Tarea\_2

Miranda Belmonte Hairo

18 de septiembre de 2016

help(arima.sim)  
library(forecast)

Modelo a utilizar, propiedades: arima.sim(model, n, rand.gen = rnorm, innov = rand.gen(n, ...), n.start = NA, start.innov = rand.gen(n.start, ...), ...)

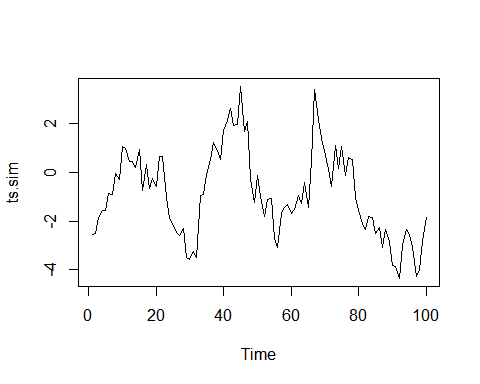
arima.sim(n = 63, list(ar = c(0.8897, -0.4858), ma = c(-0.2279, 0.2488)), 3sd = sqrt(0.1796))

An ARIMA simulation ts.sim <- arima.sim(list(order = c(1,1,0), ar = 0.7), n = 200) ts.plot(ts.sim)

Ejercicio 1: Simulacisn de ARIMA(p,d,q)

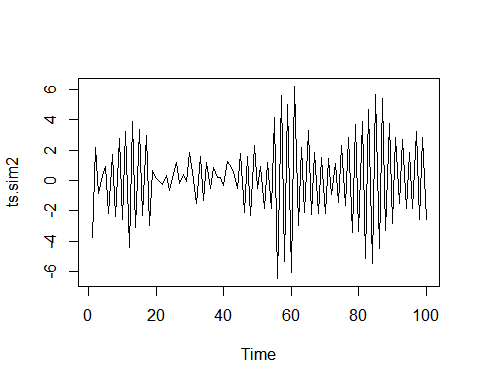
# -----1

par(mfrow=c(1,1))  
ts.sim<-arima.sim(n=100,model=list(ar=.9))  
ts.plot(ts.sim)



Se modela una simulacisn con n de 100, y coeficiente phi de .9. En el Plot de la serie se observa el comportamiento no determinista.

ts.sim2<-arima.sim(n=100,model=list(ar=-.9))  
ts.plot(ts.sim2)



Se simulo un AR(1), con coeficiente de -.9, se observa que la serie parece una caminata aleatoria.

nota coeficientes phi: # ar >- 1 ocila y diverge # -1 < ar < 0 ocila y converge # 0 < ar < 1 cae moderadamente y converge # ar < 1 explota y diverge

# -----2

par(mfrow=c(1,1))  
ts.sim3<-arima.sim(n=100,model=list(ar=1))

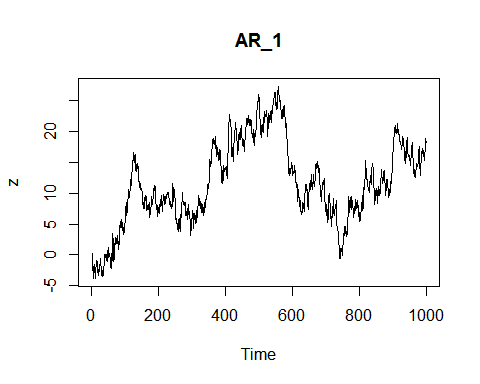
# coeficiente del AR no es estacionariol.

Como en la nota anterior sobre los valores del coeficiente, un valor mayor a uno tiende a explotar y a no converger, esto hace que el proceso de la seria no sea estacionaria.

par(mfrow=c(1,2))

Simulando un AR con un rnorm

ar.sim1<-function(n.sim,phi){  
 x<-rep(0,n.sim)  
 w=rnorm(x)  
 z=w  
 for (t in 2:n.sim) z[t] <- phi\*z[t - 1] + w[t]  
 par(mfrow=c(1,1))  
 ts.plot(z,main="AR\_1")  
}  
ar.sim1(1000,1)

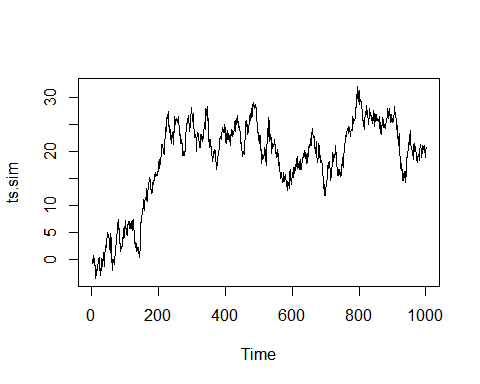


# -------3

Se programa una function utilizando el comando "arima.sim".

ar.sim<-function(n.sim){  
 ts.sim<-arima.sim(n= n.sim , model=list(order=c(1,1,1),ar=0,ma=0))  
 ts.plot(ts.sim)  
}   
ar.sim(1000)

## Warning in min(Mod(polyroot(c(1, -model$ar)))): no non-missing arguments to  
## min; returning Inf

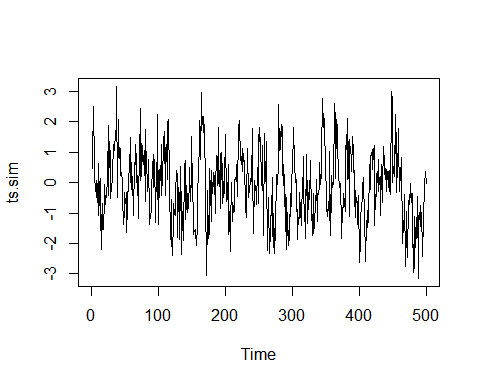


Los output fueron 1000 simulaciones y un coeficiente phi de .5, como resultado obtuvimos un na serie de tendencia no determinista, a primera vista se pudiera identificar como una caminata aleatoria.

