

Introducción a series de tiempo con R

LAAN

25/11/2019

Contents

Procesos estacionarios en covarianza	3
Función de autocorrelación (ACF o correlograma)	4
Ruido blanco	4
Inferencia	5
Autocorrelación serial individual	5
Autocorrelación serial conjunta	5
Pruebas de autocorrelación serial conjunta: pruebas portmanteau	6
Simulación	7
Procesos de media móvil MA(q)	10
Operador de rezago	11
Características de un proceso MA(q)	11
Media incondicional	11
Varianza incondicional	11
Covarianzas	12
Invertibilidad de un proceso MA(q)	12
Invertibilidad y unicidad de un proceso MA(q)	13
Correlogramas y memoria del proceso	13
Procesos autoregresivos AR(p)	14
Estacionariedad de un modelo AR(p)	14
Ecuación característica	14
Estacionariedad y ecuaciones en diferencias (Opcional)	14
Ejemplo 1	15
Ejemplo 2	16
Ejemplo 3	16
Ejemplos	17
Ejemplo 1	17
Ejemplo 2	18
Ejemplo 3 – Caminata aleatoria	18
Integración	24
Ejemplo 4 (Caminata aleatoria)	24
Ejemplo 5	24
Sobrediferenciación	25
Pruebas de raíz unitaria; de estacionariedad	25

Prueba de raíz unitaria	25
Pruebas de raíz unitaria: falta de potencia	26
Prueba de estacionariedad	28
Análisis de confirmación	29
Características de un proceso AR(p)	29
Teorema de Wold	29
Media incondicional	30
Varianza incondicional	30
Correlaciones seriales	31
Función de autocorrelación parcial (PACF)	32
Correlogramas y memoria de un proceso AR; comparaciones ACF, PACF para procesos MA, AR	32
Ejemplo	33
Modelos ARMA, ARIMA	36
Sobrep parametrización	36
Ejemplo	36
Características de un proceso ARMA(p,q)	37
La metodología de Box-Jenkins	37
Criterios de Información	38
Pronósticos	39
Método de pronóstico simple	39
Pronóstico mediante suavización exponencial	40
Pronóstico para MA(q)	40
Pronóstico para AR(p)	41
Pronósticos con R	42
Ejemplo 1 (Simulación)	42
Ejemplo 2	49
Ejemplo 3	55
Heterocedasticidad condicional y volatilidad	60
Características de la volatilidad	60
Medida de volatilidad	61
Modelación ARCH	62
Procedimiento	62
Ejemplo	63
Pronóstico	71
Modelo ARCH-Innovaciones no gaussianas	73
Innovación t-Student	74
Modelos GARCH	76
Fortalezas y debilidades de la modelación GARCH	77
Estimación GARCH de INTC	77

Diagnóstico	79
ACF, PACF	80
Ljung-Box, lag = 10	82
Ljung-Box, lag = 20	83
Ljung-Box, lag=30	83
Normalidad de las innovaciones	83
Intervalo predictivo alrededor de la media	84
Modelo (G)ARCH con serie ARMA(p,q)	87
Estimación de las ecuaciones de media	88
Detección de heterocedasticidad condicional	89
Estimación de la volatilidad	93
Gráficos con intervalos predictivos	99
Otros modelos	103
Modelo GARCH-M "GARCH in the mean"	103
Modelos EGARCH(1,1)	106
Modelos EGARCH(1,1) con prima de riesgo	109
Regresiones espurias; cointegración	111
Regresiones espurias y estacionariedad	111
Modelos VAR (Vectoriales Autoregresivos)	116
Ventajas y desventajas de la modelación VAR	116
Después de la estimación	117
Estimación de un modelo VAR	117
Determinación del orden y estimación	118
Diagnóstico	120
Causalidad según Granger	125
Respuesta a impulso unitario	128
Descomposición del error del error de pronóstico	133
Bibliografía	135

Procesos estacionarios en covarianza

Se dice que una serie (proceso) y_t es **estacionario en covarianza** si

Media $E[y_t] = \mu$

Varianza $E[(y_t - \mu)^2] = \sigma^2$

Autocovarianza $E[(y_{t_1} - \mu)(y_{t_2} - \mu)] = \gamma_{t_1-t_2}$, para todo t_1, t_2 .

Si

$$t_1 = t_2 + s$$

(s : rezago, “lag”, en inglés), entonces

$$E[(y_{t_2+s} - \mu)(y_{t_2} - \mu)] = \gamma_s$$

es una función del rezago.

El término γ_s es la **función de autocovarianza**.

Nótese que la varianza

$$\text{var}(y_t) = E[y_t^2] - (E[y_t])^2 = \gamma_0.$$

Comentario La autocovarianza γ_s depende de las unidades de medición usadas, por lo que es más conveniente usar la **serie de autocorrelaciones**

$$\tau_s = \frac{\gamma_s}{\gamma_0}, \quad s = 1, 2, \dots$$

Función de autocorrelación (ACF o correlograma)

La **función de autocorrelación** (ACF) es la gráfica de las autocorrelaciones τ_s contra s , $s = 0, 1, \dots$

Por definición, $\tau_0 = 1$.

La función de autocorrelación también se conoce como **correlograma**.

Ruido blanco

Un **ruido blanco**, u_t , es un proceso estocástico (serie de tiempo) con las siguientes propiedades:

1. $E[u_t] = \mu$, para todo $t > 0$.
2. (**Homocedasticidad**) $\text{var}(u_t) = \sigma^2$, para todo $t > 0$.
3. $\gamma_{t_1-t_2} = \begin{cases} \sigma^2, & \text{si } t_1 = t_2; \\ 0, & \text{en cualquier otro caso.} \end{cases}$

Comentarios Se presentan a continuación unas consecuencias de las propiedades de un ruido blanco.

- Para un ruido blanco, si $s = t_1 - t_2$, entonces un ruido blanco tiene función de autocorrelación

$$\tau_s = \begin{cases} 1, & s = 0, \\ 0, & s \neq 0 \end{cases}$$

- Si la media $\mu = 0$, el proceso se conoce como un ruido blanco con media cero. En este caso,

$$\text{corr}(u_t, u_s) = E[u_t u_s] = 0.$$

Además,

$$\text{var}(u_t) = \sigma^2 = E[u_t^2],$$

para todo $t > 0$.

- Si el ruido blanco también tiene distribución normal, entonces los coeficientes de autocorrelación muestral

$$\hat{\tau}_s \stackrel{aprox}{\sim} N(0, 1/T)$$

(T es el tamaño muestral).

Inferencia

Si un ruido blanco también tiene distribución normal, entonces se pueden determinar intervalos de confianza para los valores estimados $\hat{\tau}_s$ de la función de autocorrelación.

Autocorrelación serial individual

Los intervalos de confianza se pueden usar para probar la hipótesis de autocorrelación serial individual

$$H_0 : \tau_s = 0$$

$$H_1 : \tau_s \neq 0$$

con nivel de significancia α .

En lo que sigue, se usará $\alpha = .05$ en todas las pruebas de hipótesis.

La prueba de hipótesis es equivalente a

$$H_0 : \text{corr}(u_t, u_{t+s}) = 0$$

$$H_1 : \text{corr}(u_t, u_{t+s}) \neq 0$$

$$s = 1, 2, \dots$$

En la prueba de hipótesis, el estadístico de prueba es el estadístico t usual:

$$T = \frac{\hat{\tau}_s}{\text{ee}(\hat{\tau}_s)}$$

donde $\text{ee}(\hat{\tau}_s)$ es el error estándar del estimador $\hat{\tau}_s$.

Autocorrelación serial conjunta

Una prueba de autocorrelación serial está basada en la siguiente prueba de hipótesis

$$H_0 : \tau_1 = 0, \dots, \tau_m = 0$$

$$H_1 : \tau_1 \neq 0, \dots, \tau_m \neq 0$$

con $m \geq 2$.

En un diagnóstico de ruido blanco más adelante, la prueba de hipótesis es equivalente a

$$\begin{aligned}H_0 : u_t \text{ es ruido blanco} \\ H_1 : u_t \text{ no es ruido blanco}\end{aligned}$$

Pruebas de autocorrelación serial conjunta: pruebas portmanteau

Pruebas de autocorrelación serial conjunta reciben el nombre de *pruebas portmanteau* y se conocen por los nombres de los diferentes autores que la descubrieron.

Las pruebas que usaremos son las de

- Box-Pierce [BP]
- Ljung-Box [LB]

Una discusión acerca de pruebas portmanteau a casos más generales se encuentra en [M].

Prueba de Box-Pierce

Si T es el tamaño muestral, el *estadístico de Box-Pierce* es

$$Q = T \sum_{k=1}^m \hat{\tau}_k^2 \stackrel{aprox}{\sim} \chi_m^2$$

El estadístico de Box-Pierce no parece dar resultados buenos con muestras pequeñas.

Prueba de Ljung-Box

Un ajuste a muestras pequeñas es el *estadístico de Ljung-Box*

$$Q^* = T(T+2) \sum_{k=1}^m \frac{\hat{\tau}_k^2}{T-k} \stackrel{aprox}{\sim} \chi_m^2$$

Asintóticamente, $T \rightarrow \infty$, los dos estadísticos Q , Q^* son comparables.

Ejemplo

Usando $T = 100$ observaciones, se estimaron los primeros cinco coeficientes de autocorrelación:

s	1	2	3	4	5
$\hat{\tau}_s$.227	-.013	.0805	.004	-.024

Prueba la significancia estadística individual y conjunta ($m = 5$) de estos coeficientes de autocorrelación. Usar $\alpha = .05$.

Solución: Significancia estadística individual El intervalo de confianza asociado es

$$\pm 1.96 \times \frac{1}{\sqrt{T}} \Rightarrow (-.196, .196)$$

Si $\hat{\tau}_s$ se encuentra en el intervalo de confianza, entonces **no** es significativo; de lo contrario, sí es significativo.

De los valores estimados, solo el primero es significativo.

Solución: significancia estadística en conjunto El estadístico de Box-Pierce

$$Q = 100 \cdot [(.227)^2 + (-.013)^2 + (.0805)^2 + (.004)^2 + (-.024)^2] = 5.877025.$$

El estadístico de Ljung-Box da

$$Q^* = (100 \cdot 102) \cdot [(.227)^2/(100 - 1) + (-.013)^2/(100 - 2) + (.0805)^2/(100 - 3) + (.004)^2/(100 - 4) + (-.024)^2/(100 - 5)] = 6.0716.$$

Compararemos el valor de cada estadístico de prueba con el valor crítico de la distribución χ^2_5 :

```
qchisq(.05, 5, lower.tail = F)
```

```
## [1] 11.1
```

Ambos estadísticos son menores que el valor crítico, por lo que no se puede rechazar la hipótesis nula. Concluimos que

$$\tau_1 = 0, \tau_2 = 0, \tau_3 = 0, \tau_4 = 0, \tau_5 = 0$$

Simulación

Simularemos $T = 100$ observaciones de un ruido blanco con distribución normal con media 0 y varianza 1.

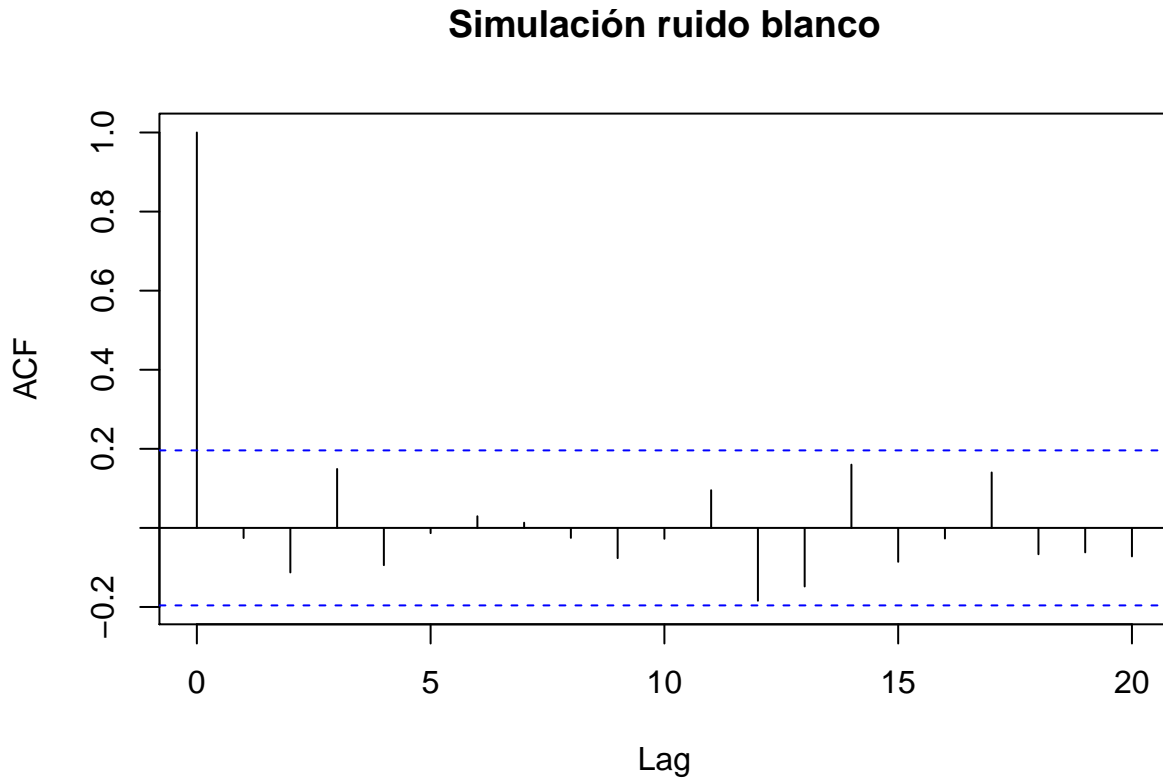
- Se determinará la ACF.
- Se probará la significancia individual de las primeras 20 observaciones.
- Se probará la significancia conjunta con, sucesivamente, $m = 10$; $m = 20$, $m = 50$, usando Q , Q^* .

El código se da a continuación. Se usará `set.seed(123)` para obtener resultados reproducibles. El comando `x <- rnorm(100)` genera 100 muestras de una distribución normal estándar:

```
set.seed(123)
x <- rnorm(100)
```

La prueba de significancia estadística individual se lleva a cabo con la función de autocorrelación `acf(as.ts(x))`, en donde se pasa el argumento `as.ts(x)` como serie de tiempo:

```
acf(as.ts(x), main = "Simulación ruido blanco")
```



El parámetro opcional `main = "Simulación ruido blanco"` permite poner un título a la gráfica.

En el correlograma (AFC), $\tau_0 = 1$ siempre.

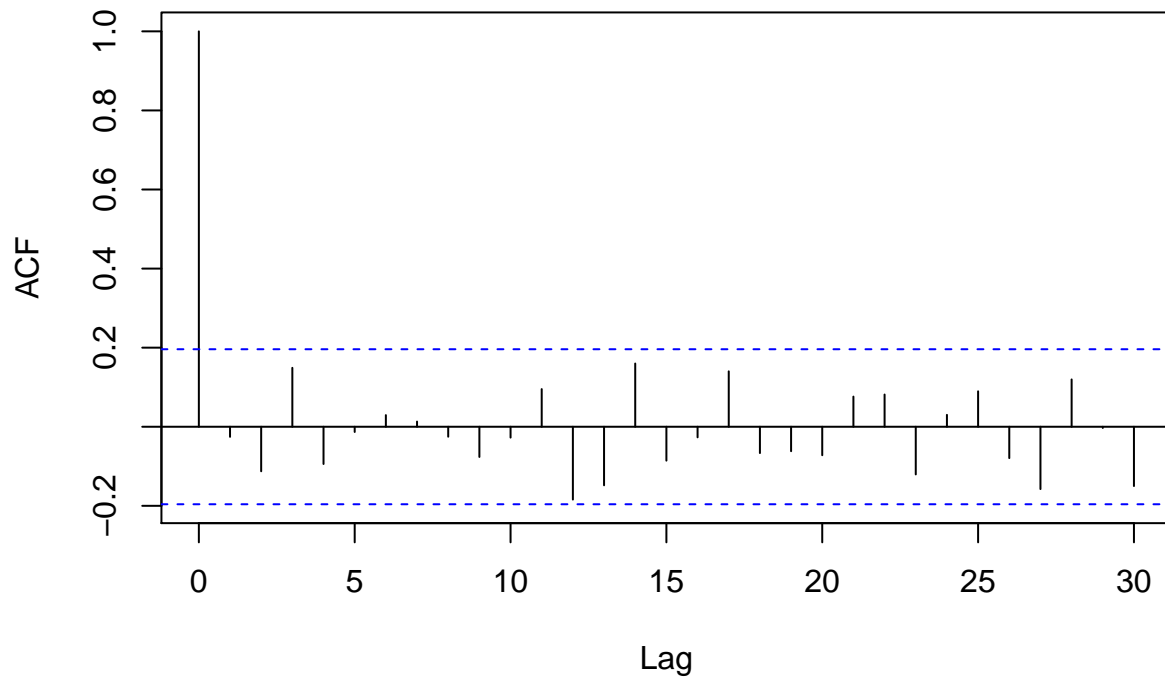
El intervalo de confianza, al 95%, está delimitado por las líneas punteadas en azul. Si un valor $\hat{\tau}_s$ rebasa el intervalo de confianza, entonces $\tau_s \neq 0$.

Podemos ver que $\tau_s = 0$, $s = 1, 2, \dots, 20$.

Lo mismo, pero con más rezagos ($m = 30$):

```
acf(as.ts(x), lag.max = 30, main = "Simulación ruido blanco")
```


Simulación ruido blanco



También en este caso, $\tau_s = 0$, $s = 1, 2, \dots, 30$.

La prueba de significancia conjunta se realiza usando la función `Box.test()`; se especifica el valor del rezago que se quiere probar ($m = 10$; $m = 20$, $m = 50$, en este ejemplo) y el estadístico de prueba (default: Box-Pierce):

```
Box.test(as.ts(x), lag=10)
```

```
##
## Box-Pierce test
##
## data:  as.ts(x)
## X-squared = 5, df = 10, p-value = 0.9
```

```
Box.test(as.ts(x), lag=20)
```

```
##
## Box-Pierce test
##
## data:  as.ts(x)
## X-squared = 18, df = 20, p-value = 0.6
```

```
Box.test(as.ts(x), lag=50)
```

```
##
## Box-Pierce test
##
## data:  as.ts(x)
## X-squared = 37, df = 50, p-value = 0.9
Box.test(as.ts(x), lag=10, type = "Ljung-Box")
```

```
##
## Box-Ljung test
##
## data:  as.ts(x)
## X-squared = 6, df = 10, p-value = 0.8
Box.test(as.ts(x), lag=20, type = "Ljung-Box")
```

```
##
## Box-Ljung test
##
## data:  as.ts(x)
## X-squared = 21, df = 20, p-value = 0.4
Box.test(as.ts(x), lag=50, type = "Ljung-Box")
```

```
##
## Box-Ljung test
##
## data:  as.ts(x)
## X-squared = 50, df = 50, p-value = 0.5
```

Concluimos que no existe evidencia de autocorrelación serial, con $m = 10$, $m = 20$, $m = 50$, respectivamente.

Procesos de media móvil MA(q)

Sea u_t un ruido blanco con media $E(u_t) = 0$ y varianza $var(u_t) = \sigma^2$.

La serie y_t es un proceso de **media móvil de orden q** , si

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q}.$$

Una serie de media móvil de orden q se denota con $MA(q)$.

Comentario En ciertos cálculos algebraicos se puede considerar, sin pérdida de generalidad, $\mu = 0$.

Operador de rezago

El **operador de rezago** L (para un periodo) se define

$$Ly_t = y_{t-1}.$$

Para periodos sucesivos, se define de manera inductiva mediante

$$L^i y_t = y_{t-i},$$

de manera que una serie $MA(q)$ se puede escribir

$$y_t = \mu + \sum_{i=1}^q \theta_i L^i u_t + u_t = \mu + \theta(L)u_t,$$

con $\theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$.

Características de un proceso MA(q)

Sea $y_t = \mu + u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q}$ un proceso $MA(q)$. A continuación se probarán las propiedades más básicas de este proceso.

Media incondicional

$$E(y_t) = \mu$$

Solución El valor esperado es lineal, por lo que

$$\begin{aligned} E[y_t] &= E[\mu + u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q}] \\ &= E[\mu] + E[u_t] + \theta_1 E[u_{t-1}] + \cdots + \theta_q E[u_{t-q}] \\ &= \mu \end{aligned}$$

siendo u_t un ruido blanco con media cero.

Varianza incondicional

$$var(y_t) = \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma^2$$

Solución Sin pérdida de generalidad, se puede suponer que $\mu = E[y_t] = 0$. En este caso, la varianza

$$var(y_t) = E[y_t^2] - (E[y_t])^2 = E[y_t^2]$$

El cuadrado del proceso

$$\begin{aligned} y_t^2 &= (u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q})^2 \\ &= u_t^2 + \theta_1^2 u_{t-1}^2 + \cdots + \theta_q^2 u_{t-q}^2 + \text{productos cruzados} \end{aligned}$$

donde los productos cruzados consisten de términos del tipo $u_{t-j}u_{t-k}$, $j \neq k$. El valor esperado de los productos cruzados es cero, por la tercera propiedad de un ruido blanco. Por lo tanto,

$$\begin{aligned} \text{var}(y_t) &= E[y_t^2] = E[u_t^2 + \theta_1^2 u_{t-1}^2 + \cdots + \theta_q^2 u_{t-q}^2 + \text{productos cruzados}] \\ &= E[u_t^2] + \theta_1^2 E[u_{t-1}^2] + \cdots + \theta_q^2 E[u_{t-q}^2] \\ &= \sigma^2 + \theta_1^2 \sigma^2 + \cdots + \theta_q^2 \sigma^2 \\ &= \sigma^2 (1 + \theta_1^2 + \cdots + \theta_q^2) \end{aligned}$$

Para pasar de la segunda línea a la tercera arriba se usa la propiedad de *homocedasticidad* de un ruido blanco (Propiedad 2).

Covarianzas

Las covarianzas

$$\gamma_s = \begin{cases} (\theta_s + \theta_{s+1}\theta_1 + \cdots + \theta_q\theta_{q-s})\sigma^2, & s = 1, 2, \dots, q \\ 0, & \text{para } s > q \end{cases}$$

La prueba de esta propiedad se encuentra en el libro de Brooks [Br].

Invertibilidad de un proceso MA(q)

Considérese un modelo $MA(q)$

$$y_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q}.$$

Usando el operador de rezago,

$$\begin{aligned} y_t &= u_t + \theta_1 L u_t + \theta_2 L^2 u_t + \cdots + \theta_q L^q u_t \\ &= (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q) u_t = \Theta_q(L) u_t \end{aligned}$$

El polinomio característico asociado es

$$\Theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q = 0$$

Se dice que el proceso es **invertible** si todas las raíces (posiblemente complejas conjugadas), z_i , $i = 1, \dots, q$, del polinomio característico tienen módulo mayor a uno

$$|z_i| > 1, \quad i = 1, \dots, q.$$

Si un proceso $MA(q)$ es invertible, entonces este proceso se puede expresar mediante un proceso $AR(\infty)$ convergente (véase abajo para procesos autoregresivos)

$$y_t = u_t + c_1 y_{t-1} + c_2 y_{t-2} + \cdots$$

Considérese el caso de un proceso $MA(1)$

$$y_t = u_t + \theta_1 y_{t-1} = (1 + \theta_1 L)u_t$$

El polinomio característico es $1 + \theta_1 z = 0$ y el proceso es invertible si $|\theta_1| < 1$.

Usando la expresión $y_t = (1 + \theta_1 L)u_t$, aplicando formalmente el operador inverso

$$(1 + \theta_1 L)^{-1} y_t = u_t,$$

y expandiendo en series de potencia,

$$\begin{aligned} \left[\sum_{i=0}^{\infty} (-\theta_1)^i L^i \right] y_t &= u_t \\ y_t + \left[\sum_{i=1}^{\infty} (-\theta_1)^i L^i \right] y_t &= u_t \end{aligned}$$

que da la expansión $AR(\infty)$ requerida

$$y_t = u_t + c_1 y_{t-1} + c_2 y_{t-2} + \cdots$$

Invertibilidad y unicidad de un proceso $MA(q)$

La invertibilidad es también importante para identificar un proceso de media móvil.

Considérese, por ejemplo, los siguientes dos procesos de media móvil

$$\begin{aligned} x_t &= u_t + \frac{1}{2} u_{t-1}, \quad u_t \sim N(0, 4) \\ y_t &= v_t + 2v_{t-1}, \quad v_t \sim N(0, 1) \end{aligned}$$

Nótese que u_t, v_t son ambos ruidos blancos. Además, no es difícil verificar que *la función de autocorrelación ACF es la misma para ambos procesos*.

Entre los dos procesos, el primero es invertible y el segundo no lo es.

Correlogramas y memoria del proceso

- La "memoria" de un proceso $MA(q)$ es de q periodos, por lo que cualquier innovación o "shock" desaparecerá para rezagos $s > q$.
- Como consecuencia del inciso anterior, se podrá pronosticar sólo por q periodos adelante.
- La función de autocorrelación de un proceso $MA(q)$ es diferente de cero sólo para los primeros q rezagos (q : orden del proceso).
- Si el proceso $MA(q)$ es invertible, entonces la función de autocorrelación parcial (PACF) tendrá un decaimiento geométrico (exponencial), con rezagos crecientes. Véase más adelante para la definición de función de autocorrelación parcial.

Procesos autoregresivos AR(p)

Un proceso y_t **autoregresivo de orden p** se expresa como

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + u_t$$

donde u_t es el término de error (ruido blanco con media cero).

Usando el operador de rezago, se puede escribir el proceso autoregresivo mediante

$$y_t = \mu + \sum_{i=1}^p L^i y_t + u_t \quad \Leftrightarrow \quad \left(1 - \sum_{i=1}^p L^i\right) y_t = \mu + u_t \quad \Leftrightarrow \quad \Phi(L)y_t = \mu + u_t$$

donde $\Phi(L) = (1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)$.

Comentario Sin pérdida de generalidad, se puede poner $\mu = 0$.

Estacionariedad de un modelo AR(p)

Se dice que la serie AR(p) y_t es *estacionaria* si

$$y_t = \Phi^{-1}(L)u_t$$

es convergente a cero.

Ecuación característica

La **ecuación característica** asociada al proceso $\Phi(L)y_t = \mu + u_t$ es

$$\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0$$

Resulta que el proceso $AR(p)$ es estacionario si todas las raíces z_i de la ecuación característica

$$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0$$

cumplen $|z_i| > 1$.

Estacionariedad y ecuaciones en diferencias (Opcional)

La fundamentación matemática de las series de tiempo autoregresiva se encuentra en las ecuaciones en diferencia.

