Temas selectos de Econometría y Finanzas. Modulo de series de tiempo

José Antonio García Ramírez

Tarea 2 , 15/09/2018

1. *Modificar la prueba adf.test de tal forma que el usuario pueda especificar si desea hacer la regresión sin constante (none), con constante (c) y con constante y tendencia (both) y que seleccione el número óptimo de rezagos de acuedo al criterio BIC.*

Decidí incluir en la implementación el caso en que hay tendencia pero no constante, si bien como vimos en clase la prueba t no es directamente comparable con la distribución del estadístico de la prueba ADF decidí en la implementación solo considerar las regresiones en donde todos los coeficientes son significativos individualmente con la prueba t. En la implementación actual no se compara verifica la significancia conjunto de los parámetros de la regresión primero porque se estaría comparando ahora con una distribución F y segundo porque cuando se implementó los resultados no eran consistentes en las simulaciones. El codigo es el siguiente:

```
adf.test.custom <- function(y, option='none')
{
  # y (numeric): vector con los datos de la serie de tiempo univariada
  y <- ts(y)
  lag <- floor(log(length(y))) + 1 #acotamos el numero de lags por el que siguiere el
                                    #texto de Chan Ngai
  datos <- data.frame(y1 = diff(y))
  for (i in 2:lag) #aumentamos las columnas de lag´s
  {
    datos[, as.character(paste0('y',i))] <- c(diff(y, lag=i), rep(NA, i-1))
  }
  names(datos) <- c('y1', names(datos)[2:lag])
  if (option == 'none')
  {
    #aplicamos el test para cada lag
    resultado <- mapply(function(x)
    {
      formula <- paste(names(datos)[x], collapse = '+')
      formula <- as.formula(paste0('y1 ~ ', formula, '-1'))
      modelo <- lm(formula , data = datos )
      resumen <- summary(modelo)
        # nos fijamos si todos los coeficientes de la regresion
        # son significativos individualmente
      coeficientes.significativos <- resumen$coefficients[, 'Pr(>|t|)']
      coeficientes.significativos <- coeficientes.significativos <= 0.05
      if(sum(coeficientes.significativos) == 1)
      {
        big <- BIC(modelo)
        # en caso de que todos los coeficientes sean significativos regresamos
        # el BIC de la regresion
        return(big)
      } else {return(Inf)} #si un coeficiente al menos es no significativo
                          #regresamos un BIC infinito
    }, 2:lag)
```

```r
}

if (option == 'c')
{
  datos[, 'c'] <- rep(1, dim(datos)[1] )
  #aplicamos el test para cada lag
  resultado <- mapply(function(x)
  {
    formula <- paste(names(datos)[x], collapse = '+')
    formula <- as.formula(paste0('y1 ~ ', formula))
    modelo <- lm(formula , data = datos )
    resumen <- summary(modelo)
    # nos fijamos si todos los coeficientes de la regresion
    # son significativos individualmente
    coeficientes.significativos <- resumen$coefficients[, 'Pr(>|t|)']
    coeficientes.significativos <- coeficientes.significativos <= 0.05
    if(sum(coeficientes.significativos) == 2)
    {
      big <- BIC(modelo)
      # en caso de que todos los coeficientes sean significativos regresamos
      #el BIC de la regresion
      return(big)
    }else {return(Inf)} #si un coeficiente al menos es no significativo
    #regresamos un BIC infinito
  }, 2:lag)
}
if (option == 't')
{
  datos[, 't'] <- cumsum(1:dim(datos)[1])
  #aplicamos el test para cada lag
  resultado <- mapply(function(x)
  {
    formula <- paste(c(names(datos)[x], 't'), collapse = '+')
    formula <- as.formula(paste0('y1 ~ ', formula, '-1'))
    modelo <- lm(formula , data = datos )
    resumen <- summary(modelo)
    # nos fijamos si todos los coeficientes de la regresion
    # son significativos individualmente
    coeficientes.significativos <- resumen$coefficients[, 'Pr(>|t|)']
    coeficientes.significativos <- coeficientes.significativos <= 0.05
    if(sum(coeficientes.significativos) == 2)
    {
      big <- BIC(modelo)
      # en caso de que todos los coeficientes sean significativos regresamos
      #el BIC de la regresion
      return(big)
    }else { return(Inf)}  #si un coeficiente al menos es no significativo
    #regresamos un BIC infinito
  }, 2:lag)
}
if (option == 'both')
{
  datos[, 't'] <- cumsum(1:dim(datos)[1])
```

```r
    #aplicamos el test para cada lag
    resultado <- mapply(function(x)
    {
      formula <- paste(c(names(datos)[2:(x)], 't'), collapse = '+')
      formula <- as.formula(paste0('y1 ~ ', formula))
      modelo <- lm(formula , data = datos )
      resumen <- summary(modelo)
      # nos fijamos si todos los coeficientes de la regresion
      # son significativos individualmente
      coeficientes.significativos <- resumen$coefficients[, 'Pr(>|t|)']
      coeficientes.significativos <- coeficientes.significativos <= 0.05
      if(sum(coeficientes.significativos) == 3)
      {
        big <- BIC(modelo)
        # en caso de que todos los coeficientes sean significativos regresamos
        #el BIC de la regresion
        return(big)
      } else { return(Inf)}  #si un coeficiente al menos es no significativo
      #regresamos un BIC infinito
    }, 2:lag)
  }
  parsimonia <- which.min(resultado)
  names(parsimonia) <- 'Lag optimo'
  return(parsimonia)
}
```

Las pruebas que realice para el caso de raíces unitarias sin constante ni tendencia, cuya salida coincide con la del test adf.test fueron: los datos del IGAE con los que hemos estado trabajando, un proceso de la forma $Y_t = Y_{t-1} + \epsilon_t$, el siguiente es un proceso de la forma $Y_t = Y_{t-2} + \epsilon_t$, y el tercer test fue con un proceso de la forma $Y_t = Y_{t-5} + \epsilon_t$, con ϵ_t como ruido blanco.

```r
library(tseries)
igae <- read.csv( "IGAE.csv", row.names = 1)
y <- igae
adf.test.custom(igae)
```

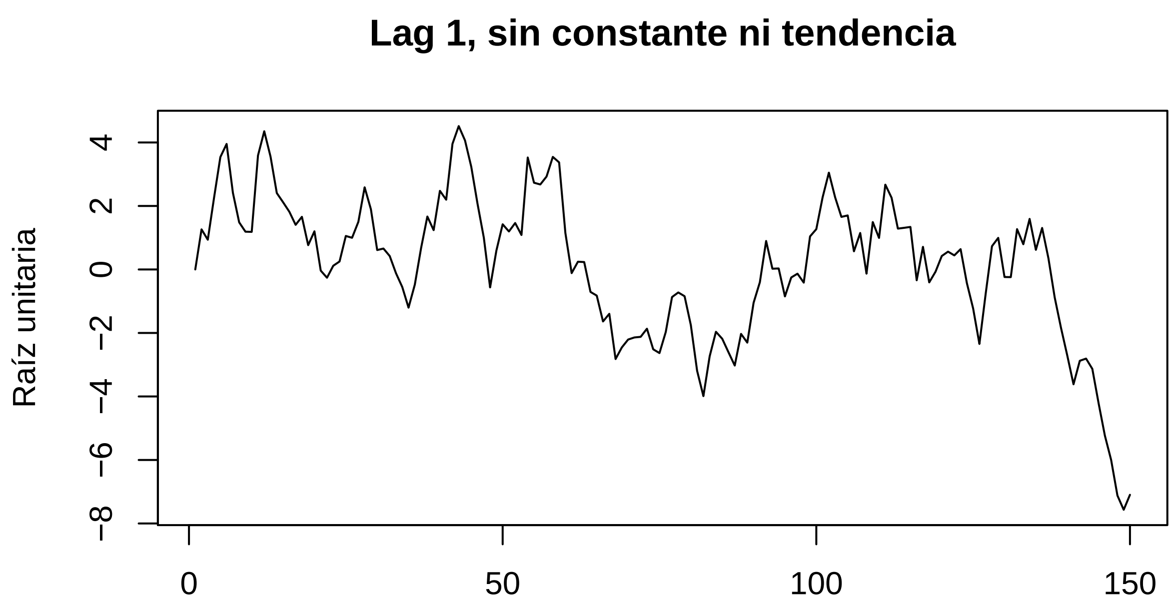
```
## Lag optimo
##          1
```

```r
adf.test(igae$IGAE, k=1)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  igae$IGAE
## Dickey-Fuller = -2.3534, Lag order = 1, p-value = 0.4274
## alternative hypothesis: stationary
```

```r
##### simulacion de una raiz aleatoria son constante ni tendencia de con lag=1
set.seed(0)
n <- 150
y <- rep(0, n)
for (i in 2:n)
{
  y[i] <- y[i-1] +rnorm(1)
}
```

```r
plot(y, type='l', main = 'Lag 1, sin constante ni tendencia', xlab='',
     ylab='Raíz unitaria' )
```



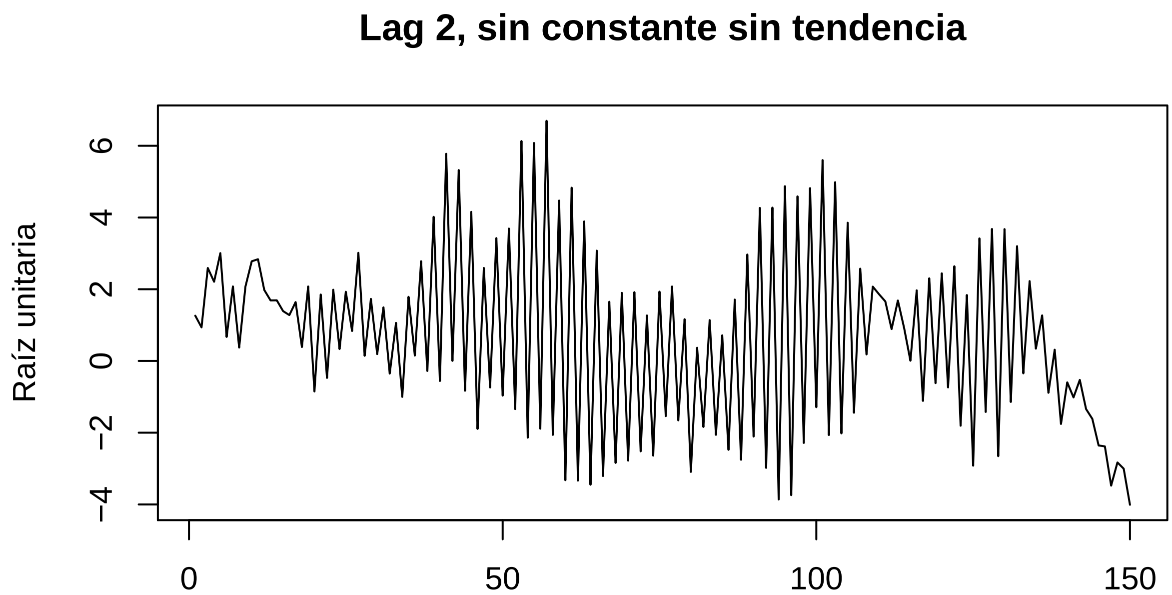

```r
adf.test.custom(y)
```

```
## Lag optimo
##          1
```

```r
adf.test(y, k=1)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -2.4852, Lag order = 1, p-value = 0.3743
## alternative hypothesis: stationary
```

```r
##### simulacion de una raiz aleatoria son constante ni tendencia de con lag=2
set.seed(0)
y <- rep(0, n)
y[1] <- rnorm(1)
y[2] <- y[1] + rnorm(1)
for (i in 3:n)
{
  y[i] <-  y[i-2] + rnorm(1)
}
plot(y, type='l', main='Lag 2, sin constante sin tendencia', xlab='',
     ylab='Raíz unitaria')
```



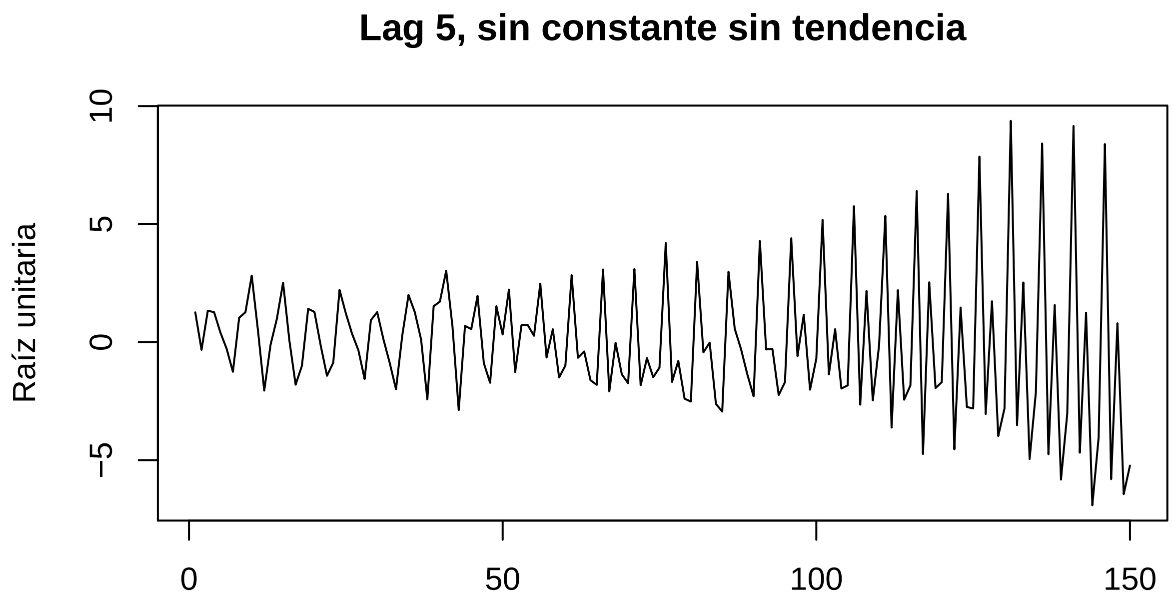

```
adf.test.custom(y)
```

```
## Lag optimo
##          2
```

```
adf.test(y, k=2)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -2.0605, Lag order = 2, p-value = 0.5512
## alternative hypothesis: stationary
```

```
##### simulacion de una raiz aleatoria son constante ni tendencia de con lag=5
set.seed(0)
y <- rep(0, n)
y[1:5] <- rnorm(5)
for (i in 6:n)
{
  y[i] <-  y[i-5] + rnorm(1)
}
plot(y, type='l', main='Lag 5, sin constante sin tendencia', xlab='',
     ylab='Raíz unitaria')
```




```r
adf.test.custom(y)
```

```
## Lag optimo
##          5
```

```r
adf.test(y, k=5)
```

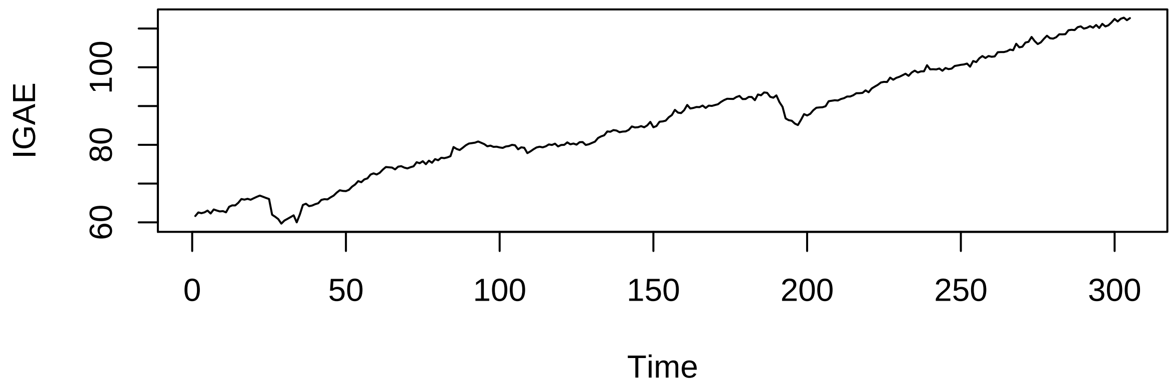
```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -1.9764, Lag order = 5, p-value = 0.5862
## alternative hypothesis: stationary
```

Los casos de prueba para el caso de raíces unitarias con constante y sin tendencia, cuya salida coincide con la del test adf.test fueron los mismos que en el inciso anterior.

```r
########################################
# test con constante
#######################################
adf.test.custom(igae, option = 'c')
```

```
## Lag optimo
##          1
```

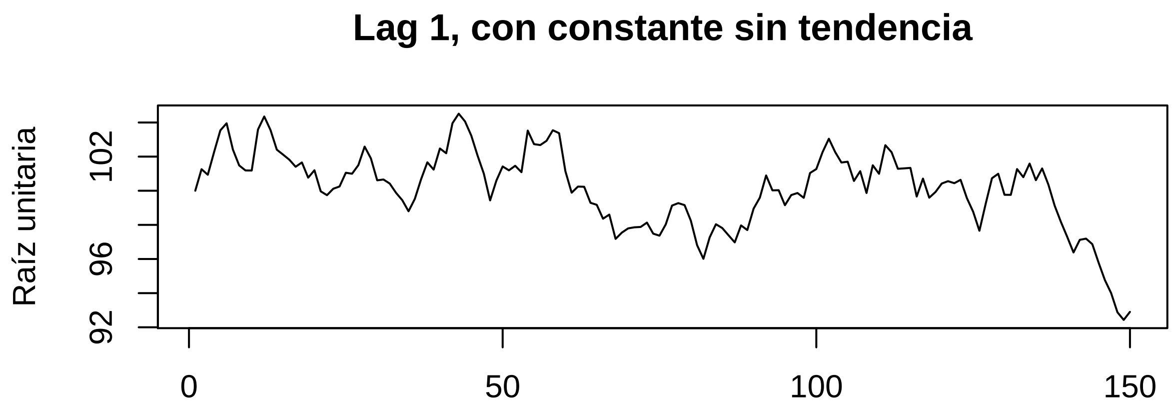
```r
plot.ts(igae)
```




```r
adf.test(igae$IGAE, k=1)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  igae$IGAE
## Dickey-Fuller = -2.3534, Lag order = 1, p-value = 0.4274
## alternative hypothesis: stationary
```

```r
set.seed(0)
n <- 150
y <- rep(0, n)
y[1] <-  100
for (i in 2:n)
{
  y[i] <- y[i-1] +rnorm(1)
}
plot(y, type='l', main = 'Lag 1, con constante sin tendencia', xlab='',
     ylab='Raíz unitaria')
```