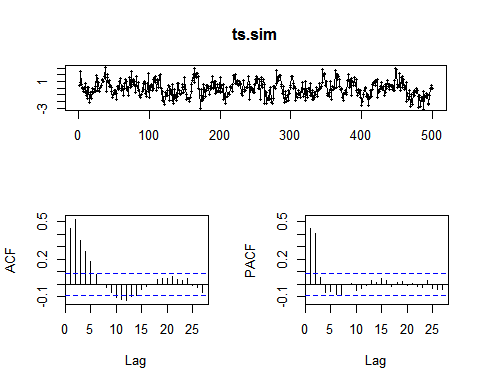
# -------4

set.seed(12)  
ts.sim<-arima.sim(n= 500 , model = list(ar=c(.3, .4) ))

ts.plot(ts.sim)



library(forecast)  
tsdisplay(ts.sim)



utilizando PACF se puede observar como los dos primeros periodos hay dos picos y a partir del teercer periodo la PACF cae a la sona de no correlacisn. caso contrario ACF tiende a una caida geometrica. esto nos infiere un modelo de un AR(2), serma una opcisn para modelar la serie,

# -------4 haciendo function para un AR2

par(mfrow=c(1,1))

ar.sim<-function(n.sim,phi,phi2){  
 ts.sim<-arima.sim(n= n.sim , model = list(ar=c(phi, phi2) ))  
 ts.plot(ts.sim)  
}   
ar.sim(1000,.7,.2)

A modo de ejemplo utilizamos 1000 simulaciones, con los coeficientes de .7 y .2, algo a destacar es que el sistema a tener dos o mas coeficientes que sumados den uno o mas, nos recuerda la no estacionariedad. El resultado de la simulacisn del AR(2), es un proceso no determinista.

##### -FIN DE EJERCICIO 1

## Ejercicio 2: Ajustando modelos

y1<-function(n.sim,phi){  
x<-rnorm(n.sim)  
z<-rnorm(1,0,1)  
y <- phi\*x+50+z  
ts.plot(y, main="AR (1)")  
}  
y1(100,.5)

Para ajustar el modelos se crea un funcisn de un AR(1), utilizamos 1000 simulacisnes y un coeficiente de .5, como la serie sigue un proceso estocastico.

x<-rnorm(100)  
z<-rnorm(1,0,1)  
y<- .9\*x +50 +z

plot(y,type="l")

Simulamos serie sin programar una funcisn.

1. Cual es la media teorica para el proceso {Yt}? #----50 por el intercepto
2. En la forma Yt = 5+0.9Yt-1 +et, cual es el valor que deberma tomar mu? #-----

mean(y)

# ---49.79266 en promedio

# 3--

x<-rnorm(100)  
z<-rnorm(1,0,1)  
y<- .9\*x +50 +z  
ar1<-arima(y,order=c(1,0,0))  
ar1

##   
## Call:  
## arima(x = y, order = c(1, 0, 0))  
##   
## Coefficients:  
## ar1 intercept  
## 0.0205 51.0242  
## s.e. 0.0997 0.0886  
##   
## sigma^2 estimated as 0.7535: log likelihood = -127.74, aic = 261.49

Utilizando un ARMA(1,0,1) con la funcisn "arima"

Interpretacisn de los parametros: -Un cambio de una unidad del periodo pasado , tendra un efecto positivo en el valor actual de un 0.0205 en promedio, siendo asm que el pasado del ar tiene efectos positivos sobre valores del periodo actual. -El coeficiente del intercepto de 51.0242 en promedio, lo cual nos indica que independiente mente el valor promedio actual sera de 51.0242 en promedio. -Con significancia solamente el el intercepto.

# ----4

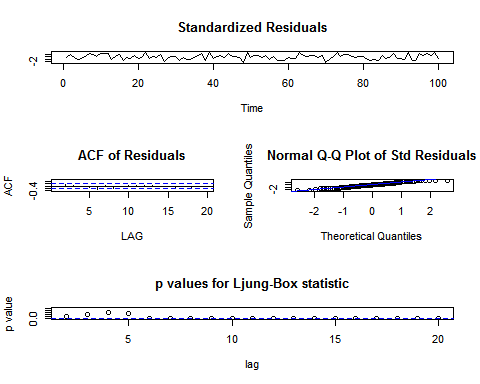
# El modelo elaborado con la funcion arima no describe en su totalidad por la variacisn de los coeficientes del autorregresivos.

# ----5

dev.off()  
par("mar")  
par(mar=c(1,1,1,1))

library(astsa)

sa<-sarima(y, 1,0,0)



sa[1]

## $fit  
##   
## Call:  
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,   
## Q), period = S), xreg = xmean, include.mean = FALSE, optim.control = list(trace = trc,   
## REPORT = 1, reltol = tol))  
##   
## Coefficients:  
## ar1 xmean  
## 0.0205 51.0242  
## s.e. 0.0997 0.0886  
##   
## sigma^2 estimated as 0.7535: log likelihood = -127.74, aic = 261.49

Explicando el modelo: -El grafico de los errores estandarizados nos mostraron un comportamiento de tendencia estocastis, lo cual se busca en los residuales. -La funcisn de autocorrelacisn nos indica la ausencia de correlacisn para los 20 periodos analizados. -Para el grafico de las desviacisnes estandart de los residuales se aprecia como tienden a seguir la tendencia. -Por zltimo la gafica de Ljuang Box p value, nos muestra como los primeros 5 periodos no se tiene correlacisn. estando por arriba de la banda de significancia lo cual hace que no se rechace la hipstesis nula de no correlacisn.