Una ecuación en diferencia de orden m se escribe como sigue

$$y_t = \alpha + \beta_1 y_{t-1} + \cdots + \beta_m y_{t-m}$$

La solución general se encuentra por medio del *principio de superposición*:

$$y_t = y_p + y_{0,t},$$

donde y_p es una solución particular y $y_{0,t}$ es la solución general de la *ecuación homogénea asociada*:

$$y_{0,t} = \beta_1 y_{t-1} + \cdots + \beta_m y_{t-m}$$

La solución particular es bastante sencilla en los casos de interés de esta monografía; es suficiente considerar $y_p = \text{constante}$, lo que implica

$$y_p = \frac{a}{1 - \beta_1 - \cdots - \beta_m}.$$

Un método para buscar soluciones de la *ecuación homogénea asociada* es mediante el *ansatz*

$$y_t = b^t$$

con b un valor incógnito. Sustituyendo, se encuentra el *polinomio característico* de grado m

$$b^{t-m} (b^m - \beta_1 b^{m-1} - \cdots - \beta_m) = 0$$

Las raíces b_1, \dots, b_m (contadas según su multiplicidad) del polinomio característico determinan los valores de b , que determinan la solución de la ecuación homogénea asociada.

Nótese que $b_i = 1/z_i$, donde z_i es la raíz del polinomio característico de la serie de tiempo

$$y_t = \alpha + \beta_1 y_{t-1} + \cdots + \beta_m y_{t-m}$$

Se ilustrará el método de solución usando ejemplos específicos.

Ejemplo 1

Resolver la ecuación en diferencia

$$y_t = .34 + .12y_{t-1}$$

Solución La solución particular $y_p = \text{const}$, por lo que

$$y_p = .34 + .12y_p$$

Despejando, se encuentra $y_p = .34/(1 - .12) = .386364$.

La solución de la ecuación homogénea asociada es $y_{0,t} = b^t$; sustituyendo en

$$y_y = .12y_{t-1}$$

se encuentra $b^t = .12b^{t-1}$; factorizando,

$$b^{t-1}(b - .12) = 0$$

que da $b = .12$.

La solución general es entonces

$$y_t = y_p + y_{0,t} = \frac{.34}{1 - .12} + A(.12)^t$$

La solución general depende de una constante arbitraria, que se determina con una condición inicial apropiada.

Nótese que

$$\lim_{t \rightarrow \infty} y_t = \frac{.34}{1 - .12}.$$

Se determinará más adelante que la serie de tiempo asociada es estacionaria.

Ejemplo 2

Resolver la ecuación en diferencia

$$y_t = \frac{1}{12}y_{t-1} + \frac{1}{12}y_{t-2}$$

Solución En este ejemplo no hay un intercepto, por lo que $y_p = 0$.

La solución de la ecuación homogénea asociada es del tipo $y_{0,t} = b^t$, por lo que

$$b^t = \frac{1}{12}b^{t-1} + \frac{1}{12}b^{t-2} \quad \Rightarrow \quad b^{t-2}(b^2 - (1/12)b - (1/12)) = 0$$

Las raíces $b_1 = 1/4$, $b_2 = 1/3$, por lo que la solución general es

$$y_t = A_1 \left(\frac{1}{4}\right)^t + A_2 \left(\frac{1}{3}\right)^t$$

También en este ejemplo la solución general converge a la solución particular $y_p = 0$:

$$\lim_{t \rightarrow \infty} y_t = 0.$$

Los valores A_1 , A_2 se determinan por medio de *dos* condiciones iniciales oportunas.

También en este caso, se encontrará que la serie de tiempo asociada es estacionaria.

Ejemplo 3

Resolver la ecuación en diferencia

$$y_t = 4 + y_{t-1}$$

Solución En este caso, la solución particular no puede ser constante porque, de otra manera, se encuentra una imposibilidad $0 = 4$.

Se busca entonces una solución lineal,

$$y_{p,t} = at$$

con a constante a determinar. En este caso,

$$at = 4 + a(t - 1) \Rightarrow a = 4,$$

por lo que $y_{p,t} = 4t$.

La solución de la ecuación homogénea asociada es este caso es una constante arbitraria, $y_0 = A$, que se determinará por medio de una condición inicial oportuna.

La solución general es

$$y_t = A + 4t.$$

Nótese que en este caso,

$$\lim_{t \rightarrow \infty} y_t = \infty$$

Veremos más adelante que la serie asociada es una caminata aleatoria, que tiene una tendencia lineal, con perturbaciones alrededor de la recta de tendencia.

Ejemplos

Ejemplo 1

Considera una serie de tiempo autoregresiva $AR(1)$:

$$y_t = \phi_1 y_{t-1} + u_t$$

Usando el operador de rezago, $Ly_t = y_{t-1}$, la serie

$$y_t = \phi_1 Ly_t + u_t \Rightarrow (1 - \phi_1 L)y_t = u_t$$

El polinomio característico asociado es

$$1 - \phi_1 z = 0$$

La serie es estacionaria si el módulo de la raíz del polinomio característico asociado

$$|z| = \left| \frac{1}{\phi_1} \right| > 1.$$

Por ejemplo, el proceso

$$y_t = .34 + .12y_{t-1} + u_t$$

es estacionario, porque la raíz del polinomio característico

$$1 - .12z = 0$$

es $z = 1/.12 = 100/12 > 1$.

Por otro lado, el proceso

$$y_t = .34 + 3y_{t-1} + u_t$$

no es estacionario, porque la raíz del polinomio característico

$$1 - 3z = 0$$

es $z = 1/3 < 1$.

Ejemplo 2

La serie

$$y_t = \frac{1}{12}y_{t-1} + \frac{1}{12}y_{t-2} + u_t$$

es estacionaria.

La ecuación característica asociada

$$1 - \frac{1}{12}z - \frac{1}{12}z^2 = 0$$

tiene raíces $z_1 = 3$, $z_2 = -4$.

Ejemplo 3 – Caminata aleatoria

Un caso importante de proceso *no estacionario* es una caminata aleatoria, modelada por

$$y_t = \mu + y_{t-1} + u_t,$$

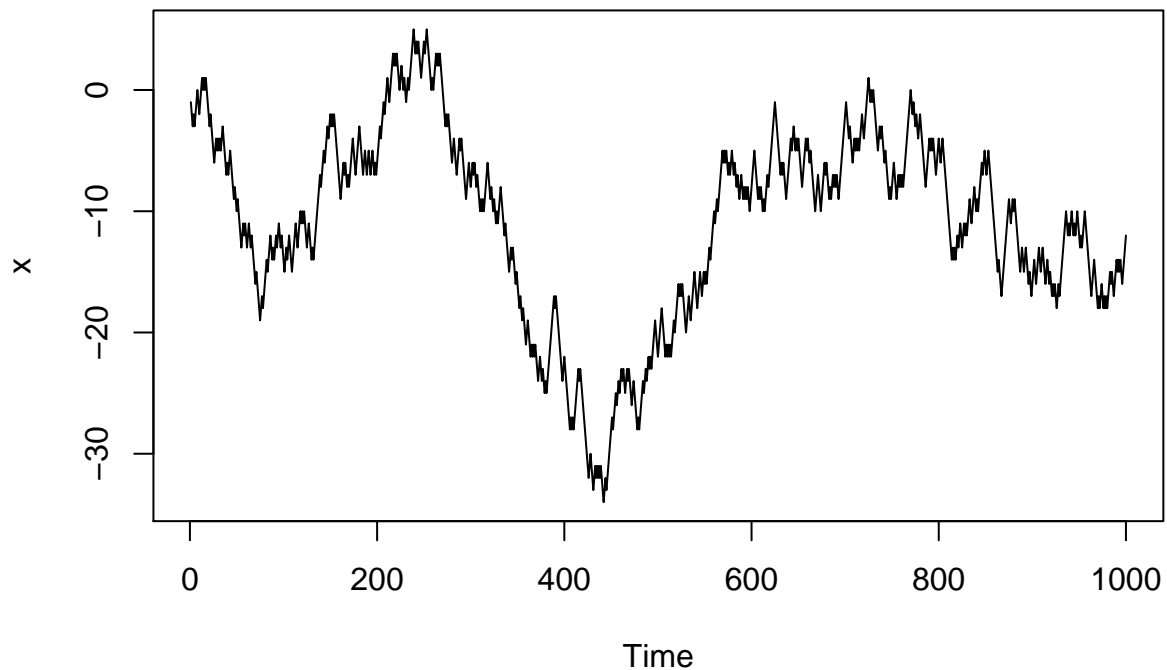
donde u_t es un ruido blanco con media cero. Si $\mu \neq 0$, el modelo es una caminata aleatoria con deriva (“drift”, en inglés).

Simulaciones de caminatas aleatorias

Caminata aleatoria simétrica

La simulación más sencilla consiste en dar un paso a la izquierda (valor = -1) con probabilidad p y un paso a la derecha (valor 1) con probabilidad $1 - p$. A continuación se presenta el caso $p = .5$:

```
set.seed(123)
n <- 1000
x <- cumsum(sample(c(-1, 1), n, replace = TRUE))
plot.ts(x)
```



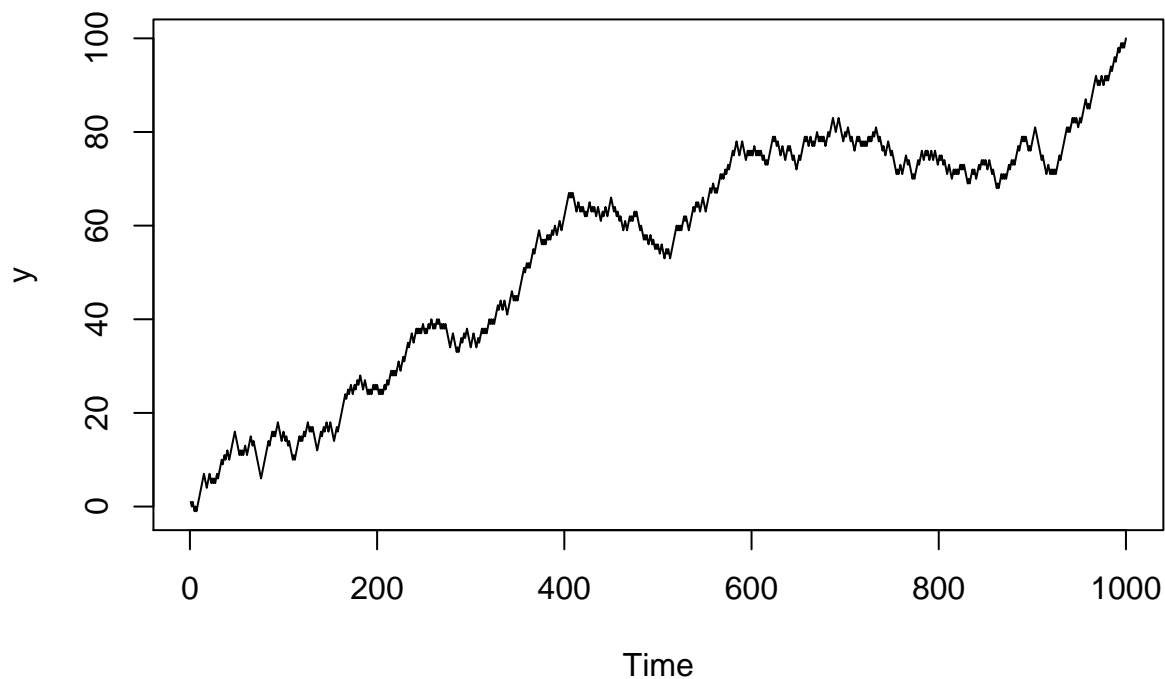
La función `sample()` toma un valor, con remplazo (`replace = TRUE`), del conjunto $\{-1, 1\}$ con probabilidad $p = .5$ (default).

En este caso, el proceso no tiene deriva, $\mu = 0$.

Caminata aleatoria asimétrica

En este ejemplo, usaremos $\Pr(y_t = -1) = .45$.

```
n <- 1000
y <- cumsum(sample(c(-1, 1), n, prob = c(.45, .55), replace = TRUE))
plot.ts(y)
```



En este ejemplo, el valor esperado estimado es

$$E[y_t] = .45(-1) + .55(1) = .1$$

por lo que se concluye que el proceso tiene una deriva estimada de .1.

Simulación de una caminata aleatoria con distribución no bernoulliana

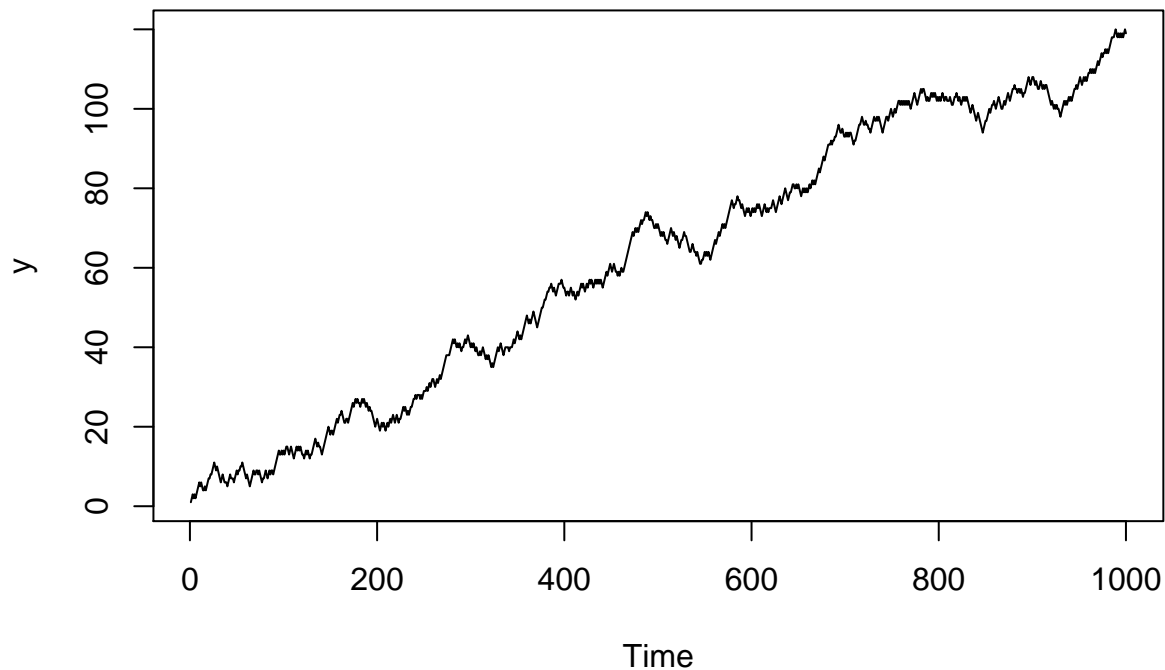
Caminata aleatoria con tres valores

Se usará la función de probabilidad

$$y_t = \begin{cases} -1, & \text{con probabilidad} = .4 \\ 0, & \text{con probabilidad} = .1 \\ 1, & \text{con probabilidad} = .5 \end{cases}$$

con probabilidad positiva de quedarse en el valor actual.

```
n <- 1000
y <- cumsum(sample(c(-1,0, 1), n, prob = c(.4, .1, .5), replace = TRUE))
plot.ts(y)
```



El valor esperado estimado es

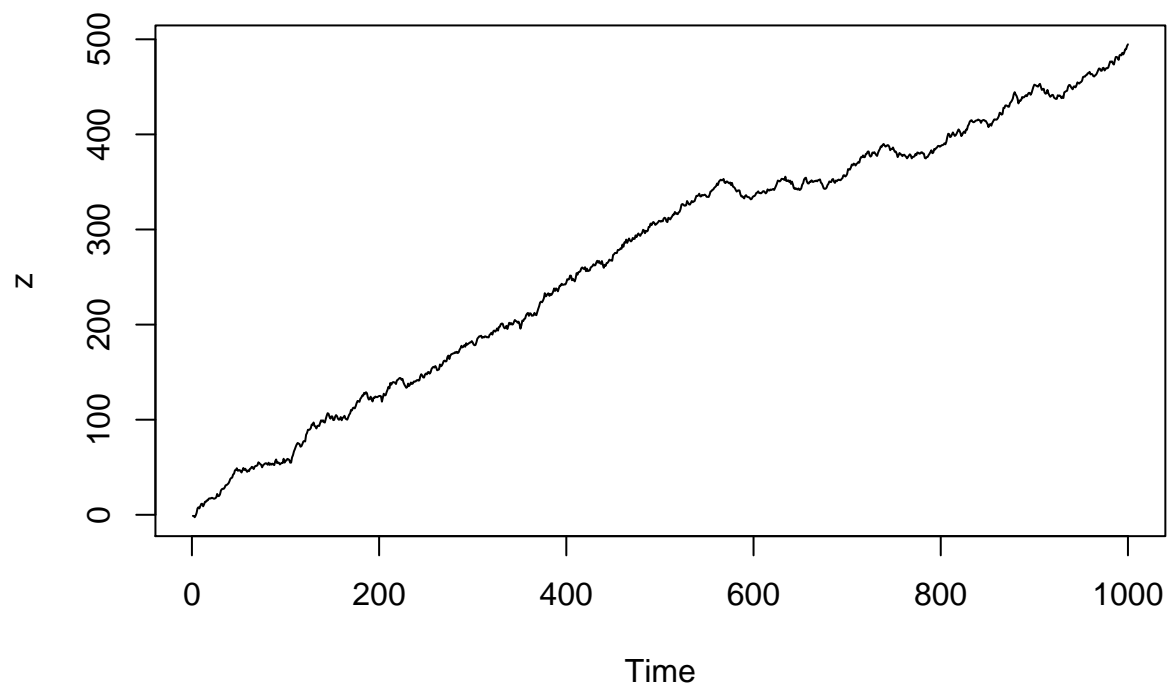
$$E[y_t] = .4(-1) + 0 + .5(1) = .1$$

por lo que se concluye que el proceso tiene una deriva estimada de .1.

Caminata aleatoria con distribución normal

Se puede simular una caminata aleatoria con deriva tomando muestras de una distribución como la normal. En la simulación a continuación, se muestrea una distribución normal, con media .5 y varianza 4. Nótese que en R se pasa el valor correspondiente de la desviación estándar.

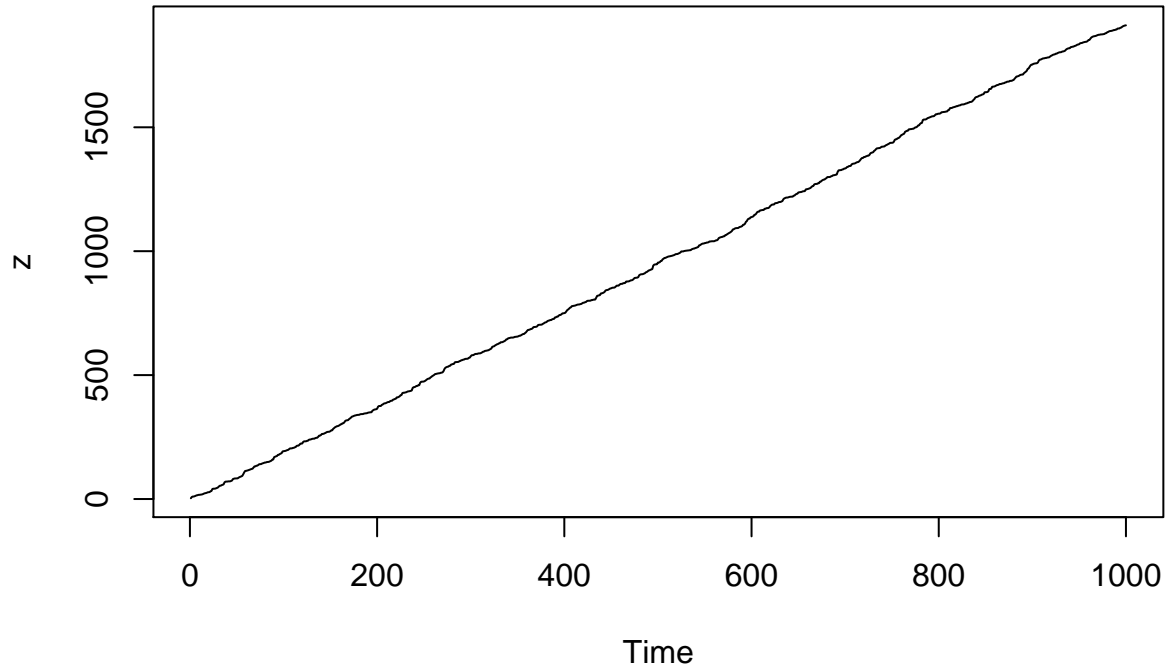
```
n <- 1000
z <- cumsum(rnorm(n, mean=.5, sd = 2) )
plot.ts(z)
```



Caminata aleatoria con distribución exponencial

En el siguiente ejemplo, usaremos la distribución exponencial con media $\lambda = 2$ (en R se pasa la tasa $1/\lambda = .5$)

```
n <- 1000  
z <- cumsum(rexp(n, rate = .5) )  
plot.ts(z)
```



Se pueden usar otras distribuciones, cambiando de manera oportuna la distribución simulada.

Tendencia lineal de una caminata aleatoria

En los tres ejemplos con deriva se puede observar como la tendencia es lineal. Efectivamente, usando sustituciones repetidas en el proceso

$$\begin{aligned}
 y_t &= \mu + y_{t-1} + u_t \\
 &= \mu + (\mu + y_{t-2} + u_{t-1}) + u_t \\
 &= 2\mu + (\mu + y_{t-3} + u_{t-2}) + u_{t-1} + u_t \\
 &\vdots \\
 &= t\mu + y_0 + \sum_{i=0}^t u_{t-i}
 \end{aligned}$$

se encuentra que la tendencia es lineal. El término y_0 corresponde al valor inicial ($t = 0$) del proceso.

Integración

Si una serie no estacionaria y_t se diferencia d veces y el resultado es una serie estacionaria, entonces se dice que y_t es integrada de orden d :

$$y_t \sim I(d) \Rightarrow \Delta^d y_t \sim I(0).$$

Si $d = 0$, entonces la serie es estacionaria. Si $y_t \sim I(1)$, entonces la serie tiene una raíz unitaria; si $y_t \sim I(2)$ tiene dos raíces unitarias y así sucesivamente.

El operador de diferencia $\Delta^d y_t$ se define inductivamente

$$\Delta^d y_t = \Delta (\Delta^{d-1} y_t)$$

Por ejemplo,

$$\Delta^2 y_t = \Delta (\Delta y_t) = \Delta (y_t - y_{t-1}) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}.$$

Ejemplo 4 (Caminata aleatoria)

La caminata aleatoria

$$y_t = y_{t-1} + u_t \quad \Delta y_t = u_t$$

por lo que la caminata aleatoria $\sim I(1)$.

Ejemplo 5

Considera $y_t = 3y_{t-1} + 4y_{t-2} + u_t$. La serie tiene polinomio característico

$$1 - 3z - 4z^2 = 0$$

que tiene raíces

```
polyroot(c(1, -3, -4))
```

```
## [1] 0.25+0i -1.00-0i
```

esto es, $|z_i| \leq 1$, por lo que la serie no es estacionaria.

Usando las raíces del polinomio característico, la serie dada se puede escribir como

$$(1 - 4L)(1 + L)y_t = u_t$$

Sobrediferenciación

Si una serie diferenciada es estacionaria, diferenciarla una, o más veces, puede causar una estructura no invertible $MA(q)$ de las innovaciones.

Por ejemplo, considera la caminata aleatoria $y_t = y_{t-1} + u_t$. La serie diferenciada

$$\Delta y_t = u_t$$

es estacionaria; diferenciando una vez más,

$$\Delta(\Delta y_t) = \Delta u_t = u_t - u_{t-1}$$

que es un proceso $MA(1)$ no invertible.

Pruebas de raíz unitaria; de estacionariedad

Es importante disponer de pruebas de hipótesis de estacionariedad de una serie de tiempo. Hay dos tipos de enfoque:

1. Pruebas de raíz unitaria: prueba de Dickey-Fuller aumentada (ADF); prueba de Phillips-Perron (PP)
2. Pruebas de estacionariedad: prueba KPSS (Kwiatkowski-Phillips-Schmidt-Shin)

Prueba de raíz unitaria

En la prueba de raíz unitaria, la prueba de hipótesis asociada es

H_0 : la serie tiene una raíz unitaria

H_1 : la serie no tiene una raíz unitaria

En una prueba de hipótesis de raíz unitaria, cuando se rechaza H_0 se determina que la serie es estacionaria.

La idea atrás de una prueba de raíz unitaria es la siguiente; usaremos una serie $AR(1)$ por simplicidad. Considera

$$y_t = \phi y_{t-1} + u_t.$$

El objetivo es determinar si $\phi = 1$ (presencia de raíz unitaria, es decir, serie no estacionaria), contra la alternativa $\phi < 1$ (serie estacionaria).

Restando y_{t-1} a ambos lados

$$\Delta y_t = \overbrace{(\phi - 1)}^{\psi =} y_{t-1} + u_t,$$

por lo que

$$\begin{aligned} H_0 : \phi &= 1 & \Leftrightarrow & \psi = 0 \\ H_1 : \phi &< 1 & \Leftrightarrow & \psi < 0 \end{aligned}$$

Comentarios

- Es posible que exista autocorrelación serial en la estimación de la regresión lineal por MCO.
- Para eliminar autocorrelación serial de primer orden, se usa la prueba de Dickey-Fuller aumentada.
- La prueba de Dickey-Fuller puede considerar una serie de tiempo con tendencia lineal

$$y_t = \phi y_{t-1} + \alpha + \beta t + u_t.$$

- La prueba de Phillips-Perron es semejante a la ADF, pero incluye correcciones para autocorrelación serial residual de orden superior al primero.
- Las dos pruebas de raíces unitarias tienen resultados semejante y tienen la misma desventaja, esto es, falta de *potencia*¹, especialmente con muestra pequeñas y con raíces cercana a uno.

Pruebas de raíz unitaria: falta de potencia

Por cuestiones didácticas, definiremos una función de R para realizar la simulación.

```
power.unit.root <- function(phi, nobs){  
  z <- arima.sim(nobs, model = list(ar = phi))  
  adf = adf.test(z)  
  pp = pp.test(z)  
  ans = list(adf = adf, pp = pp)  
  ans  
}
```

Muestras pequeñas

```
library(tseries)  
set.seed(123)  
power.unit.root(.95,20)
```

```
## $adf  
##  
## Augmented Dickey-Fuller Test
```

¹La potencia de una prueba de hipótesis es la probabilidad de rechazar H_0 cuando es falsa.

```
##
## data:  z
## Dickey-Fuller = -2, Lag order = 2, p-value = 0.4
## alternative hypothesis: stationary
##
##
## $pp
##
##  Phillips-Perron Unit Root Test
##
## data:  z
## Dickey-Fuller Z(alpha) = -13, Truncation lag parameter = 2,
## p-value = 0.3
## alternative hypothesis: stationary
```

Se simula una serie estacionaria, pero ambas pruebas no lo reconocen. Esto es, la potencia de la prueba no es buena.