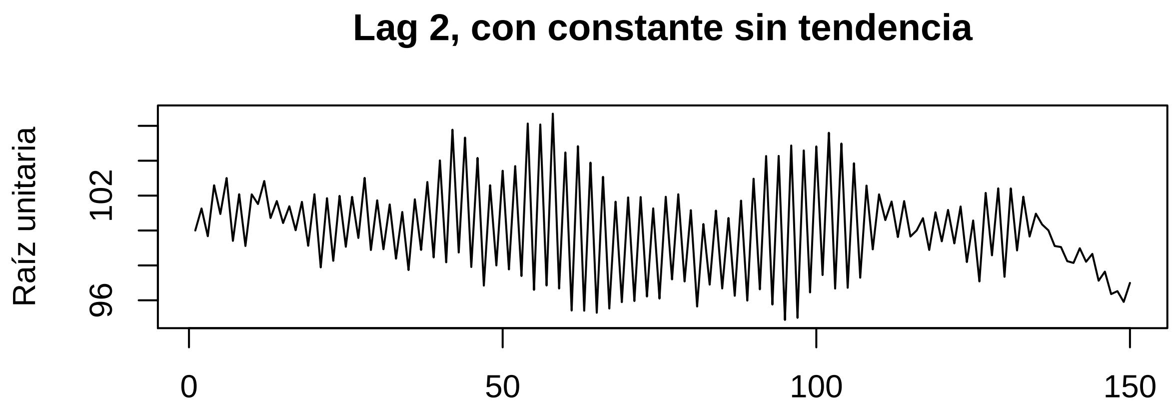

```
adf.test.custom(y, option = 'c')
```

```
## Lag optimo
##          1
```

```
adf.test(y, k=1)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -2.4852, Lag order = 1, p-value = 0.3743
## alternative hypothesis: stationary
```

```
##############
set.seed(0)
y <- rep(0, n)
y[1] <- 100
y[2] <- y[1] + rnorm(1)
for (i in 3:n)
{
  y[i] <-  y[i-2] + rnorm(1)
}
plot(y, type='l', main = 'Lag 2, con constante sin tendencia', xlab='',
     ylab='Raíz unitaria')
```



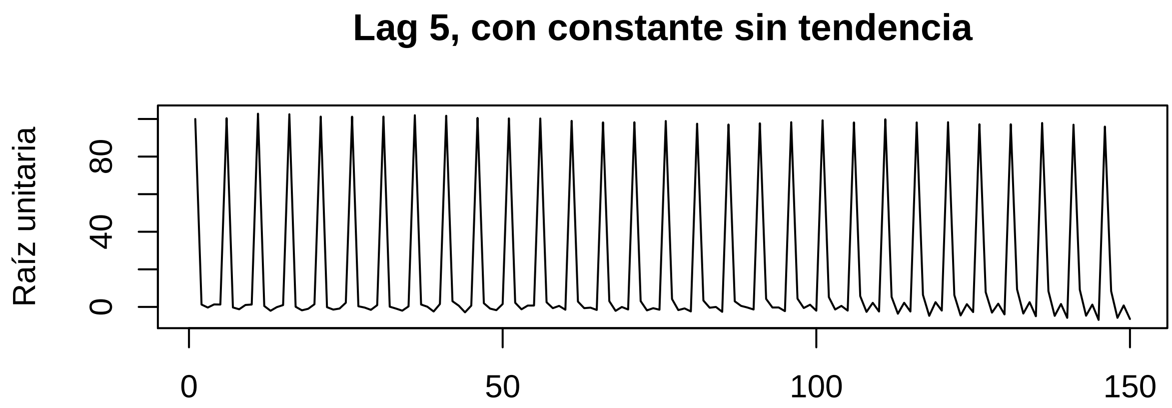

```
adf.test.custom(y)
```

```
## Lag optimo
##          2
```

```
adf.test(y, k=2)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -2.5053, Lag order = 2, p-value = 0.3659
## alternative hypothesis: stationary
```

```
###################
set.seed(0)
y <- rep(0, n)
y[1] <- 100
y[2:5] <- rnorm(4)
for (i in 6:n)
{
  y[i] <-  y[i-5] + rnorm(1)
}
plot(y, type='l', main = 'Lag 5, con constante sin tendencia', xlab='',
     ylab='Raíz unitaria')
```




```r
adf.test.custom(y)
```

```
## Lag optimo
##          5
```

```r
adf.test(y, k=5)
```

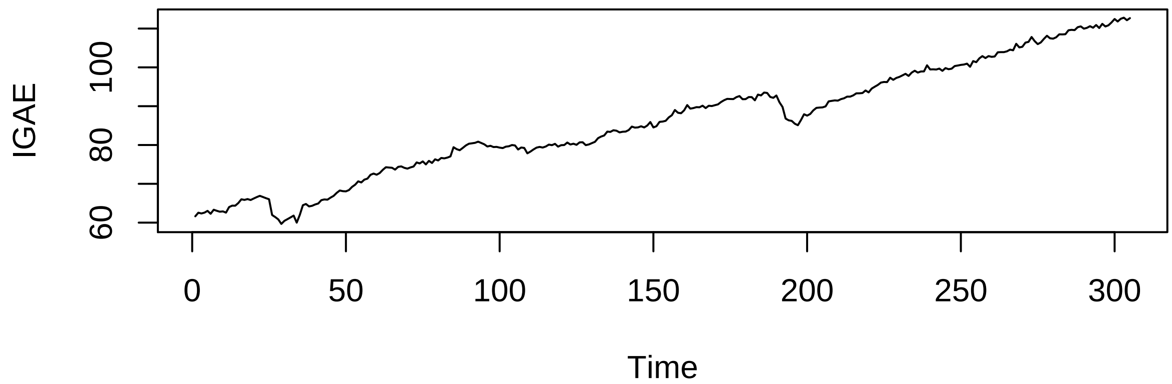
```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -2.4447, Lag order = 5, p-value = 0.3911
## alternative hypothesis: stationary
```

Los casos de prueba para el caso de raíces unitarias sin constante y con tendencia, cuya salida también coincide con la del test adf.test fueron los mismos.

```r
########################################
# test con tendencia
#######################################
adf.test.custom(igae, option = 't')
```

```
## Lag optimo
##          1
```

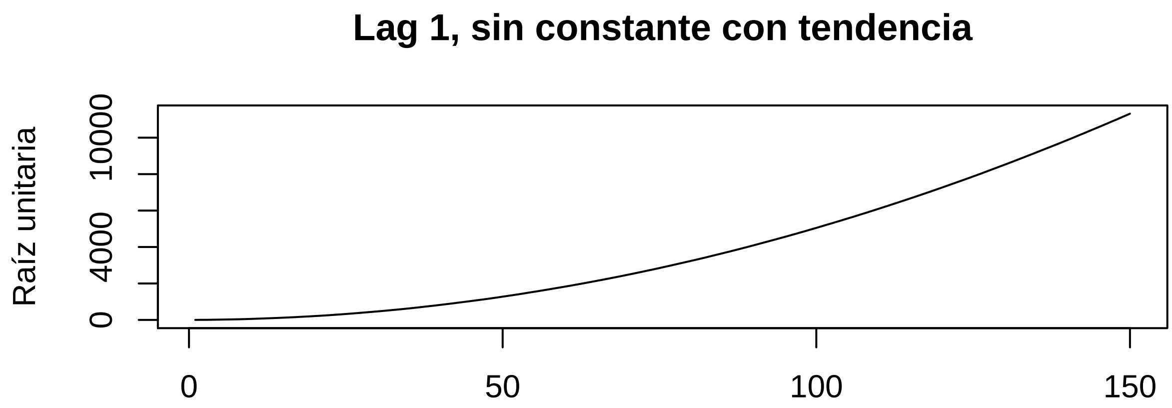
```r
plot.ts(igae)
```




```r
adf.test(igae$IGAE, k=1)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  igae$IGAE
## Dickey-Fuller = -2.3534, Lag order = 1, p-value = 0.4274
## alternative hypothesis: stationary
```

```r
set.seed(0)
n <- 150
y <- rep(0, n)
y[1] <-  0
for (i in 2:n)
{
  y[i] <- i + y[i-1] +rnorm(1)
}
plot(y, type='l', main = 'Lag 1, sin constante con tendencia', xlab='',
     ylab='Raíz unitaria')
```



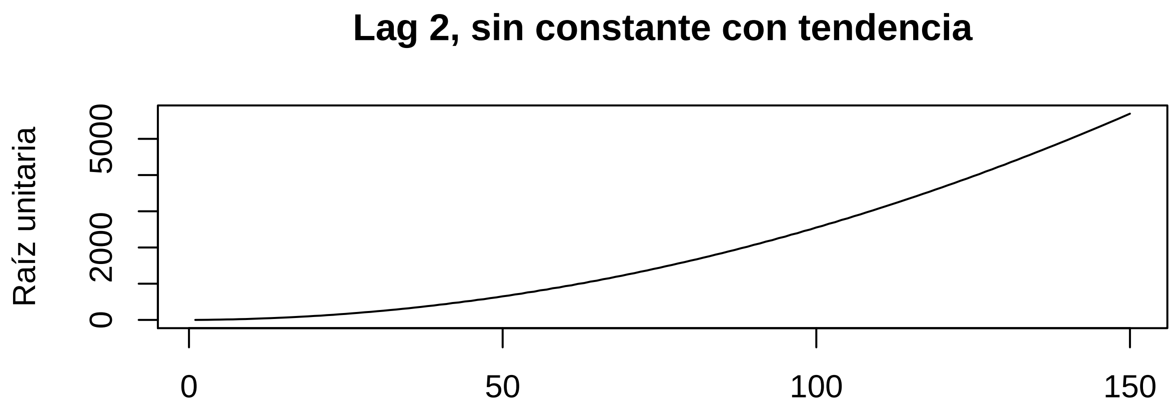

```
adf.test.custom(y, option = 't')
```

```
## Lag optimo
##          1
```

```
adf.test(y, k=1)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -0.88291, Lag order = 1, p-value = 0.9521
## alternative hypothesis: stationary
```

```
##############
set.seed(0)
y <- rep(0, n)
y[1] <- 0
y[2] <- y[1] + rnorm(1)
for (i in 3:n)
{
  y[i] <-i +  y[i-2] + rnorm(1)
}
plot(y, type='l', main = 'Lag 2, sin constante con tendencia', xlab='',
     ylab='Raíz unitaria')
```



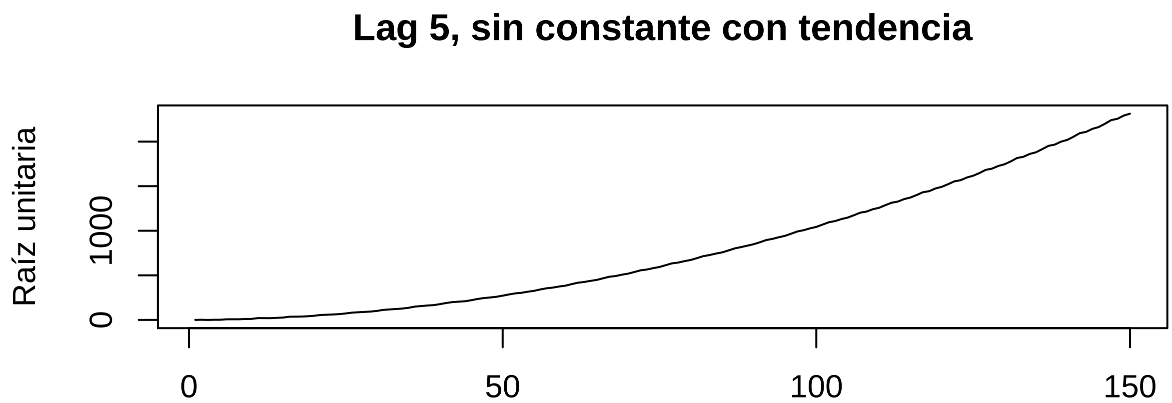

```r
adf.test.custom(y)
```

```
## Lag optimo
##          2
```

```r
adf.test(y, k=2)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -0.79653, Lag order = 2, p-value = 0.9598
## alternative hypothesis: stationary
```

```r
###################
set.seed(0)
y <- rep(0, n)
y[1] <- 0
y[2:5] <- rnorm(4)
for (i in 6:n)
{
  y[i] <- i +  y[i-5] + rnorm(1)
}
plot(y, type='l', main = 'Lag 5, sin constante con tendencia', xlab='',
     ylab='Raíz unitaria')
```




```r
adf.test.custom(y)
```

```
## Lag optimo
##          5
```

```r
adf.test(y, k=5)
```

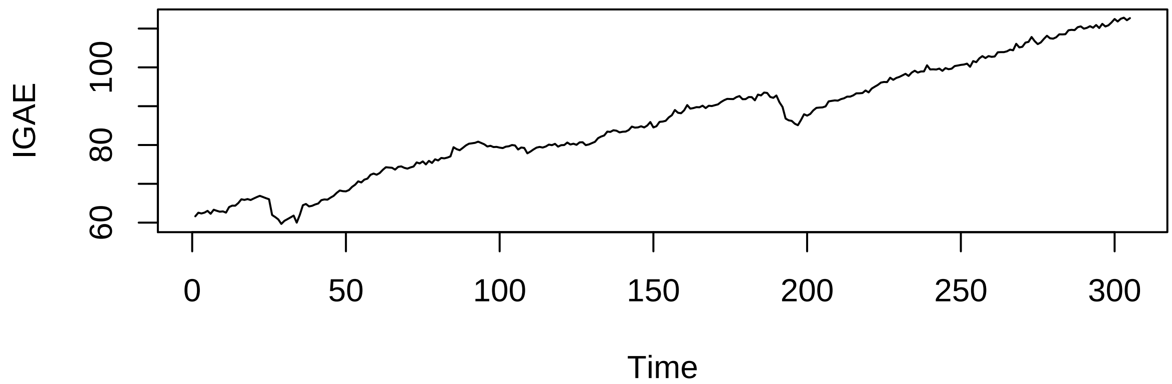
```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -1.4625, Lag order = 5, p-value = 0.8003
## alternative hypothesis: stationary
```

Finalmente los casos de prueba para el caso de raíces unitarias con constante y con tendencia, cuya salida también coincide con la del test adf.test fueron los mismos (cuando menos para estas simulaciones sabemos el lag exactoq eu genera al proceso por lo que el criterio bayesiano eligé bien).

```r
########################################
# test con ambos
#######################################
adf.test.custom(igae, option = 'both')
```

```
## Lag optimo
##          1
```

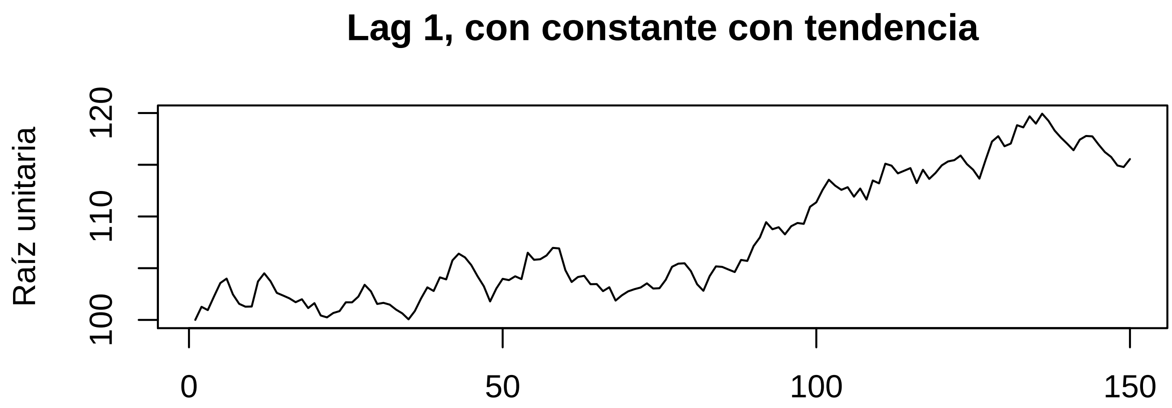
```r
plot.ts(igae)
```




```r
adf.test(igae$IGAE, k=1)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  igae$IGAE
## Dickey-Fuller = -2.3534, Lag order = 1, p-value = 0.4274
## alternative hypothesis: stationary
```

```r
set.seed(0)
n <- 150
y <- rep(0, n)
y[1] <-  100
for (i in 2:n)
{
  y[i] <- i/500 + y[i-1] +rnorm(1)
}
plot(y, type='l', main = 'Lag 1, con constante con tendencia', xlab='',
     ylab='Raíz unitaria')
```



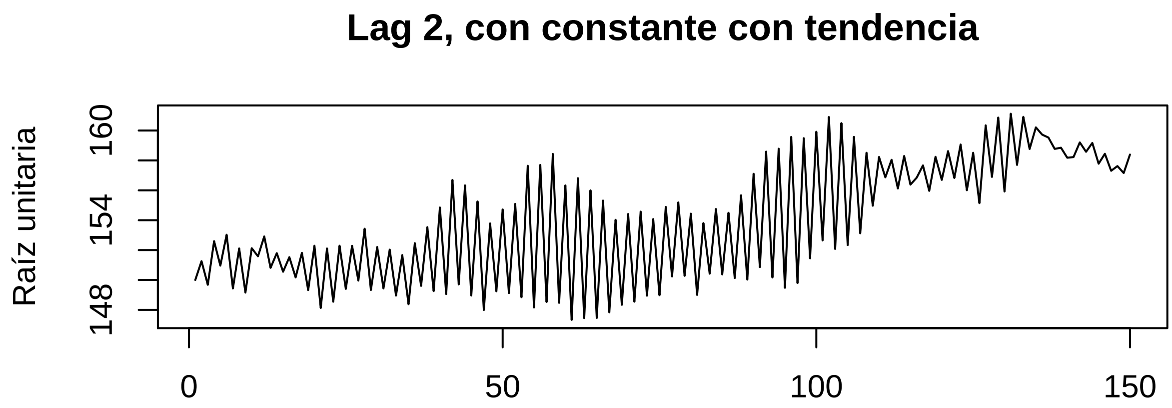

```r
adf.test.custom(y, option = 'both')
```

```
## Lag optimo
##          1
```

```r
adf.test(y, k=1)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -2.4607, Lag order = 1, p-value = 0.3845
## alternative hypothesis: stationary
```

```r
##############
set.seed(0)
y <- rep(0, n)
y[1] <- 150
y[2] <-  y[1] + rnorm(1)
for (i in 3:n)
{
  y[i] <-i/500 +  y[i-2] + rnorm(1)
}
plot(y, type='l', main = 'Lag 2, con constante con tendencia', xlab='',
     ylab='Raíz unitaria')
```



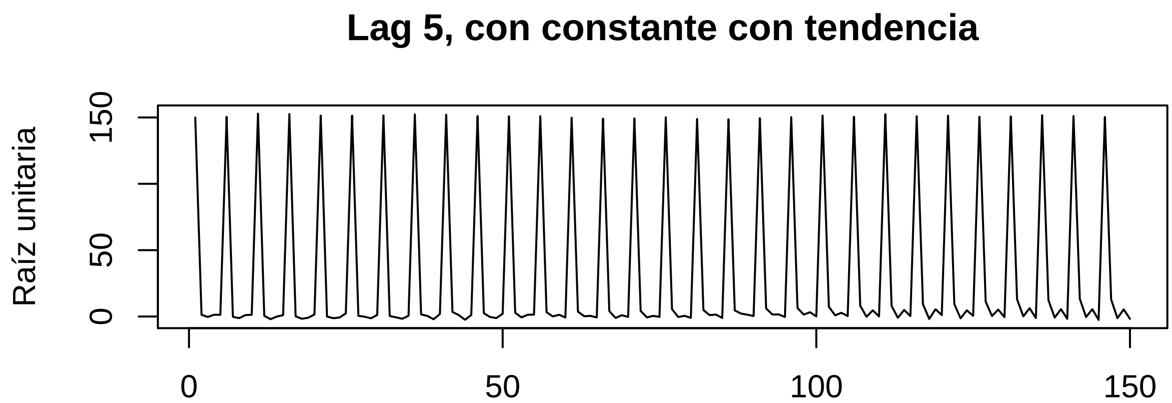

```r
adf.test.custom(y)
```

```
## Lag optimo
##          2
```

```r
adf.test(y, k=2)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -2.4034, Lag order = 2, p-value = 0.4083
## alternative hypothesis: stationary
```

```r
###################
set.seed(0)
y <- rep(0, n)
y[1] <- 150
y[2:5] <- rnorm(4)
for (i in 6:n)
{
  y[i] <- i/500 +  y[i-5] + rnorm(1)
}
plot(y, type='l', main = 'Lag 5, con constante con tendencia', xlab='',
     ylab='Raíz unitaria')
```




```r
adf.test.custom(y)
```

```
## Lag optimo
##          5
```

```r
adf.test(y, k=5)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -2.8455, Lag order = 5, p-value = 0.2241
## alternative hypothesis: stationary
```

2. *Programar la prueba de Busetti-Harvey basado en el paper* ***TESTING FOR THE PRESENCE OF A RANDOM WALK IN SERIES WITH STRUCTURAL BREAKS***

Después de leer el paper mencionado, es de notar que el test que se propone en la sección 5 es más versátil sin embargo su implementación requiere de un proceso de minimización el cual puede ser costoso además de distinguir entre cuatro casos (los casos para los modelos 1, 2, 2a y 2b).