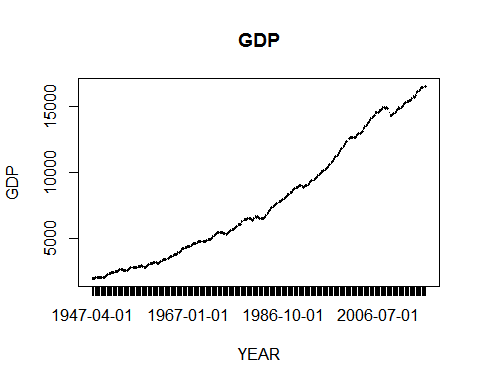
Conclusisn del ehercicio: Me parece mas clara la arima, dado a que arroja a primera instancia los coeficientes, pero a manera de analisis un sarima te da mas pruebas sobre la consistencia del modelo.

##### -FIN DE EJERCICIO 2

## --Ejercicio 3

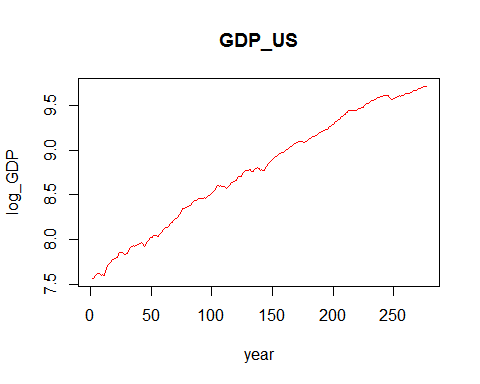
dir()  
ser<-read.csv("GDPC1.csv" ,header=FALSE,skip = 2)  
View(ser)  
names(ser)[1] <- "YEAR"  
names(ser)[2] <- "GDP"  
dev.off()  
class(ser$GDP)

ser<-read.csv("GDPC1.csv" ,header=FALSE,skip = 2)  
names(ser)[1] <- "YEAR"  
names(ser)[2] <- "GDP"  
plot(ser,type="l", main="GDP")



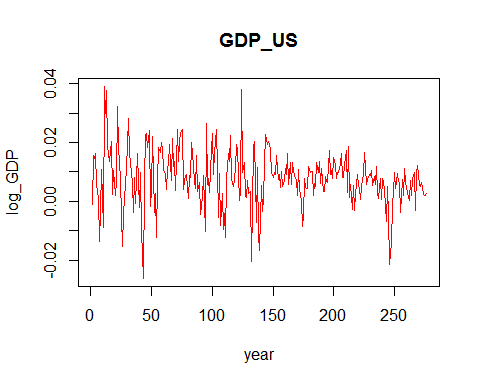
Se grafico la serie del GDP de los Estados Unidas de 1947 al 2016, con Datos de la Fred de US. La serie se grafica en niveles.

ser<-read.csv("GDPC1.csv" ,header=FALSE,skip = 2)  
names(ser)[1] <- "YEAR"  
names(ser)[2] <- "GDP"  
plot(log(ser$GDP),type="l",col="red",main="GDP\_US",ylab = "log\_GDP",xlab = "year")



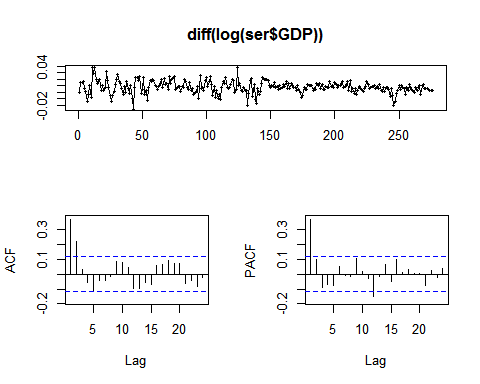
Utilizamos Logs dado a su comprtamiento tendencial.

ser<-read.csv("GDPC1.csv" ,header=FALSE,skip = 2)  
names(ser)[1] <- "YEAR"  
names(ser)[2] <- "GDP"  
plot(diff(log(ser$GDP)),type="l",col="red",main="GDP\_US",ylab = "log\_GDP",xlab = "year")



Se diferencia la serie en logs para tener tasas porcentuales, on ellas pudimos observar como la serie tiende a parecer estacionaria.

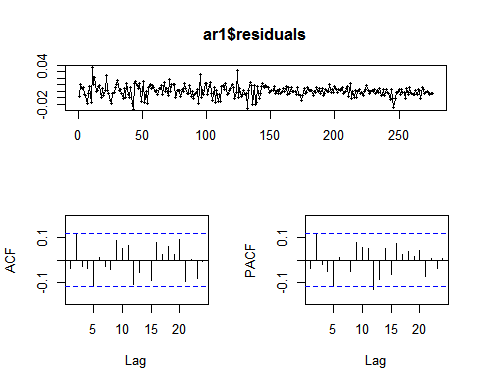
library(forecast)  
tsdisplay(diff(log(ser$GDP)))



Utilizando la funcisn "tsdisplay" en la serie diferenciada (estacionaria), se puede determinar con las funciones ACF Y PACF las opciones para modelar la media.