Muestras grandes

```
library(tseries)
set.seed(123)
power.unit.root(.95,1000)

## Warning in adf.test(z): p-value smaller than printed p-value
## Warning in pp.test(z): p-value smaller than printed p-value
## $adf
##
##  Augmented Dickey-Fuller Test
##
## data:  z
## Dickey-Fuller = -4, Lag order = 9, p-value = 0.01
## alternative hypothesis: stationary
##
##
## $pp
##
##  Phillips-Perron Unit Root Test
##
## data:  z
## Dickey-Fuller Z(alpha) = -53, Truncation lag parameter = 7,
## p-value = 0.01
## alternative hypothesis: stationary
```

Con muestras grandes, ambas pruebas reconocen que el proceso es estacionario.

Prueba de estacionariedad

La prueba de estacionariedad más conocida es la prueba de Kwiatkowski, Phillips, Schmidt, and Y. Shin ([**KPSS**]).

El planteamiento de esta prueba se basa en la decomposición de la serie y_t mediante

$$y_t = \xi t + r_t + u_t,$$

donde el término ξt es un término de tendencia y

$$r_t = r_{t-1} + v_t, \quad v_t \sim N(0, \sigma_v^2) \quad (\text{iid});$$

el valor inicial, r_0 , sirve como intercepto en la expresión de la serie y_t .

Si denotamos con σ_v^2 la varianza del proceso v_t , entonces la hipótesis de estacionariedad de la serie y_t se puede expresar mediante

$$\begin{aligned} H_0 : \sigma_v^2 &= 0 \\ H_1 : \sigma_v^2 &> 0 \end{aligned}$$

Tomando en cuenta que u_t es un ruido blanco (proceso estacionario), entonces la serie y_t , bajo H_0 , es estacionaria (en tendencia).

Se puede también considerar el caso especial $\xi = 0$; en este caso, la serie y_t es estacionaria alrededor de un valor constante r_0 .

La librería **tseries** incluye una implementación de la prueba KPSS.

Como comparación con los resultados de las pruebas de raíz unitaria realizadas arriba, se usarán las mismas series simuladas (muestra pequeña; muestra grande):

```
library(tseries)
set.seed(123)
z20 <- arima.sim(20, model = list(ar = .95))
kpss.test(z20)
```

```
## Warning in kpss.test(z20): p-value greater than printed p-value
```

```
##
## KPSS Test for Level Stationarity
##
## data: z20
## KPSS Level = 0.3, Truncation lag parameter = 2, p-value = 0.1
```

```
library(tseries)
set.seed(123)
z1000 <- arima.sim(1000, model = list(ar = .95))
kpss.test(z1000)
```

```
##
## KPSS Test for Level Stationarity
##
## data: z1000
## KPSS Level = 0.4, Truncation lag parameter = 7, p-value = 0.07
```

En ambas simulaciones no se rechaza la hipótesis nula de estacionariedad.

Análisis de confirmación

El análisis de confirmación (“confirmatory data analysis”, en inglés) es el uso simultáneo de una prueba de raíz unitaria y de estacionariedad. Siguiendo la notación en el libro de Brooks, p. 365, denotaremos con $y_t \sim I(1)$ una serie no estacionaria y con $y_t \sim I(0)$ una estacionaria.

La hipótesis nula y alternativa de cada enfoque son

ADF/PP	KPSS
$H_0 : y_t \sim I(1)$	$H_0 : y_t \sim I(0)$
$H_1 : y_t \sim I(0)$	$H_0 : y_t \sim I(1)$

Hay cuatro posibles resultados:

ADF/PP	KPSS
Rechazar H_0	No rechazar H_0
No rechazar H_0	Rechazar H_0
Rechazar H_0	Rechazar H_0
No rechazar H_0	No rechazar H_0

Un resultado robusto es el primer par (estacionariedad) o el segundo (no estacionariedad). Los últimos dos pares dan resultados conflictivos.

Características de un proceso AR(p)

Teorema de Wold

Cualquier proceso estacionario se puede escribir mediante la suma de una parte determinística y de un proceso estocástico, este último como $MA(\infty)$.

El Teorema de Wold es intuitivamente claro en el caso de procesos autoregresivos. Considérese, por ejemplo, el caso $AR(1)$,

$$y_t = \phi_1 y_{t-1} + u_t$$

donde, si perder de generalidad, se puede suponer $\mu = 0$. Entonces,

$$\begin{aligned}y_t &= \phi_1 y_{t-1} + u_t \\y_{t-1} &= \phi_1 y_{t-2} + u_{t-1}.\end{aligned}$$

Sustituyendo, se encuentra

$$y_t = \phi_1(\phi_1 y_{t-2} + u_{t-1}) + u_t = \phi_1^2 y_{t-2} + \phi_1 u_{t-1} + u_t$$

Expresando el proceso y_t mediante sus rezagos sucesivos, y pasando al límite, se encuentra que el proceso $AR(1)$ se puede expresar mediante una serie de rezagos sucesivos de un ruido blanco con media cero.

A continuación se probarán las propiedades más básicas de un proceso $AR(p)$. Para simplificar el procedimiento, se ilustrará el caso $p = 1$ y después se dará el caso general $p > 1$.

Media incondicional

Se procede por pasos sucesivos.

$$\begin{aligned}E[y_t] &= E[\mu + \phi_1 y_{t-1} + u_t] \\&= \mu + \phi_1 E[y_{t-1}] = \mu + \phi_1 E[\mu + \phi_1 y_{t-2} + u_{t-1}] \\&= (\mu + \phi_1 \mu) + \phi_1^2 E[y_{t-2}]\end{aligned}$$

Después de n rezagos,

$$E[y_t] = (1 + \phi_1 + \phi_1^2 + \cdots + \phi_1^n) \mu + \phi_1^n E[y_{t-n}]$$

El modelo es estacionario, $|\phi_1| < 1$, por lo que $\phi_1^n \rightarrow 0$, cuando $n \rightarrow \infty$; por ende,

$$E[y_t] = \lim_{n \rightarrow \infty} \mu(1 + \cdots + \phi_1^n) = \mu \sum_{n=1}^{\infty} \phi_1^n = \frac{\mu}{1 - \phi_1}.$$

En el caso general,

$$E[y_t] = \frac{\mu}{1 - \phi_1 - \phi_2 \cdots - \phi_p}.$$

Varianza incondicional

Sin pérdida de generalidad, se supondrá $\mu = E[y_t] = 0$, por lo que

$$\text{var}(y_t) = E[y_t^2].$$

El Teorema de Wold implica

$$\begin{aligned}y_t &= \phi_1 y_{t-1} + u_t \\&= u_t + \phi_1 u_{t-1} + \phi_1^2 u_{t-2} + \cdots\end{aligned}$$

por lo que

$$y_t^2 = u_t^2 + \phi_1^2 u_{t-1}^2 + \phi_1^4 u_{t-2}^2 + \cdots + \text{productos cruzados}.$$

Entonces,

$$\begin{aligned} \text{var}(y_t) &= (1 + \phi_1^2 + \phi_1^4 + \cdots) \sigma^2 = \sum_{i=0}^{\infty} (\phi_1^2)^i \sigma^2 \\ &= \frac{\sigma^2}{1 - \phi_1^2}. \end{aligned}$$

Denotaremos $\text{var}(y_t) = \gamma_0$.

El caso general es semejante.

Correlaciones seriales

Las correlaciones seriales τ_s , $s = 1, 2, \dots$, son los coeficientes de la función de autocorrelación (ACF) y están definidas

$$\tau_s = \frac{\gamma_s}{\gamma_0} = \frac{\text{cov}(y_t, y_{t-s})}{\text{var}(y_t)}.$$

Sin falta de generalidad, se supondrá $\mu = 0 = E[y_t]$.

Analizaremos primeramente el caso $s = 1$.

Usando el Teorema de Wold,

$$\begin{aligned} y_t &= u_t + \phi_1 u_{t-1} + \phi_1^2 u_{t-2} + \cdots \\ y_{t-1} &= u_{t-1} + \phi_1 u_{t-2} + \phi_1^2 u_{t-3} + \cdots \end{aligned}$$

de manera que

$$y_t y_{t-1} = \phi_1 u_{t-1}^2 + \phi_1^3 u_{t-2}^2 + \phi_1^5 u_{t-3}^2 + \cdots + \text{productos cruzados}$$

Calculando los valores esperados,

$$\begin{aligned} E[y_t y_{t-1}] &= E[\phi_1 u_{t-1}^2] + E[\phi_1^3 u_{t-2}^2] + E[\phi_1^5 u_{t-3}^2] \\ &\quad + \cdots + E[\text{productos cruzados}] \end{aligned}$$

por lo que

$$\text{cov}(y_t, y_{t-1}) = \phi_1 \sigma^2 \left[\sum_{i=0}^{\infty} (\phi_1^2)^i \right] = \frac{\phi_1 \sigma^2}{1 - \phi_1^2}$$

Por lo que $\tau_1 = \phi_1$.

El caso $s > 1$ se analiza de una manera semejante y se encuentra que

$$\tau_s = \phi_1^s$$

Función de autocorrelación parcial (PACF)

Usaremos la definición que se encuentra en el libro de [T], Secc. 2.4.2, pp. 59-61.

La PACF (función de autocorrelación parcial) de una serie de tiempo estacionaria es una función de la ACF y se puede introducir de la siguiente manera:

$$\begin{aligned}y_t &= \phi_{0,1} + \phi_{1,1}y_{t-1} + e_{1t} \\y_t &= \phi_{0,2} + \phi_{1,2}y_{t-1} + \phi_{2,2}y_{t-2} + e_{2t} \\y_t &= \phi_{0,3} + \phi_{1,3}y_{t-1} + \phi_{2,3}y_{t-2} + \phi_{3,3}y_{t-3} + e_{3t} \\&\vdots = \vdots\end{aligned}$$

La PACF muestral con rezago s , $\hat{\phi}_{s,s}$, da la contribución adicional del término de rezago, y_{t-s} , a la serie y_t sobre el modelo $AR(s-1)$.

Por lo tanto, en un modelo $AR(p)$, la PACF muestral $\hat{\phi}_{j,j} \approx 0$, para todo $j > p$.

Se puede mostrar, para un modelo $AR(p)$ con errores normales, que

- $\hat{\phi}_{p,p} \rightarrow \phi_p$, si el tamaño muestral, $T \rightarrow \infty$.
- $\hat{\phi}_{j,j} \rightarrow 0$, para todo $j > p$.
- La varianza asintótica de $\hat{\phi}_{j,j} = 1/T$, para $j > p$.

En conclusión, la PACF muestral de un modelo $AR(p)$ es igual a cero, al rezago j , $j > p$.

Correlogramas y memoria de un proceso AR; comparaciones ACF, PACF para procesos MA, AR

Si $y_t = \mu + \sum_{i=1}^p L^i y_t + u_t$ es un proceso $AR(p)$. Entonces

- un modelo $AR(p)$ estacionario tendrá un correlograma que tiende a cero de manera geométrica (exponencial; posiblemente de carácter oscilatorio);
- el Teorema de Wold permite expresar un modelo $AR(p)$ por medio de un $MA(\infty)$, por lo que la "memoria" de las innovaciones u_t persiste de manera indefinida; en el caso de un proceso estacionario, la memoria decaerá de manera geométrica.
- La función de autocorrelación parcial decaerá a cero después de p rezagos (p : orden del proceso autoregresivo), porque existen relaciones entre y_t y sus primeros p rezagos pero y_t no tendrá ninguna relación con y_{t-s} , $s > p$.
- Considerando las relaciones entre procesos $AR(p)$ estacionarios con $MA(\infty)$ y de procesos $MA(q)$ invertibles con $AR(\infty)$ se puede concluir que la ACF de un proceso AR tiene la misma forma de una PACF de un proceso MA ; la ACF de un proceso MA tiene la misma forma de una PACF de un proceso AR .

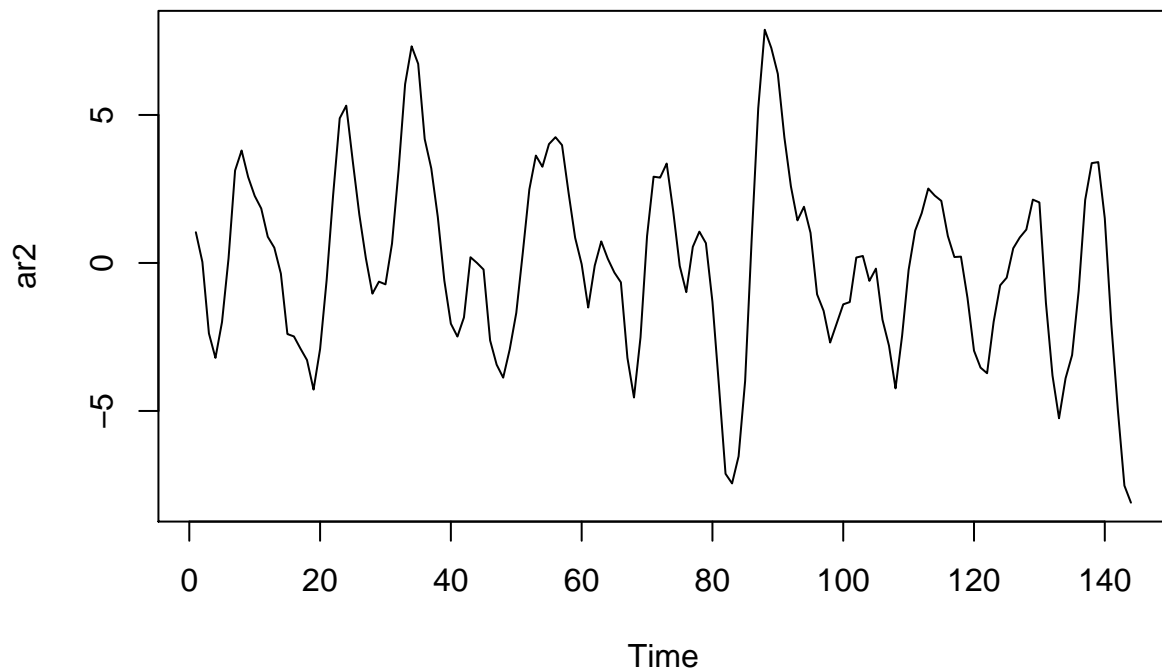
Ejemplo

Se presenta a continuación una simulación de la serie $AR(2)$,

$$y_t = 1.5y_{t-1} - .75y_{t-2} + u_t.$$

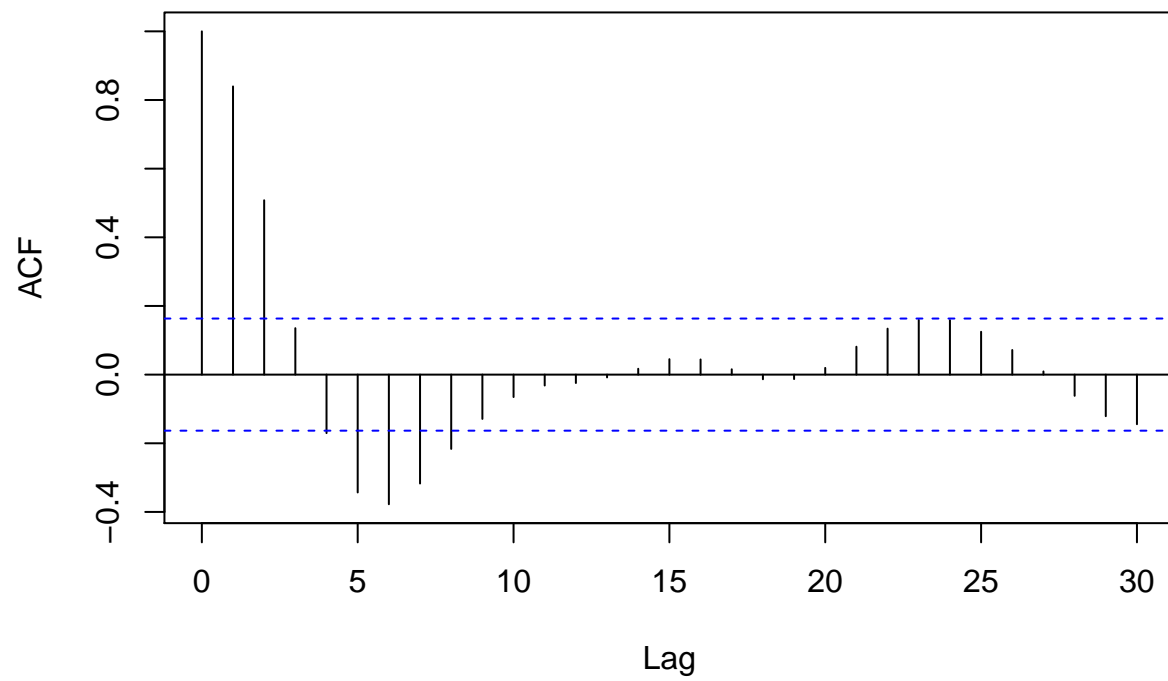
Esta serie presenta las características “ideales” mencionadas arriba; además, la serie presenta un carácter oscilatorio y se mostrará como se pueden usar las raíces complejas conjugadas del polinomio característico para estimar el ciclo.

```
set.seed(12321)
ar2 <- arima.sim(n=144, model = list(ar=c(1.5, -.75)))
plot.ts(ar2)
```



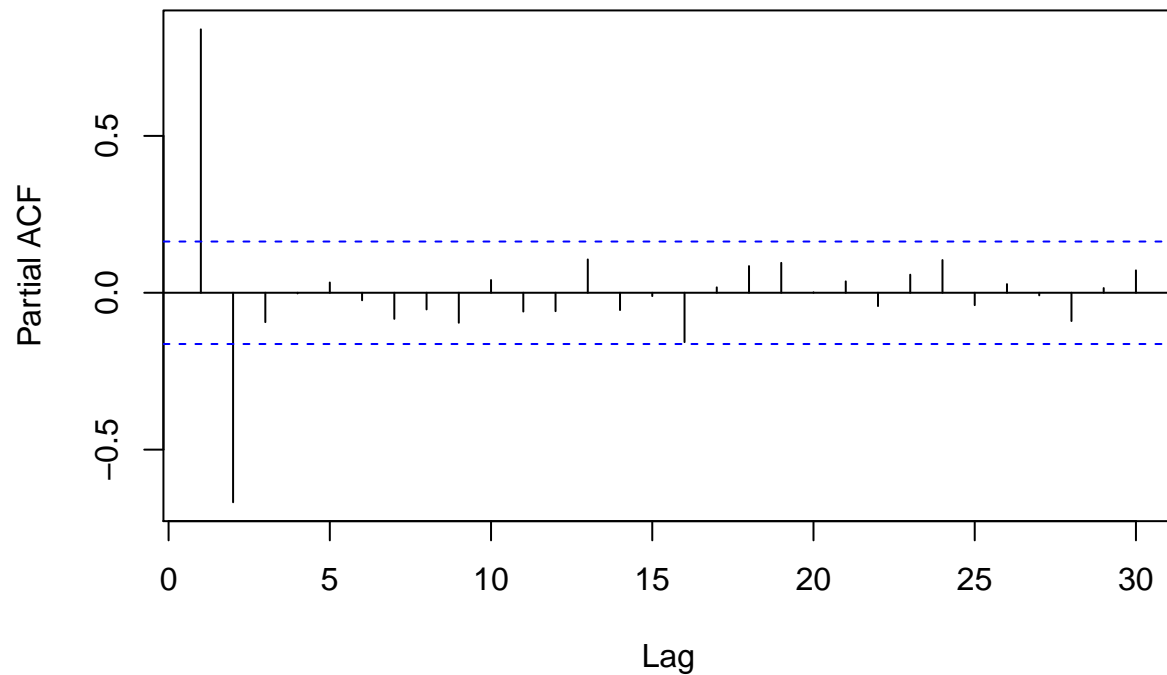
```
acf(ar2, lag.max = 30)
```

Series ar2



```
pacf(ar2, lag.max = 30)
```

Series ar2



El cálculo de las raíces del polinomio característico asociado

$$\Phi(z) = 1 - 1.5z + .75z^2 = 0$$

```
z = c(1, -1.5, .75)
root <- polyroot(z)
```

El ciclo se estima usando el *argumento* de una de las raíces; el argumento está en radianes por unidad de tiempo, por lo que lo transformamos en ciclos

```
Arg(root[1]) # argumento en radianes/tiempo
```

```
## [1] 0.524
```

```
arg <- Arg(root[1])/(2*pi) # argumento en ciclos/tiempo
1/arg # periodo estimado
```

```
## [1] 12
```

Modelos ARMA, ARIMA

Un proceso $ARMA(p, q)$ es una serie de tiempo con características estacionaria autoregresivas y de media móvil:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q} + u_t$$

Usando el operador de rezago, se puede escribir el proceso $ARMA(p, q)$ como

$$\Phi(L)y_t = \mu + \Theta(L)u_t$$

Todas las raíces de la ecuación característica asociada a $\Phi(L)$ tienen módulo menor a 1.

También se pedirá que la parte de media móvil del proceso sea invertible; véase la sección de invertibilidad y unicidad de un proceso $MA(q)$.

Sobrep parametrización

El material de esta sección se puede también encontrar en el libro de [ShS], Sección 3.1.

La sobrep parametrización de un proceso ARMA ocurre cuando las ecuaciones características asociadas a $\Phi(L)$, $\Theta(L)$ comparten una raíz.

Ejemplo

Considèrese el modelo $ARMA(1, 1)$

$$y_t = -.8y_{t-1} + .8u_{t-1} + u_t$$

Usando el operador de rezago, podemos escribir

$$(1 + .8L)y_t = (1 + .8L)u_t$$

que es lo mismo que un proceso de ruido blanco

$$y_t = u_t$$

Siguiendo [ShS], se generará una muestra $n = 1000$ de un ruido blanco $u_t \sim N(10, 1)$ y se ajustarán los valores a un modelo $ARMA(1, 1)$:

```
set.seed(1111)
x <- rnorm(1000, mean = 10)
arima(x, order = c(1,0,1))
```

```
##
## Call:
## arima(x = x, order = c(1, 0, 1))
##
## Coefficients:
##          ar1          ma1  intercept
##          0.990   -0.977           9.992
## s.e.      0.011    0.015         0.073
##
## sigma^2 estimated as 0.975:  log likelihood = -1406,  aic = 2820
```

La estimación del modelo es

$$(1 - .9901L)y_t = (1 - .9768)u_t + 9.9923$$

Se puede observar como los valores estimados $\hat{\phi}_1, \hat{\theta}_1$, se encuentran en el intervalo de confianza (95%) del otro, por lo que podemos deducir que se trata de un modelo sobreparametrizado.

Observando la estimación, vale la pena mencionar que los que se indica como **intercept** (intercepto) es en realidad la estimación de la media μ .

Características de un proceso ARMA(p,q)

Las características de un proceso $ARMA(p, q)$ son una combinación de un proceso autoregresivo y de media móvil.

- El valor esperado (media) del proceso es

$$E(y_t) = \frac{\mu}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}.$$

- La función de autocorrelación ACF tendrá un patrón decreciente geométricamente.
- La función de autocorrelación parcial PACF tendrá un patrón decreciente geométricamente.
- Después de q rezagos (q : orden de la parte de media móvil) la parte AR dominará el proceso, especialmente las funciones ACF, PACF.

La metodología de Box-Jenkins

Este procedimiento estructura la estimación de los procesos $ARMA(p, q)$:

1. Identificación
2. Estimación

3. Diagnóstico

Identificación En esta fase, se usan la forma de los correlogramas (AFC; PACF) para determinar el posible modelo ($AR(p)$, $MA(q)$, $ARMA(p, q)$), como se explicó arriba.

Estimación Una vez identificado el modelo, se estima el modelo usando técnicas como MCO o máxima verosimilitud. En R, hay también una función `auto.arima()` (librería `forecast`) que permite estimar un modelo minimizando uno de los criterios de información disponibles (véase abajo para la definición de los criterios de información más usados). El uso de esta función puede tener un efecto de sobreestimación del modelo, que no es necesariamente deseable; véase también la fase de diagnóstico.

Diagnóstico Una vez estimado el modelo, con valores p , q apropiados, se determina si el modelo se puede reducir, considerando la significancia estadística de los coeficientes estimados.

Para determinar si el modelo necesita términos adicionales, se puede usar el método de “overfitting”, que consiste en la estimación de un modelo con orden mayor; en la comparación se usarán algunos de los criterios de información y la significancia estadística de los parámetros adicionales.

Una vez determinado, y estimado, el modelo apropiado, se determinará si los residuales tienen autocorrelación serial, individualmente y en conjunto. Cuando se lleva a cabo una prueba portmanteau en R, usando la función `Box.test()`, es importante pasar el parámetro adicional `fitdf = p+q` para restar el número de grados de libertad del modelo estimado (con los valores apropiados p , q).

Puede ser conveniente usar pruebas de normalidad de los errores, aún si, en muestras grandes, se pueden usar versiones del Teorema Central del Límite.

Una regla general es que el modelo sea *parsimonioso*, de manera que capture la *señal* y no el *ruido*.

Para los propósitos de estas notas, se usarán modelos estacionarios, por lo que no se usarán modelos $ARIMA(p, d, q)$ explícitamente.

Criterios de Información

La idea del criterio de información es *penalizar* el uso excesivo de parámetros a estimar; en este sentido, es una herramienta más sofisticada del R^2 ajustado que se usa en regresión lineal.

Su uso es también justificado cuando se usan datos reales “ruidosos” porque el uso de los correlogramas puede no ser tan claros, como menciona la teoría.

La idea general de un criterio de información es de incluir en una fórmula el número de parámetros del modelo y *minimizar* esta fórmula.

A continuación se presentan los criterios de información más usados: AIC (criterio de información de Akaike), (S)BIC (criterio de información bayesiano de Schwarz) y HQIC

(criterio de información de Hannah-Quinn).

En todos, T es el tamaño muestral; $k = p + q + 1$ es el número de parámetros a estimar; $\hat{\sigma}^2$ es la varianza residual, $\hat{\sigma}^2 = SRC/T$:

$$\begin{aligned} AIC &= \ln(\hat{\sigma}^2) + \frac{2k}{T} \\ (S)BIC &= \ln(\hat{\sigma}^2) + \frac{k}{T} \ln(T) \\ HQIC &= \ln(\hat{\sigma}^2) + \frac{2k}{T} \ln(\ln(T)) \end{aligned}$$

De los tres criterios, el que más penaliza el número de parámetros a estimar es el SBIC y el que penaliza menos es el AIC; el HQIC se encuentra en el medio.

En términos estadísticos, (S)BIC consistente, pero ineficiente; AIC es, por lo general, eficiente, pero no es consistente.