Se optó por implementar la prueba simplificada referida en la sección IV que requiere saber de la posición de las interrupciones o saltos en la tendencia, sea que ésta sea originada por un salto en el intercepto del modelo 1 o bien un cambio en el intercepto o la pendiente del modelo 2.

Si bien **la distribución del estadístico $\epsilon_i^*(k)$ del test simplificado** (que se aborda en la sección 4 y se define en la ecuación 4.5) **depende de la localización de las posiciones del salto** o supuesto cambio estructural **su distribución asintótica no cambia** y distingue dos casos cuando se tiene el modelo 1 o el modelo 2. Dada esta bonita propiedad opte por implementar este test simplificado apoyado por lo mencionado al final de la cuarta sección del paper ''The conclusions are similar to those reached for the case of a single break, with the simplified test having a size close to the nominal and power comparable with the Latest", además de que en la práctica pueden identificarse los candidatos a puntos de salto en una serie y que el modelo 2 engloba a los modelos 2a y 2b.

Un punto importante de la implementación es **la cota que existe referente al número de posibles saltos estructurales**, que el paper denota como parámetro k, que debe de ser menor o igual a 4 en vista de que **no se pudieron replicar los valores críticos de la tabla II del paper** (pues al usar la definición 4.2 que estipula la suma de una serie de infinitos quantiles de variables aleatorias χ^2 con k grados de libertad no se pudieron reproducir los resultados de la tabla II, ni considerando el mismo cuantil para todas las v.a. ni cambiando este cuantil por los centiles hasta el p-valor requerido, los valores obtenidos no son proporcionales a los de la tabla II).

A continuación muestro la implementación de la prueba simplificada de la sección IV:

```
Busetti.Harvey <- function(y, option='both', k = k,
                           posicion=posicion,
                           p=.95)
{
  # y (numeric): vector con la serie de tiempo
  # K (int): numero de posibles saltos estructurales 1<=k<=4
  # posiciones (int): vector con los indices en donde la serie se sospecha que
            #presenta cambios estructurales
  # p (double): confianza a la que se requiere el test
  if(k >4) stop()
  serie <- y
  posicion <- posicion
  k <- k
   #creamos un dataframe con las posiciones para particionar la serie
  particion <- data.frame(start = c(1, posicion+1),
                          stop = c(posicion, length(serie)))
  e <- serie
    # a continuacion particionamos la serie con el data.frame 'particion'
  muestras <- lapply(X=1:dim(particion)[1],
                     function(x)
                       {
                          return(e[particion$start[x]:particion$stop[x]])

                     })
  estadistico.particion <- function(parte)
  {
    #calculo del numerados del estadistico dado por la ecuacion (4.5) del paper
```

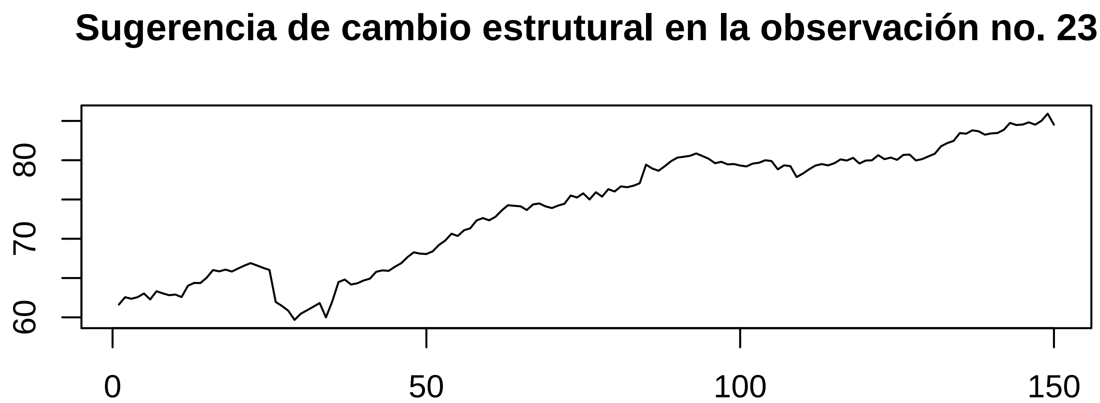
```r
    media.parte <- mean(unlist(parte), na.rm=TRUE )
    e.s <- sum((cumsum( parte- media.parte))**2)
    return(e.s/(length(parte)**2))
  }
  errores <- lapply(muestras, estadistico.particion)
  sigma <- var(serie) #para ambos casos la varianza se calcula igual
  # se termina calculo del estadistico de la ecuacion (4.5)
  estadistico <- sum(unlist(errores))/sigma
    #tabla de valores de los valores criticos para el modelo de la forma 1
  tabla1 <- data.frame(k = 1:4,
                      p0.9 =c(0.347, 0.607, 0.841, 1.063),
                      p0.95= c(0.461, 0.748, 1.000, 1.237  ),
                      p0.99 = c(0.743, 1.074, 1.359, 1.623 ))

  #tabla de valores de los valores criticos para el modelo de la forma 2
  tabla2 <- data.frame(k=1:4,
                      p0.9= c(0.119, 0.211, 0.296, 0.377),
                      p0.95= c(0.149, 0.247, 0.332, 0.423 ),
                      p0.99 = c(0.218, 0.329, 0.428, 0.521  ) )
  if(option == 'c') #determinacion del valor critico
  {
    valor.critico <- tabla1[ k , paste0('p',p)]
  } else {
    valor.critico <- tabla2[ k , paste0('p',p)]
  }
  a <- ifelse(estadistico >= valor.critico, 'Se rechaza H0, ie sí hay cambio estructural',
              'No se rechaza H0, ie no hay cambio estructural' )
  a <- paste0(a, ' en las posiciones: ', posicion, ' con confianza de: ', p)
  return(a)}  #regresamos un mensaje imformativo
```

Los casos de prueba para la función fueron el índice IGAE con el que hemos estado trabajando probando primero 1 cambio estructural en su constante en la posición 23 y acotando la serie a sus primeras 150 observaciones, como hemos anteriormente la prueba ADF señala la existencia de raíz unitaria en esta serie.

```r
y <- igae$IGAE[1:150]
adf.test(y)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -2.3316, Lag order = 5, p-value = 0.4382
## alternative hypothesis: stationary
```

```r
ts.plot(y, xlab='', ylab='', main='Sugerencia de cambio estrutural en la observación no. 23')
```




```r
Busetti.Harvey(y, option='c', k = 1, posicion = c(23), p = 0.95)
```

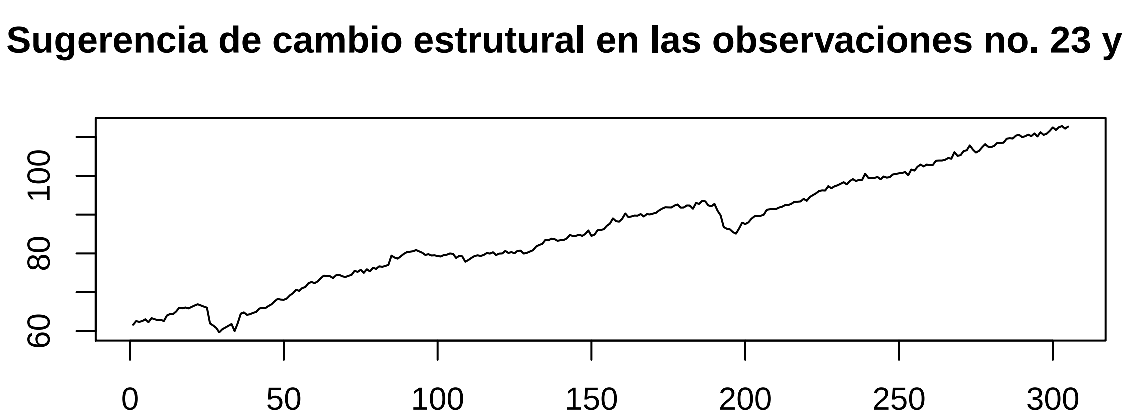
```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 23 con confianza de: 0.95"
```

El siguiente caso de prueba fue el test de dos saltos estructurales en las observaciones 23 y 193 del mismo indicador IGAE

```r
y <- igae$IGAE
adf.test(y)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -3.4103, Lag order = 6, p-value = 0.05294
## alternative hypothesis: stationary
```

```r
ts.plot(y, xlab='', ylab='', main='Sugerencia de cambio estrutural en las observaciones no. 23 y 193')
```

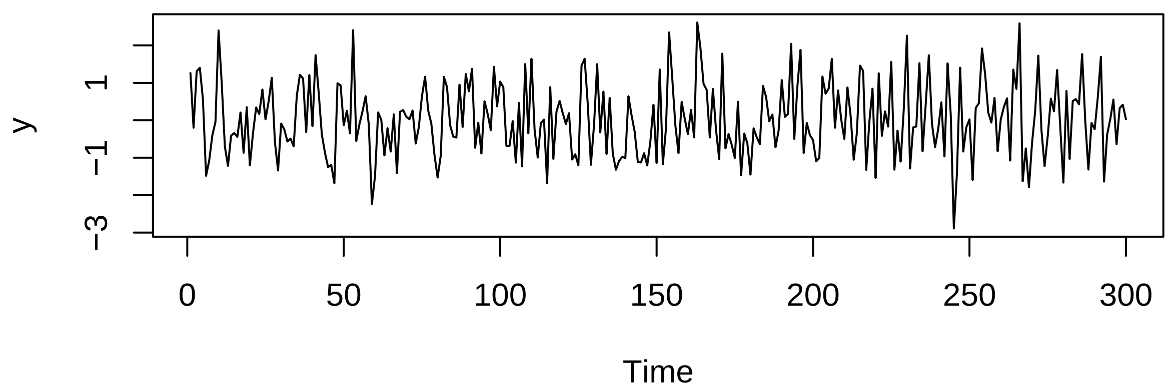


```r
Busetti.Harvey(y, option='c', k = 2, posicion = c(23), p = 0.95)
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 23 con confianza de: 0.95"
```

Un caso de prueba fue generar un modelo $AR(1)$ donde sabemos que no hay cambios estructurales en su valor medio sin tendencia, el test ADF nos dice que no hay presencia de raíz unitaria:

```r
set.seed(0)
n <- 300
y <- rep(0,n)
y[1] <- rnorm(1)
for (i in 2:n)
{
  y[i] <- 0.1*y[i-1] + rnorm(1)
}
ts.plot(y)
```




```r
adf.test(y)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -6.9435, Lag order = 6, p-value = 0.01
## alternative hypothesis: stationary
```

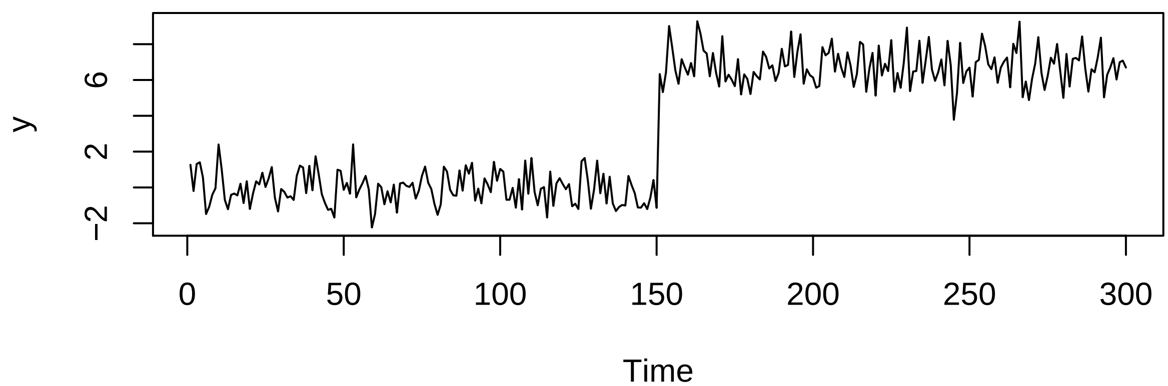
```r
Busetti.Harvey(y, option='c', k=1, posicion=150, p=.95)
```

```
## [1] "No se rechaza H0, ie no hay cambio estructural en las posiciones: 150 con confianza de: 0.95"
```

Y también lo probamos con un modleo $AR(1)$ donde si hay cambio en el valor medio a partir de la observación no. 150, y el test ADF señala la presencia de una raíz unitaria.

```r
set.seed(0)
n <- 300
y <- rep(0,n)
y[1] <- rnorm(1)
for (i in 2:150)
```

```r
{
  y[i] <- 0.1*y[i-1] + rnorm(1)
}
y[151] <- y[150] + 6 + rnorm(1)
for (i in 152:n)
{
  y[i] <- 0.1*y[i-1] + rnorm(1) + 6
}
ts.plot(y)
```




```r
adf.test(y)
```

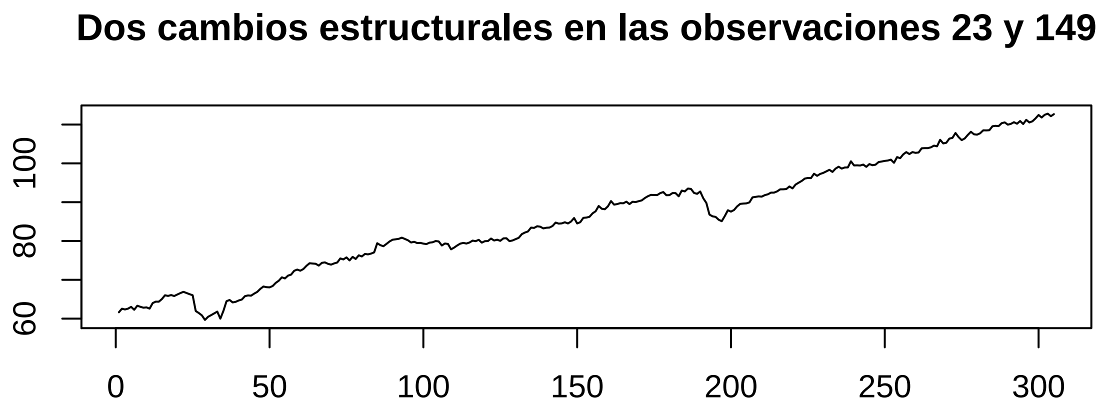
```
##
##  Augmented Dickey-Fuller Test
##
## data:  y
## Dickey-Fuller = -2.2644, Lag order = 6, p-value = 0.4649
## alternative hypothesis: stationary
```

```r
Busetti.Harvey(y, posicion = 159, k = 1, p = 0.99)
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 159 con confianza de: 0.99"
```

Para probar la implemntación en el caso de salto en la tendencia que depende del tiempo se consideraron las mismas series, primero el indicador IGAE con el que hemos trabajado y dos supuestos saltos estructurales en la media puntual o en la tendencia a través del tiempo:

```r
y <- igae$IGAE
ts.plot(y,
        main='Dos cambios estructurales en las observaciones 23 y 149',
        xlab='', ylab='')
```

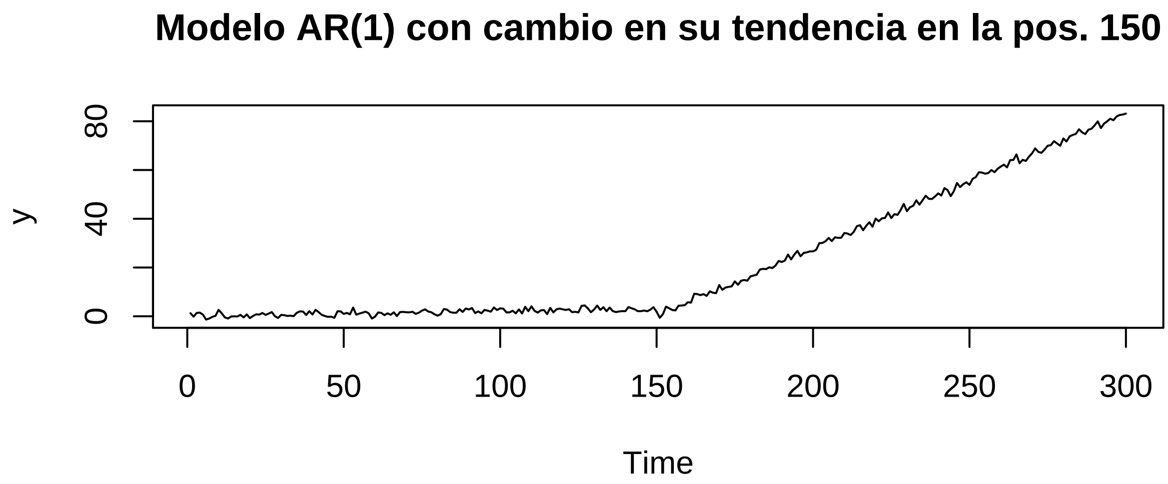



```r
Busetti.Harvey(y, option='both', k = 2, posicion = c(23,149), p = 0.95)
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 23 con confianza de: 0.95"
## [2] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 149 con confianza de: 0.95"
```