La funcion de aurocorrelacisn se observa como tiende a caer, despuis de los dos primeros periodos, en tanto la PACF vemos que sin incluir el primer periodo las autocorrelaciones caen en el area de significancia.

library(forecast)  
ar1<-arima(diff(log(ser$GDP)),order=c(1,0,0))  
tsdisplay(ar1$residuals)

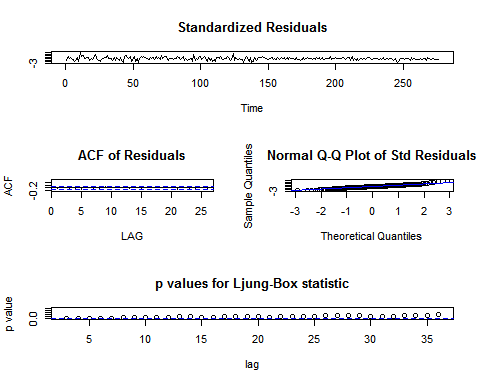


El modelos que se eligio fue un AR(1)

Al utilizar acf y pacf para los residuales de nuestro modelos, podemos concluir la normalidad de los errores, no detectando autocorrelacisn

par("mar")  
par(mar=c(1,1,1,1))

library(forecast)  
sa1<-sarima(diff(log(ser$GDP)),1,0,0,1,0,0,12)



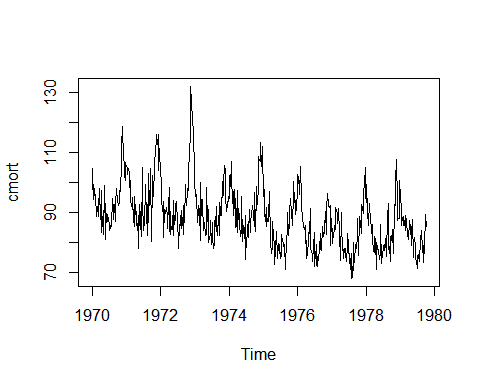
En el modelo sarima, se concluye de manera similar a los resultados que se obtienen con el modelo arima, los coeficientes varian muy poco y los errores no son autocorrelacionados, se puede concluir si utilizamos el aic como crmterio de eleccisn del modelo -sar1\_aic=-1824.28 -ar1\_aic=-1822.28 el modelo arima al tener un aic mayor es el mejor mejor modelo entre estos dos.

##### -FIN DE EJERCICIO 3

##### ---Ejercicio 4

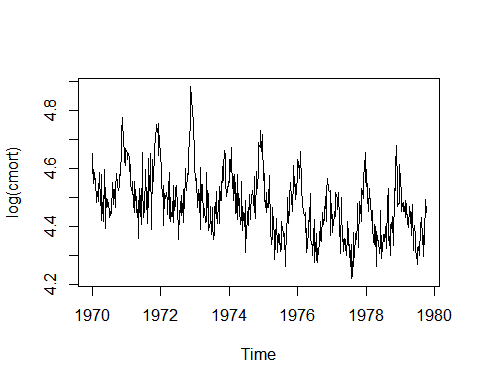
help(cmort)  
dev.off()

plot(cmort,type="l")

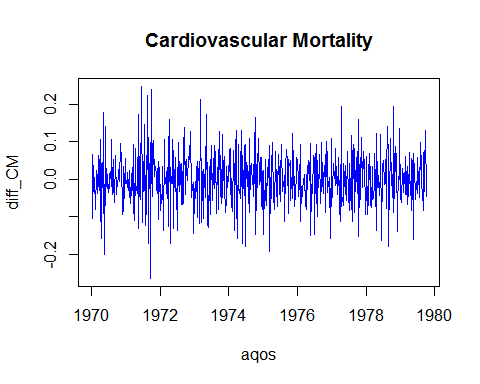


Graficando la serie Cardiovascular Mortality a niveles, esta serie es extraida del estudio que se hace en LA Polution study, cuyo datos son promedios demanales se 10 aqos (1970-1979).

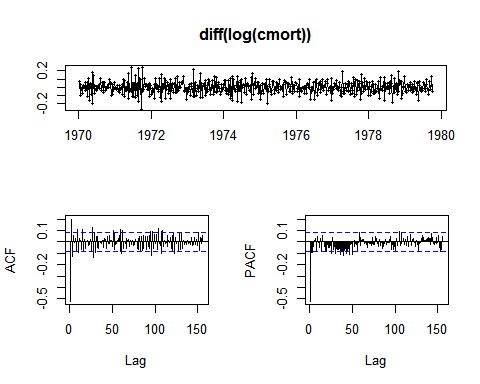
plot(log(cmort),type="l")



plot(diff(log(cmort)),  
 type="l",col="blue",  
 main="Cardiovascular Mortality",  
 xlab="aqos",ylab="diff\_CM")

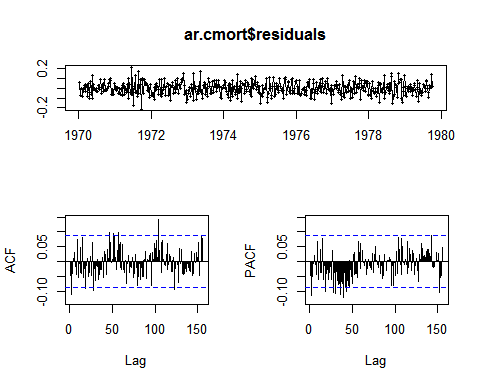


tsdisplay(diff(log(cmort)))



Se diferencia la serie en logs y se grafica. Se utilizo un Modelo AR(1), dado aque el ACF decae geometricamente, y el Pacf tiene un pico que sale de los intervalos y despues caen, esto quiere decir que esta autoccorrelacionada con un periodo atras.

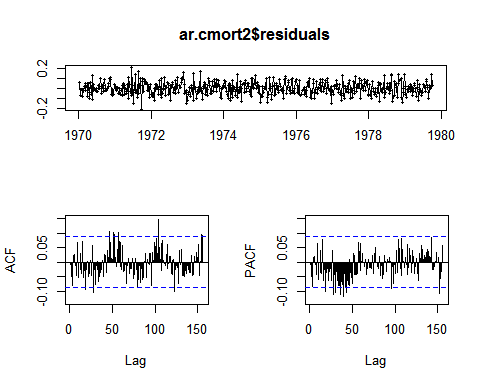
cmort1<-diff(log(cmort))  
ar.cmort<-arima(cmort1,order = c(1,0,0))  
tsdisplay(ar.cmort$residuals)



si se tienen problemas de autocorrelacisn en los residuales con el modelo seleccionado.