La varianza de las estimaciones será, por lo general, mayor usando (S)BIC que AIC.

En término de elección del modelo, cuando $T \rightarrow \infty$ (muestras muy grandes) (S)BIC proporciona el modelo correcto y AIC proporciona un modelo no parsimonioso. Si aplica la Ley de los Grandes Números, (S)BIC puede ser la mejor elección, porque la varianza, de todas maneras tiende a cero, cuando $T \rightarrow \infty$.

Pronósticos

Indicamos con Ω_t toda la información disponible hasta el tiempo t (incluido).

Un *pronóstico*, al tiempo $t + h$, se genera por medio de un valor esperado, condicionado sobre Ω_t :

$$f_{t,h} = E[y_{t+h}|\Omega_t]$$

Método de pronóstico simple

1. Si se supone que no habrá cambios, entonces un pronóstico simple, h pasos en adelante, es de usar el valor actual

$$E[y_{t+h}|\Omega_t] = y_t$$

Este tipo de pronóstico es óptimo, si el proceso es una caminata aleatoria:

$$y_{t+1} = y_t + u_t.$$

porque, en este caso,

$$E[y_{t+1}|\Omega_t] = E[y_t + u_t|\Omega_t] = E[y_t|\Omega_t] + E[u_t|\Omega_t] = y_t$$

2. Si la serie es estacionaria (“mean-reverting”), otro método de pronóstico simple es usar la media incondicionada.

Pronóstico mediante suavización exponencial

La técnica de suavización exponencial considera que se puede determinar un pronóstico usando una combinación de las últimas observaciones; la pregunta es cuánto peso se debe asignar a estas observaciones.

Una desventaja de una modelación ARMA (además de la determinación del número de parámetros a estimar), es que un modelo $MA(q)$ tiene una “memoria” limitada sólo a los últimos q rezagos, mientras que un modelo $AR(p)$ tiene una memoria indefinida. De hecho, un suavizado exponencial es un caso especial de un modelo $ARIMA(0, 1, 1)$, con coeficientes MA iguales a $(1 - \alpha)$.

La ventaja de una modelación de suavización exponencial es que permite dar un peso a observaciones en el pasado, sin limitarse necesariamente a las últimas más recientes, como en un modelo de media móvil. El peso se asigna mediante una relación recursiva, de la siguiente manera. Si S_t es el valor actual suavizado, entonces

$$S_t = \alpha y_t + (1 - \alpha)S_{t-1},$$

con $0 < \alpha < 1$.

El valor actual S_t depende de una manera geoméricamente decreciente de los valores pasados; sustituyendo hacia atrás

$$\begin{aligned} S_t &= \alpha y_t + (1 - \alpha)S_{t-1} \\ &= \alpha y_t + (1 - \alpha) [\alpha y_{t-1} + (1 - \alpha)S_{t-2}] \\ &= \alpha [y_t + (1 - \alpha)y_{t-1}] + (1 - \alpha)^2 S_{t-2} \\ &= \alpha [y_t + (1 - \alpha)y_{t-1} + (1 - \alpha)^2 y_{t-2}] + (1 - \alpha)^3 S_{t-3} \\ &\vdots \quad (\text{después de } n \text{ sustituciones hacia atrás}) \\ &= \alpha \left[\sum_{i=0}^n (1 - \alpha)^i y_{t-i} \right] + (1 - \alpha)^{n+1} S_{t-(n+1)}. \end{aligned}$$

Cuando $n \rightarrow \infty$, el valor suavizado actual se expresa mediante una combinación geométrica de sus valores anteriores.

El pronóstico de una suavización exponencial es, entonces,

$$f_{t,s} = S_t, \quad s = 1, 2, \text{ dots}$$

El método de suavización exponencial se puede extender a series con tendencias (método de Holt) o estacionalidades (método de Winter).

Pronóstico para MA(q)

Por simplicidad, se considera el caso $q = 2$; el caso general se deriva de manera directa.

Sea

$$y_t = \mu + \theta_1 u_{t-1} + \theta_2 u_{t-2} + u_t$$

Entonces,

$$\begin{aligned} f_{t,1} &= \mu + \theta_1 u_t + \theta_2 u_{t-1} \\ f_{t,2} &= \mu + \theta_2 u_t \\ f_{t,3} &= f_{t,s} = \mu, \quad s > 3. \end{aligned}$$

La prueba es directa, ya que

$$y_{t+1} = \mu + \theta_1 u_t + \theta_2 u_{t-1} + u_{t+1}$$

por lo que

$$E[y_{t+1}|\Omega_t] = E[\mu + \theta_1 u_t + \theta_2 u_{t-1} + u_{t+1}|\Omega_t] = \mu + \theta_1 u_t + \theta_2 u_{t-1},$$

ya que u_{t+1} es independiente de u_s , $s \leq t$. Un paso más,

$$E[y_{t+2}|\Omega_t] = E[\mu + \theta_1 u_{t+1} + \theta_2 u_t + u_{t+2}|\Omega_t] = \mu + \theta_2 u_t.$$

Finalmente,

$$E[y_{t+s}|\Omega_t] = E[\mu + \theta_1 u_{t+s-1} + \theta_2 u_{t+s-2} + u_{t+s}|\Omega_t] = \mu,$$

para todo $s \geq 3$.

Como ya se comentó arriba, un modelo $MA(q)$ sólo tiene memoria de los últimos q rezagos y esto se ve reflejado también en el pronóstico.

Pronóstico para AR(p)

Por simplicidad, también se considera el caso $p = 2$.

Sea

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t$$

Entonces,

$$\begin{aligned} f_{t,1} &= \mu + \phi_1 y_t + \phi_2 y_{t-1} \\ f_{t,2} &= \mu + \phi_1 f_{t,1} + \phi_2 y_t \\ f_{t,3} &= \mu + \phi_1 f_{t,2} + \phi_2 f_{t,1} \\ f_{t,s} &= \mu + \phi_1 f_{t,s-1} + \phi_2 f_{t,s-2} \end{aligned}$$

La prueba es directa, ya que

$$y_{t+1} = \mu + \phi_1 y_t + \phi_2 y_{t-1} + u_{t+1}$$

por lo que

$$E[y_{t+1}|\Omega_t] = E[\mu + \phi_1 y_t + \phi_2 y_{t-1} + u_{t+1}|\Omega_t] = \mu + \phi_1 y_t + \phi_2 y_{t-1}.$$

Un paso más,

$$\begin{aligned} E[y_{t+2}|\Omega_t] &= E[\mu + \phi_1 y_{t+1} + \phi_2 y_t + u_{t+2}|\Omega_t] \\ &= \mu + \phi_1 E[y_{t+1}|\Omega_t] + \phi_2 y_t \\ &= \mu + \phi_1 f_{t,1} + \phi_2 y_t, \end{aligned}$$

ya que $f_{t,1} = E[y_{t+1}|\Omega_t]$. Finalmente,

$$E[y_{t+s}|\Omega_t] = E[\mu + \phi_1 u_{t+s-1} + \phi_2 u_{t+s-2} + u_{t+s}|\Omega_t] = \mu + \phi_1 f_{t,s-1} + \phi_2 f_{t,s-2},$$

para todo $s \geq 3$.

Como se ve también en el pronóstico, un proceso autoregresivo tiene una memoria “larga”.

Pronósticos con R

Ejemplo 1 (Simulación)

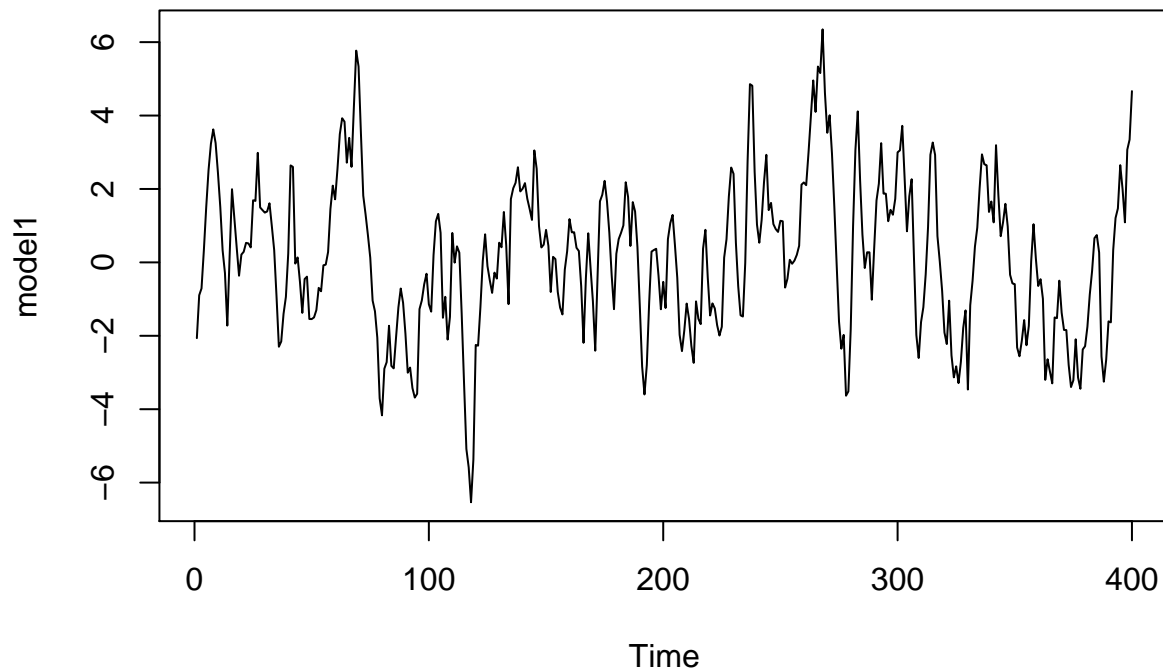
A continuación presentaremos una implementación en R del pronóstico de una serie simulada $AR(2)$.

Para ilustrar también la metodología Box-Jenkins, se seguirán los pasos de este procedimiento y se estimará la serie simulada. En el caso de una serie simulada, el procedimiento de Box-Jenkins funciona de manera ideal. En el caso de series de tiempo reales, las cosas son un poco más complicadas, por lo que podría ser útil el uso de herramientas como `auto.arima()` (librería `forecast`), `eacf()` (librería `TSA`) para ayudar en la selección del orden apropiado.

Para determinar la bondad del pronóstico, se usarán el 80% de las observaciones para generar el pronóstico y se usará el restante 20% (“holdout sample”) para construir pronósticos fuera de muestra.

```
set.seed(123)
library(tseries) # tseries::adf.test
library(forecast) # forecast::forecast()
library(TSA) # tsa::eacf() extended ACF
model1 <- arima.sim(model=list(ar=c(1.2,-.32)),n=400)

plot.ts(model1) # the series looks stationary
```



La serie parece estacionaria; probaremos esta hipótesis mediante una prueba de Dickey-Fuller y una KPSS:

```
adf.test(model1)
```

```
## Warning in adf.test(model1): p-value smaller than printed p-value
##
## Augmented Dickey-Fuller Test
##
## data: model1
## Dickey-Fuller = -5, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary
```

```
kpss.test(model1)
```

```
## Warning in kpss.test(model1): p-value greater than printed p-value
##
## KPSS Test for Level Stationarity
##
## data: model1
## KPSS Level = 0.1, Truncation lag parameter = 5, p-value = 0.1
```

Efectivamente, la serie es estacionaria. Para determinar si la serie tiene media cero

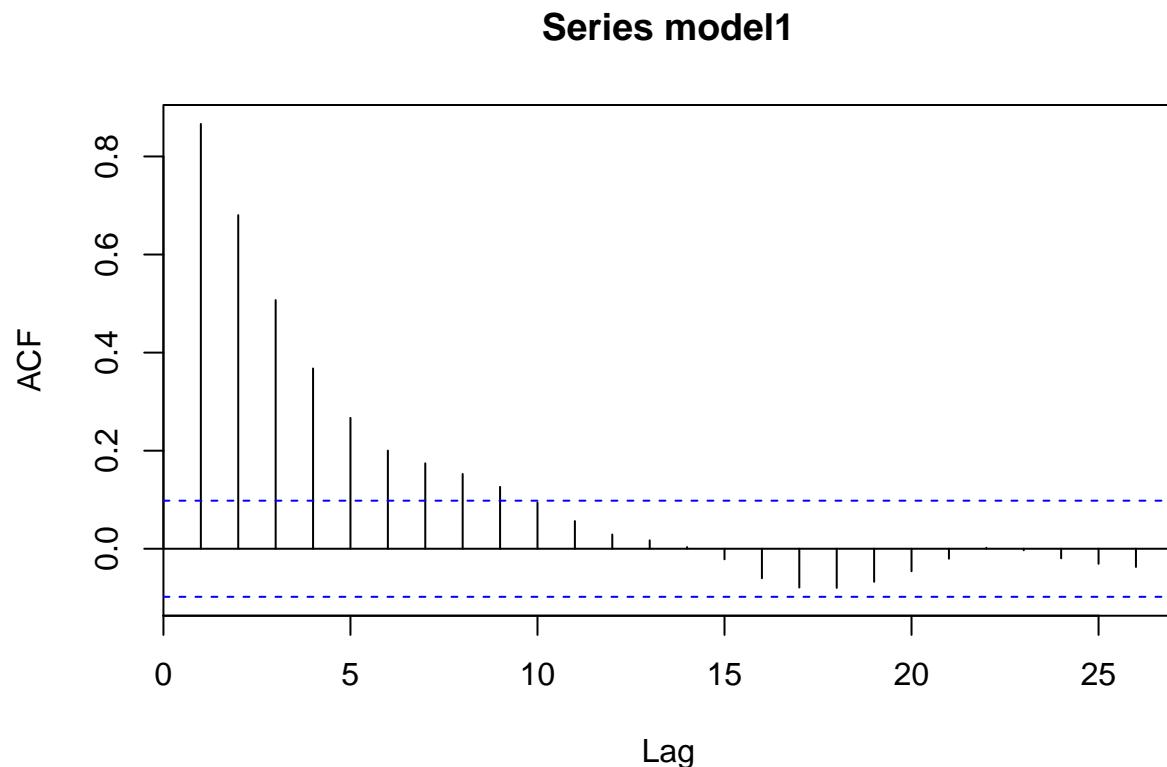
```
media <- mean(model1)
t.test(model1)
```

```
##
## One Sample t-test
##
## data:  model1
## t = 1, df = 399, p-value = 0.2
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## -0.074  0.336
## sample estimates:
## mean of x
##      0.131
```

La serie tiene media cero, por lo que se eliminará la media en la estimación sucesiva.

Pasamos a la identificación del modelo. Primeramente, la función de autocorrelación:

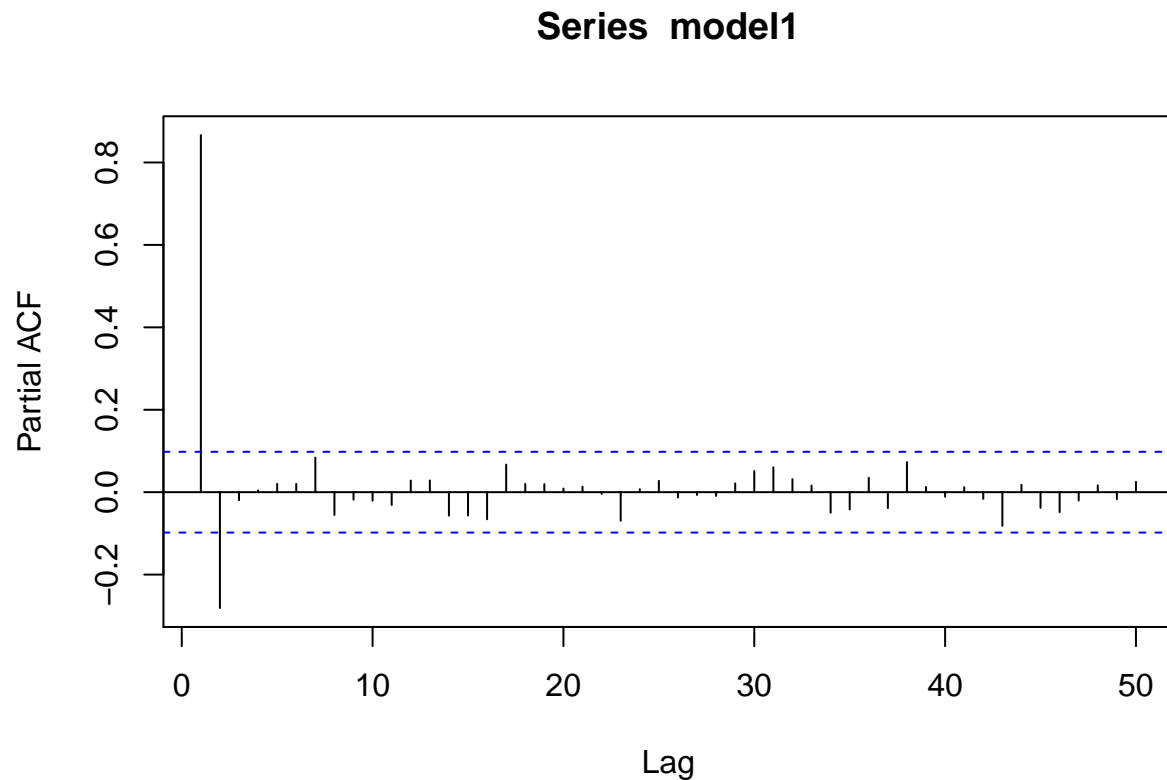
```
acf(model1)
```



que muestra un patrón geoméricamente decreciente, consistente con una serie autoregresiva.

El orden se puede determinar usando la función de autocorrelación parcial, PACF:

```
pacf(model1, lag.max = 50)
```



La PACF sugiere $p = 2$.

Pasamos a la estimación del modelo $AR(2)$ sobre el 80% de los datos.

```
model1.train <- Arima(model1[1:320], order = c(2,0,0), include.mean = F)
model1.train
```

```
## Series: model1[1:320]
## ARIMA(2,0,0) with zero mean
##
## Coefficients:
##          ar1      ar2
##          1.186  -0.361
## s.e.    0.052   0.052
##
## sigma^2 estimated as 0.91:  log likelihood=-439
## AIC=884   AICc=884   BIC=895
```

Para el diagnóstico se usarán la función ACF (para determinar si existe autocorrelación serial individual); pruebas de Ljung-Box (con diferentes rezagos, para pruebalos de autocorrelación serial conjunta) y pruebas de normalidad (Pruebas de Shapiro-Wilk; Jarque-Bera).

```

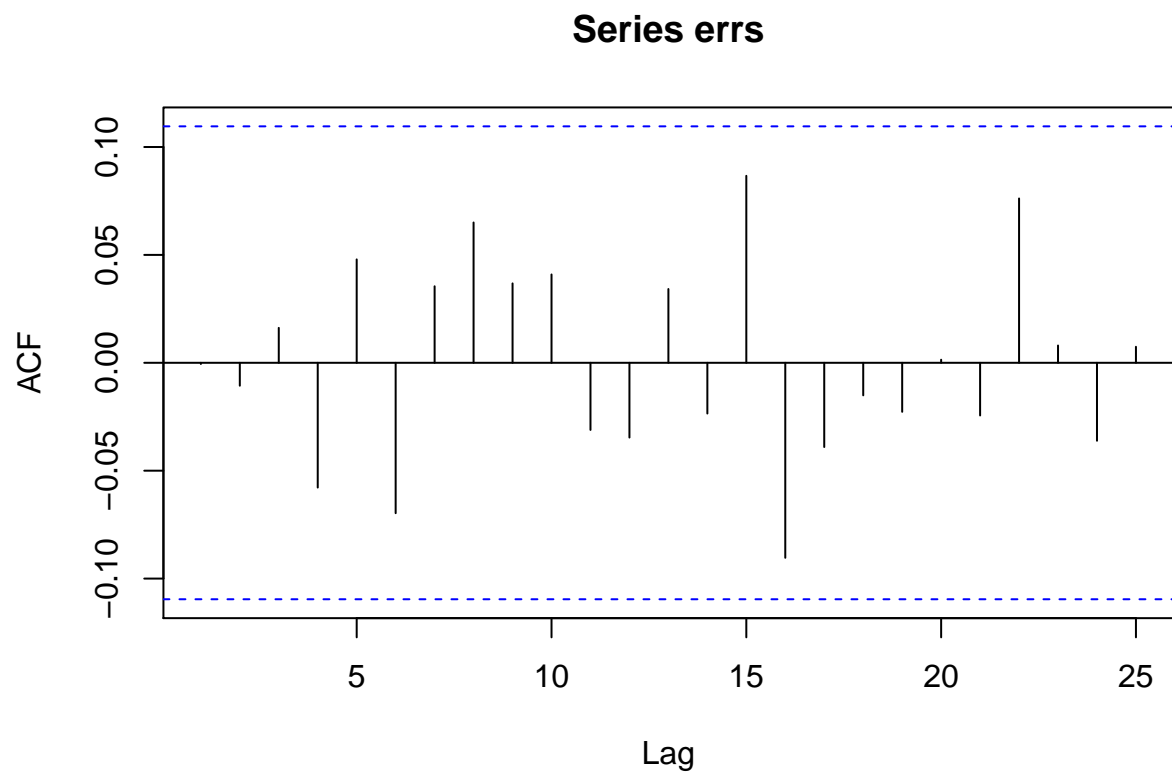
errs <- model1.train$residuals
Box.test(errs, lag = 20, type = "Ljung-Box", fitdf = 2)

##
## Box-Ljung test
##
## data:  errs
## X-squared = 14, df = 18, p-value = 0.7
Box.test(errs, lag = 25, type = "Ljung-Box", fitdf = 2)

##
## Box-Ljung test
##
## data:  errs
## X-squared = 16, df = 23, p-value = 0.8
Box.test(errs, lag = 30, type = "Ljung-Box", fitdf = 2)

##
## Box-Ljung test
##
## data:  errs
## X-squared = 17, df = 28, p-value = 0.9
acf(errs) # individual significance

```



```
shapiro.test(errs)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  errs
## W = 1, p-value = 0.1
```

```
jarque.bera.test(errs)
```

```
##
##  Jarque Bera Test
##
## data:  errs
## X-squared = 4, df = 2, p-value = 0.1
```

No existe evidencia de autocorrelación serial; además, los errores también parecen tener distribución normal.

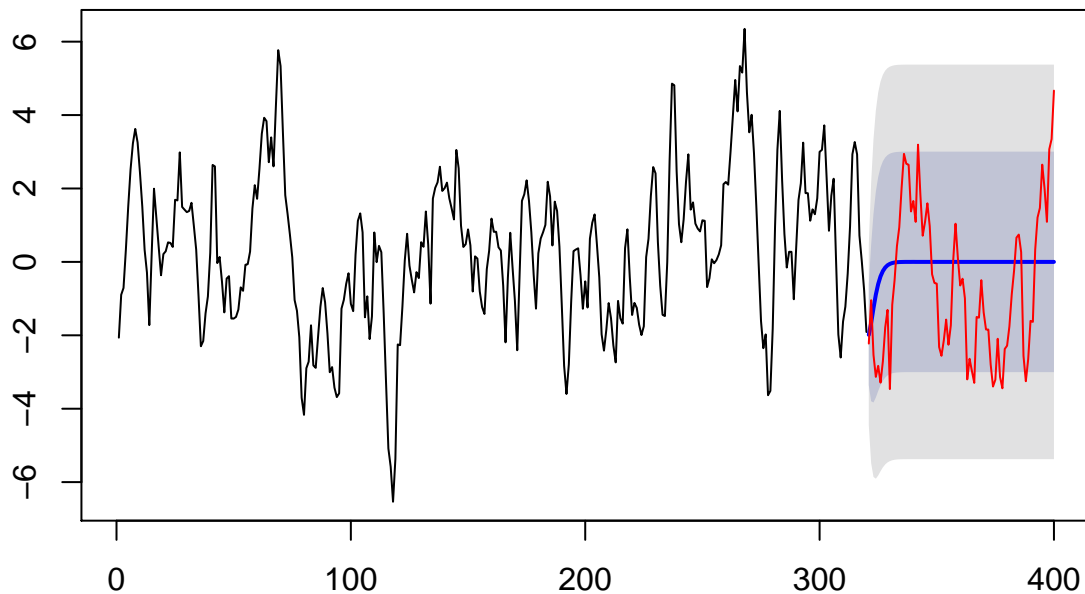
Después, se usará este modelo estimado para crear el pronóstico con intervalos de confianza 85%, 99% (para propósitos ilustrativos; por defecto, los intervalos de confianza son de 80%, 95%):

```
in.sample <- forecast(model1.train, h =80, level = c(85, 99))
```

Después, se creará la serie fuera de muestra para determinar la bondad del pronóstico:

```
model1.test = window(ts(model1), start = 321)
plot(in.sample)
lines(model1.test, col="red")
```

Forecasts from ARIMA(2,0,0) with zero mean



La totalidad de las observaciones fuera de muestra se encuentran en el intervalo de confianza al 99%. El pronóstico parece ser adecuado.

Otras medidas de desempeño son *MAE* (Mean Absolute Error) y *MAPE* (Mean Absolute Percentage Error) para medir de manera más cuantitativa la discrepancia entre los valores pronosticados con los valores fuera de muestra:

$$MAE = \frac{1}{T} \sum_{i=1}^T | \text{pronóstico} - \text{valor fuera de muestra} |$$

$$MAPE = \frac{1}{T} \sum_{i=1}^T \left| \frac{\text{pronóstico} - \text{valor fuera de muestra}}{\text{valor fuera de muestra}} \right|$$

El objeto `in.sample` contiene los valores promedios del pronóstico; extraemos la media y la comparamos con los valores fuera de muestra:


```
(1/400)*sum(abs((in.sample$mean - model1.test))) # MAE
```

```
## [1] 0.355
```

```
(1/400)*sum(abs((in.sample$mean - model1.test)/model1.test)) # MAPE
```

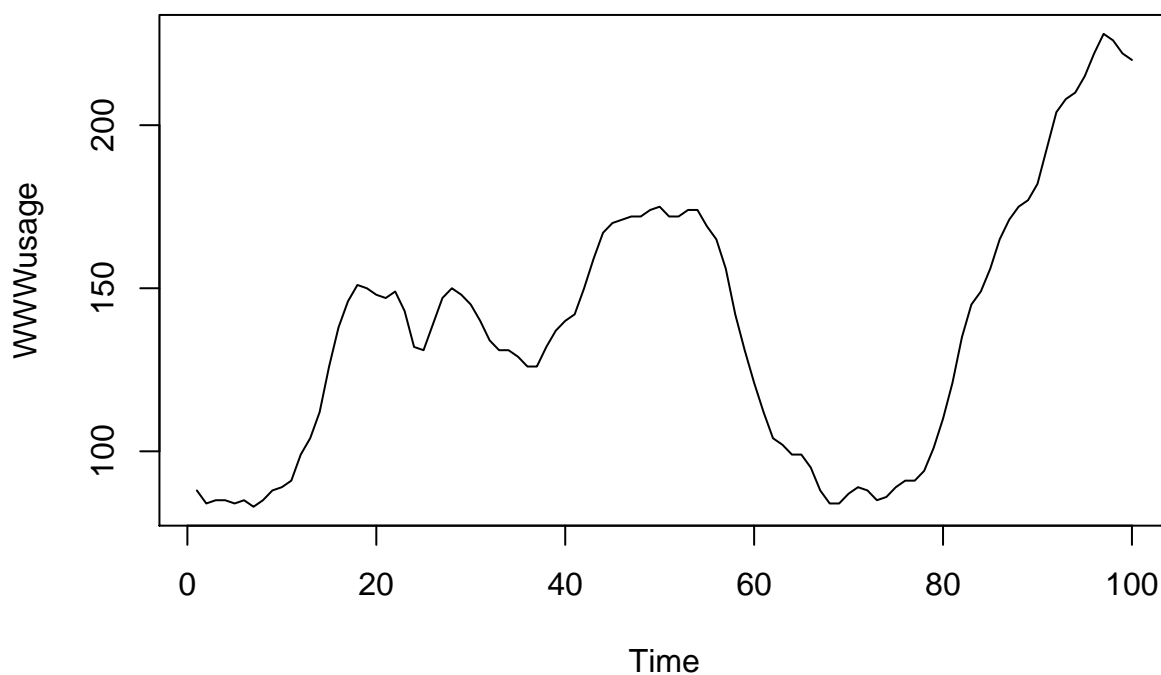
```
## [1] 0.193
```

Ejemplo 2

Para la autonomía del documento, se usará la serie `WWWusage`, que cuenta el número de usuarios conectado a Internet por minuto; véase la ayuda `?WWWusage`.