Nuevamente probamos con un modelo $AR(1)$ con cambio en su tendencia y continuo a trozos:

```r
set.seed(0)
y <- rep(0,n)
y[1] <- rnorm(1)
for (i in 2:150)
{
  y[i] <- 0.1*y[i-1] + rnorm(1) + i/50
}
for (i in 150:n)
{
  y[i] <- 0.1*y[i-1] + rnorm(1) + (i-150)/2
}
ts.plot(y, main='Modelo AR(1) con cambio en su tendencia en la pos. 150')
```

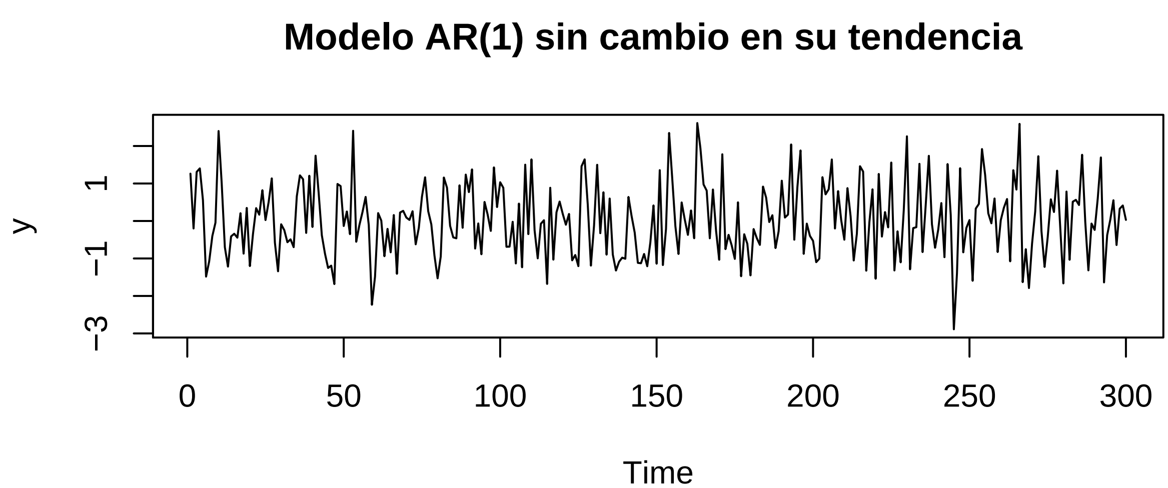



```r
Busetti.Harvey(y, posicion = 150, k = 1, p = 0.95, option='both')
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 150 con confianza de: 0.95"
```

Y lo probamos con el mismo modelo $AR(1)$ pero sin cambios en su tendencia:

```r
set.seed(0)
y <- rep(0,n)
y[1] <- rnorm(1)
for (i in 2:n)
{
  y[i] <- 0.1*y[i-1] + rnorm(1)
}
ts.plot(y, main='Modelo AR(1) sin cambio en su tendencia ')
```




```r
Busetti.Harvey(y, posicion = c(50), k = 1, p = 0.99, option='both')
```

```
## [1] "No se rechaza H0, ie no hay cambio estructural en las posiciones: 50 con confianza de: 0.99"
```

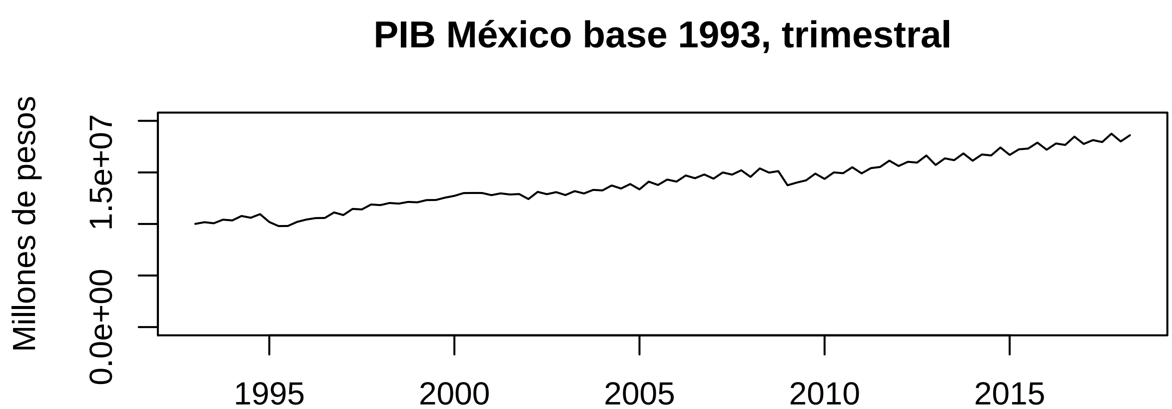
3. *Realizar contrastes de hipótesis de estacionariedad usando las pruebas programadas anteriormente para las siguientes series y ajustarles un modelo $ARIMA$ (especificar las pruebas de acuerdo a criterios estadísticos, gráficos y econométricos):*

a. *Producto Interno Bruto Trimestral de México, base 2013 (desestacionalizado).*

Los datos se obtuvieron de http://www.inegi.org.mx/sistemas/bie/?idserPadre=10200034#D10200034

Primero graficamos la serie del PIB con base en 1993, notamos que existe tendencia y según el test ADF existe una raíz unitaria (como también nuestro test de tendencia lineal lo sugiere y usaremos el lag recomendado por nuestra prueba con base en el criterio BIC), al realizar primeras diferencias (con lag propuesto por nuestra prueba implementada) la raíz unitaria desaparece.

```
PIB <- read.csv('BIE_BIE20180916230447.csv', skip = 2)
PIB <- PIB[,1:2]
names(PIB) <- c('Periodo', 'PIB,a.precios.de.mercado')
PIB <- ts(PIB$`PIB,a.precios.de.mercado`, start = 1993, frequency = 4)
PIB <- na.omit(PIB)
ts.plot(PIB, ylim=c(0,20000000), xlab='', ylab='Millones de pesos', main='PIB México base 1993, trimest
```




```
adf.test(unlist(PIB)) # raiz unitaria
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  unlist(PIB)
## Dickey-Fuller = -3.4258, Lag order = 4, p-value = 0.05401
## alternative hypothesis: stationary
```

```
adf.test.custom(PIB, option = 't')
```

```
## Lag optimo
##          4
```

```
adf.test(diff(unlist(PIB), k=4)) #no hay raiz
```

```
##
##  Augmented Dickey-Fuller Test
##
```

```
## data:  diff(unlist(PIB), k = 4)
## Dickey-Fuller = -4.6468, Lag order = 4, p-value = 0.01
## alternative hypothesis: stationary
```

Procedemos a realizar el test de cambios estructurales en tendencia en las observaciones 7 y 64:

```
Busetti.Harvey(unlist(PIB), option='both', k = 1, posicion = 64, p=0.95)
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 64 con confianza de: 0.95"
```

```
#ts.plot(PIB[1:20], ylim=c(0,20000000), xlab='', ylab='Millones de pesos', main='PIB México base 1993,
#posible cambio en 7
Busetti.Harvey(unlist(PIB), option='both', k = 1, posicion = 7, p=0.95)
```

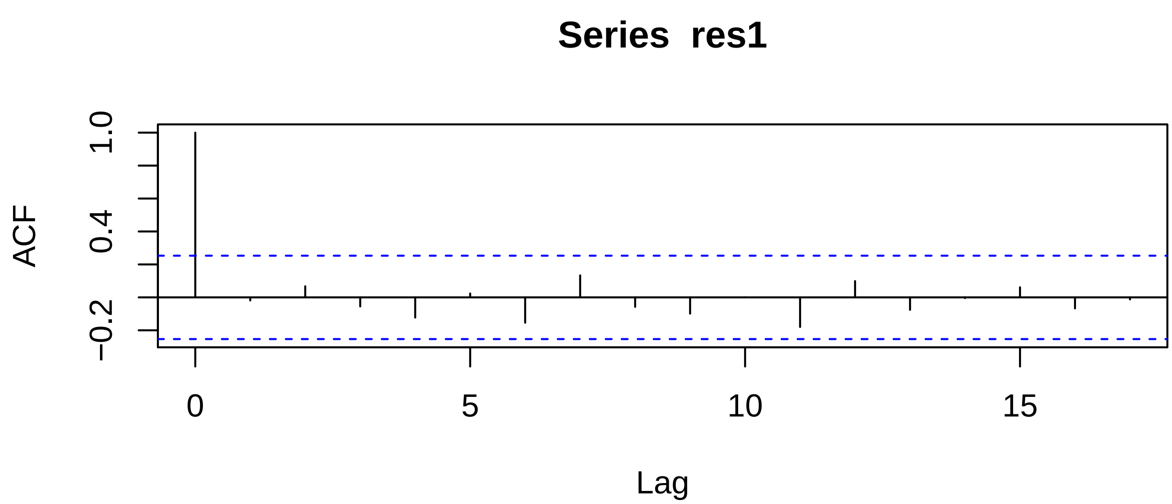
```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 7 con confianza de: 0.95"
```

```
Busetti.Harvey(unlist(PIB), option='both', k = 2, posicion = c(7,64), p=0.95)
```

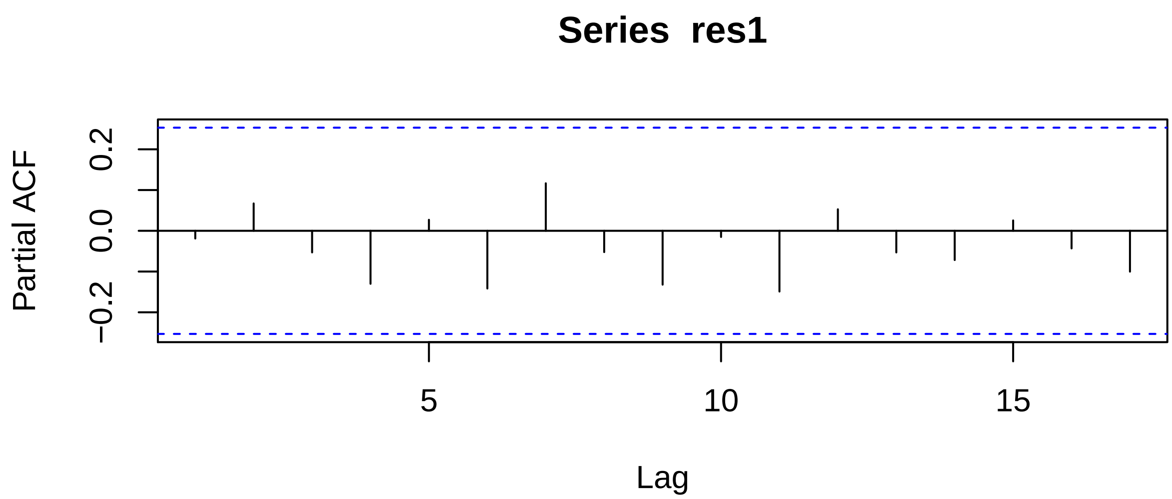
```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 7 con confianza de: 0.95"
## [2] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 64 con confianza de: 0.95"
```

En vista de que se presentan cambios estructurales en varias observaciones ajustamos solo dos modelos el primero para las primeras 64 observaciones, el cual presenta residuos autocorrelacionados pero son normales:

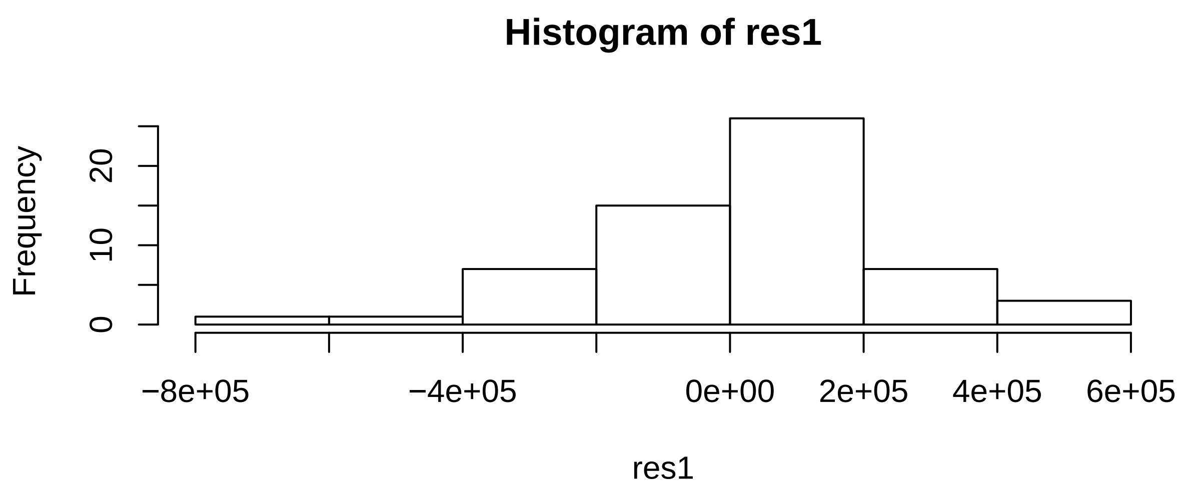
```
#auto.arima(diff(PIB[1:64], 4))
arima1 <- arima(diff(PIB[1:64], 4), order= c(0,0,3) )
#summary(arima1)
res1 <- arima1$residuals
acf(res1)
```




```
pacf(res1)
```




```r
hist(res1)
```



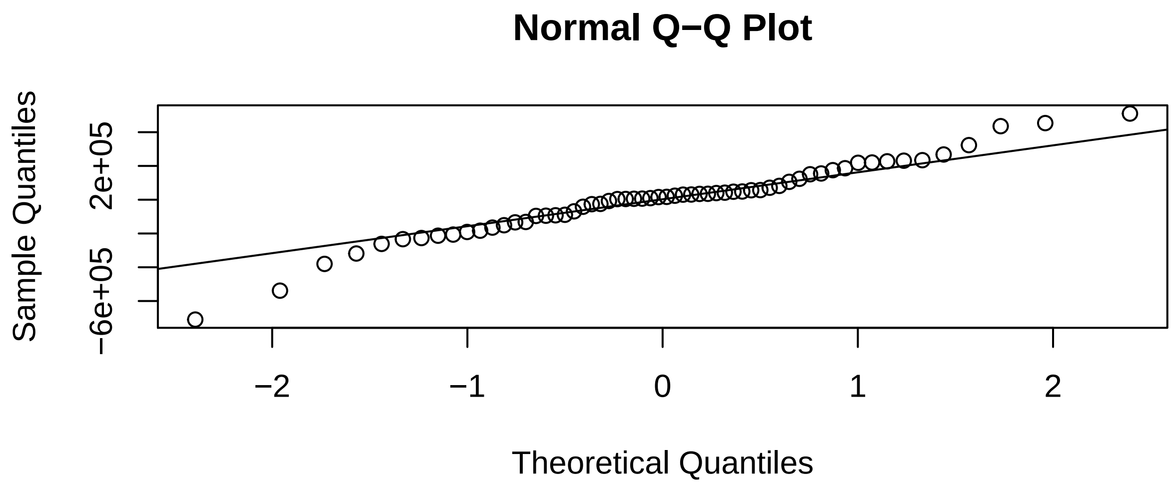

```r
shapiro.test(res1)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res1
## W = 0.9636, p-value = 0.07065
```

```r
BIC(arima1)
```

```
## [1] 1667.339
```

```r
qqnorm(res1)
qqline(res1)
```

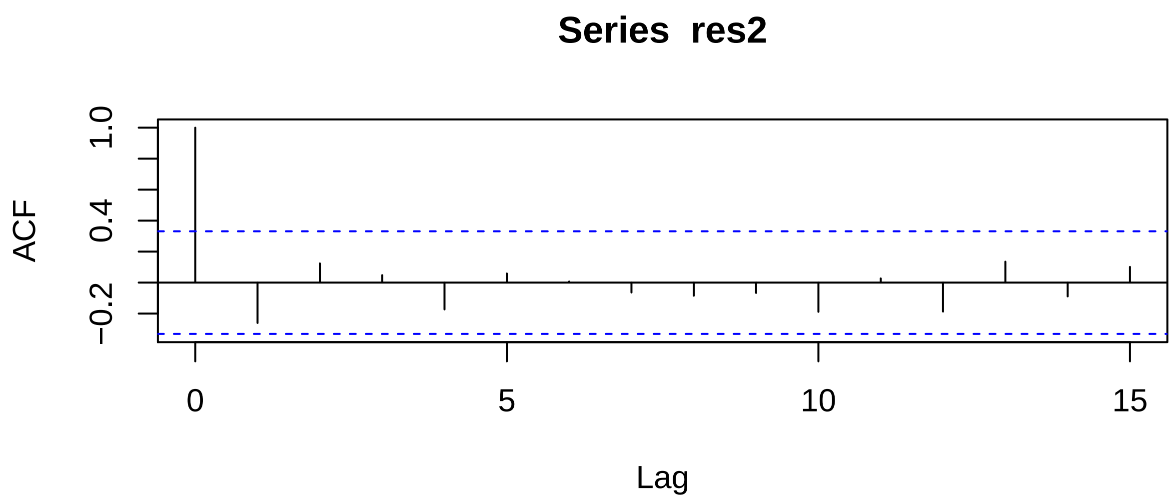



```
summary(lm(res1[-60]~ diff(res1)))
```

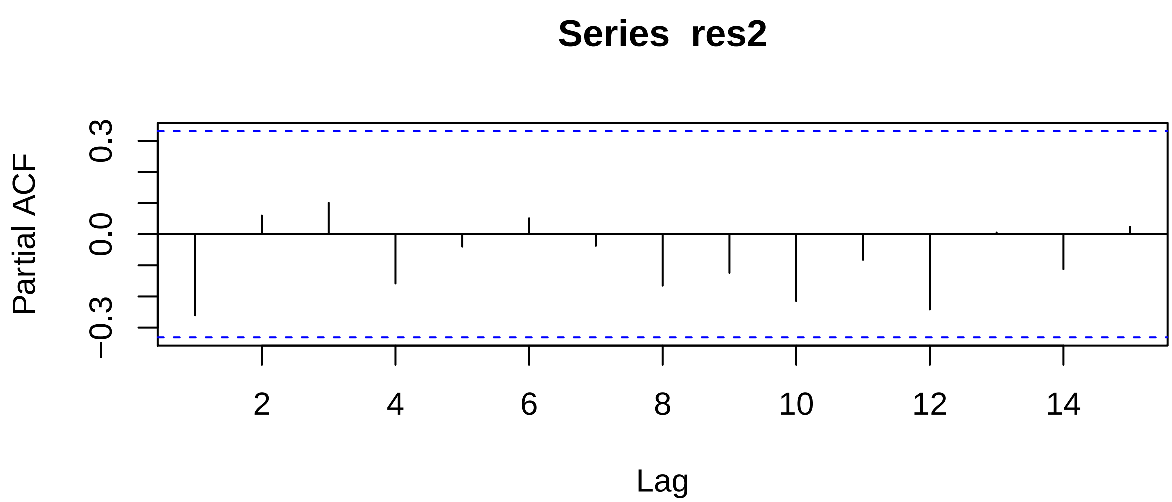
```
##
## Call:
## lm(formula = res1[-60] ~ diff(res1))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -626702  -63410    6988  100363  278957
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.146e+03  1.987e+04   0.209    0.835
## diff(res1)  -4.906e-01  6.495e-02  -7.553 3.81e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 152600 on 57 degrees of freedom
## Multiple R-squared:  0.5002, Adjusted R-squared:  0.4915
## F-statistic: 57.05 on 1 and 57 DF,  p-value: 3.806e-10
```