## ajustando a un ar(2)

ar.cmort2<-arima(cmort1,order = c(2,0,0))  
tsdisplay(ar.cmort2$residuals)

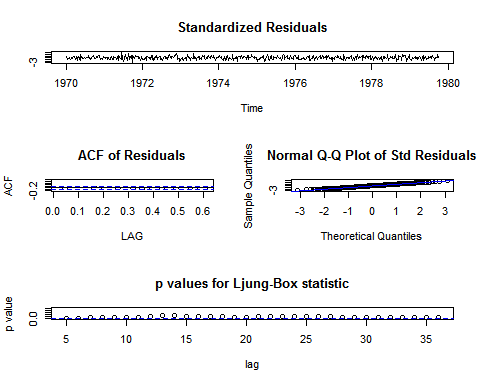


se sigue teniendo problemas de correlacisn con los residuales, como observamos los ACF,PACF.

# ajustandolo a un sarima

dev.off()  
par("mar")  
par(mar=c(1,1,1,1))

sarima(cmort1,2,0,0,2,0,0,12)

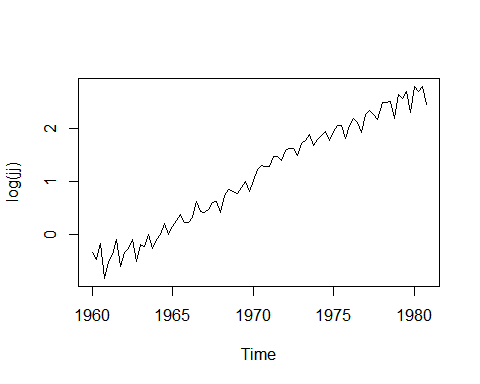


## $fit  
##   
## Call:  
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,   
## Q), period = S), xreg = xmean, include.mean = FALSE, optim.control = list(trace = trc,   
## REPORT = 1, reltol = tol))  
##   
## Coefficients:  
## ar1 ar2 sar1 sar2 xmean  
## -0.5654 -0.0872 0.0315 0.0041 -0.0003  
## s.e. 0.0444 0.0454 0.0452 0.0457 0.0018  
##   
## sigma^2 estimated as 0.004152: log likelihood = 670.68, aic = -1329.36  
##   
## $degrees\_of\_freedom  
## [1] 502  
##   
## $ttable  
## Estimate SE t.value p.value  
## ar1 -0.5654 0.0444 -12.7367 0.0000  
## ar2 -0.0872 0.0454 -1.9201 0.0554  
## sar1 0.0315 0.0452 0.6974 0.4859  
## sar2 0.0041 0.0457 0.0903 0.9281  
## xmean -0.0003 0.0018 -0.1572 0.8752  
##   
## $AIC  
## [1] -4.464515  
##   
## $AICc  
## [1] -4.460239  
##   
## $BIC  
## [1] -5.422813

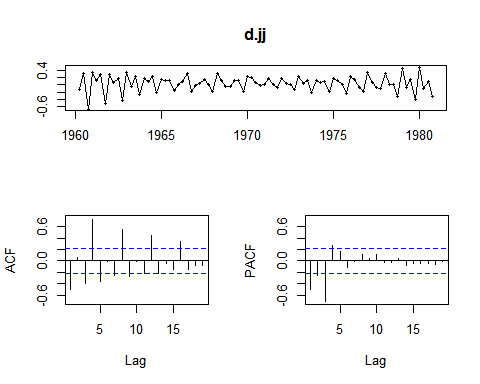
modelo ajustado , sin errores correlacionados Analizando graficos: -El standarized residuals, nos muestra un comportamiento de tendencia estocastica en los residuales. -ACF residuals nos indica la no correlacisn en los 20 rezagos analizados. -El normal Q-Q plot of Std Residual, podemos ver que los plots se hacercan a la linea de tendencia. -Los pvalues del Ljuang Box nos reportaron problemas de correlacisn en los dos primeros periodos.

#### serie jj

d.jj<-diff(log(jj))  
plot(log(jj),type="l")



tsdisplay(d.jj)



la serie pareceria que siguiera una caminata aleatoria por las funciones de autocorrelacisn, la ACF pareciera estar constantes o tener periodos en los cuales brinca(estacionales).

dev.off()

## null device   
## 1

ar\_jj<-arima(d.jj,order = c(4,0,0))  
tsdisplay(ar\_jj$residuals)

Modelamos la media con un AR(4), observando los residuales nos damos cuenta que en el periodo 7 se tiene significancia en la correlacisn.

par("mar")  
par(mar=c(1,1,1,1))  
sarima(d.jj,3,0,0,3,0,0,4)

nota: jbox-ho=no autocorrelacisn, busco aceptar con pvalue mayor a .05

El mejor modelo lo determina el ar(4), y con un sarima (3,0,0,3,0,0,4), se presenta un pequeqo pico en el periodo 7, el cual influye a un pequeqo problema con la normalidad de los residuales. Para el modelo sarima la grafica de residuales estandarizados nos muestra un comportamiento no determinista, lo cual es lo que bucamos, el acf al igual que los residuales estimados con el arima nos muestra un pico en el preriodo 7 mmnimo, el normal qq plot nos muestra un comportamiento aceotable siguiendo la linea de tendencia. el Ljuang -box por el contrario se tiene problemas en los primeros dos coeficientes estando en el area de significancia, lo cual nos indica correlacisn.