La gráfica indica una posible no estacionariedad de la serie de tiempo asociada,

```
plot.ts(WWWusage)
```



por lo que aplicamos las pruebas de estacionariedad

```
adf.test(WWWusage)
```

```
##
```

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

```
##
```

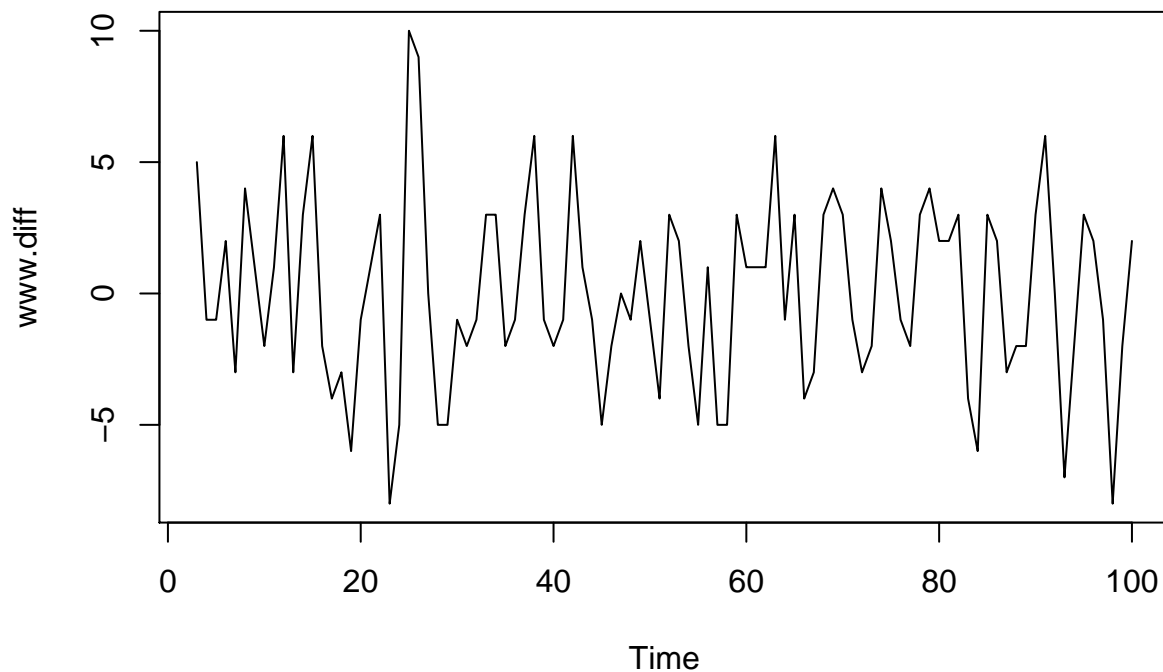
```
## data: WWWusage
## Dickey-Fuller = -3, Lag order = 4, p-value = 0.3
## alternative hypothesis: stationary
```

```
kpss.test(WWWusage)
```

```
##
## KPSS Test for Level Stationarity
##
## data: WWWusage
## KPSS Level = 0.5, Truncation lag parameter = 4, p-value = 0.05
```

Existe cierta evidencia de no estacionariedad, por lo que pasamos a considerar la serie diferenciada dos veces (diferenciando una vez, no se encuentra evidencia robusta de estacionariedad)

```
www.diff <- as.ts(diff(diff(WWWusage)))
plot.ts(www.diff)
```



```
adf.test(www.diff)
```

```
## Warning in adf.test(www.diff): p-value smaller than printed p-value
##
## Augmented Dickey-Fuller Test
```

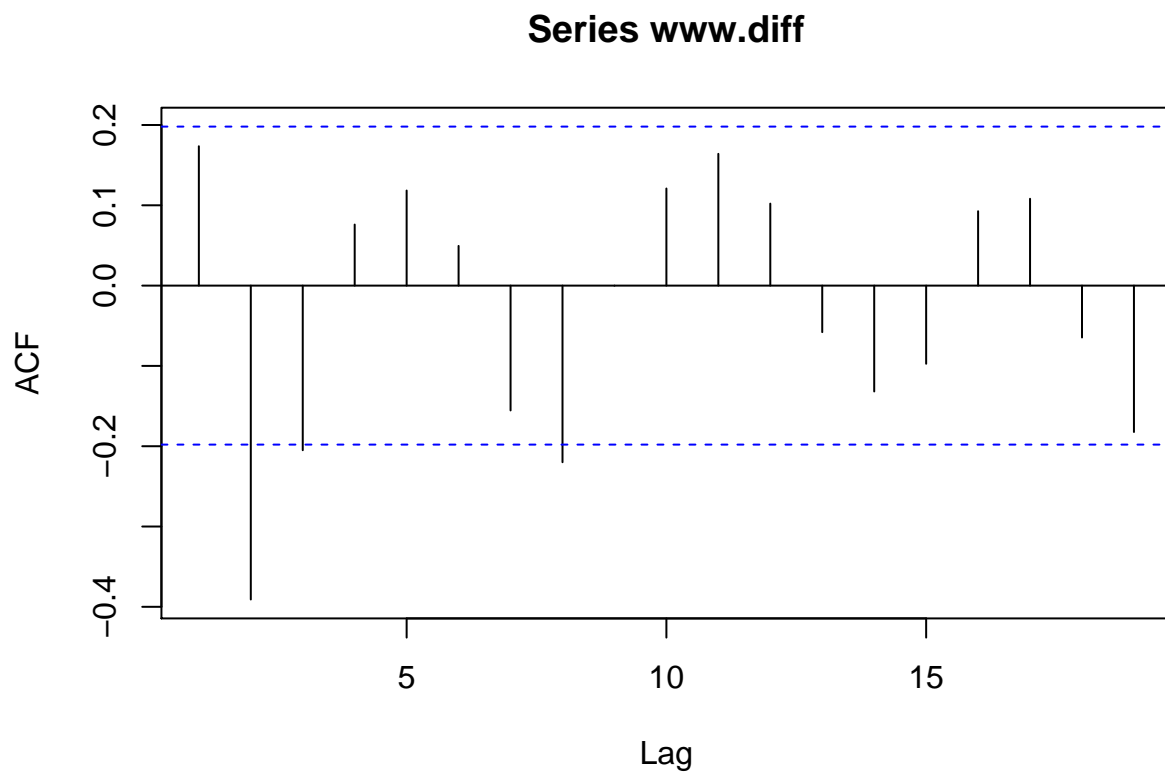
```
##
## data:  www.diff
## Dickey-Fuller = -5, Lag order = 4, p-value = 0.01
## alternative hypothesis: stationary

kpss.test(www.diff)

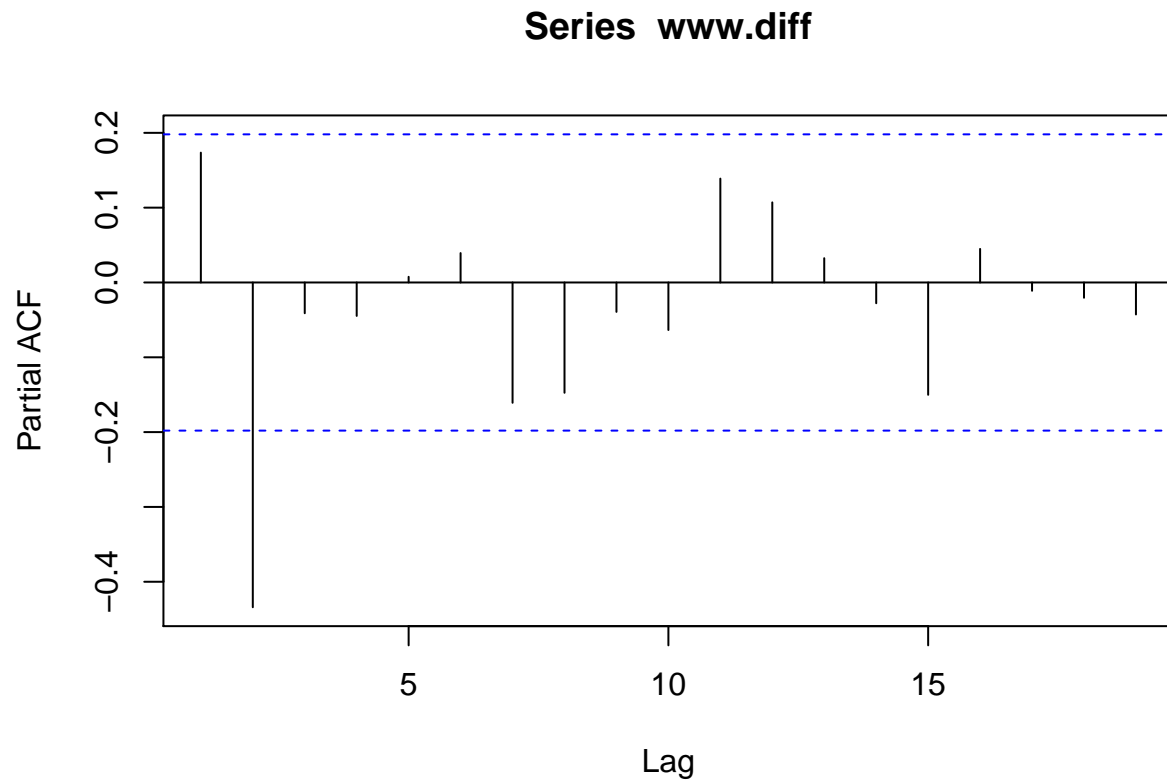
## Warning in kpss.test(www.diff): p-value greater than printed p-value
##
## KPSS Test for Level Stationarity
##
## data:  www.diff
## KPSS Level = 0.05, Truncation lag parameter = 3, p-value = 0.1
```

Los correlogramas, ACF, PACF, muestran que un posible modelo $AR(2)$ es apropiado:

```
acf(www.diff)
```



```
pacf(www.diff)
```



Para determinar si incluir, o no, una media, se prueba la hipótesis de media nula:

```
t.test(www.diff)
```

```
##
##  One Sample t-test
##
## data:  www.diff
## t = 0.06, df = 97, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.706  0.747
## sample estimates:
## mean of x
##  0.0204
```

No hay evidencia que la media sea diferente de cero.

Estimamos el modelo $AR(2)$:

```
www.modelo1.fit <- Arima(www.diff, include.mean = F, order = c(2,0,0))
www.modelo1.fit
```

```
## Series: www.diff
## ARIMA(2,0,0) with zero mean
```

```
##
## Coefficients:
##          ar1      ar2
##          0.258 -0.441
## s.e.    0.092   0.091
##
## sigma^2 estimated as 10.3:  log likelihood=-253
## AIC=511   AICc=512   BIC=519
```

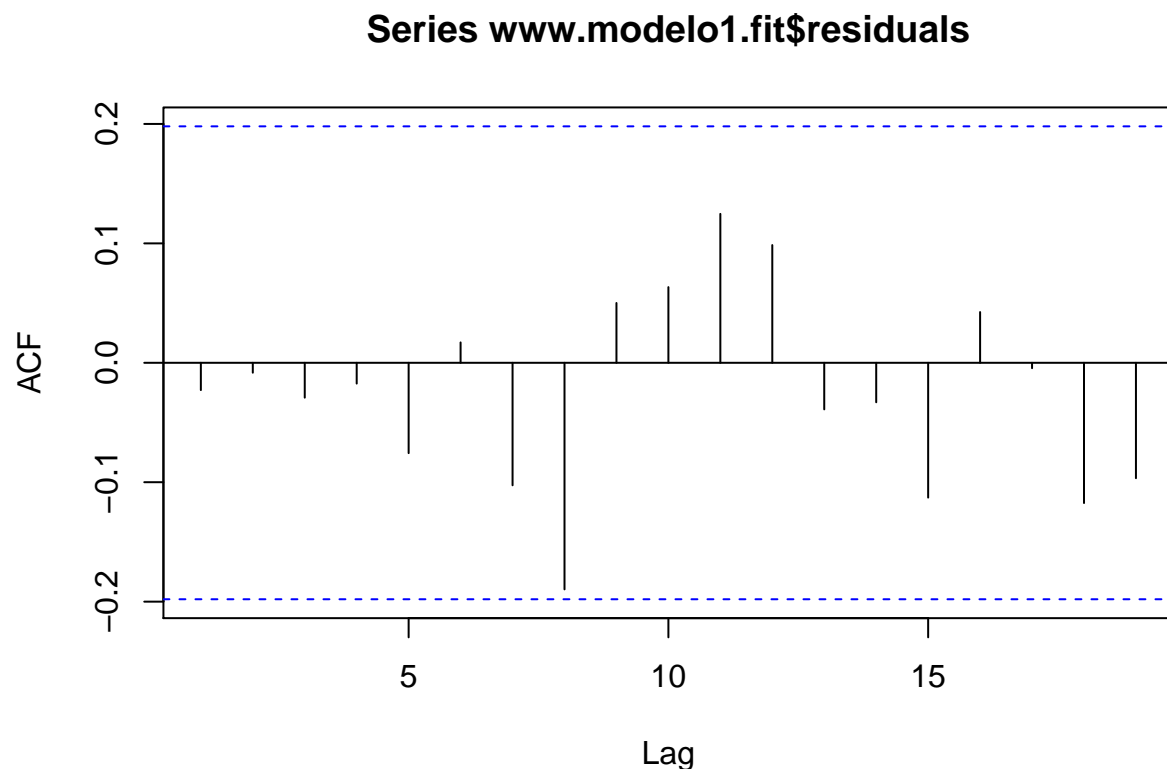
Corremos también la función `auto.arima()` que identifica el mismo modelo:

```
www.modelo2.fit <- auto.arima(www.diff, max.p = 6,
                             max.q = 6, allowmean = F,
                             seasonal = F, stationary = T)
www.modelo2.fit
```

```
## Series: www.diff
## ARIMA(2,0,0) with zero mean
##
## Coefficients:
##          ar1      ar2
##          0.258 -0.441
## s.e.    0.092   0.091
##
## sigma^2 estimated as 10.3:  log likelihood=-253
## AIC=511   AICc=512   BIC=519
```

Se realiza el diagnóstico de autocorrelación serial, individual

```
acf(www.modelo1.fit$residuals)
```



y en conjunto, con diferentes rezagos:

```
Box.test(www.modelo1.fit$residuals, fitdf = 2, lag=12)
```

```
##
## Box-Pierce test
##
## data:  www.modelo1.fit$residuals
## X-squared = 8, df = 10, p-value = 0.6
```

```
Box.test(www.modelo1.fit$residuals, fitdf = 2, lag=20)
```

```
##
## Box-Pierce test
##
## data:  www.modelo1.fit$residuals
## X-squared = 14, df = 18, p-value = 0.7
```

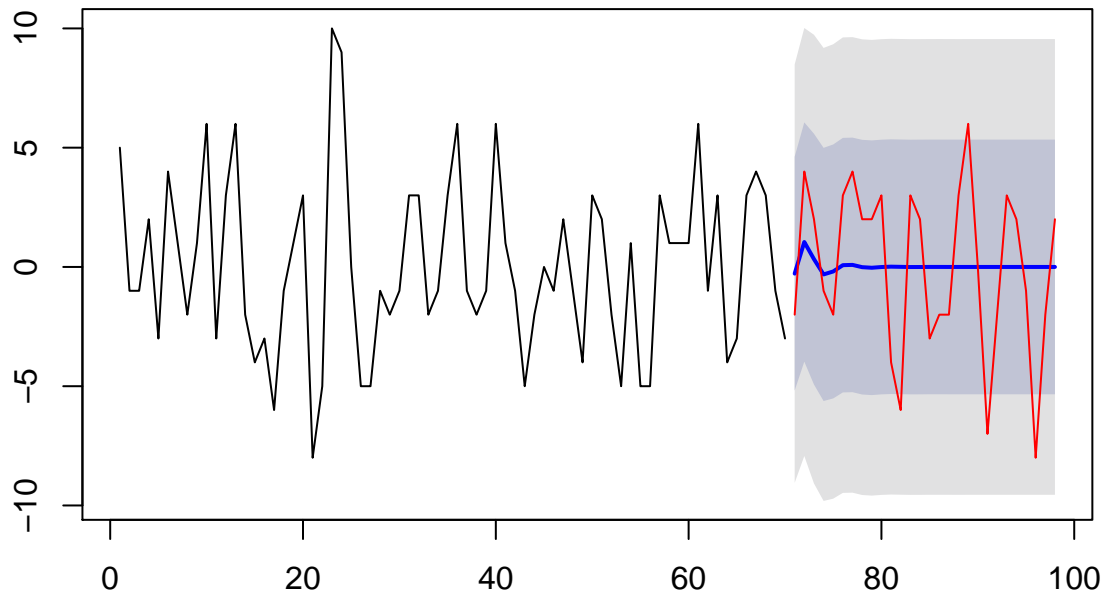
No hay evidencia de autocorrelación serial.

El pronóstico se

```
model1.train <- Arima(www.diff[1:70], order = c(2,0,0), include.mean = F)
in.sample <- forecast(model1.train, h = length(www.diff)-70, level = c(85, 99))
www.test = window(ts(www.diff), start = 71)
```

```
plot(in.sample)
lines(www.test, col="red")
```

Forecasts from ARIMA(2,0,0) with zero mean



El pronóstico parece razonable.

EL cálculo de los parámetros MAE , $MAPE$, se da a continuación:

```
(1/length(www.diff))*sum(abs((in.sample$mean - www.test))) # MAE
```

```
## [1] 0.824
```

```
(1/length(www.diff))*sum(abs((in.sample$mean - www.test)/www.test)) # MAPE
```

```
## [1] Inf
```

El cálculo del parámetro $MAPE$ se ve afectado por una división entre cero.

Ejemplo 3

En este tercer ejemplo, se consideran los *retornos mensuales* de Intel Corporation, (ticker INTC), bajados de Yahoo! Finance, desde el 1ero de marzo 1980 hasta el 1ero de febrero 2019. Más adelante analizaremos la volatilidad de estos retornos.

Se prueba que los retornos son estacionario mediante las pruebas de Dickey-Fuller, KPSS:

```
# INTC <- read.csv("/Users/laan/pCloud Sync/2019/EcoFin2/Data/INTC.csv") OSX
# INTC <- read.csv("/home/lino/pCloudDrive/pCloud Sync/2019/EcoFin2/Data/INTC.csv")
INTC <- read.csv("INTC.csv")
intc <- ts(INTC$Close, frequency = 12, start = c(1980, 03)) # frecuencia mensual
intc.rtn <- diff(log(intc))
adf.test(intc.rtn)
```

```
## Warning in adf.test(intc.rtn): p-value smaller than printed p-value
```

```
##
## Augmented Dickey-Fuller Test
##
## data: intc.rtn
## Dickey-Fuller = -8, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary
```

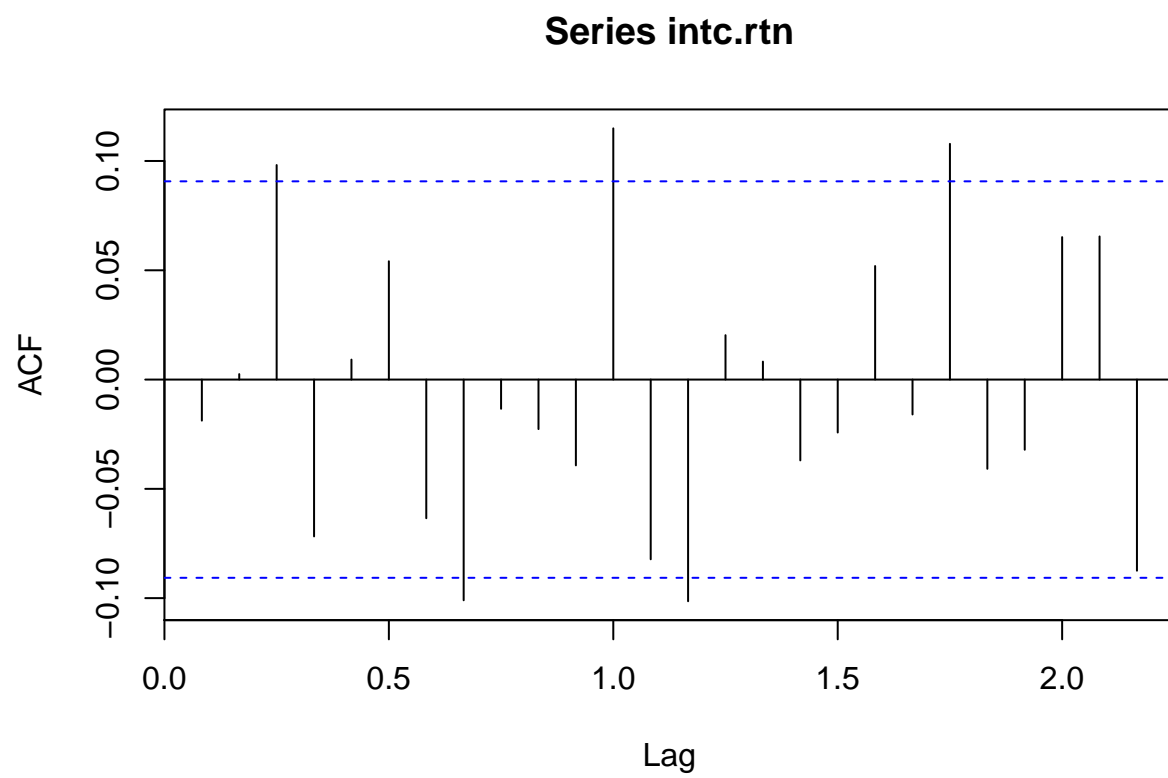
```
kpss.test(intc.rtn)
```

```
## Warning in kpss.test(intc.rtn): p-value greater than printed p-value
```

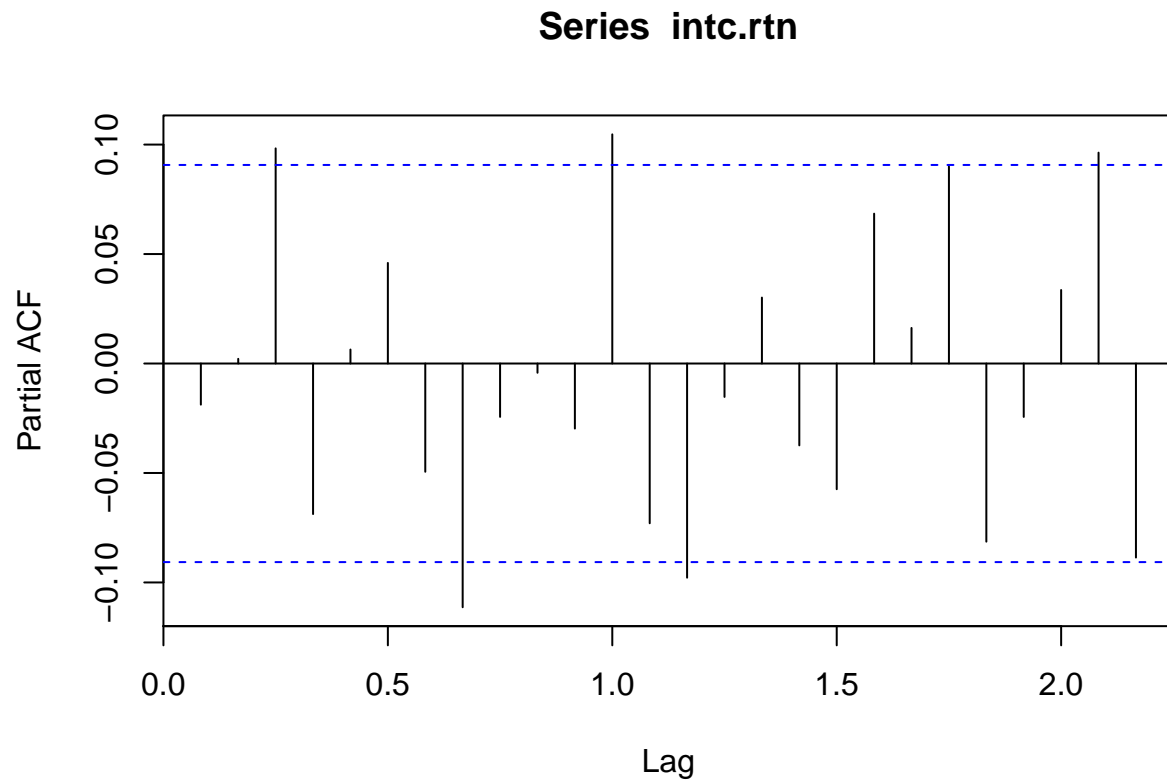
```
##
## KPSS Test for Level Stationarity
##
## data: intc.rtn
## KPSS Level = 0.2, Truncation lag parameter = 5, p-value = 0.1
```

Los correlogramas sugieren un modelo mixto $ARMA(p, q)$:

```
acf(intc.rtn)
```

```
pacf(intc.rtn)
```



Para determinar si incluir la media en la estimación del modelo, probaremos la hipótesis de media nula:

```
t.test(intc.rtn)
```

```
##
##  One Sample t-test
##
## data:  intc.rtn
## t = 2, df = 466, p-value = 0.03
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.00106 0.02059
## sample estimates:
## mean of x
##  0.0108
```

Hay evidencia que la media es diferente de cero, con valor estimado $\hat{\mu} = 0.011$.