El segundo modelo para las observaciones restantes, prsenta residuos correlacionados y no normales

```
#auto.arima(diff(PIB[64:102], 4))
arima2 <- arima(diff(PIB[64:102], 4), order= c(0,0,1) )
#summary(arima2)
res2 <- arima2$residuals
acf(res2)
```




```r
pacf(res2)
```




```r
hist(res2)
```

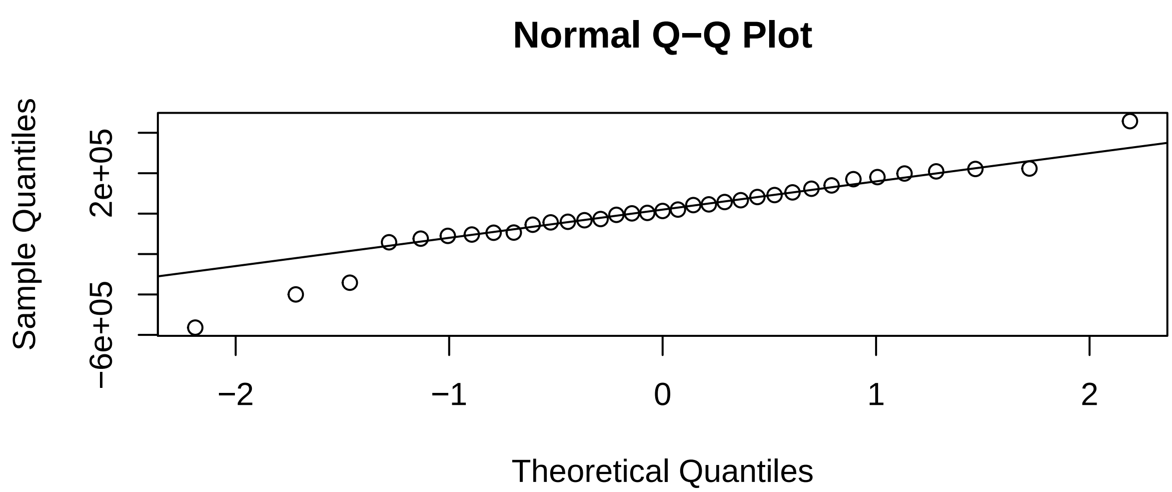

```r
shapiro.test(res2)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res2
## W = 0.93296, p-value = 0.0343
```

```r
BIC(arima2)
```

```
## [1] 959.8881
```

```r
qqnorm(res2)
qqline(res2)
```



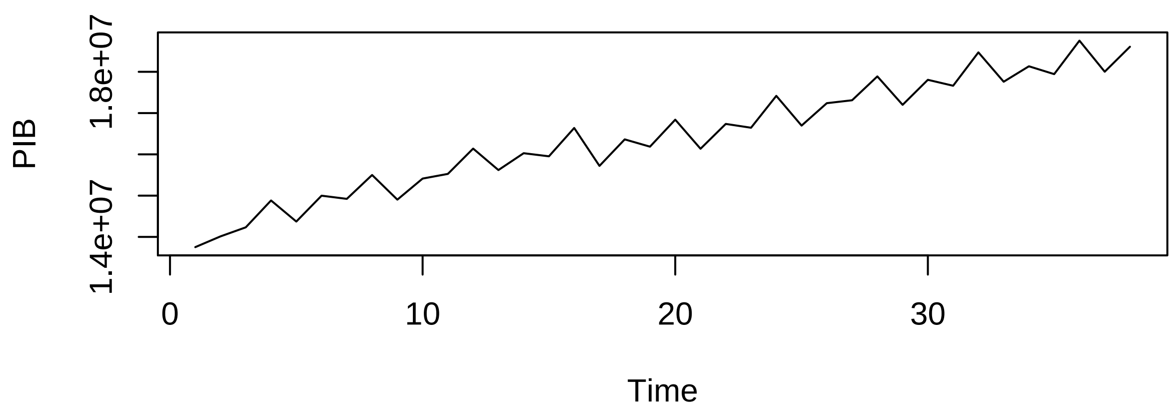

```r
summary(lm(res2[-35]~ diff(res2)))
```

```
##
## Call:
```

```
## lm(formula = res2[-35] ~ diff(res2))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -210584  -68986    5032   55103  309238
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.583e+04  1.793e+04   0.883    0.384
## diff(res2)  -5.603e-01  6.373e-02  -8.791 4.81e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 104300 on 32 degrees of freedom
## Multiple R-squared:  0.7072, Adjusted R-squared:  0.698
## F-statistic: 77.28 on 1 and 32 DF,  p-value: 4.806e-10
```

Sin embargo al considerar la importancia del cambio estructural en la observación no. 64 procedemos a un nuevo análisis, checamos primero si posee una raíz unitaria, y como el test *ADF* lo indica al considerar segundas diferencias tenemos un proceso estacionario.

```
PIB <- PIB[65:102]
ts.plot(PIB)
```




```
adf.test(PIB)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  PIB
## Dickey-Fuller = -2.244, Lag order = 3, p-value = 0.4779
## alternative hypothesis: stationary
```

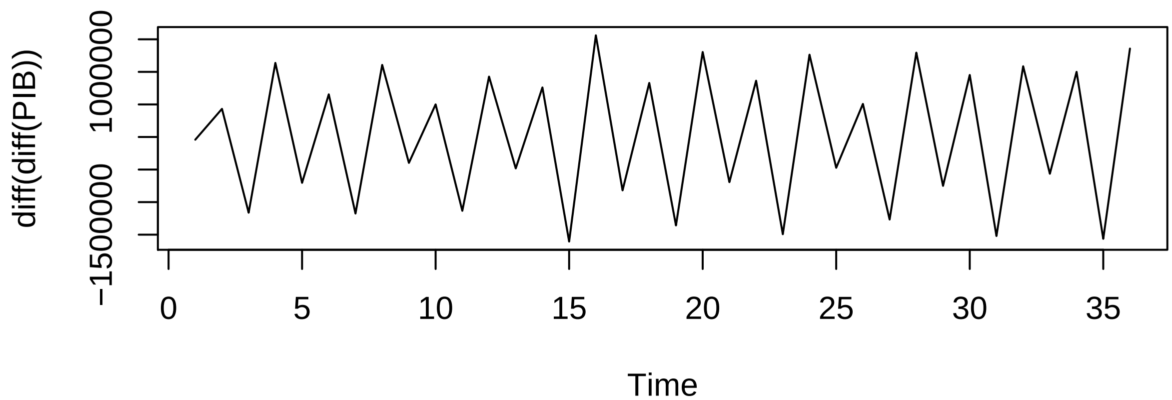
```
adf.test.custom(diff(diff(PIB)), option='t')
```

```
## Lag optimo
##          1
```

```r
adf.test(diff(diff(PIB)))
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(diff(PIB))
## Dickey-Fuller = -7.9441, Lag order = 3, p-value = 0.01
## alternative hypothesis: stationary
```

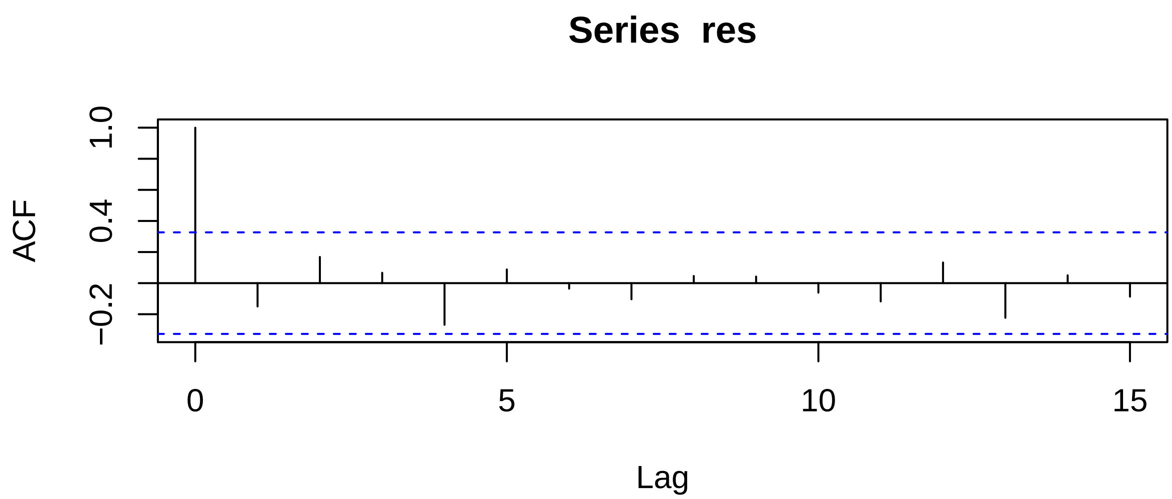
```r
ts.plot(diff(diff(PIB)))
```



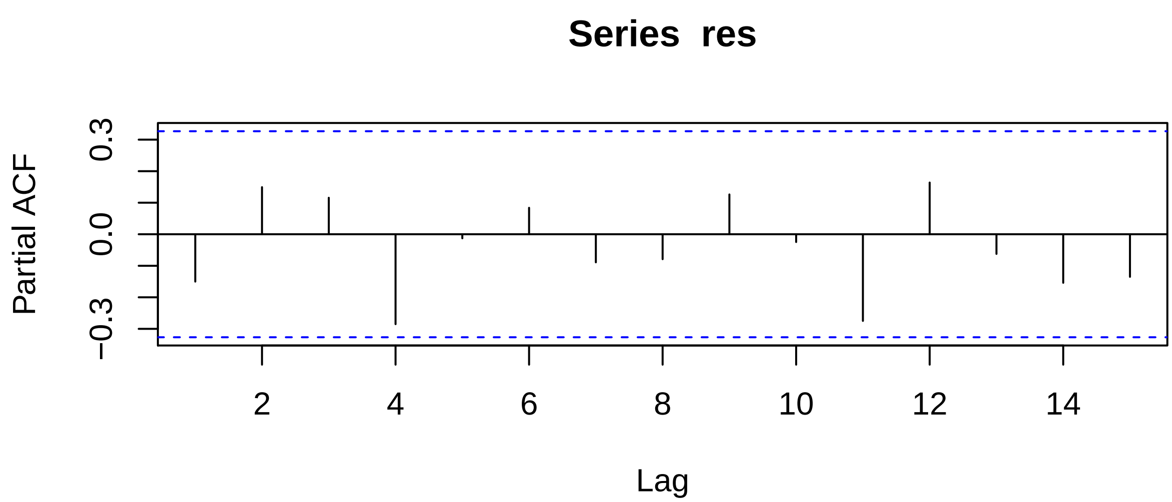

```r
#auto.arima(diff(diff(PIB)))
modelo.pronostico <- arima(diff(diff(PIB)), order=c(4,0,2))
res <- modelo.pronostico$residuals
```

Que después de ajustar un modelo ARIMA, encontramos que aunque la muestra es pequeña sus residuos son normales aunque autocorrelacionados.

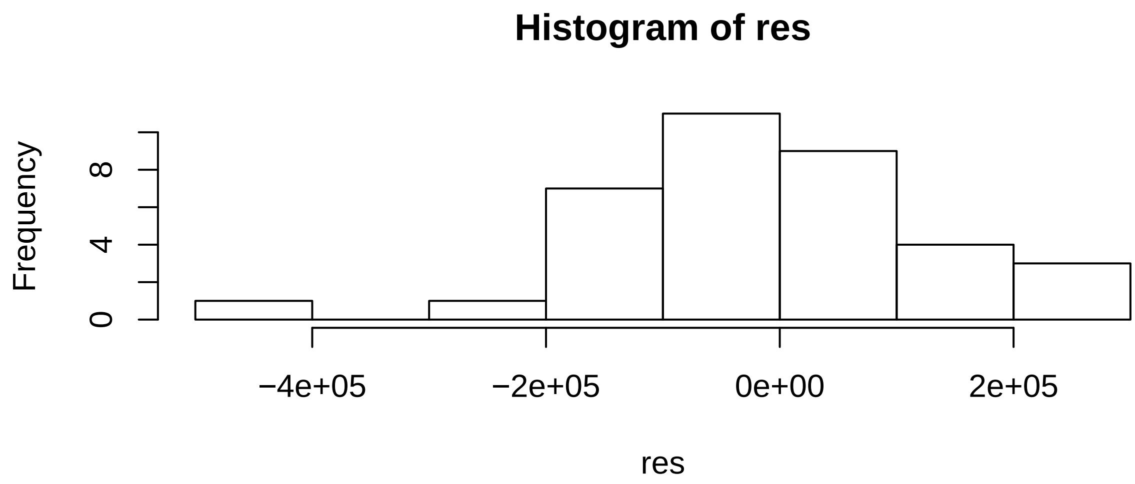
```r
acf(res)
```




```r
pacf(res)
```




```r
hist(res)
```



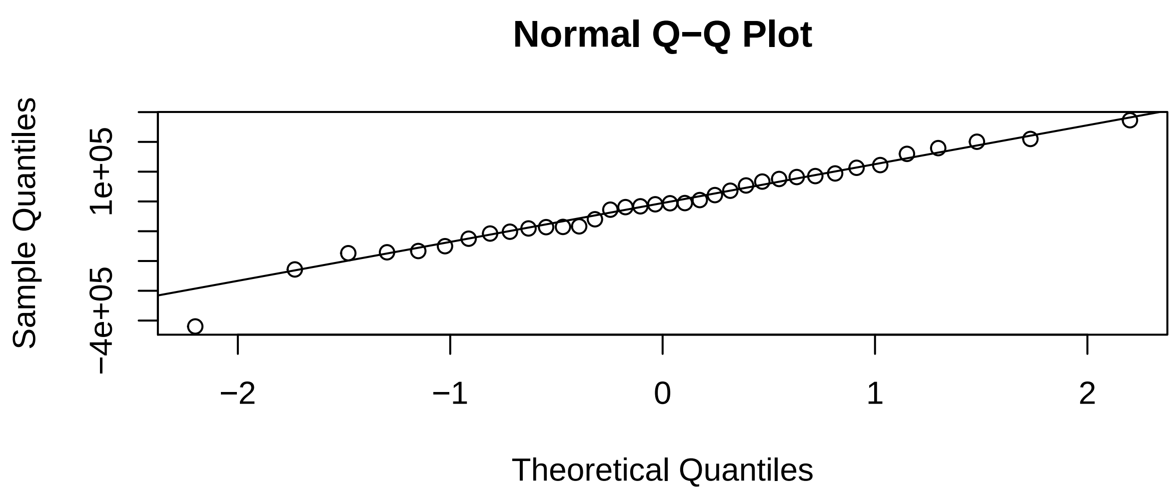

```
shapiro.test(res)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res
## W = 0.97952, p-value = 0.7293
```

```
BIC(modelo.pronostico)
```

```
## [1] 994.1333
```

```
qqnorm(res)
qqline(res)
```




```
summary(lm(res[-36]~ diff(res)))
```

```
##
```

```
## Call:
## lm(formula = res[-36] ~ diff(res))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -204405  -53654    5889   65030  145717
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.245e+04  1.571e+04  -0.792    0.434
## diff(res)   -4.905e-01  7.454e-02  -6.580 1.77e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 92940 on 33 degrees of freedom
## Multiple R-squared:  0.5675, Adjusted R-squared:  0.5544
## F-statistic:  43.3 on 1 and 33 DF,  p-value: 1.768e-07
```

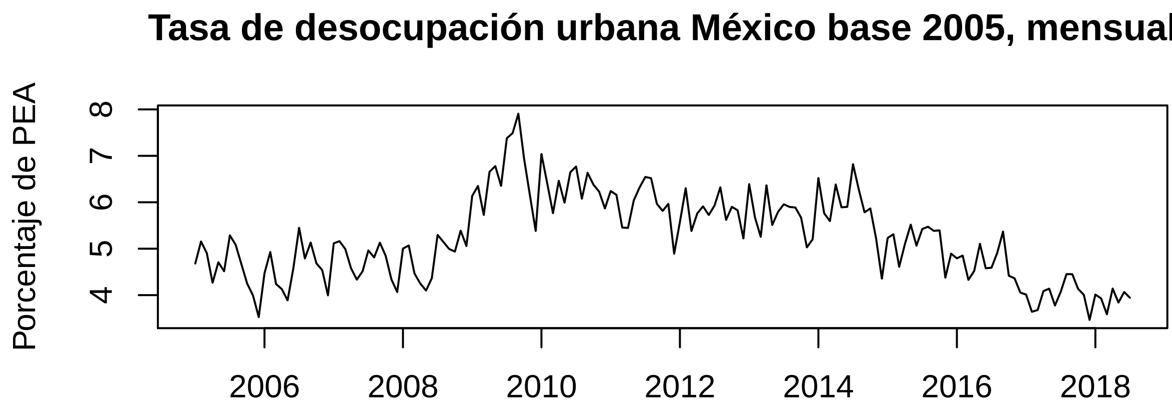
La conclusión de este ejercicio fue la importancia del cambio estructural en la observación 64, la cual corresponde al primer mes del año 2008, hecho que se corresponde con la crisis internacional originada en EU y que como podemos ver afectado el PIB de México. Un aspecto muy importante a recalcar es que este cambio estructural hizo que la serie a partir de 2008 fuese de un orden de integración mayor. En los ejercicios siguientes de igual manera se prestará atención a el último cambio estructural pues como es costumbre una de las aplicaciones de series de tiempo es el pronóstico hacia adelante.

b. *Tasa de desocupación urbana (desestacionalizada).*

Los datos se obtuvieron de [http://www.inegi.org.mx/sistemas/bie/]http://www.inegi.org.mx/sistemas/bie/)

Primero graficamos la serie de la tasa de desocupación urbana con base en 2005, notamos que existe tendencia alrededor de una media fija y según el test *ADF* existe una raíz unitaria (como también nuestro test de tendencia media lo sugiere y usaremos el lag recomendado por nuestra prueba con base en el criterio BIC), al realizar primeras diferencias (con lag propuesto por nuestra prueba implementada) la raíz unitaria desaparece.

```
des <- read.csv('BIE_BIE20180917024302.csv', skip = 3)
des <- des[,1:2]
names(des) <- c('Periodo', 'Tasa.desocupacion')
des <- ts(des$Tasa.desocupacion, start = 2005, frequency = 12)
des <- na.omit(des)
ts.plot(des, xlab='', ylab='Porcentaje de PEA', main='Tasa de desocupación urbana México base 2005, mens
```




```
adf.test(unlist(des)) # raiz unitaria
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  unlist(des)
## Dickey-Fuller = -1.3464, Lag order = 5, p-value = 0.849
## alternative hypothesis: stationary
```