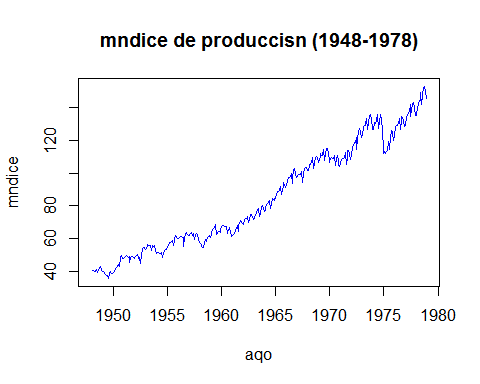
##### -FIN DE EJERCICIO 4

##### -Ejercicio 5: Una probadita...

dev.off()

### 1

plot(prodn, main="mndice de produccisn (1948-1978)",  
 xlab="aqo",ylab="mndice",col="blue")



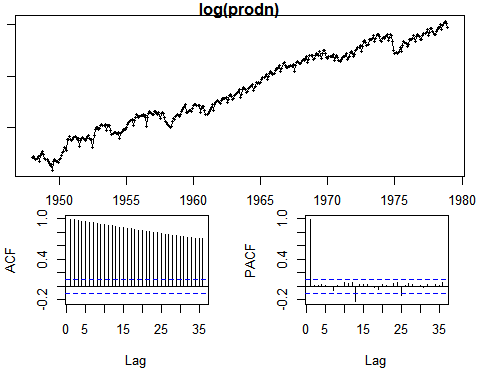
Graficando la serie del indice de produccisn mensual a niveles.

### 2

par("mar")

## [1] 5.1 4.1 4.1 2.1

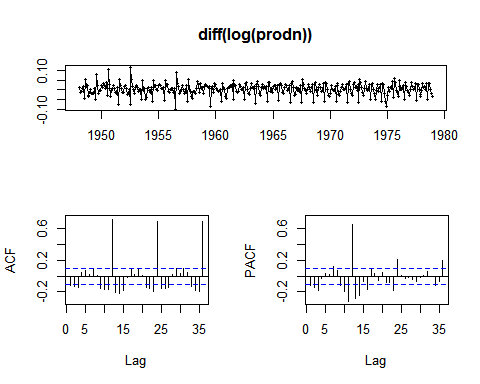
par(mar=c(1,1,1,1))  
tsdisplay(log(prodn))



Se grafico la serie en logaritmos, tambiin usamos la funcisn "tsdisplay" para obtener los ACF y PACF se observo el comportamiento de las correlacisnes en logaritmos.

### 3

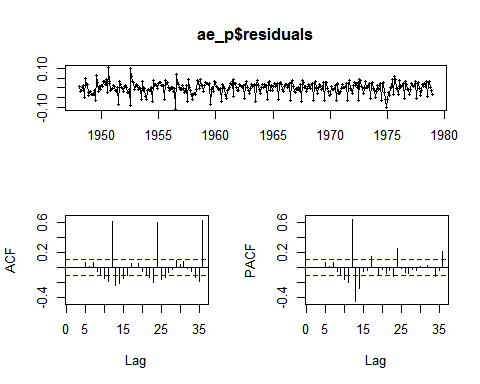
tsdisplay(diff(log(prodn)),lag.max = 36)



Se diferencis la serie en logaritmos para analizar las funciones de correlacisn.

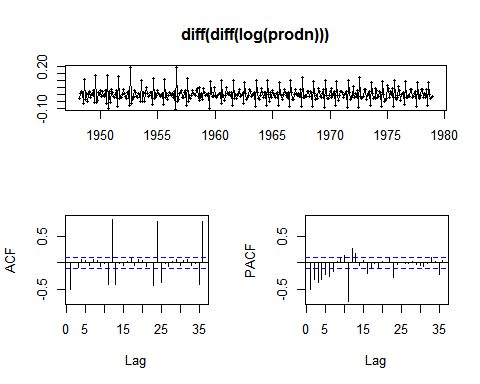
#### 4

ae\_p<-arima(diff(log(prodn)),  
 order=c(3,0,0))  
tsdisplay(ae\_p$residuals)



Modelando la media con un AR(3), y tomando los residuales para observar si los residuales estan correlacionados, observamos problemos con los periodos 12. 24 y 36.

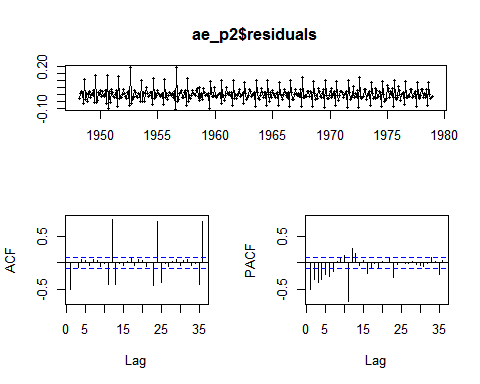
tsdisplay(diff(diff(log(prodn))),lag.max = 36)



Se diferencio la serie estacionaria por diferencia, se observa estacionalidad por los brincos periodicos en las funciones de autocorrelacisn.

##### 5

ae\_p2<-arima(diff(diff(log(prodn)),  
 order=c(3,0,0)))  
tsdisplay(ae\_p2$residuals)



Modelando la serie con las segundas diferiencias.