Para la estimación usamos la función `auto.arima()`:

```
intc.fit <- auto.arima(intc.rtn, max.p = 10, max.q = 10, stationary = T, seasonal = F)
intc.fit
```

```
## Series: intc.rtn
```

```
## ARIMA(0,0,0) with non-zero mean
##
## Coefficients:
##      mean
##      0.011
## s.e.  0.005
##
## sigma^2 estimated as 0.0115:  log likelihood=380
## AIC=-756   AICc=-756   BIC=-748
```

que resulta en un modelo

$$intc.rtn = \mu + u_t.$$

(El log de los precios mensuales son una caminata aleatoria.)

El diagnóstico,

```
Box.test(intc.fit$residuals, lag=12)
```

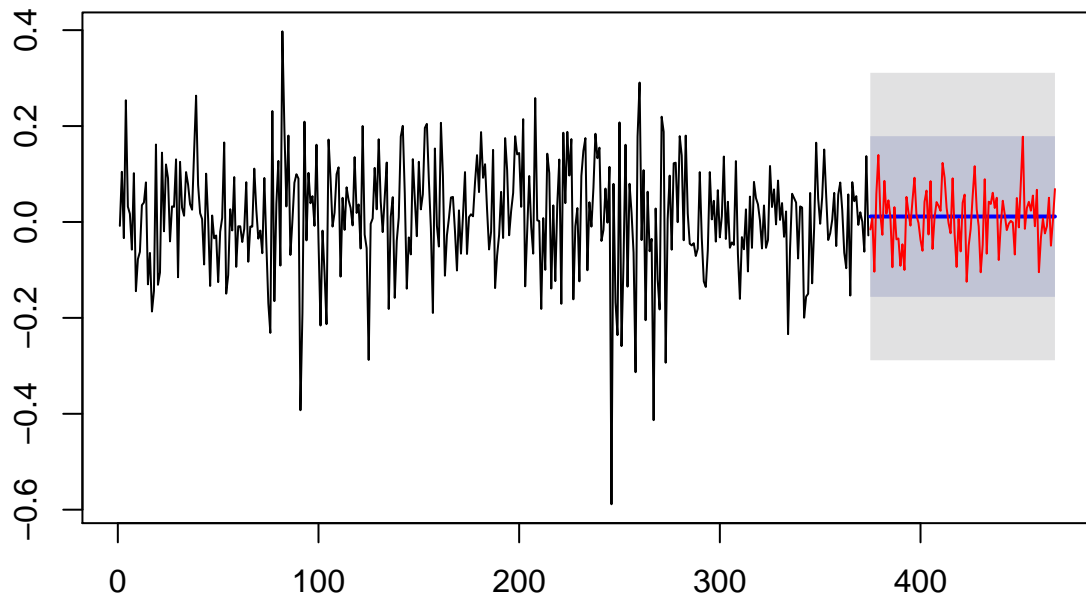
```
##
## Box-Pierce test
##
## data:  intc.fit$residuals
## X-squared = 22, df = 12, p-value = 0.03
```

todavía muestra cierta autocorrelación serial residual que no consideraremos.

Para el pronóstico, usamos el 80% de las observaciones para el pronóstico dentro de muestra y el remanente para el diagnóstico fuera de muestra.

```
intc.train <- Arima(intc.rtn[1:374], order = c(0,0,0), include.mean = T)
in.sample <- forecast(intc.train, h = 467-374, level = c(85, 99))
intc.test = window(ts(intc.rtn), start = 375)
plot(in.sample)
lines(intc.test, col="red")
```

Forecasts from ARIMA(0,0,0) with non-zero mean



El pronóstico parece bueno.

El cálculo de *MAE*, *MAPE* da

```
(1/length(intc.rtn))*sum(abs((in.sample$mean - intc.test))) # MAE
```

```
## [1] 0.00927
```

```
(1/length(intc.rtn))*sum(abs((in.sample$mean - intc.test)/intc.test)) # MAPE
```

```
## [1] 0.262
```

Heterocedasticidad condicional y volatilidad

Seguiremos, por lo general, el Cap. 4 del libro de R. Tsay [T].

La volatilidad se expresa, estadísticamente, en término de desviación estándar del activo bajo consideración.

Características de la volatilidad

Hay ciertas características de la volatilidad que se observan en los datos financieros.

- La volatilidad se agrupa: gran volatilidad en ciertos periodos de tiempo; baja volatilidad en otros.
- La volatilidad evolue de manera continua, por la mayor parte.
- Se observa un **efecto de apalancamiento**. La volatilidad reacciona de manera diferente a fuertes subida y a fuertes caídas de precios: el mayor impacto sobre la volatilidad se percibe durante las caídas de los precios.
 - Para modelar este efecto de apalancamiento, se desarrollaron los modelos EGARCH (Exponential Generalized AutoRegressive Heterocedastic), TGARCH (Threshold Generalized AutoRegressive Heterocedastic).

Medida de volatilidad

La volatilidad se puede medir de diferentes maneras:

1. La volatilidad de un activo se estima usando los precios del mismo, de sus derivados o de ambos.
2. La volatilidad se observa mediante
 - (a) retornos diarios del activo;
 - (b) datos de alta frecuencia (high-frequency data);
3. precios de opciones relacionadas al activo.

Estas diferentes manera de medición dan lugar a diferentes definiciones de volatilidad.

La definición que se usará es la *desviación estándar condicional de los retornos diarios*.

Otra medida de volatilidad es la *volatilidad implícita*, definida mediante la fórmula de Black-Scholes para opciones. El índice VIX del CBOE (Chicago Board Options Exchange) mide la volatilidad implícita. En México, la BMV, a través de MexDer, provee el índice de volatilidad VIMEX.

Para observar un efecto de volatilidad (efecto ARCH, autoregresivo, condicional heterocedástico) sobre un activo, se muestra que para los retornos del activo

1. la autocorrelación de los retornos es débil, o no existente;
2. los retornos no son independientes;
3. la media de la serie es diferente de cero.

La dependencia funcional típicamente se expresa mediante una función cuadrática (véase la modelación ARCH más abajo).

Para propósitos ilustrativos, seguiremos usando los retornos mensuales de Intel Corporation. Ya determinamos que la media de los retornos es diferente de cero.

Modelación ARCH

Un modelo $ARCH(M)$ consiste de una ecuación de media y de varianza como se da a continuación.

La *ecuación de la media* para los retornos, r_t , está dada por

$$\mu_t = E[r_t | \Omega_{t-1}],$$

típicamente un modelo $ARMA(p, q)$ con media.

La *ecuación de la varianza* mide el grado de dependencia estadística de los retornos:

$$\begin{aligned} u_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= Var(r_t | \Omega_{t-1}) = E[(r_t - \mu_t)^2 | \Omega_{t-1}] \\ &= \alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_m u_{t-m}^2, \end{aligned}$$

con $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \dots, m$.

En la ecuación de la varianza, el término ε_t es una serie de variables aleatorias independientes, idénticamente distribuidas, con media cero y varianza uno. Distribuciones usadas para ε_t son la normal estándar; la t -Student estandarizada; la GED (Generalized Error Distribution). En datos financieros las colas son, por lo general, más gruesas que las colas de una normal estándar.

Procedimiento

1. Estimar la ecuación de la media para los retornos; si un modelo $ARMA(0, 0)$, con media, todavía tiene dependencia serial residua, entonces usar un modelo $ARMA(p, q)$ apropiado. Empíricamente, la ecuación de la media, $\mu_t = E[r_t | \Omega_{t-1}]$ es sencilla: un modelo $ARMA(1, 1)$ (con media) en muchos casos es suficiente.
2. Usar los residuos de la ecuación de la media para detectar efectos ARCH: se puede usar la prueba de Ljung-Box sobre los residuales al cuadrado, por ejemplo; también la prueba ML (multiplicadores de Lagrange) de Engle. En la prueba de Engle, el estadístico de prueba es el estadístico F de significancia del modelo, con valor-p asociado.
3. Si se observan efectos ARCH, se puede usar la PACF de u_t^2 para determinar el orden m del modelo ARCH.
4. Una vez detectados efectos ARCH, se estima la ecuación de la media y de varianza simultáneamente. La estimación de un modelo ARCH se realiza maximizando la verosimilitud, mediante la distribución de ε_t .

Ejemplo

Usaremos los retornos mensuales r_t de Intel Corporation; ya vimos en el Ejemplo 3, Sección Pronóstico, que estos retornos satisfacen la ecuación

$$r_t = \mu + u_t.$$

Para determinar la existencia del efecto heterocedástico autoregresivo usamos la prueba de Ljung-Box sobre los residuales cuadrados:

```
Box.test((intc.rtn - mean(intc.rtn))^2, lag = 12)
```

```
##
## Box-Pierce test
##
## data: (intc.rtn - mean(intc.rtn))^2
## X-squared = 46, df = 12, p-value = 6e-06
```

Usaremos también el script `archTest.R` de R.~Tsay para usar la prueba ML de Engle:

```
source("/Users/laan/pCloud Sync/0/EcoFin2/Data/archTest.R") # OSX
# source("~/pCloudDrive/pCloud Sync/0/EcoFin2/Data/archTest.R")
# source("/home/lino/pCloudDrive/pCloud Sync/2019/EcoFin2/Data/archTest.R")
archTest(intc.rtn - mean(intc.rtn), 12)
```

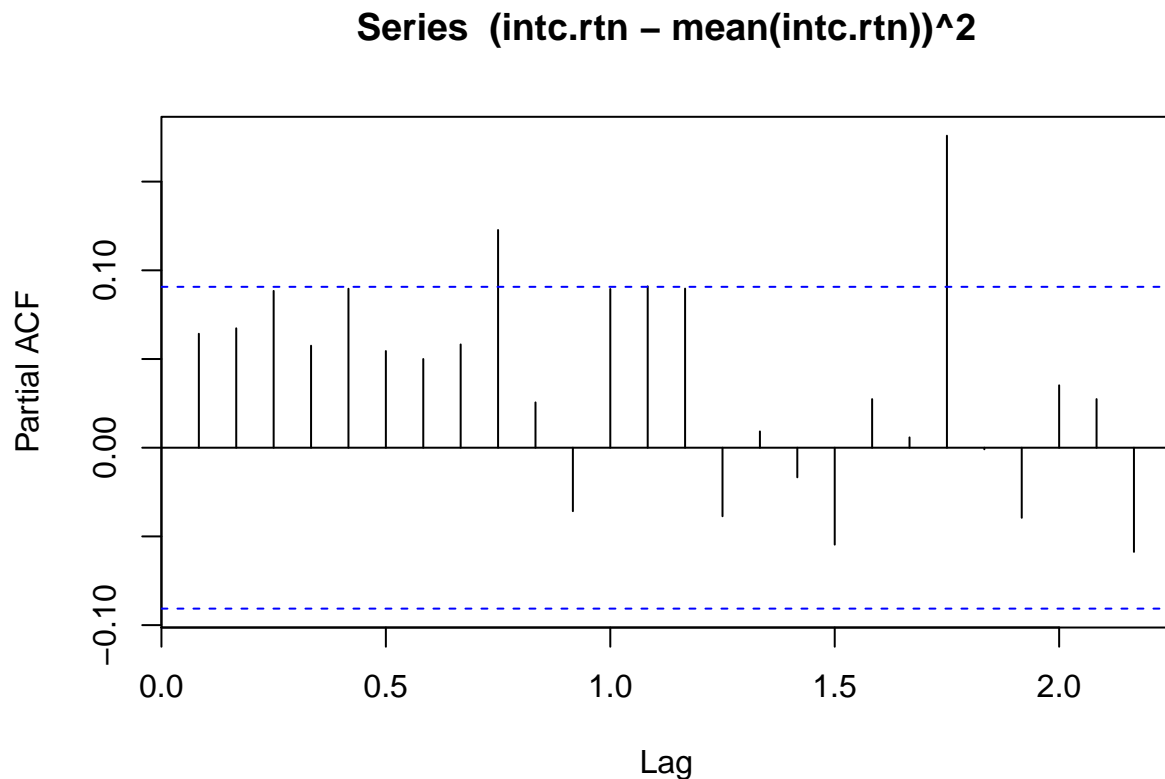
```
##
## Call:
## lm(formula = atsq ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.0537 -0.0084 -0.0056  0.0015  0.3489
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.00523    0.00178   2.94  0.0034 **
## x1           0.02926    0.04736   0.62  0.5370
## x2           0.03679    0.04735   0.78  0.4375
## x3           0.05532    0.04737   1.17  0.2434
## x4           0.02902    0.04710   0.62  0.5382
## x5           0.06771    0.04706   1.44  0.1509
## x6           0.03948    0.04713   0.84  0.4027
## x7           0.03667    0.04713   0.78  0.4370
## x8           0.05516    0.04704   1.17  0.2415
## x9           0.11746    0.04689   2.51  0.0126 *
## x10          0.02291    0.04716   0.49  0.6273
## x11         -0.03818    0.04714  -0.81  0.4185
```

```
## x12          0.08987    0.04715    1.91    0.0573 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0248 on 442 degrees of freedom
## Multiple R-squared:  0.0614, Adjusted R-squared:  0.0359
## F-statistic: 2.41 on 12 and 442 DF,  p-value: 0.00499
```

El valor-p de la prueba de Engle es = .004986.

Existe evidencia de efecto ARCH. Para estimar el orden del modelo $ARCH(m)$ usamos la PACF con los residuales

```
pacf((intc.rtn - mean(intc.rtn))^2)
```



Se considera que $m = 5$ puede ser adecuado.

Para la estimación del modelo $ARCH(5)$ usamos la función `garchFit()` de la librería `fGarch`.

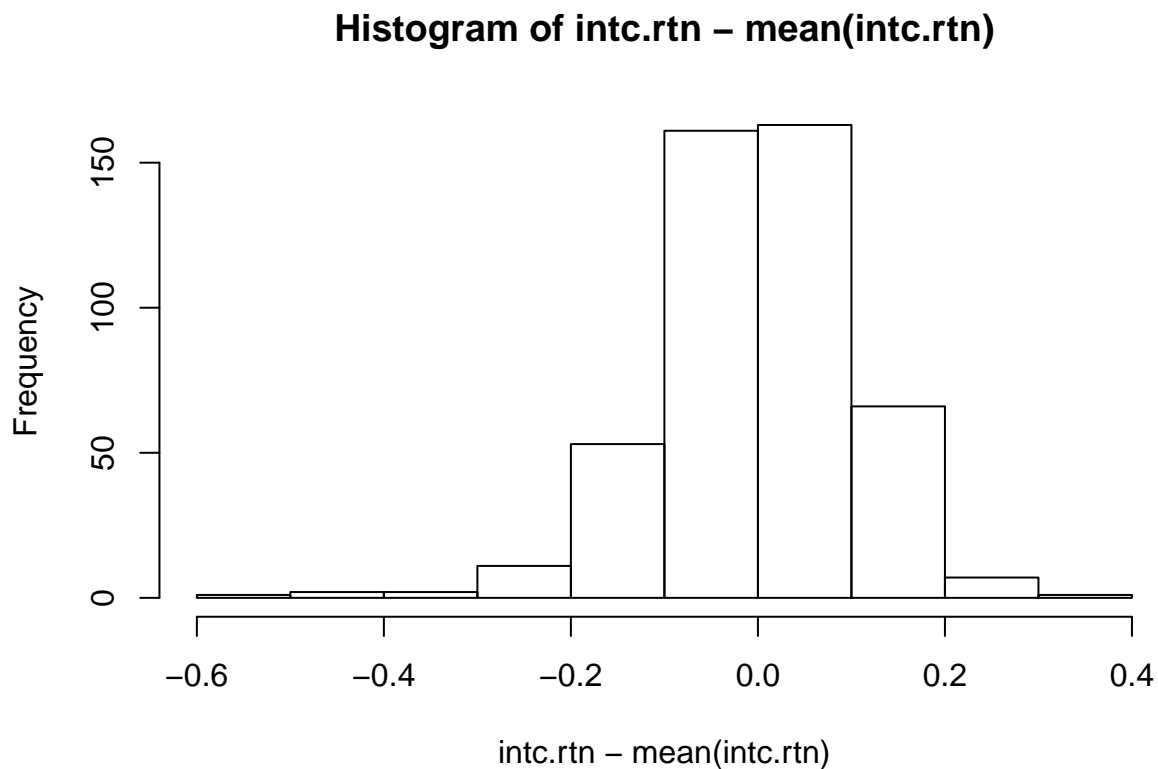
Nótese que el término α_0 en la ecuación de la varianza se denota con ω en la estimación con R.

Para determinar la distribución condicional, probaremos la hipótesis de normalidad de los errores:


```
shapiro.test((intc.rtn - mean(intc.rtn)))
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  (intc.rtn - mean(intc.rtn))  
## W = 1, p-value = 4e-08
```

```
hist(intc.rtn - mean(intc.rtn))
```



Los errores no son normales; usando la información del histograma, se usará una distribución condicional *t*-Student asimétrica.

```
library(fGarch)  
intc.archfit1 <- garchFit(~ 1 + garch(5,0),  
                        data = intc.rtn, trace = F, cond.dist = "sstd")  
intc.archfit1
```

```
##  
## Title:  
##  GARCH Modelling  
##  
## Call:  
##  garchFit(formula = ~1 + garch(5, 0), data = intc.rtn, cond.dist = "sstd",
```

```
##      trace = F)
##
## Mean and Variance Equation:
## data ~ 1 + garch(5, 0)
## <environment: 0x7f89b114c688>
## [data = intc.rtn]
##
## Conditional Distribution:
## sstd
##
## Coefficient(s):
##      mu      omega    alpha1    alpha2    alpha3    alpha4
## 0.0106938 0.0057569 0.1706409 0.0500237 0.0949283 0.1054819
##      alpha5      skew      shape
## 0.1076036 0.9253258 6.3448763
##
## Std. Errors:
## based on Hessian
##
## Error Analysis:
##      Estimate Std. Error t value Pr(>|t|)
## mu      0.010694    0.004534   2.358 0.018355 *
## omega   0.005757    0.001118   5.148 2.63e-07 ***
## alpha1  0.170641    0.089317   1.911 0.056069 .
## alpha2  0.050024    0.047918   1.044 0.296511
## alpha3  0.094928    0.058758   1.616 0.106183
## alpha4  0.105482    0.071994   1.465 0.142883
## alpha5  0.107604    0.067463   1.595 0.110713
## skew    0.925326    0.061410  15.068 < 2e-16 ***
## shape   6.344876    1.722128   3.684 0.000229 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
## 411      normalized: 0.88
##
## Description:
## Mon Nov 25 14:07:58 2019 by user:
```

Observamos que α_2 , α_3 , α_4 , α_5 no son significativos, por lo que consideramos que un modelo $ARCH(1)$ puede ser más apropiado:

```
intc.archfit2 <- garchFit(~ 1 + garch(1,0),
                        data = intc.rtn, trace = F, cond.dist = "sstd")
intc.archfit2
```

```
##
## Title:
## GARCH Modelling
##
## Call:
## garchFit(formula = ~1 + garch(1, 0), data = intc.rtn, cond.dist = "sstd",
## trace = F)
##
## Mean and Variance Equation:
## data ~ 1 + garch(1, 0)
## <environment: 0x7f89b3454008>
## [data = intc.rtn]
##
## Conditional Distribution:
## sstd
##
## Coefficient(s):
##      mu      omega    alpha1      skew      shape
## 0.0117473 0.0089556 0.2416264 0.9320647 6.3052580
##
## Std. Errors:
## based on Hessian
##
## Error Analysis:
##      Estimate Std. Error t value Pr(>|t|)
## mu      0.011747  0.004609   2.549 0.010807 *
## omega   0.008956  0.001084   8.260 2.22e-16 ***
## alpha1  0.241626  0.095949   2.518 0.011793 *
## skew    0.932065  0.058882  15.829 < 2e-16 ***
## shape   6.305258  1.752387   3.598 0.000321 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
## 402      normalized: 0.861
##
## Description:
## Mon Nov 25 14:07:58 2019 by user:
```

El modelo estimado es

$$\hat{\mu} = .01174 + u_t$$

$$\hat{\sigma}_t^2 = .00895 + .24162u_{t-1}^2$$

y cada término es significativo. El retorno esperado estimado es $\hat{\mu} = .01174 = 1.174\%$.

Diagnóstico No es necesario usar pruebas de normalidad ya que se modeló con innovaciones no normales. El diagnóstico en este caso consiste en verificar que no hay heterocedasticidad residua.

```
sresi=intc.archfit2@residuals/intc.archfit2@sigma.t
Box.test(sresi,lag=12,type='Ljung')
```

```
##
## Box-Ljung test
##
## data: sresi
## X-squared = 18, df = 12, p-value = 0.1
```

```
archTest(sresi,12)
```

```
##
## Call:
## lm(formula = atsqr ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.050 -0.778 -0.520  0.166 30.235
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.5551     0.1722   3.22  0.0014 **
## x1            -0.0547     0.0474  -1.15  0.2489
## x2             0.0258     0.0474   0.54  0.5867
## x3             0.0466     0.0474   0.98  0.3256
## x4             0.0233     0.0470   0.50  0.6194
## x5             0.0367     0.0469   0.78  0.4346
## x6             0.0392     0.0469   0.84  0.4040
## x7             0.0389     0.0469   0.83  0.4078
## x8             0.0595     0.0469   1.27  0.2056
## x9             0.1394     0.0467   2.98  0.0030 **
## x10            0.0425     0.0471   0.90  0.3674
## x11            -0.0351     0.0472  -0.74  0.4574
## x12            0.0851     0.0471   1.81  0.0713 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.23 on 442 degrees of freedom
## Multiple R-squared:  0.0503, Adjusted R-squared:  0.0245
## F-statistic: 1.95 on 12 and 442 DF, p-value: 0.0273
```

Hay poca evidencia de heterocedasticidad residua. Para eliminar el efecto ARCH residuo, se considera un modelo $ARCH(2)$.

```

intc.archfit3 <- garchFit(~ 1 + garch(2,0),
                        data = intc.rtn, trace = F, cond.dist = "sstd")
intc.archfit3

##
## Title:
##  GARCH Modelling
##
## Call:
##  garchFit(formula = ~1 + garch(2, 0), data = intc.rtn, cond.dist = "sstd",
##    trace = F)
##
## Mean and Variance Equation:
##  data ~ 1 + garch(2, 0)
## <environment: 0x7f89949bedd0>
##  [data = intc.rtn]
##
## Conditional Distribution:
##  sstd
##
## Coefficient(s):
##      mu      omega    alpha1    alpha2      skew      shape
## 0.0119997 0.0083835 0.2079315 0.0790665 0.9313925 6.2019613
##
## Std. Errors:
##  based on Hessian
##
## Error Analysis:
##      Estimate Std. Error t value Pr(>|t|)
## mu      0.012000   0.004617   2.599 0.009346 **
## omega   0.008384   0.001095   7.659 1.89e-14 ***
## alpha1  0.207931   0.096430   2.156 0.031061 *
## alpha2  0.079066   0.061753   1.280 0.200415
## skew    0.931392   0.060047  15.511 < 2e-16 ***
## shape   6.201961   1.677563   3.697 0.000218 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
## 404      normalized: 0.865
##
## Description:
##  Mon Nov 25 14:07:58 2019 by user:

```

```
sresi3=intc.archfit3$residuals/intc.archfit3$sigma.t
Box.test(sresi3,lag=12,type='Ljung')
```

```
##
## Box-Ljung test
##
## data: sresi3
## X-squared = 17, df = 12, p-value = 0.2
```

```
archTest(sresi3,12)
```

```
##
## Call:
## lm(formula = atsqr ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.13  -0.80  -0.55   0.21  32.88
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.6161    0.1834    3.36 0.00085 ***
## x1            -0.0433    0.0473   -0.91 0.36098
## x2            -0.0333    0.0473   -0.70 0.48213
## x3             0.0425    0.0473    0.90 0.36927
## x4             0.0230    0.0469    0.49 0.62452
## x5             0.0353    0.0469    0.75 0.45202
## x6             0.0308    0.0469    0.66 0.51164
## x7             0.0306    0.0469    0.65 0.51412
## x8             0.0642    0.0469    1.37 0.17169
## x9             0.1349    0.0467    2.89 0.00407 **
## x10            0.0361    0.0471    0.77 0.44432
## x11            -0.0294    0.0471   -0.63 0.53226
## x12            0.0992    0.0471    2.11 0.03565 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.34 on 442 degrees of freedom
## Multiple R-squared:  0.0461, Adjusted R-squared:  0.0202
## F-statistic: 1.78 on 12 and 442 DF, p-value: 0.0491
```

El modelo $ARCH(2)$ elimina el efecto ARCH residuo. El retorno esperado estimado es un poco superior al anterior.

Pronóstico

El pronóstico de un modelo ARCH no es muy diferente del pronóstico de un modelo AR. Los detalles se encuentran en el libro de R. Tsay, Sección 4.5.3. Consideramos, para propósitos explicativos, el caso $ARCH(1)$; supón que, el momento $h = 4$, se quiere pronosticar $\sigma_{h=4}^2(1)$, el valor de la volatilidad 1 periodo en adelante. Entonces,

$$\sigma_{h=4}^2(1) = \alpha_0 + \alpha_1 u_{4+1-1}^2 = \alpha_0 + \alpha_1 u_4^2;$$

el pronóstico 2 periodos en adelante,

$$\sigma_{h=4}^2(2) = \alpha_0 + \alpha_1 \sigma_4^2(1)$$

y así, sucesivamente. Véase la fórmula (4.11) en el libro de Tsay.

Se presentarán diferentes situaciones del pronóstico del modelo estimado anteriormente $ARCH(1)$, 10 periodos (meses) en el futuro.