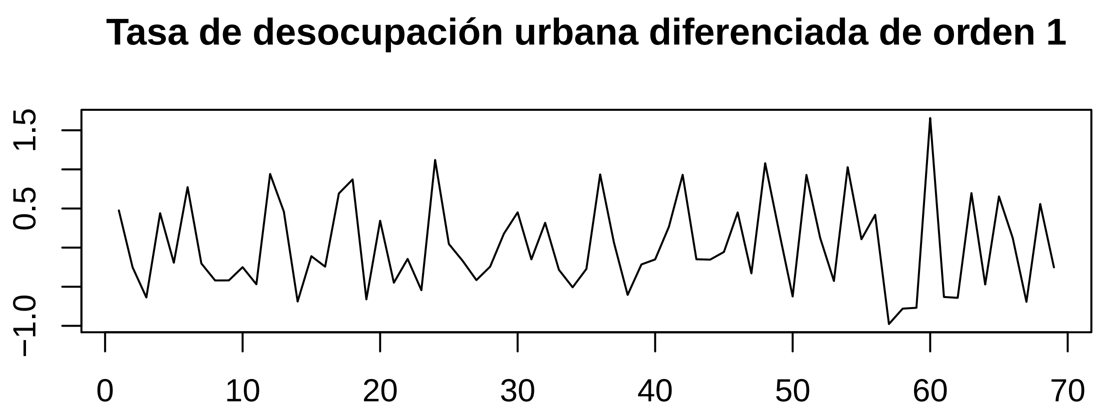
```
adf.test.custom(unlist(des), option = 'both')
```

```
## Lag optimo
##          1
```

```
adf.test(diff(unlist(des), k=1)) #no hay raiz
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(unlist(des), k = 1)
## Dickey-Fuller = -6.3512, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

```
ts.plot(diff(des[1:70]), xlab='', ylab='', main='Tasa de desocupación urbana diferenciada de orden 1')
```



Procedemos a realizar el test de cambios estructurales en tendencia alrededor de la observacion 60 correspondiente:

```
Busetti.Harvey(unlist(des), option='c', k = 1, posicion = 60, p=0.99)
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 60 con confianza de: 0.99"
```

```
Busetti.Harvey(unlist(des), option='both', k = 1, posicion = 65, p=0.99)
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 65 con confianza de: 0.99"
```

```
Busetti.Harvey(unlist(des), option='both', k = 1, posicion = c(55), p=0.99)
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 55 con confianza de: 0.99"
```

En vista de que se presentan cambios estructurales alrededor del inicio del 2010 ajustamos solo dos modelos el primero para las primeras 60 observaciones, el cual presenta residuos autocorrelacionados pero son normales:

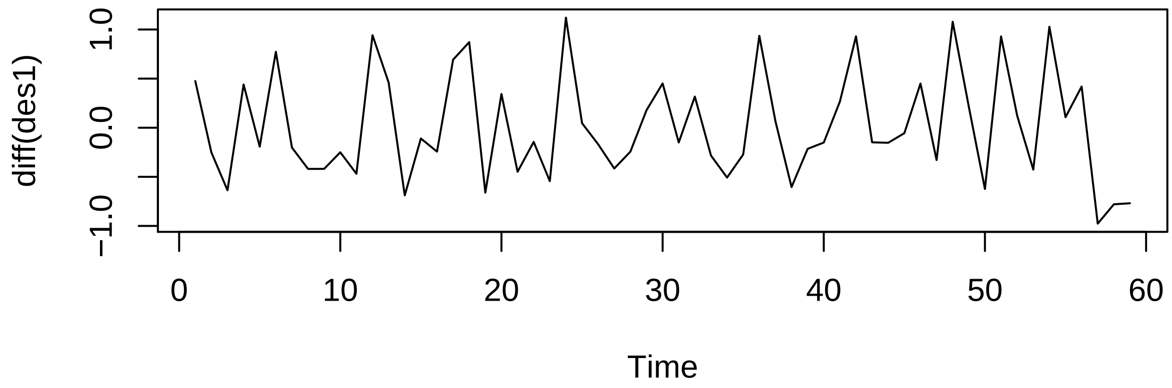
```
des1 <- des[1:60]
adf.test(diff(des1))
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(des1)
## Dickey-Fuller = -5.5227, Lag order = 3, p-value = 0.01
## alternative hypothesis: stationary
```

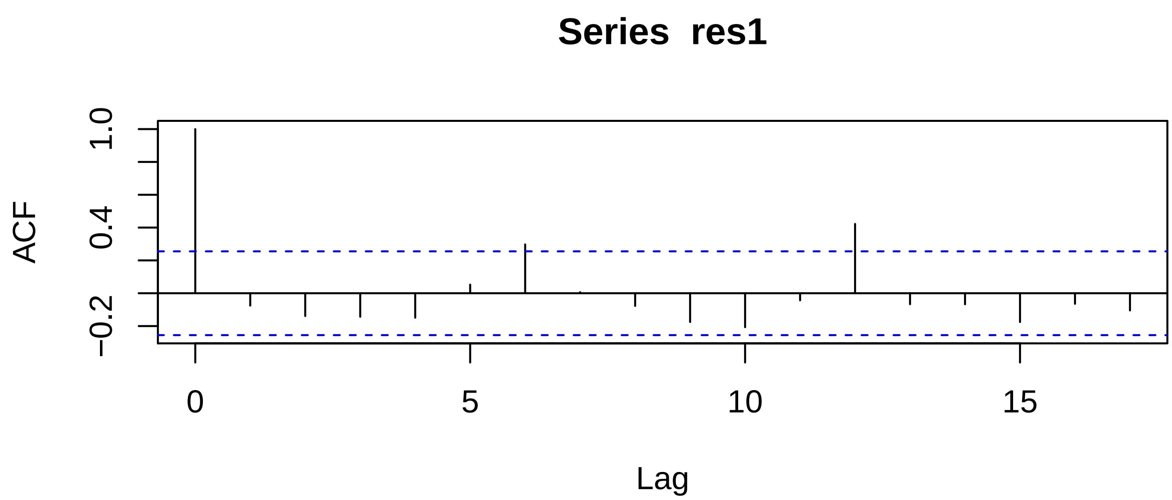
```
adf.test.custom(des1, option = 'both')
```

```
## Lag optimo
##          1
```

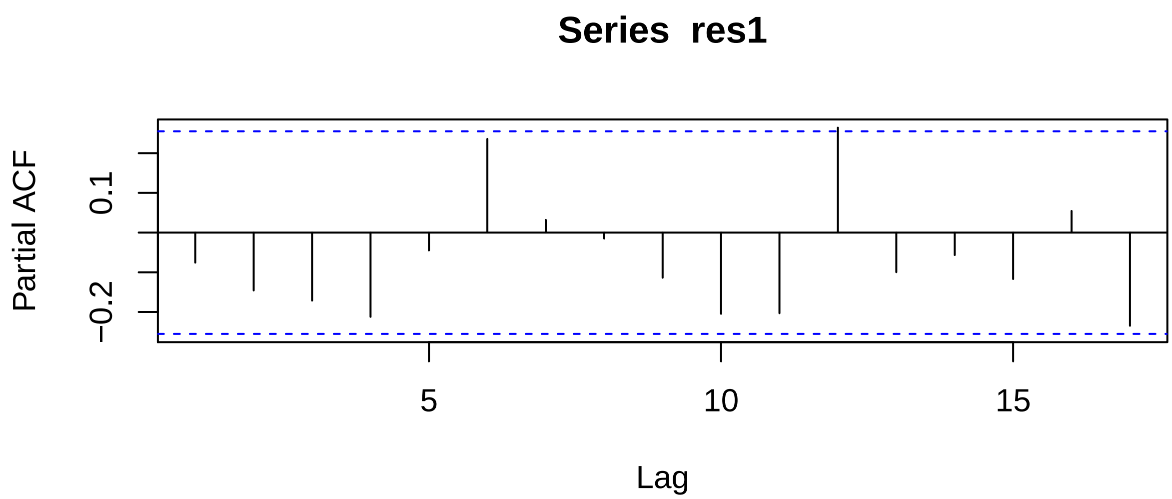
```
ts.plot(diff(des1))
```



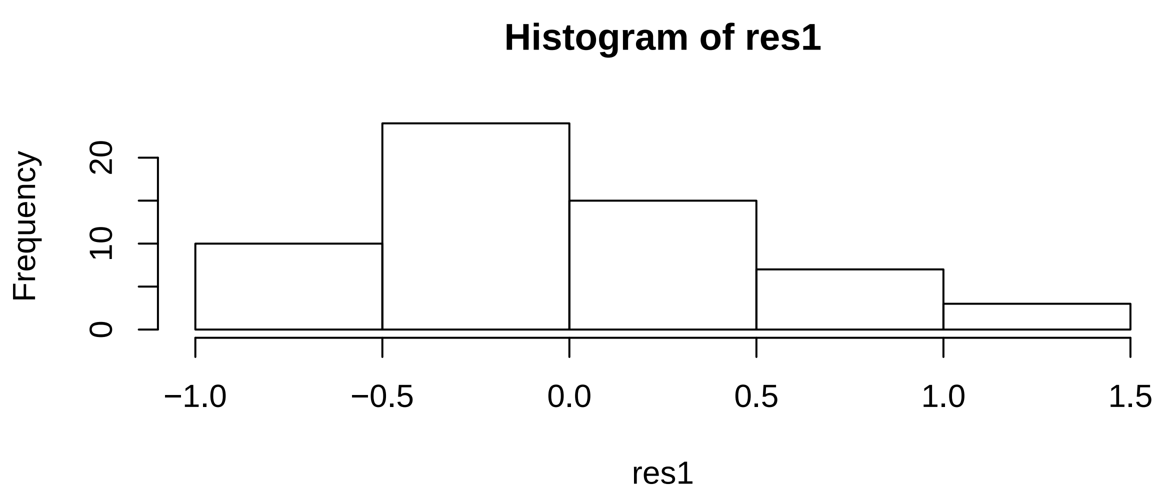

```r
#auto.arima(diff(des1))
arima1 <- arima(diff(des1), order= c(0,0,0) )
#summary(arima1)
res1 <- arima1$residuals
res1 <- na.omit(res1)
acf(res1)
```




```r
pacf(res1)
```




```r
hist(res1)
```



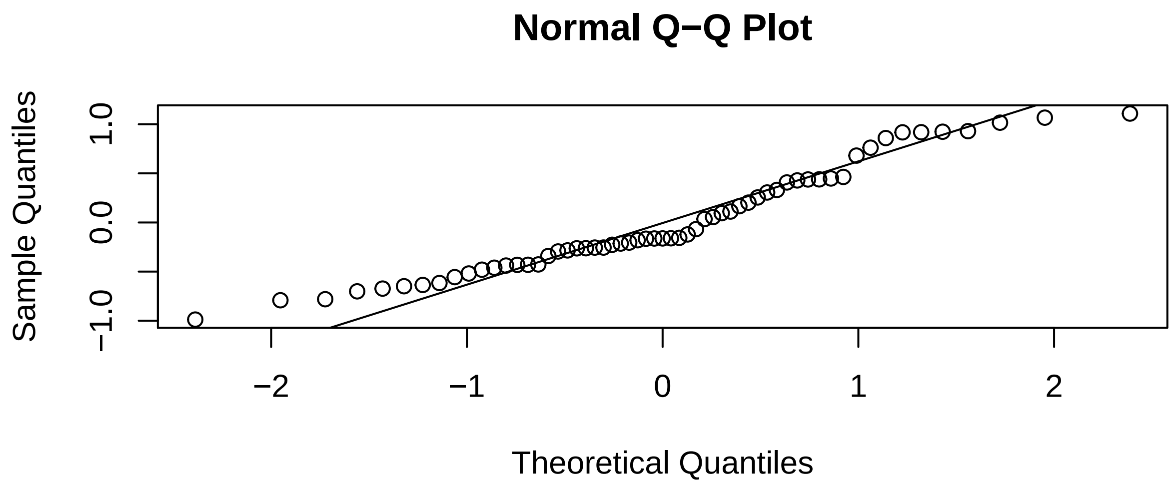

```r
shapiro.test(res1)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res1
## W = 0.94743, p-value = 0.01281
```

```r
BIC(arima1)
```

```
## [1] 102.551
```

```r
qqnorm(res1)
qqline(res1)
```




```r
summary(lm(res1[-59]~ diff(res1)))
```

```
##
## Call:
## lm(formula = res1[-59] ~ diff(res1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.89415 -0.30052 -0.02228  0.26861  0.76632
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.002855   0.048707   0.059    0.953
## diff(res1)  -0.494411   0.061842  -7.995 7.89e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3708 on 56 degrees of freedom
## Multiple R-squared:  0.533,  Adjusted R-squared:  0.5247
## F-statistic: 63.92 on 1 and 56 DF,  p-value: 7.888e-11
```

Acontinuación ajustamos un modelo ARIMA solo para las observaciones siguientes al 2010

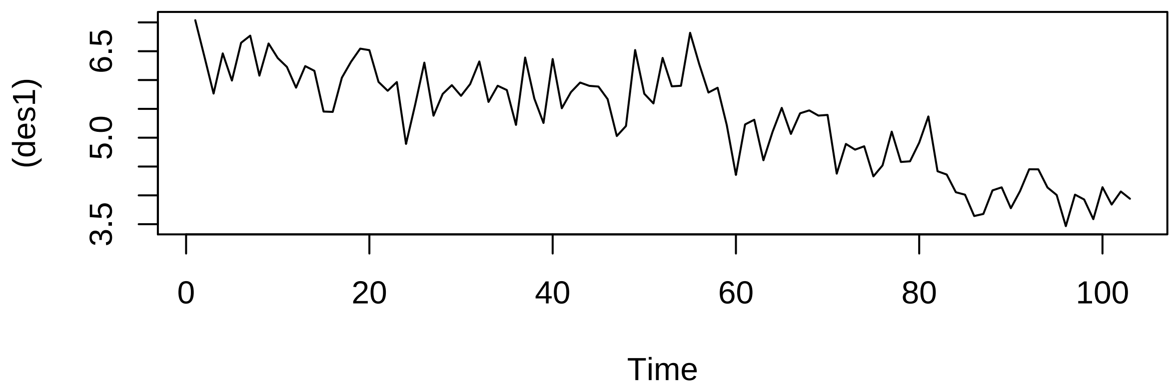
```r
des1 <- des[61:163]
adf.test((des1)) #existe raiz
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  (des1)
## Dickey-Fuller = -3.3478, Lag order = 4, p-value = 0.06698
## alternative hypothesis: stationary
```

```r
adf.test.custom(des1, option = 't')
```

```
## Lag optimo
##          1
```

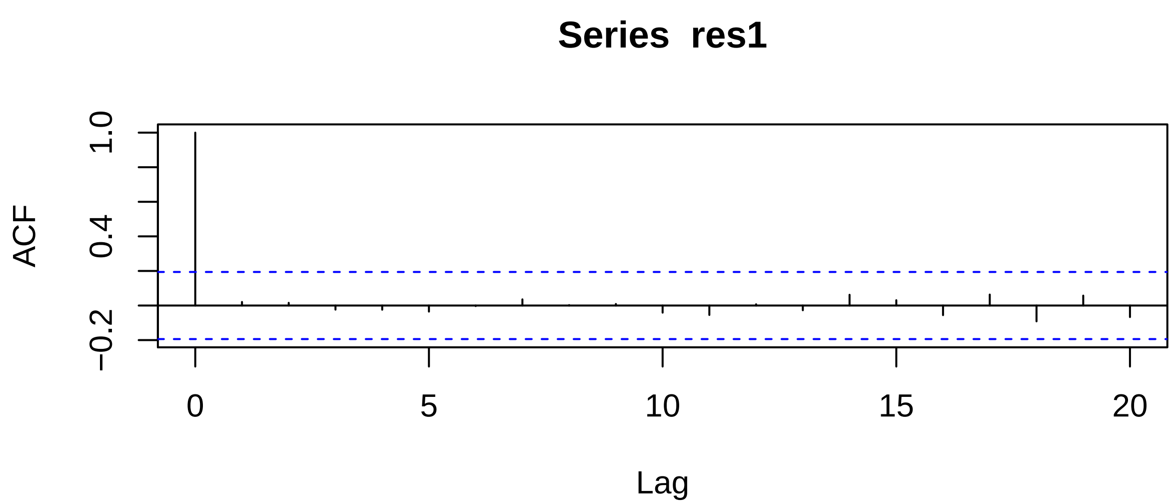
```r
ts.plot((des1)) #tendencia a la baja
```



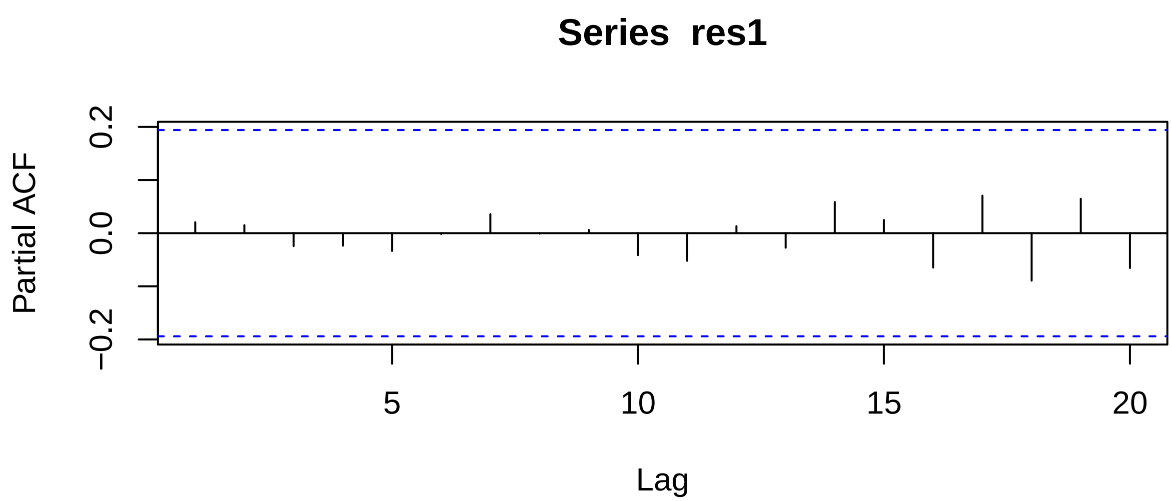