##### 6

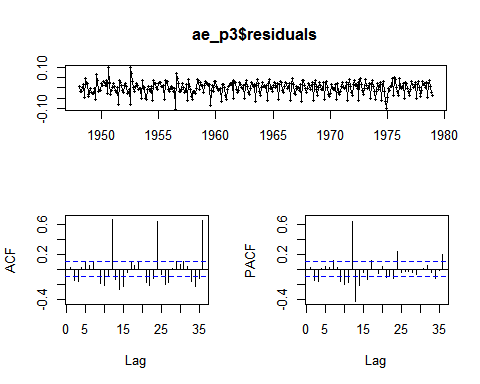
dev.off()

## null device   
## 1

par(mar=c(1,1,1,1))  
tsdisplay(diff(log(prodn)),lag.max = 11)

Se busca un modelos para la serie inter cilos

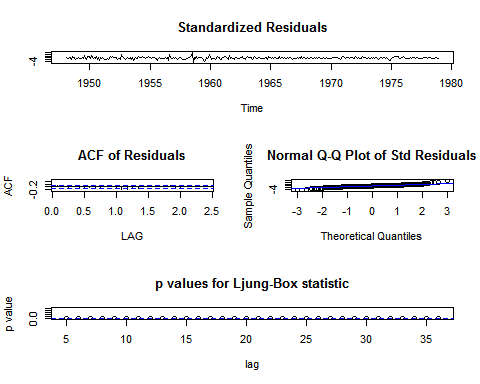
ae\_p3<-arima(diff(log(prodn)),  
 order=c(0,0,1))  
tsdisplay(ae\_p3$residuals)



El modelo arima(0,0,1), simula los acf y pacf de la serie prodn

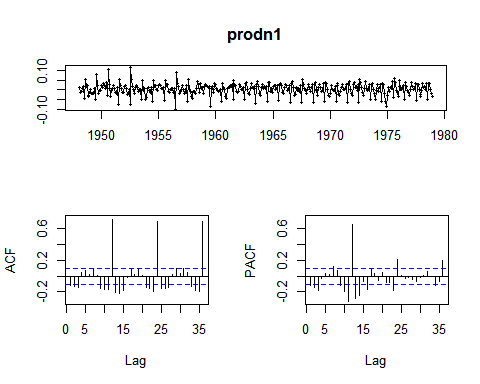
#### -7

prodn1<-diff(log(prodn))  
sar2<-sarima(prodn1,0,1,1,3,0,0,12)

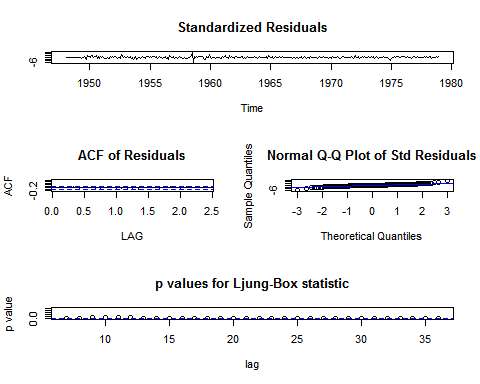


Con los parametros del inciso 5 & 6 el standarized residuals nos muestra un comportamientos no tendencia, lo cual es bueno para el modelo. -El Q-Q plot tiente a sus valores iniciales y finales a salir de su tendencia. -Los p-values se encuentran dentro de la banda de significancia, lo cual no es un buen indicador, ya que nos menciona autocorrelacisn en los residuales.

tsdisplay(prodn1)



sar2<-sarima(prodn1,3,0,0,3,1,0,12)



optarma el modelo sarima(3,0,0,3,1,0,12) Dado a cque no presenta tantos problemas en el ACF y el q-q- un comportamiento normal, el ppvalue despues del primer periodo, las correlacisnes son insignificativas que es lo que se busca.

##### --8

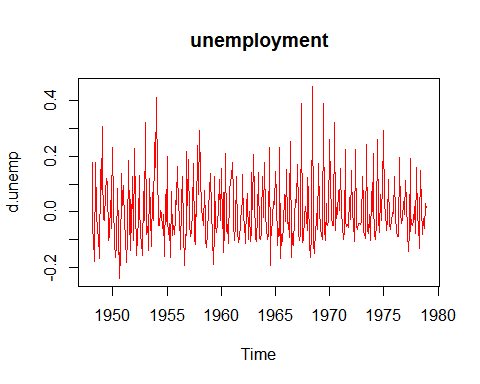
dev.off()

## null device   
## 1

library(astsa)  
plot(unemp)

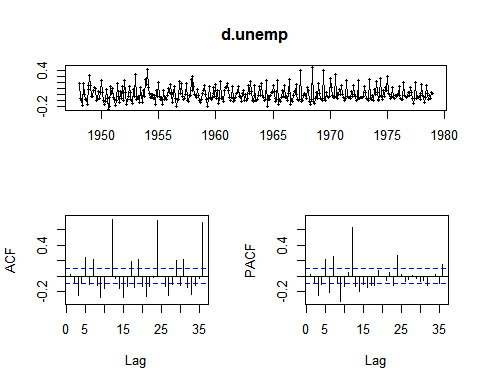
Graficando la serie unemployement de US

par(mfrow=c(1,1))  
d.unemp<-diff(log(unemp))  
plot(d.unemp,main="unemployment",col="red")



Diferenciando la serie en logaritmos y graficandola.

library(forecast)  
tsdisplay(d.unemp)



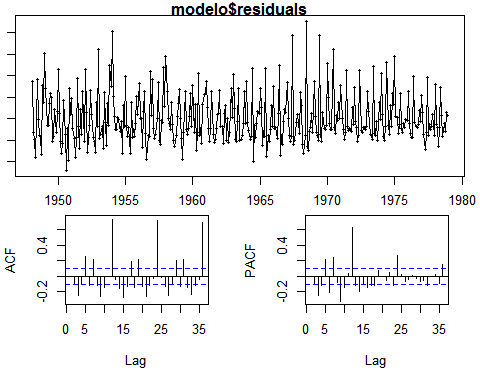
analizando autocorrelacion de la serie diferenciada. observamos saltos cada 11 periodos.

modelo que simule la afc & pacf de la serie d.unemp

modelo<-arima(d.unemp,order = c(1,0,0))  
par(mar=c(1,1,1,1))  
par("mar")

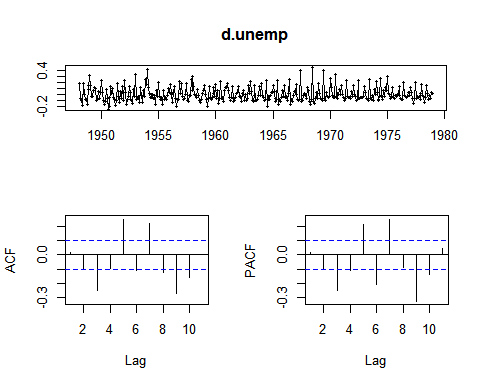
## [1] 1 1 1 1

tsdisplay(modelo$residuals)



modelo para los periodes intercilo el modeloque nos da esa forma es: los picos de estacionalidad.