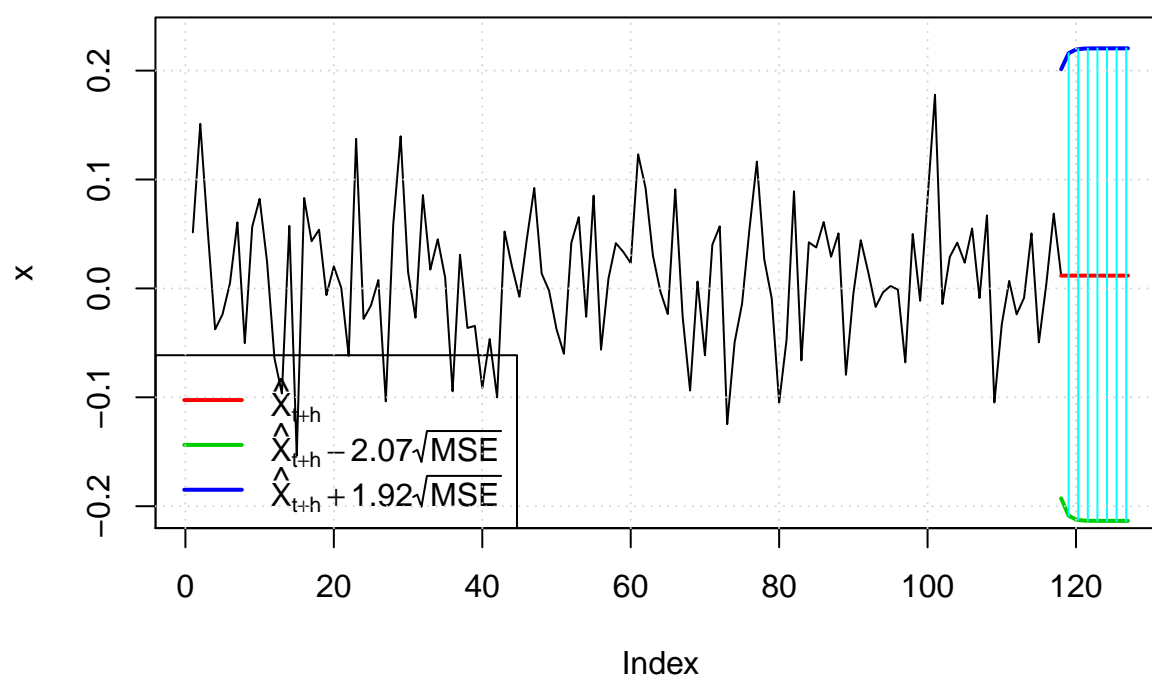
```
predict(intc.archfit2, n.ahead = 10)
```

##	meanForecast	meanError	standardDeviation
## 1	0.0117	0.0987	0.0987
## 2	0.0117	0.1063	0.1063
## 3	0.0117	0.1081	0.1081
## 4	0.0117	0.1085	0.1085
## 5	0.0117	0.1086	0.1086
## 6	0.0117	0.1087	0.1087
## 7	0.0117	0.1087	0.1087
## 8	0.0117	0.1087	0.1087
## 9	0.0117	0.1087	0.1087
## 10	0.0117	0.1087	0.1087

Visualización con ploteo:

```
predict(intc.archfit2, n.ahead = 10, plot = TRUE)
```

Prediction with confidence intervals

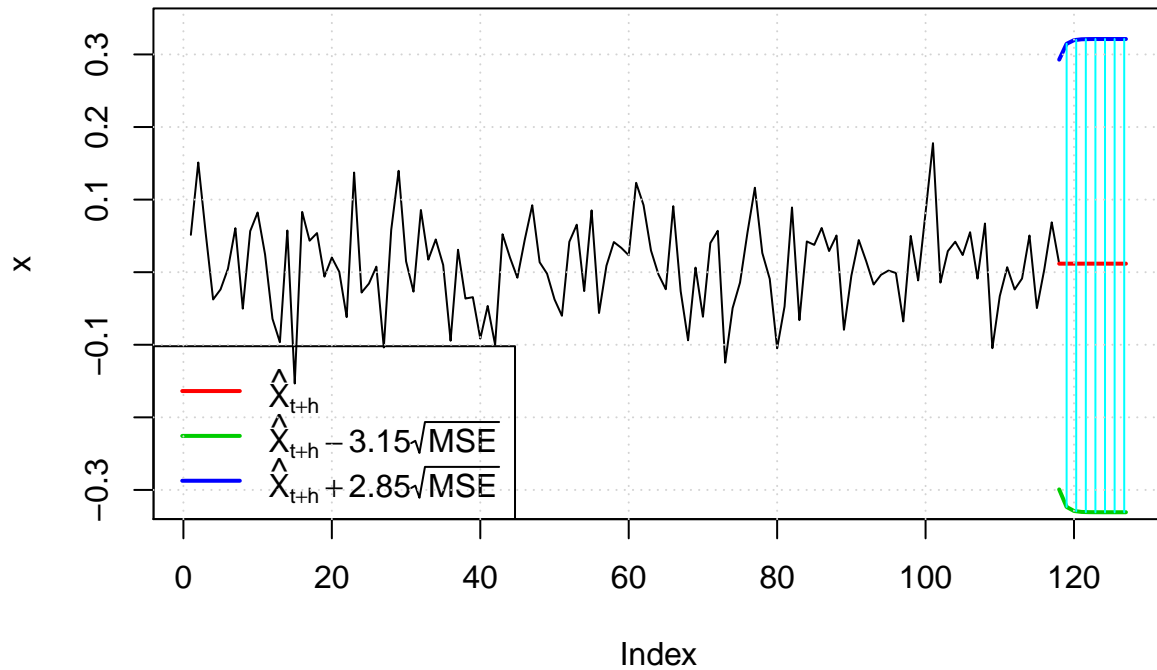


##	meanForecast	meanError	standardDeviation	lowerInterval	upperInterval
## 1	0.0117	0.0987	0.0987	-0.193	0.201
## 2	0.0117	0.1063	0.1063	-0.209	0.216
## 3	0.0117	0.1081	0.1081	-0.212	0.219
## 4	0.0117	0.1085	0.1085	-0.213	0.220
## 5	0.0117	0.1086	0.1086	-0.213	0.220
## 6	0.0117	0.1087	0.1087	-0.213	0.220
## 7	0.0117	0.1087	0.1087	-0.213	0.220
## 8	0.0117	0.1087	0.1087	-0.213	0.220
## 9	0.0117	0.1087	0.1087	-0.213	0.220
## 10	0.0117	0.1087	0.1087	-0.213	0.220

Visualización con ploteo e intervalo de confianza al 99%.

```
predict(intc.archfit2, n.ahead = 10, plot = TRUE, conf=.99)
```


Prediction with confidence intervals



##	meanForecast	meanError	standardDeviation	lowerInterval	upperInterval
## 1	0.0117	0.0987	0.0987	-0.299	0.293
## 2	0.0117	0.1063	0.1063	-0.323	0.315
## 3	0.0117	0.1081	0.1081	-0.329	0.320
## 4	0.0117	0.1085	0.1085	-0.330	0.321
## 5	0.0117	0.1086	0.1086	-0.331	0.321
## 6	0.0117	0.1087	0.1087	-0.331	0.321
## 7	0.0117	0.1087	0.1087	-0.331	0.321
## 8	0.0117	0.1087	0.1087	-0.331	0.321
## 9	0.0117	0.1087	0.1087	-0.331	0.321
## 10	0.0117	0.1087	0.1087	-0.331	0.321

Modelo ARCH–Innovaciones no gaussiana

La librería `fGarch` puede manejar innovaciones que siguen la distribución t -Student; t -Student asimétrica; GDE, y GDE asimétrica.

A continuación se presenta el modelo de volatilidad presentado arriba,

$$\begin{aligned} \text{intc.returns} &= \mu + u_t \\ u_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_m u_{t-m}^2, \end{aligned}$$

usando algunas de estas innovaciones no gaussianas.

Cada una de estas opciones se pasa usando el parámetro `cond.dist=` con la abreviación correspondiente: `std`, `sstd`, `ged`, `sged`, respectivamente. Otra posibilidad es usar QMLE (Quasi-Maximum Likelihood Estimation), que supone distribución normal, con errores estándares robustos.

Se pueden también pronosticar valores futuros usando estas distribuciones.

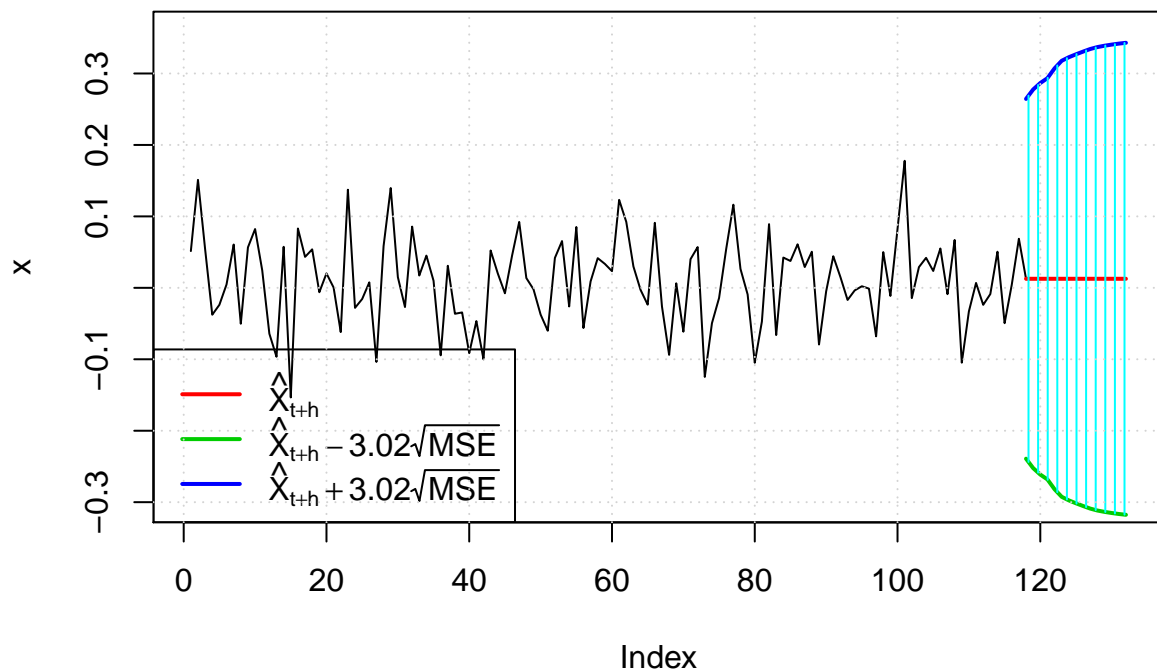
Innovación t-Student

```
intc.archfit.t <- garchFit(~ 1 + garch(5,0), data = intc.rtn, trace = F, cond.dist = "std",
intc.archfit.t

##
## Title:
##  GARCH Modelling
##
## Call:
##  garchFit(formula = ~1 + garch(5, 0), data = intc.rtn, cond.dist = "std",
##    trace = F)
##
## Mean and Variance Equation:
##  data ~ 1 + garch(5, 0)
## <environment: 0x7f89b2f12ed8>
##  [data = intc.rtn]
##
## Conditional Distribution:
##  std
##
## Coefficient(s):
##      mu      omega    alpha1    alpha2    alpha3    alpha4
## 0.0126420 0.0058145 0.1684766 0.0523688 0.1031607 0.0961664
##   alpha5    shape
## 0.1059516 6.0413008
##
## Std. Errors:
##  based on Hessian
##
```

```
## Error Analysis:
##      Estimate Std. Error  t value Pr(>|t|)
## mu      0.012642   0.004237   2.984  0.00284 **
## omega    0.005814   0.001137   5.115 3.14e-07 ***
## alpha1   0.168477   0.090190   1.868  0.06176 .
## alpha2   0.052369   0.047835   1.095  0.27361
## alpha3   0.103161   0.060639   1.701  0.08890 .
## alpha4   0.096166   0.071017   1.354  0.17570
## alpha5   0.105952   0.067763   1.564  0.11792
## shape    6.041301   1.533817   3.939 8.19e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
## 410      normalized:  0.879
##
## Description:
## Mon Nov 25 14:07:59 2019 by user:
predict(intc.archfit.t, n.ahead = 15, plot = TRUE, cond.dist = "std", conf = .99)
```

Prediction with confidence intervals



```
##      meanForecast meanError standardDeviation lowerInterval upperInterval
```

## 1	0.0126	0.0833	0.0833	-0.239	0.264
## 2	0.0126	0.0876	0.0876	-0.252	0.278
## 3	0.0126	0.0906	0.0906	-0.261	0.287
## 4	0.0126	0.0929	0.0929	-0.268	0.293
## 5	0.0126	0.0974	0.0974	-0.282	0.307
## 6	0.0126	0.1009	0.1009	-0.293	0.318
## 7	0.0126	0.1026	0.1026	-0.297	0.323
## 8	0.0126	0.1039	0.1039	-0.302	0.327
## 9	0.0126	0.1052	0.1052	-0.305	0.331
## 10	0.0126	0.1064	0.1064	-0.309	0.334
## 11	0.0126	0.1073	0.1073	-0.312	0.337
## 12	0.0126	0.1079	0.1079	-0.314	0.339
## 13	0.0126	0.1084	0.1084	-0.315	0.340
## 14	0.0126	0.1089	0.1089	-0.317	0.342
## 15	0.0126	0.1092	0.1092	-0.318	0.343

Para el diagnóstico y el pronóstico, se procede como arriba.

Modelos GARCH

La modelación ARCH puede necesitar órdenes grandes; la modelación *Generalized AutoRegressive Conditional Heteroskedasticity*, GARCH, introducida por Bollerslev (1986), no necesita, por lo general, órdenes grandes, aún si comparte algunas desventajas con la modelación ARCH.

Si $r_t = \text{intc.returns}$ es la serie de retornos, sea

$$a_t = r_t + \mu_t$$

la innovación al tiempo t .

La serie u_t sigue un modelo $GARCH(m, s)$ si

$$u_t = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i u_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2.$$

donde ε_t es una secuencia de variables aleatorias, con la misma distribución, con media 0 y varianza 1.

Una característica importante de la modelación GARCH es que la ecuación de la varianza condicional, σ_t^2 , se puede representar mediante un modelo ARMA para u_t^2 , más una función de una *martingala* (ruido blanco).

A continuación, se presentará el caso $GARCH(1, 1)$ por simplicidad de exposición:

$$u_t = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

Fortalezas y debilidades de la modelación GARCH

Las ventajas y desventajas son semejantes a las de la modelación ARCH.

Ventajas

- Se producen agrupaciones de volatilidad.
- las colas son más gruesas de una normal.
- Los pronósticos, $\sigma_h^2(\ell)$, se obtienen como los pronósticos de la modelación $ARMA(p, q)$.
A la larga,

$$\lim_{\ell \rightarrow \infty} \sigma_h^2(\ell) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1} = \text{Var}(u_t).$$

Desventajas

- El modelo $GARCH(1, 1)$ necesita la condición $\alpha_1 + \beta_1 < 1$, para tener varianza finita y positiva.
- No se puede modelar el efecto de apalancamiento.
- En estudios empíricos de volatilidad de series de tiempo con alta frecuencia, la curtosis de los modelos GARCH no parece coincidir con las observaciones.

Estimación GARCH de INTC

A diferencia de la modelación ARCH, no existe evidencia en la determinación del orden de un modelo GARCH. Por lo general, se usan modelos $GARCH(1, 1)$, $GARCH(1, 2)$, $GARCH(2, 1)$ en los estudios empíricos.

A continuación se presenta la estimación $GARCH(1, 1)$ de *intc.returns*, retornos de INTC mensual. Después, se llevará a cabo el diagnóstico, complementando las herramientas usadas en la sección anterior con

- un intervalo predictivo alrededor de la ecuación dinámica de media;
- la serie de tiempo de la volatilidad, usando *volatility* de la librería *fBasics*, cargada por defecto cuando se carga *fGarch*.

```
intc.garchfit <- garchFit(~ 1 + garch(1,1), data = intc.rtn, trace = F)
summary(intc.garchfit)
```

```
##
## Title:
##  GARCH Modelling
##
## Call:
##  garchFit(formula = ~1 + garch(1, 1), data = intc.rtn, trace = F)
##
```

```

## Mean and Variance Equation:
## data ~ 1 + garch(1, 1)
## <environment: 0x7f89aa572fb8>
## [data = intc.rtn]
##
## Conditional Distribution:
## norm
##
## Coefficient(s):
##      mu      omega      alpha1      beta1
## 8.2857e-03  7.1133e-05  5.7384e-02  9.3667e-01
##
## Std. Errors:
## based on Hessian
##
## Error Analysis:
##      Estimate  Std. Error  t value Pr(>|t|)
## mu      8.286e-03   4.212e-03   1.967 0.049140 *
## omega   7.113e-05   8.881e-05   0.801 0.423154
## alpha1  5.738e-02   1.664e-02   3.448 0.000565 ***
## beta1   9.367e-01   1.946e-02  48.125 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
## 409      normalized:  0.875
##
## Description:
## Mon Nov 25 14:07:59 2019 by user:
##
##
## Standardised Residuals Tests:
##
##      Jarque-Bera Test  R      Chi^2  112      0
##      Shapiro-Wilk Test  R      W      0.98      5.38e-06
##      Ljung-Box Test    R      Q(10)  9.23      0.51
##      Ljung-Box Test    R      Q(15)  14.8      0.464
##      Ljung-Box Test    R      Q(20)  17.8      0.599
##      Ljung-Box Test    R^2  Q(10)  3.1      0.979
##      Ljung-Box Test    R^2  Q(15)  6.01     0.98
##      Ljung-Box Test    R^2  Q(20)  7.32     0.995
##      LM Arch Test      R      TR^2  4.71     0.967
##
## Information Criterion Statistics:
##      AIC      BIC      SIC      HQIC

```

```
## -1.73 -1.70 -1.73 -1.72
```

El modelo estimado es (innovaciones gaussianas)

$$r_t = .0082492 + u_t$$
$$\sigma_t^2 = .0001072 + .0587060u_{t-1}^2 + .9324359\sigma_{t-1}^2$$

Vale la pena subrayar que α_0 no es significativo en este modelo; el retorno promedio estimado es al límite de la significancia estadística, con valor-p asociado igual a .052853.

El pronóstico a la larga estimado es

$$\lim_{\ell \rightarrow \infty} \sigma_h^2(\ell) = \frac{.0001072}{1 - .0587060 - .9324359} = 0.01210191802.$$

La varianza muestral es comparable al valor asintótico estimado:

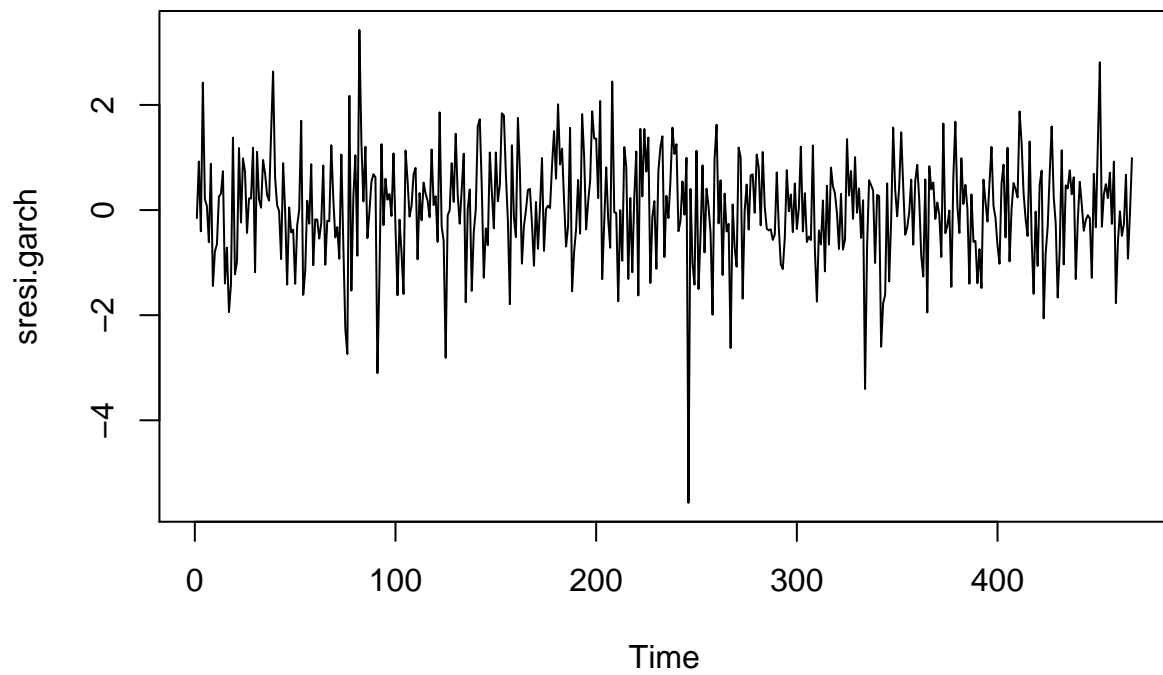
```
var(intc.rtn)
```

```
## [1] 0.0115
```

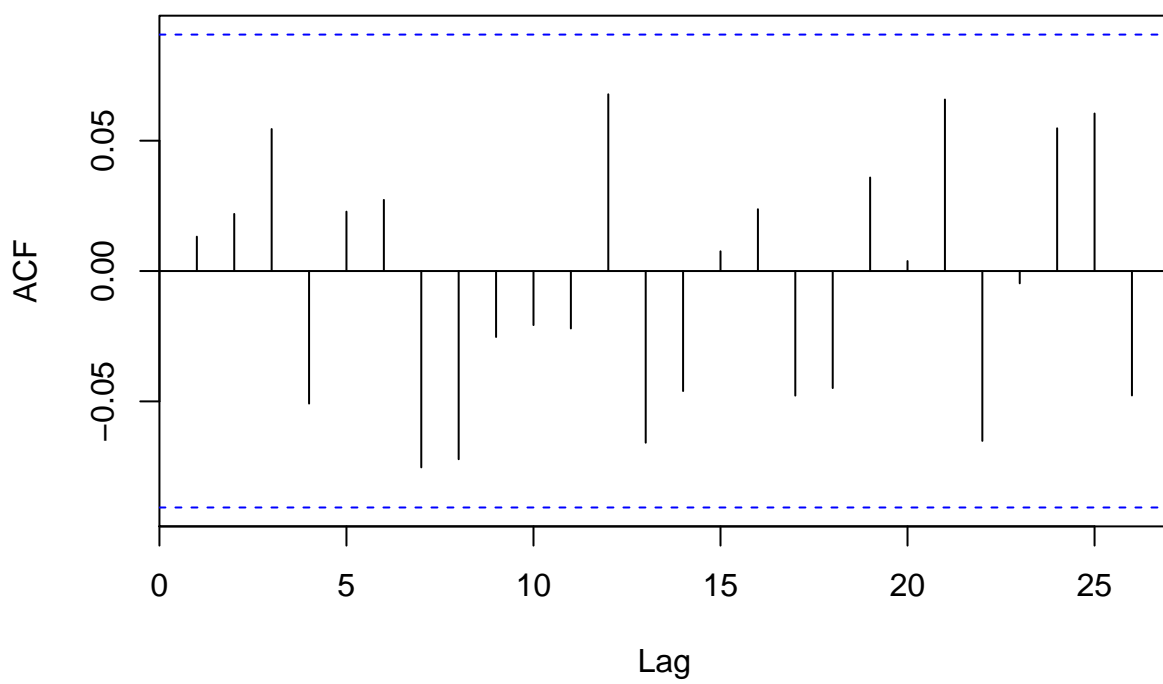
Diagnóstico

Las pruebas de hipótesis a la base del diagnóstico son las mismas que vimos en la modelación ARCH.

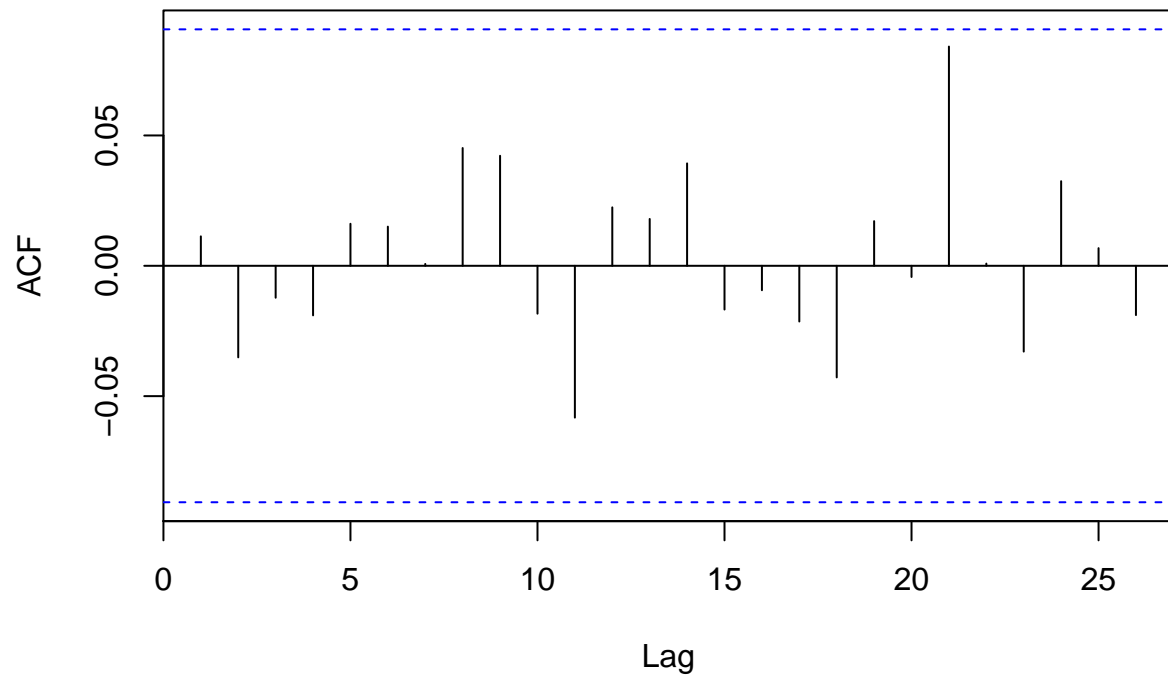
ACF, PACF



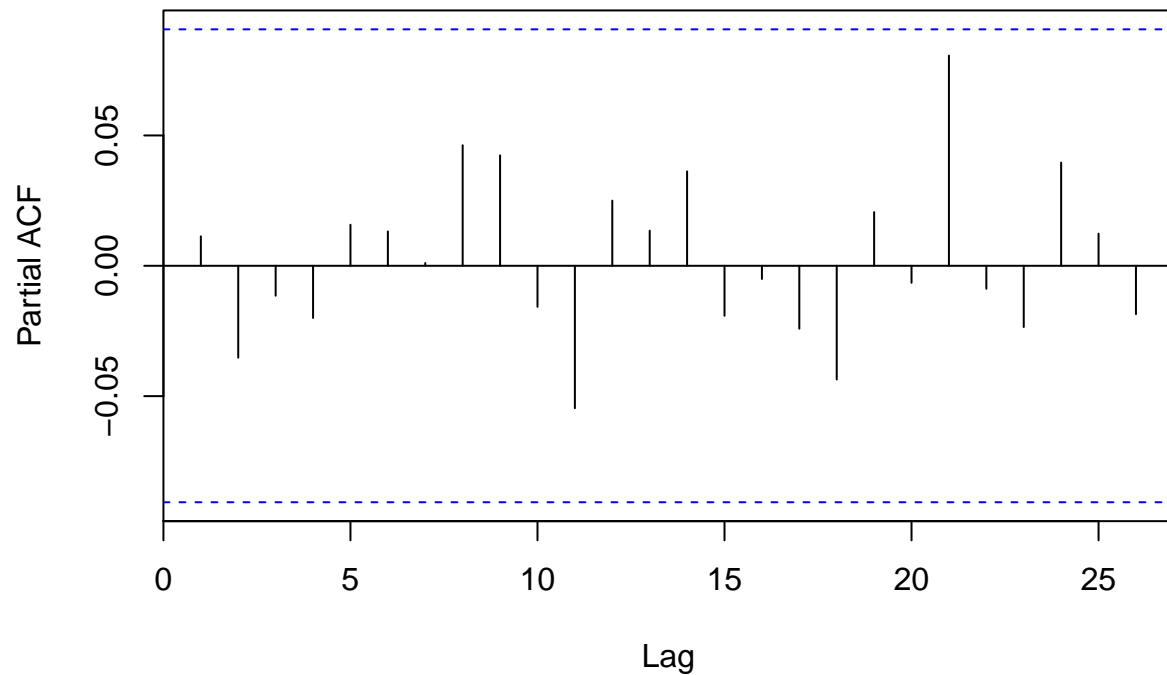
Series sresi.garch



Series sresi.garch^2



Series sresi.garch^2



Ljung-Box, lag = 10

```
Box.test(sresi.garch,lag=10,type='Ljung')
```

```
##
##  Box-Ljung test
##
## data:  sresi.garch
## X-squared = 9, df = 10, p-value = 0.5
```

```
Box.test(sresi.garch^2,lag=10,type='Ljung')
```

```
##
##  Box-Ljung test
##
## data:  sresi.garch^2
## X-squared = 3, df = 10, p-value = 1
```

Ljung-Box, lag = 20

```
Box.test(sresi.garch,lag=20,type='Ljung')

##
##  Box-Ljung test
##
## data:  sresi.garch
## X-squared = 18, df = 20, p-value = 0.6

Box.test(sresi.garch^2,lag=20,type='Ljung')

##
##  Box-Ljung test
##
## data:  sresi.garch^2
## X-squared = 7, df = 20, p-value = 1
```

Ljung-Box, lag=30

```
Box.test(sresi.garch,lag=30,type='Ljung')

##
##  Box-Ljung test
##
## data:  sresi.garch
## X-squared = 27, df = 30, p-value = 0.6

Box.test(sresi.garch^2,lag=30,type='Ljung')

##
##  Box-Ljung test
##
## data:  sresi.garch^2
## X-squared = 13, df = 30, p-value = 1
```

Normalidad de las innovaciones

Una implementación de la prueba de Jarque-Bera se encuentra en el paquete **fBasics** cargado simultáneamente con la librería **fGarch**, mientras que la prueba de Shapiro-Wilk se encuentra en la distribución **stats**, cargada por defecto.

```
jarqueberaTest(sresi.garch)
```

```
##
```

```
## Title:
##  Jarque - Bera Normalality Test
##
## Test Results:
##  STATISTIC:
##    X-squared: 112.054
##  P VALUE:
##    Asymptotic p Value: < 2.2e-16
##
## Description:
##  Mon Nov 25 14:07:59 2019 by user:
```

```
shapiroTest(sresi.garch)
```

```
##
## Title:
##  Shapiro - Wilk Normality Test
##
## Test Results:
##  STATISTIC:
##    W: 0.9802
##  P VALUE:
##    5.381e-06
##
## Description:
##  Mon Nov 25 14:07:59 2019 by user:
```

El modelo parece ser adecuado, entre los modelos de orden bajo; las innovaciones gaussianas no son las correctas.