```r
adf.test(diff(des1)) #se elimina tendencia
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(des1)
## Dickey-Fuller = -5.7921, Lag order = 4, p-value = 0.01
## alternative hypothesis: stationary
```

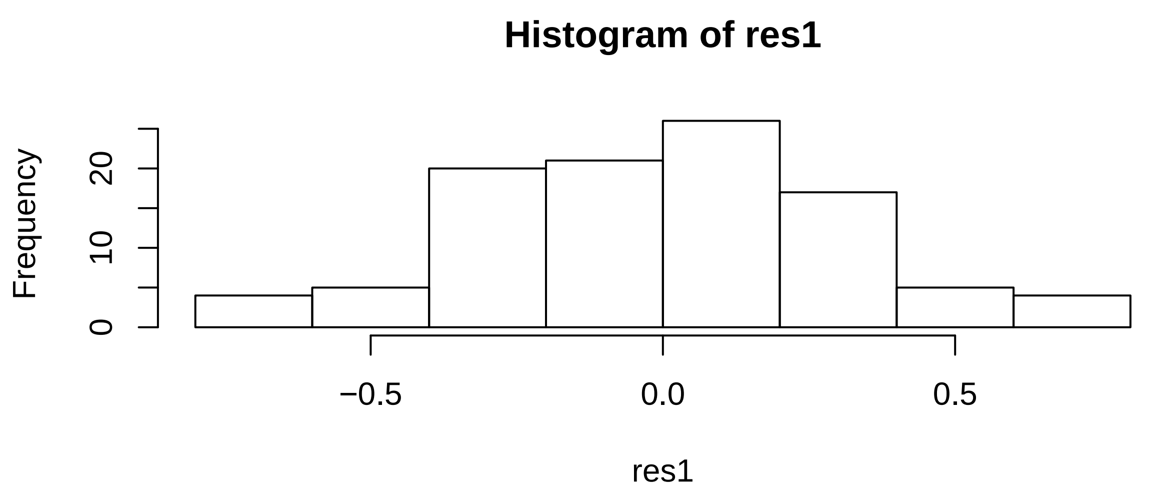
```r
#auto.arima(diff(des1))
arima1 <- arima(diff(des1), order= c(12,0,8) )
#summary(arima1)
res1 <- arima1$residuals
res1 <- na.omit(res1)
acf(res1)
```



```r
pacf(res1)
```




```r
hist(res1)
```



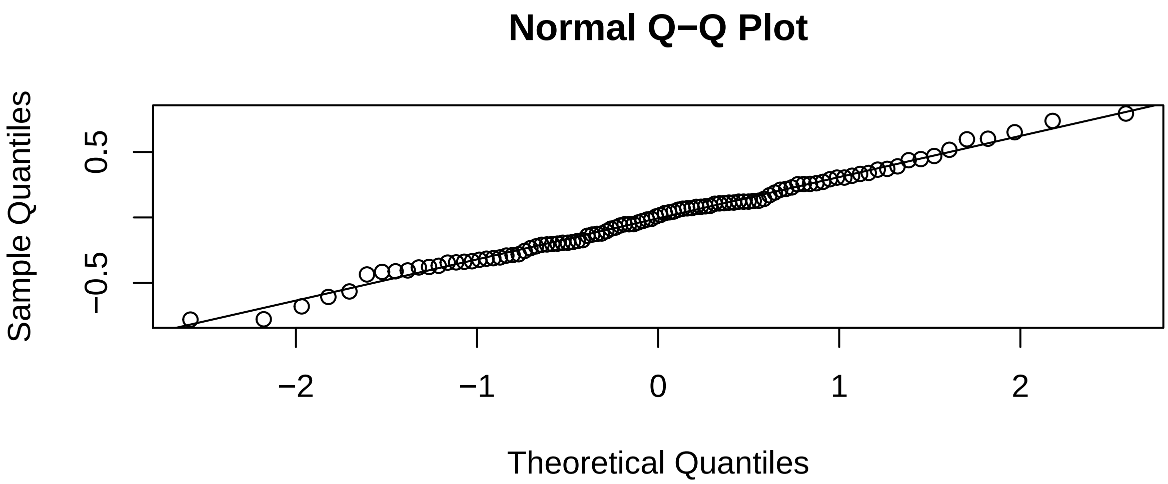

```r
shapiro.test(res1)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res1
## W = 0.99233, p-value = 0.8362
```

```r
BIC(arima1)
```

```
## [1] 164.9362
```

```r
qqnorm(res1)
qqline(res1)
```




```r
summary(lm(res1[-102]~ diff(res1)))
```

```
##
## Call:
## lm(formula = res1[-102] ~ diff(res1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.49429 -0.14968 -0.01907  0.17704  0.56217
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.004408   0.022700  -0.194    0.846
## diff(res1)  -0.503345   0.051310  -9.810 2.84e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2281 on 99 degrees of freedom
## Multiple R-squared:  0.4929, Adjusted R-squared:  0.4878
## F-statistic: 96.24 on 1 and 99 DF,  p-value: 2.842e-16
```

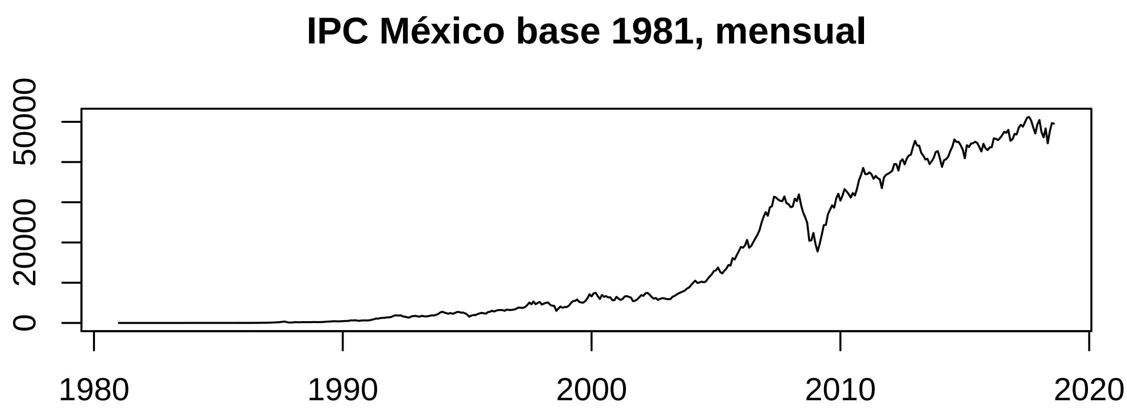
Como podemos apreciar el cambio estructural a partir del 2010 se ve reflejado en los parámetros de los modelos ARIMA mientras en las observaciones anteriores a 2010 solo se requiere de diferenciar una vez la serie para las posteriores a 2010 se requiere de la estimación de 20 parámetros adicionales, desconocemos qué fenómeno se presentó en el año 2010 que se causa de este cambio.

c. *Indice de precios y cotizaciones de la Bolsa Mexicana de Valores (último índice del mes).*

En esta ocación los datos se obtuvieron de http://www.banxico.org.mx/SieInternet/consultarDirectorioInternetAction.do?accion=consultarCuadro&idCuadro=CF57&sector=7&locale=es

Primero graficamos la serie del IPC con base en 1981, notamos que existe tendencia alrededor de una media fija y según el test *ADF* existe una raíz unitaria (como también nuestro test de tendencia media lo sugiere y usaremos el lag recomendado por nuestra prueba con base en el criterio BIC), al realizar primeras diferencias (con lag propuesto por nuestra prueba implementada) la raíz unitaria desaparece.

```r
IPC <- read.csv('Consulta_20180917-031719187.csv', skip = 17)
IPC <- IPC[,1:2]
names(IPC) <- c('Periodo', 'IPC')
IPC <- ts(IPC$IPC, start = 1981, frequency = 12)
IPC <- na.omit(IPC)
ts.plot(IPC, xlab='', ylab='', main='IPC México base 1981, mensual')
```




```r
adf.test(unlist(IPC)) # raiz unitaria
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  unlist(IPC)
## Dickey-Fuller = -1.5597, Lag order = 7, p-value = 0.764
## alternative hypothesis: stationary
```

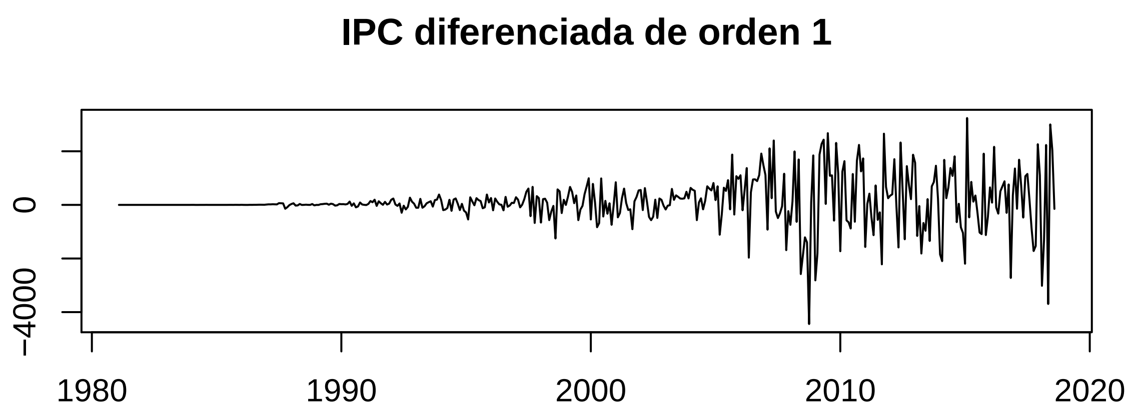
```r
adf.test.custom(unlist(IPC), option = 'c')
```

```
## Lag optimo
##          1
```

```r
adf.test(diff(unlist(IPC), k=1)) #no hay raiz
```

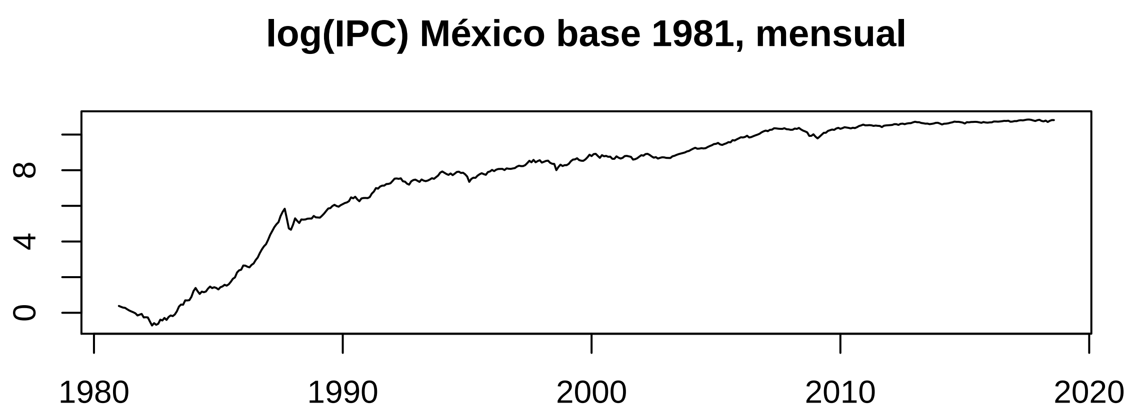
```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(unlist(IPC), k = 1)
## Dickey-Fuller = -7.7807, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary
```

```r
ts.plot(diff(IPC), xlab='', ylab='', main='IPC diferenciada de orden 1')
```



Lo anterior nos hace sospechar sobre cambios multiplicativos en la varianza de la serie por lo que rehacemos los test pero sobre los logaritmos de los datos:

```
IPC <- read.csv('Consulta_20180917-031719187.csv', skip = 17)
IPC <- IPC[,1:2]
names(IPC) <- c('Periodo', 'IPC')
IPC <- ts(log(IPC$IPC), start = 1981, frequency = 12)
IPC <- na.omit(IPC)
ts.plot(IPC, xlab='', ylab='', main='log(IPC) México base 1981, mensual')
```




```
adf.test(unlist(IPC)) # raiz unitaria
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  unlist(IPC)
## Dickey-Fuller = -1.6474, Lag order = 7, p-value = 0.7269
## alternative hypothesis: stationary
```

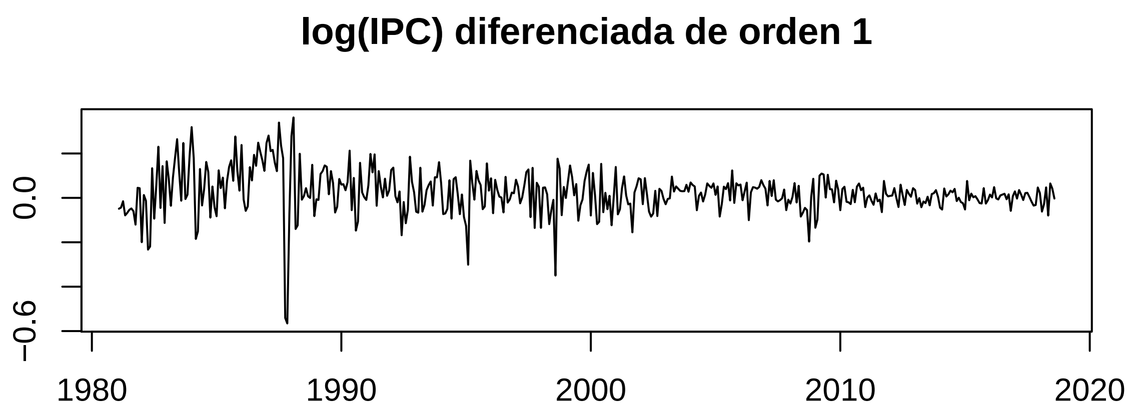
```r
adf.test.custom(unlist(IPC), option = 'c')
```

```
## Lag optimo
##          1
```

```r
adf.test(diff(unlist(IPC), k=1)) #no hay raiz
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  diff(unlist(IPC), k = 1)
## Dickey-Fuller = -6.5795, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary
```

```r
ts.plot(diff(IPC), xlab='', ylab='', main='log(IPC) diferenciada de orden 1')
```



Con lo anterior vemos que la varianza se estabiliza, y nos hace sospechar de un cambio estructural antes de 1990, en la observación no. 82 perteneciente al mes de noviembre de 1987

Procedemos a realizar el test de cambios estructurales en media alrededor de la observacion 82:

```r
Busetti.Harvey(unlist(IPC), option='c', k = 1, posicion = 82, p=0.99)
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 82 con confianza de: 0.99"
```

```r
Busetti.Harvey(unlist(des), option='c', k = 1, posicion = 80, p=0.99)
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 80 con confianza de: 0.99"
```

```r
Busetti.Harvey(unlist(des), option='both', k = 1, posicion = 82, p=0.99)
```

```
## [1] "Se rechaza H0, ie sí hay cambio estructural en las posiciones: 82 con confianza de: 0.99"
```

En vista de que se presentan cambios estructurales alrededor de la observación no. 82 ajustamos solo dos modelos el primero para las primeras 82 observaciones, el cual presenta residuos no autocorrelacionados pero no normales:

```r
des1 <- IPC[1:82]
adf.test(diff(des1))
```

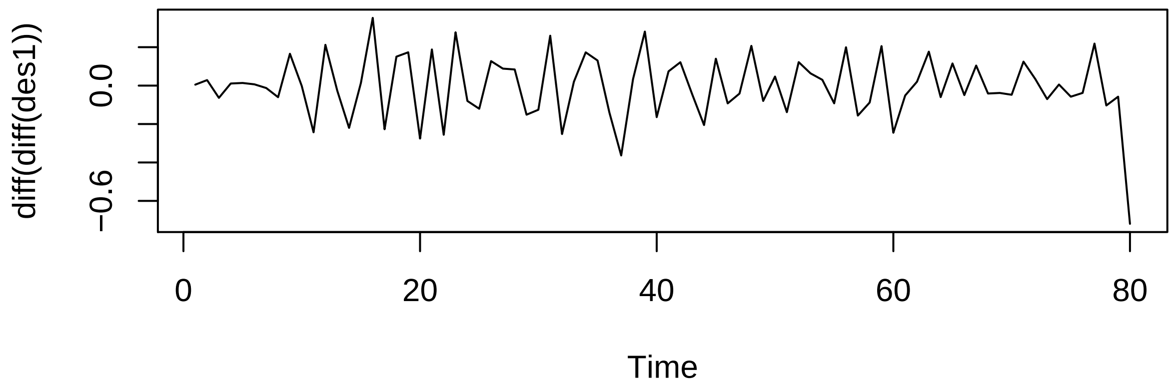
```
##
```

```
##  Augmented Dickey-Fuller Test
##
## data:  diff(des1)
## Dickey-Fuller = -3.0602, Lag order = 4, p-value = 0.1417
## alternative hypothesis: stationary
```

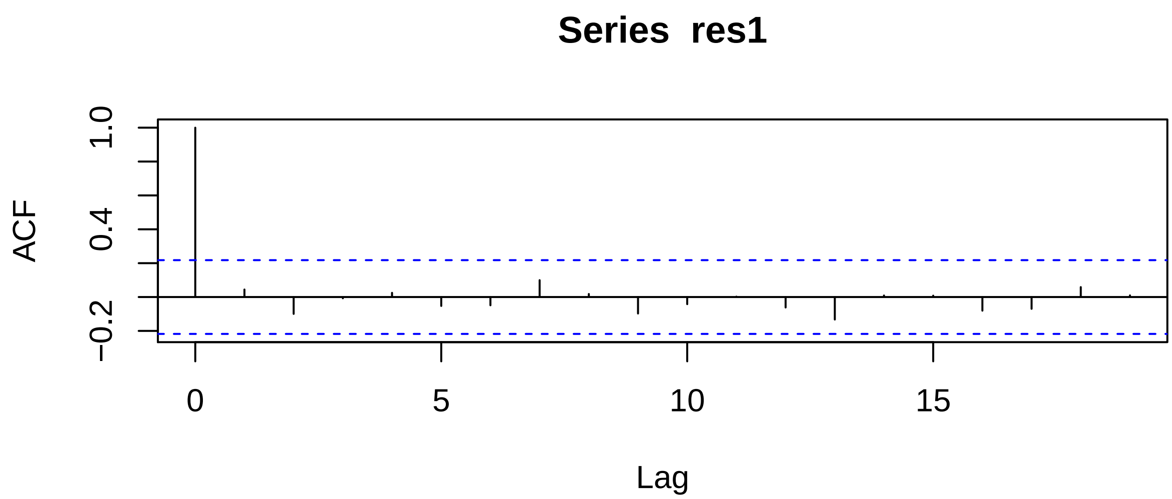
```
adf.test.custom(des1, option = 'c')
```