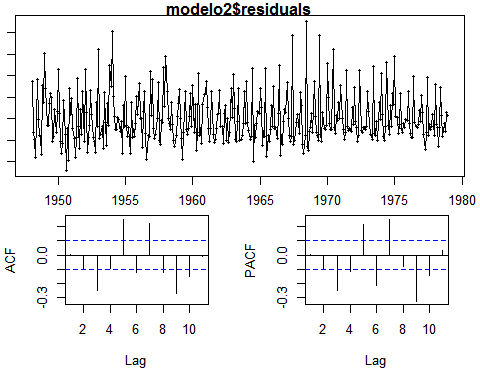
tsdisplay(d.unemp, lag.max = 11)



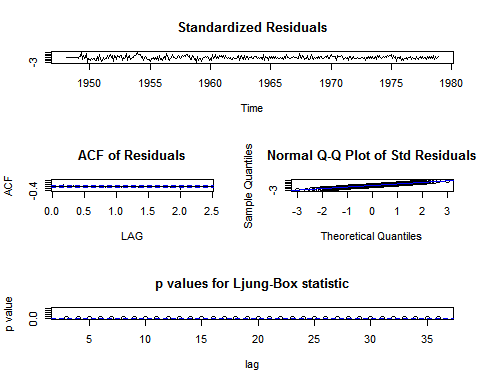
un ar(3) dado al comportamiento de PACF

modelo2<-arima(d.unemp,order = c(1,0,0))  
par(mar=c(1,1,1,1))  
par("mar")

tsdisplay(modelo2$residuals,lag.max = 11)



sarima(d.unemp,1,0,0,1,1,0,12)

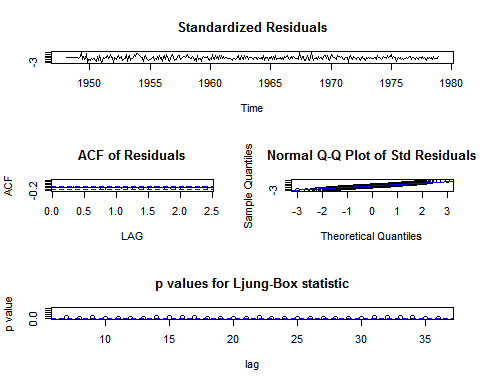


## $fit  
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,   
## Q), period = S), xreg = constant, optim.control = list(trace = trc,   
## Coefficients:  
## ar1 sar1 constant  
## 0.1342 -0.4948 -1e-04  
## s.e. 0.0524 0.0474 2e-04  
## sigma^2 estimated as 0.004524: log likelihood = 457.91, aic = -907.81  
## $ttable  
## Estimate SE t.value p.value  
## ar1 0.1342 0.0524 2.5614 0.0108  
## sar1 -0.4948 0.0474 -10.4431 0.0000  
## constant -0.0001 0.0002 -0.3867 0.6992  
## $AIC  
## [1] -4.382157  
##   
## $AICc  
## [1] -4.376471  
##   
## $BIC  
## [1] -5.350489

Se observa un comportamiento no determinista en los residuales standarizados, un acf autocorrelacionado en el segundo periodo, La gafica del normal QQ de la desviacisn de los residuales tienden a estar sobre su tendencia. pero con el p value del ljuang box tenemos una serie con correlaciones en todos sus rezagos.

cambiando modelo

sarima(d.unemp,3,0,0,3,1,0,12)



## $fit  
##   
## Call:  
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,   
## Q), period = S), xreg = constant, optim.control = list(trace = trc, REPORT = 1,   
## reltol = tol))  
##   
## Coefficients:  
## ar1 ar2 ar3 sar1 sar2 sar3 constant  
## 0.0604 0.1824 0.1012 -0.6678 -0.4279 -0.2477 0e+00  
## s.e. 0.0534 0.0527 0.0527 0.0533 0.0615 0.0560 2e-04  
##   
## sigma^2 estimated as 0.0037: log likelihood = 492.06, aic = -968.12  
##   
## $degrees\_of\_freedom  
## [1] 364  
##   
## $ttable  
## Estimate SE t.value p.value  
## ar1 0.0604 0.0534 1.1307 0.2589  
## ar2 0.1824 0.0527 3.4608 0.0006  
## ar3 0.1012 0.0527 1.9198 0.0557  
## sar1 -0.6678 0.0533 -12.5253 0.0000  
## sar2 -0.4279 0.0615 -6.9620 0.0000  
## sar3 -0.2477 0.0560 -4.4256 0.0000  
## constant 0.0000 0.0002 -0.2493 0.8033  
##   
## $AIC  
## [1] -4.561766  
##   
## $AICc  
## [1] -4.555303  
##   
## $BIC  
## [1] -5.487876

El modelo al analizar los pvalue del Ljuang Box nos muestran que los rezagos 3 y 4 son los que caen en la banda de significancia, el grafico del afc nos indica solo una corerelacion en el rezago 8 esti serma mejor modelo que el primero.

##### -FIN DE EJERCICIO 5