.947049 -2.127589 9.655297 0.7865153  
## ACF1  
## Training set 0.08213804

#### MODELO 2

plot(pron2, shaded = FALSE, xlab = "Años", ylab = "precio",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")

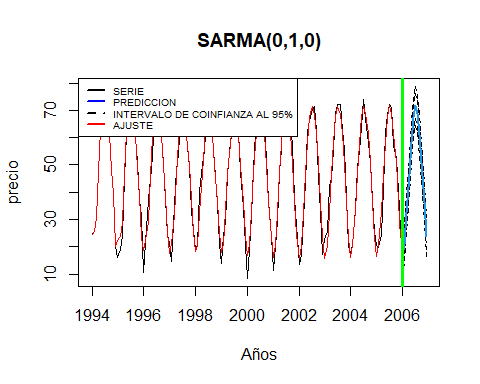


accuracy(pron2)

## ME RMSE MAE MPE MAPE MASE  
## Training set -0.1418842 3.732405 2.797088 -2.149689 8.981968 0.7464933  
## ACF1  
## Training set 0.07438451

#### *MODELO 3*

plot(pron3, shaded = FALSE, xlab = "Años", ylab = "precio",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")



accuracy(pron3)

## ME RMSE MAE MPE MAPE MASE  
## Training set -0.1628102 3.314491 2.46352 -2.580408 8.282389 0.6574699  
## ACF1  
## Training set 0.07191584

##### INTERPRETACION

Basado en los criterios de información (AIC, BIC) y las medidas de error (RMSE, MAE, MAPE, MASE), el modelo SARMA(0,1,0) es el que parece ser el mejor entre los tres evaluados. Tiene un AIC y BIC más bajos, lo cual indica un buen ajuste según estos criterios, y también presenta las menores medidas de error en términos de RMSE, MAE, y MAPE. Por lo tanto, recomendaría utilizar SARMA(0,1,0) para realizar predicciones sobre la serie temporal.

## CONCLUSION

En el contexto de las temperaturas promedio mensuales en Dubuque, EE.UU., de enero de 1994 a diciembre de 2005, los veranos son largos, calurosos y mojados, y los inviernos son helados, nevados y ventosos, con una variabilidad significativa a lo largo del año. Analizamos los datos utilizando modelos SARIMA para capturar la estructura estacional de la serie temporal.

### EVALUACIÓN DE MODELOS SARIMA

Hemos evaluado tres modelos diferentes:

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

* AIC: 761.66
* BIC: 770.3
* RMSE: 3.986116
* MAE: 2.947049
* MAPE: 9.655297
* ACF1: 0.08213804

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

* AIC: 751.23
* BIC: 762.76
* RMSE: 3.732405
* MAE: 2.797088
* MAPE: 8.981968
* ACF1: 0.07438451

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

* AIC: 736.25
* BIC: 742.02
* RMSE: 3.314491
* MAE: 2.46352
* MAPE: 8.282389
* ACF1: 0.07191584

Basado en los criterios de información (AIC, BIC) y las medidas de error (RMSE, MAE, MAPE), el modelo **SARIMA(0,1,0)** (ARIMA(0,0,0)(0,1,1)[12]) es el que parece ser el mejor entre los tres evaluados. Este modelo tiene el AIC y BIC más bajos, lo que indica un buen ajuste según estos criterios. Además, presenta las menores medidas de error en términos de RMSE, MAE y MAPE, sugiriendo que tiene el mejor rendimiento predictivo. La baja autocorrelación en los residuales (ACF1 = 0.07191584) también respalda la calidad del modelo.

Por lo tanto, escogemos el modelo **SARIMA(0,1,0)** para realizar predicciones sobre la serie temporal de temperaturas en Dubuque. Este modelo no solo captura adecuadamente la estacionalidad inherente a las temperaturas mensuales, sino que también proporciona las predicciones más precisas según las métricas evaluadas.