```
## Lag optimo
##          1
```

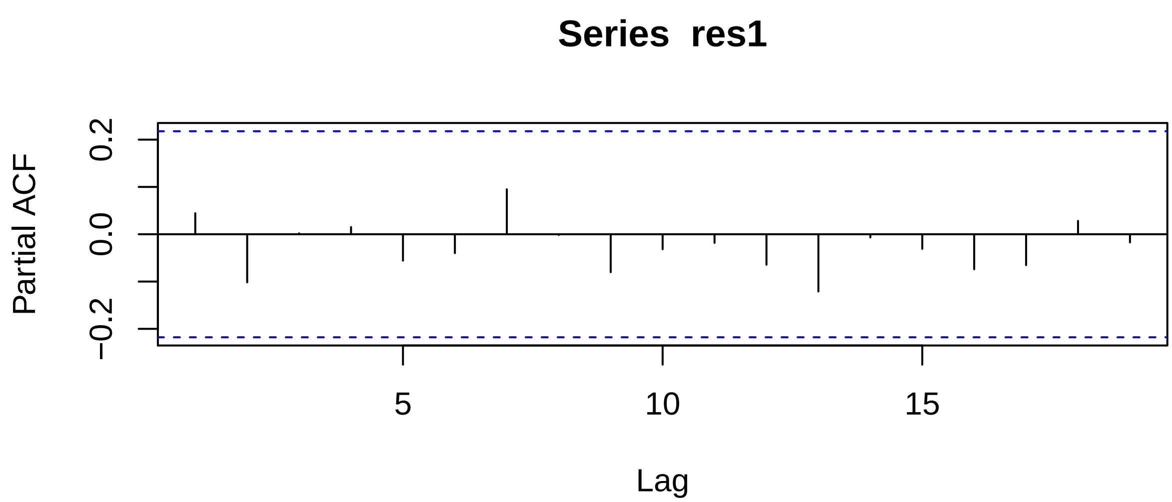
```
ts.plot(diff(diff(des1)))
```



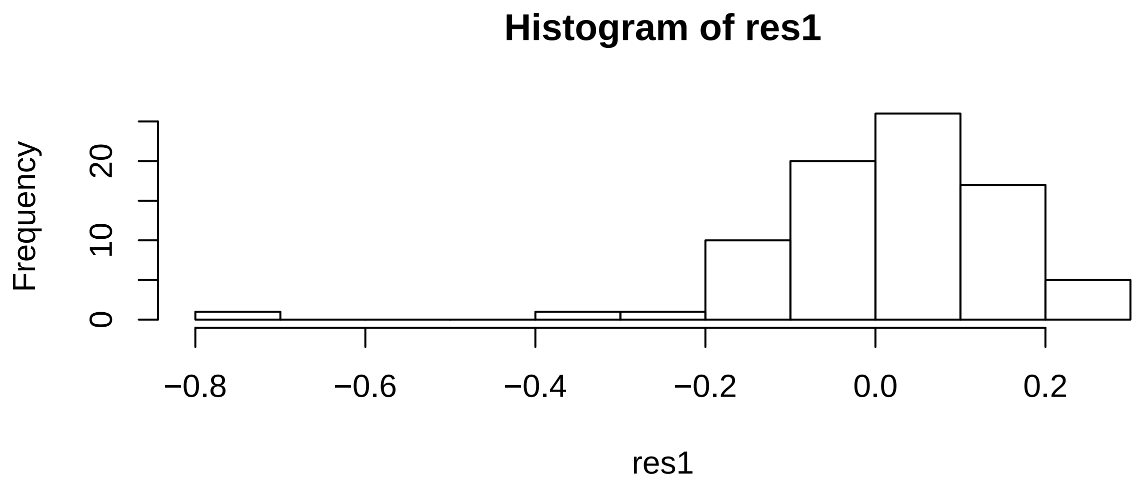

```
#auto.arima(diff(des1))
arima1 <- arima(diff(des1), order= c(0,1,1) )
#summary(arima1)
res1 <- arima1$residuals
res1 <- na.omit(res1)
acf(res1)
```




```r
pacf(res1)
```




```r
hist(res1)
```



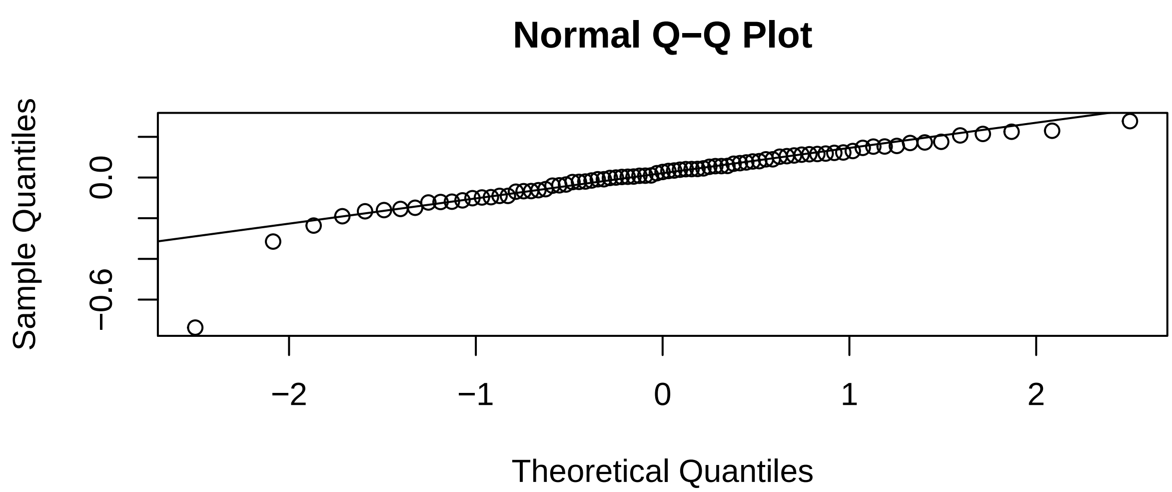

```r
shapiro.test(res1)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res1
## W = 0.87876, p-value = 1.519e-06
```

```r
BIC(arima1)
```

```
## [1] -75.10298
```

```r
qqnorm(res1)
qqline(res1)
```




```r
summary(lm(res1[-59]~ diff(res1)))
```

```
##
## Call:
```

```
## lm(formula = res1[-59] ~ diff(res1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.79490 -0.06666  0.00960  0.09183  0.24494
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.01245    0.01597   0.779    0.438
## diff(res1)  -0.06255    0.09030  -0.693    0.491
##
## Residual standard error: 0.1426 on 78 degrees of freedom
## Multiple R-squared:  0.006115,	Adjusted R-squared:  -0.006627
## F-statistic: 0.4799 on 1 and 78 DF,  p-value: 0.4905
```

Acontinuación ajustamos un modelo ARIMA para las observaciones siguientes a noviembre de 1987

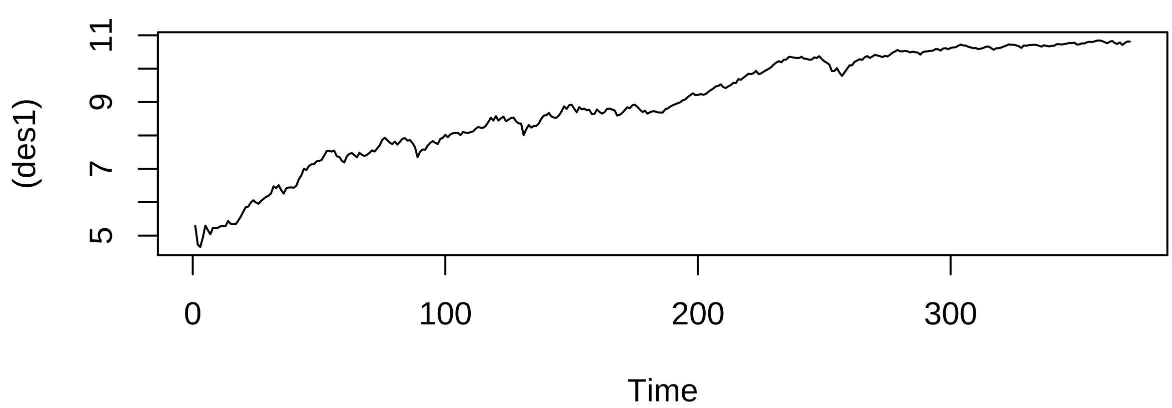
```r
des1 <- IPC[82:452]
adf.test((des1)) #existe raiz
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  (des1)
## Dickey-Fuller = -2.702, Lag order = 7, p-value = 0.2807
## alternative hypothesis: stationary
```

```r
adf.test.custom(des1, option = 'c')
```

```
## Lag optimo
##          1
```

```r
ts.plot((des1)) #tendencia a la baja
```

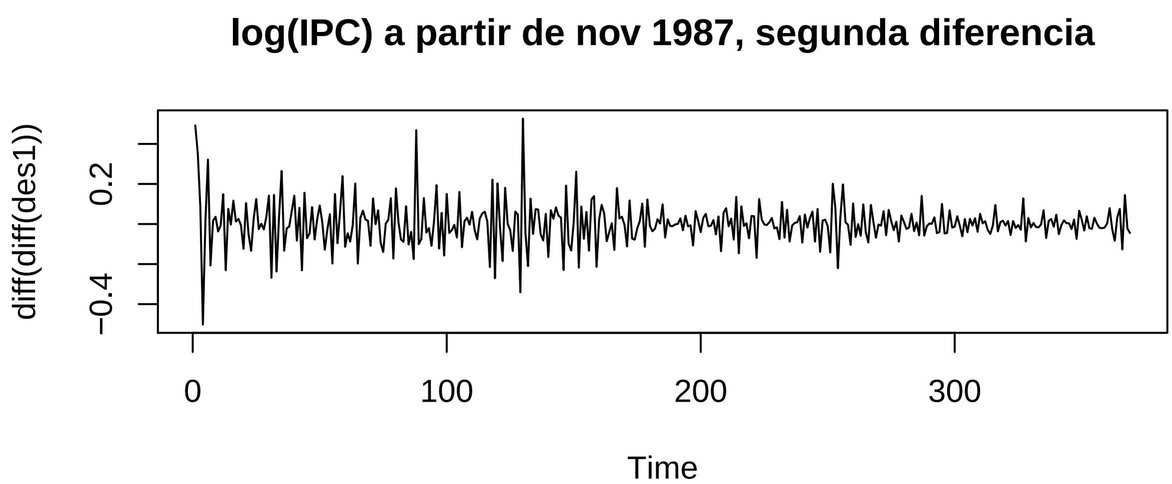



```r
adf.test(diff(des1)) #se elimina tendencia
```

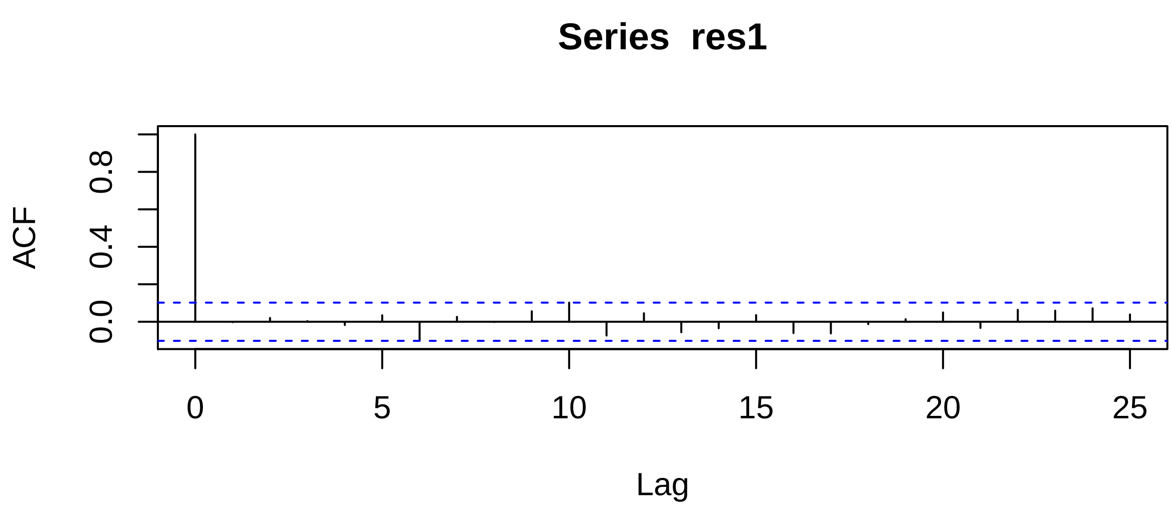
```
##
##  Augmented Dickey-Fuller Test
```

```
##
## data:  diff(des1)
## Dickey-Fuller = -7.4724, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary
```

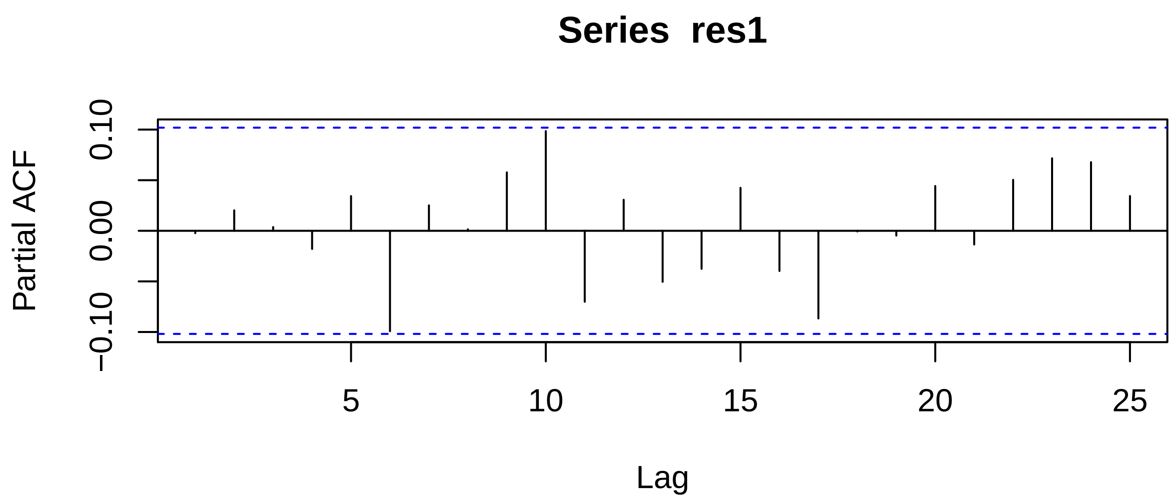
```
ts.plot(diff(diff(des1)), main='log(IPC) a partir de nov 1987, segunda diferencia')
```



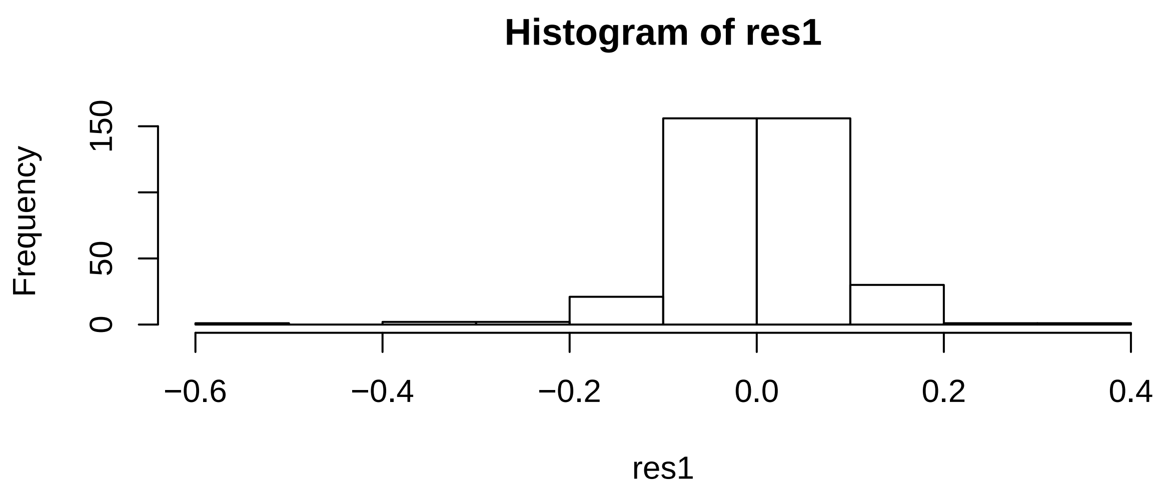

```
#auto.arima(diff(diff(des1)))
arima1 <- arima(diff(des1), order= c(4,0,3) )
#summary(arima1)
res1 <- arima1$residuals
res1 <- na.omit(res1)
acf(res1)
```




```
pacf(res1)
```




```r
hist(res1)
```



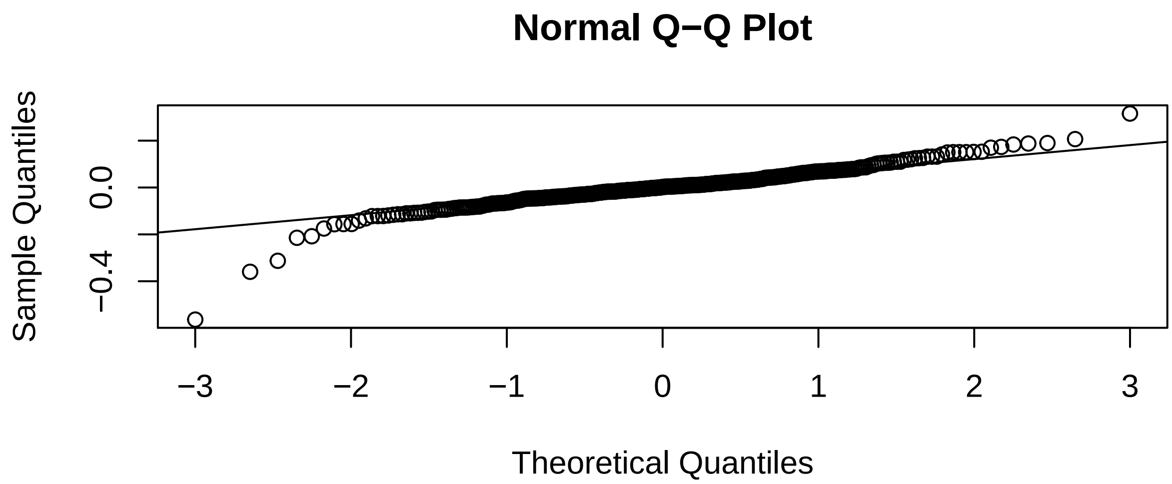

```r
shapiro.test(res1)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res1
## W = 0.91888, p-value = 3.039e-13
```

```r
BIC(arima1)
```

```
## [1] -770.3092
```

```r
qqnorm(res1)
qqline(res1)
```




```r
summary(lm(res1[-370]~ diff(res1)))
```

```
##
## Call:
## lm(formula = res1[-370] ~ diff(res1))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.291434 -0.027876 -0.000977  0.029294  0.253675
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.000689   0.002809   0.245    0.806
## diff(res1)  -0.536835   0.025962 -20.677   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05395 on 367 degrees of freedom
## Multiple R-squared:  0.5381, Adjusted R-squared:  0.5368
## F-statistic: 427.6 on 1 and 367 DF,  p-value: < 2.2e-16
```

De donde concluimos que el cambio estructural de noviembre de 1987 (del cual desconocemos su causa) incremento en un orden la integración de la serie haciendo necesaria la estimación de más parámetros. En vista de que los residuos de este modelo no son normales ni no autocorrelacionados no descartamos otro posible cambio estructural en la serie a partir de noviembre de 1987 como la ultima gráfica de la segunda diferencia de la serie sugiere un cambio alrededor del 2008.