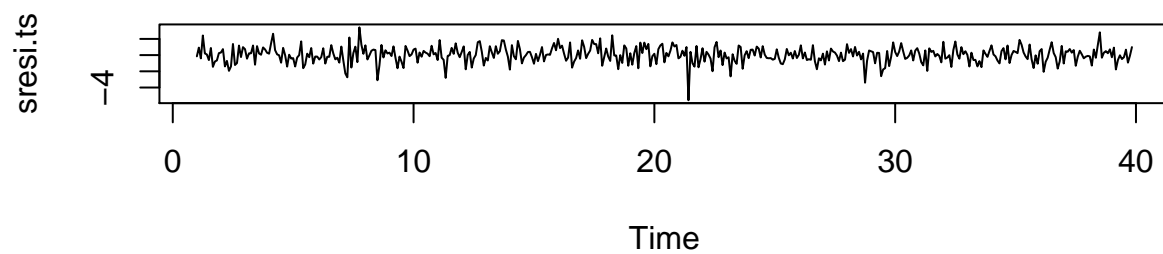
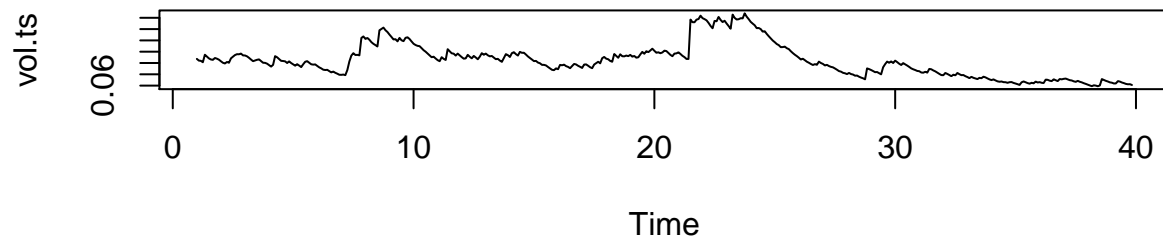
Intervalo predictivo alrededor de la media

En esta parte construiremos un intervalo de confianza, al 95%, alrededor del promedio estimado de los retornos.

Primeramente, se definirá la *serie de volatilidad* del modelo estimado.

A continuación se encuentra el código requerido:

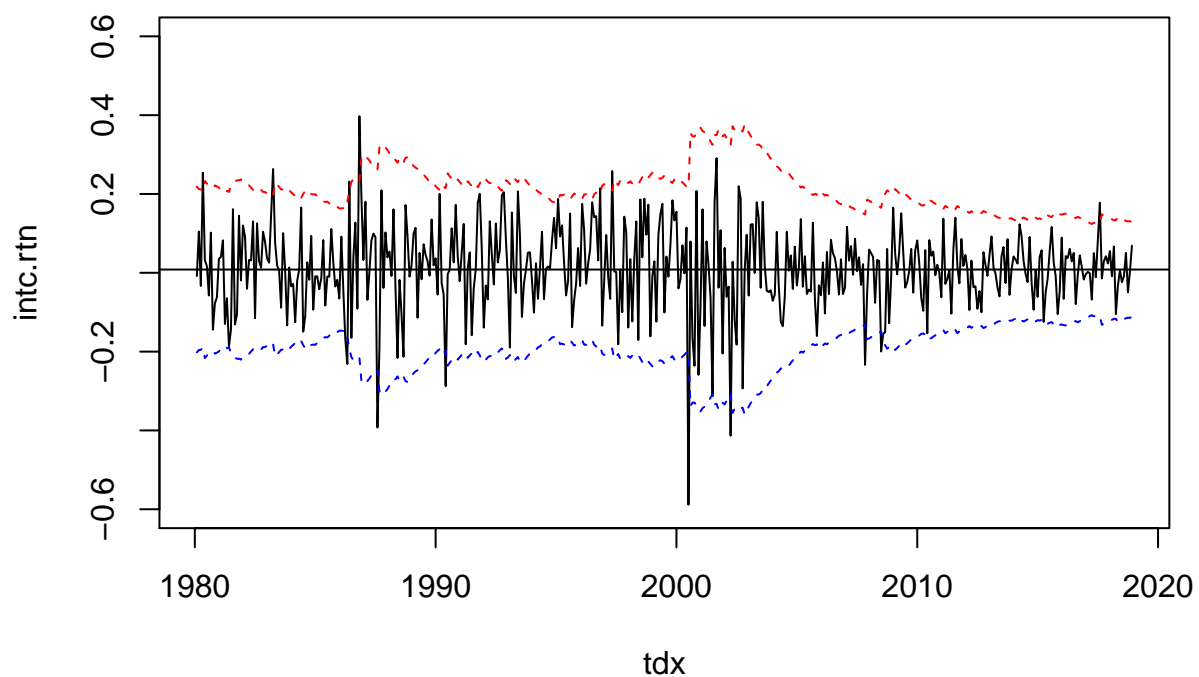
```
v1 <- volatility(intc.garchfit) # volatilidad
sresi <- residuals(intc.garchfit, standardize = T) # residuos estandarizados
vol.ts <- ts(v1, frequency = 12) # volatilidad como serie de tiempo
sresi.ts <- ts(sresi, frequency = 12) # residuos como serie de tiempo
par(mfrow = c(2,1)) # dos gráficas al mismo tiempo
plot.ts(vol.ts)
plot.ts(sresi.ts)
```



```
par(mfrow = c(1,1)) # regresar al default
```

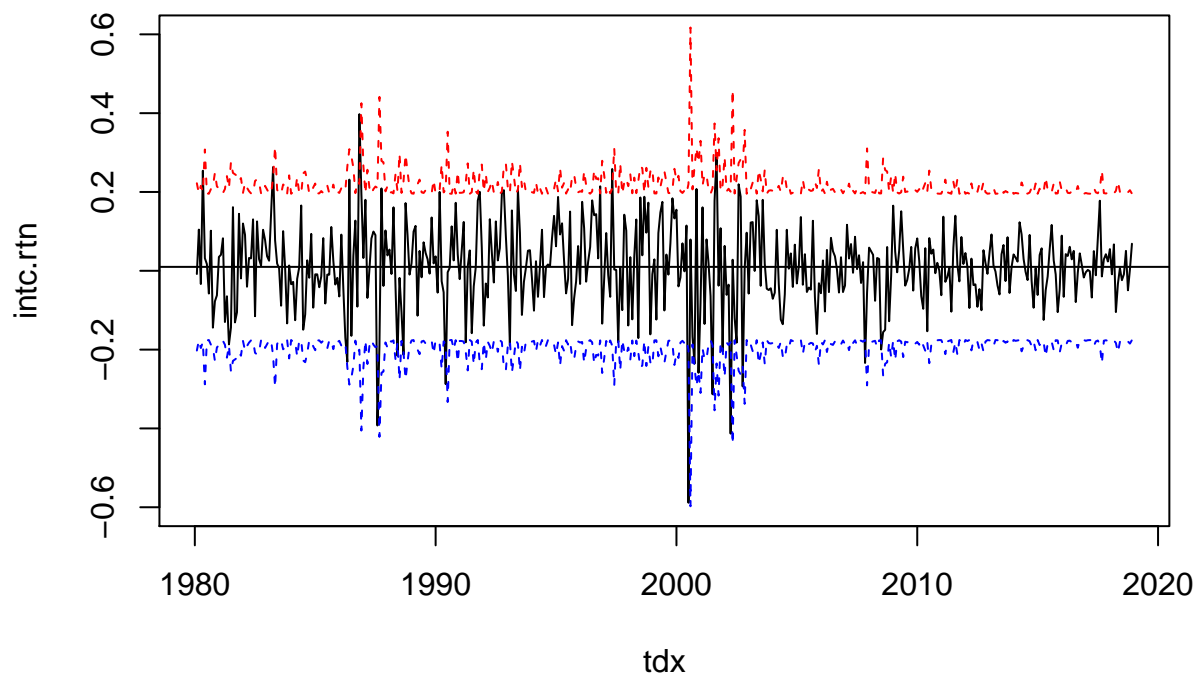
Después, se define el intervalo de confianza:

```
upper <- .0082492 + 1.96*v1
lower <- .0082492 - 1.96*v1
tdx <- c(1:467)/12+1980
plot(tdx,intc.rtn, type = "l", ylim = c(-.6,.6))
lines(tdx,upper, lty=2, col = "red")
lines(tdx,lower, lty=2, col = "blue")
abline(h=c(.0082492))
```



A comparación, se presenta el intervalo predictivo para el modelo $ARCH(1)$:

```
v2 <- volatility(intc.archfit2) # volatilidad
upper2 <- .009915 + 1.96*v2
lower2 <- .009915 - 1.96*v2
tdx <- c(1:467)/12+1980
plot(tdx,intc.rtn, type = "l", ylim = c(-.6,.6))
lines(tdx,upper2, lty=2, col = "red")
lines(tdx,lower2, lty=2, col = "blue")
abline(h=c(.009915))
```



Por lo general, la modelación $ARCH(m)$ sobreestima la volatilidad, mientras que $GARCH(1,1)$ sigue más la tendencia.

Modelo (G)ARCH con serie ARMA(p,q)

En esta parte del documento se analiza la modelación (G)ARCH del índice S&P 100 (símbolo del ticker `^OEX`) en el periodo 2/agosto/1982-4/marzo/2019, con frecuencia semanal. La serie se descargó de Yahoo! Finance, el día 6/marzo/2019.

Más específicamente, se considerarán dos modelos de ecuaciones de media:

$$\begin{aligned} r_t &= \mu + u_t & ARMA(0,0) \\ r_t &= \mu + \phi_1 r_{t-1} + \theta_1 u_{t-1} + u_t & ARMA(1,1) \end{aligned}$$

donde r_t es la serie de los retornos, al cierre, del índice S&P 100. Para cada ecuación de media, se estimará la volatilidad mediante $ARCH(m)$, $GARCH(m,s)$ y se determinará cuál de ellos mejor se ajusta a los retornos. Finalmente, se presentarán gráficas de intervalos predictivos de los modelos de volatilidad.

```
library(readr)
OEX <- read_csv("OEX.csv")
```

```
oex.ts <- ts(OEX$Close, frequency = 52)
oex.rtn <- diff(log(oex.ts))
```

Estimación de las ecuaciones de media

Modelo $ARMA(1,1)$ Se estimará el modelo usando la función `Arima` (librería `forecast`):

```
m1 <- Arima(oex.rtn, order = c(1,0,1))
m1

## Series: oex.rtn
## ARIMA(1,0,1) with non-zero mean
##
## Coefficients:
##          ar1      ma1    mean
##        -0.695   0.617   0.002
## s.e.    0.116   0.127   0.000
##
## sigma^2 estimated as 0.000514:  log likelihood=4521
## AIC=-9035   AICc=-9035   BIC=-9013
```

Modelo $ARMA(0,0)$ El modelo estimado es

```
m0 <- Arima(oex.rtn, order = c(0,0,0))
m0

## Series: oex.rtn
## ARIMA(0,0,0) with non-zero mean
##
## Coefficients:
##          mean
##         0.002
## s.e.    0.001
##
## sigma^2 estimated as 0.000519:  log likelihood=4510
## AIC=-9017   AICc=-9017   BIC=-9006
```

Nótese que la función `auto.arima()` tiene alguna inestabilidad numérica:

```
auto.arima(oex.rtn, stationary = T, seasonal = F)

## Series: oex.rtn
## ARIMA(2,0,2) with non-zero mean
##
## Coefficients:
##
## Warning in sqrt(diag(x$var.coef)): NaNs produced
```



```
##          ar1      ar2      ma1      ma2      mean
##          0.136    0.572   -0.213   -0.506    0.002
## s.e.      NaN      NaN      NaN      NaN    0.000
##
## sigma^2 estimated as 0.000515:  log likelihood=4521
## AIC=-9031   AICc=-9031   BIC=-8998
```

El mejor ajuste lo proporciona el modelo $ARMA(1,1)$ (mínimo del criterio de información).

Las ecuaciones de media estimadas son

$$r_t = .0017 + u_t$$

$$r_t = .0017 - .6951r_{t-1} + .6173u_{t-1} + u_t$$

Nótese que el retorno promedio estimado es lo mismo,

$$\hat{\mu} = .0017 = .17\% \text{ diario,}$$

en ambos modelos.

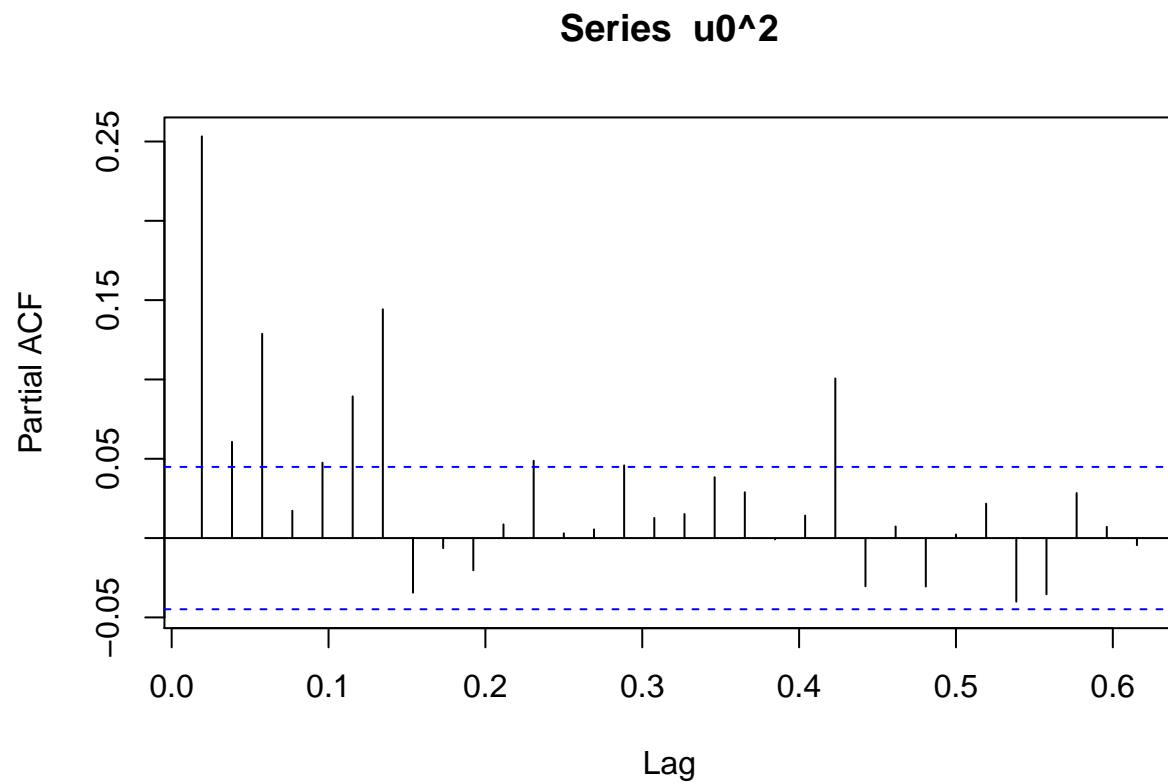
Detección de heterocedasticidad condicional

Para detectar el efecto $ARCH$, se grafica la PACF del cuadrado de los residuales de los dos modelos considerados.

$ARMA(0,0)$ Los residuales, en este caso, son la diferencia

$$u = oex.rtn - mean(oex.rtn)$$

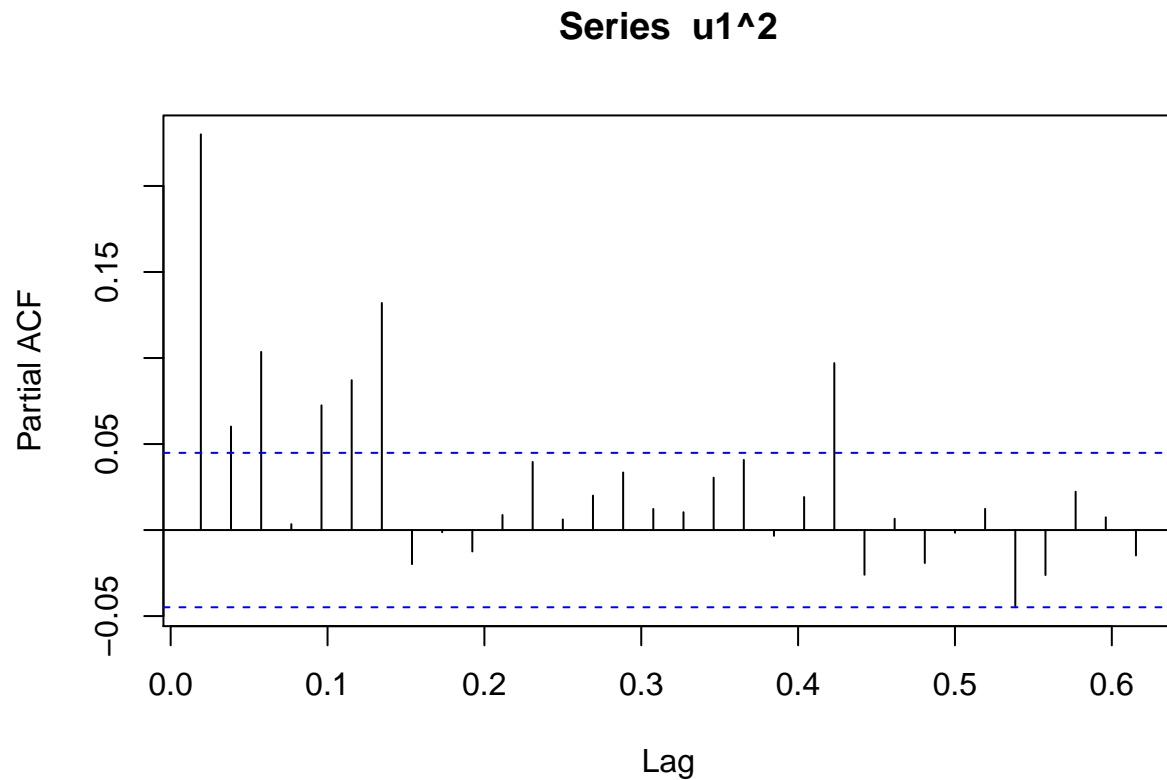
```
u0 = oex.rtn - mean(oex.rtn)
pacf(u0^2)
```



Se detecta el efecto de heterocedasticidad condicional; el orden posible es $m = 7$.

$ARMA(1,1)$ Los residuales, en este caso, se extraen de `m1`, definido arriba

```
u1 = m1$residuals  
pacf(u1^2)
```



También en este caso se detecta el efecto ARCH, con $m = 7$.

Los dos correlogramas parciales son muy semejantes.

Usando el script de R. Tsay, `archTest.R`:

```
source("/Users/laan/pCloud Sync/0/EcoFin2/Data/archTest.R") #OSX
# source("~/pCloudDrive/pCloud Sync/0/EcoFin2/Data/archTest.R")
archTest(u0)
```

```
##
## Call:
## lm(formula = atsq ~ x)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.00710	-0.00036	-0.00023	0.00004	0.03834

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.41e-04	3.84e-05	6.27	4.5e-10 ***
x1	2.21e-01	2.30e-02	9.62	< 2e-16 ***
x2	1.61e-02	2.35e-02	0.69	0.493

```
## x3          1.20e-01  2.35e-02   5.11  3.6e-07 ***
## x4          -1.06e-02  2.34e-02  -0.45   0.649
## x5          2.72e-02  2.33e-02   1.16   0.245
## x6          6.10e-02  2.33e-02   2.61   0.009 **
## x7          1.45e-01  2.34e-02   6.19  7.2e-10 ***
## x8          -2.83e-02  2.34e-02  -1.21   0.227
## x9          -2.48e-03  2.33e-02  -0.11   0.915
## x10         -2.01e-02  2.28e-02  -0.88   0.378
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.00134 on 1888 degrees of freedom
## Multiple R-squared:  0.115, Adjusted R-squared:  0.11
## F-statistic: 24.5 on 10 and 1888 DF, p-value: <2e-16
```

```
archTest(u1)
```

```
##
## Call:
## lm(formula = atsq ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.00771 -0.00036 -0.00023  0.00004  0.04069
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.47e-04  3.96e-05   6.24  5.4e-10 ***
## x1          1.99e-01  2.30e-02   8.64 < 2e-16 ***
## x2          2.16e-02  2.34e-02   0.92  0.3562
## x3          9.94e-02  2.34e-02   4.25  2.3e-05 ***
## x4         -2.55e-02  2.33e-02  -1.09  0.2743
## x5          5.20e-02  2.33e-02   2.23  0.0256 *
## x6          6.21e-02  2.33e-02   2.67  0.0076 **
## x7          1.29e-01  2.33e-02   5.54  3.5e-08 ***
## x8         -1.54e-02  2.34e-02  -0.66  0.5087
## x9          6.78e-04  2.32e-02   0.03  0.9767
## x10         -1.23e-02  2.28e-02  -0.54  0.5885
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.00138 on 1888 degrees of freedom
## Multiple R-squared:  0.0958, Adjusted R-squared:  0.091
## F-statistic: 20 on 10 and 1888 DF, p-value: <2e-16
```

se obtiene el mismo resultado.

Estimación de la volatilidad

ARCH(7)

Los modelos de volatilidad que corresponden a las ecuación de media *ARMA*(0,0), *ARMA*(1,1), respectivamente, se dan a continuación. Se usarán innovaciones gaussianas, por el momento.

```
m0.arch <- garchFit(~ 1 + garch(7,0), data = oex.rtn, trace=F)
summary(m0.arch)
```

```
##
## Title:
##  GARCH Modelling
##
## Call:
##  garchFit(formula = ~1 + garch(7, 0), data = oex.rtn, trace = F)
##
## Mean and Variance Equation:
##  data ~ 1 + garch(7, 0)
## <environment: 0x7f89a92c38c8>
##  [data = oex.rtn]
##
## Conditional Distribution:
##  norm
##
## Coefficient(s):
##           mu           omega      alpha1      alpha2      alpha3      alpha4
## 0.00253794 0.00014271 0.23320004 0.10782260 0.17069609 0.08935652
##      alpha5      alpha6      alpha7
## 0.03463438 0.06016693 0.07333878
##
## Std. Errors:
##  based on Hessian
##
## Error Analysis:
##           Estimate Std. Error  t value Pr(>|t|)
## mu      0.0025379  0.0003972   6.389 1.67e-10 ***
## omega   0.0001427  0.0000163   8.755 < 2e-16 ***
## alpha1  0.2332000  0.0335868   6.943 3.83e-12 ***
## alpha2  0.1078226  0.0303336   3.555 0.000379 ***
## alpha3  0.1706961  0.0349660   4.882 1.05e-06 ***
## alpha4  0.0893565  0.0316762   2.821 0.004788 **
## alpha5  0.0346344  0.0289507   1.196 0.231570
## alpha6  0.0601669  0.0260152   2.313 0.020736 *
## alpha7  0.0733388  0.0245706   2.985 0.002838 **
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
## 4725      normalized:  2.48
##
## Description:
## Mon Nov 25 14:08:02 2019 by user:
##
##
## Standardised Residuals Tests:
##
##               Statistic p-Value
## Jarque-Bera Test   R      Chi^2 169      0
## Shapiro-Wilk Test  R      W      0.986    1.95e-12
## Ljung-Box Test     R      Q(10) 17.3     0.0678
## Ljung-Box Test     R      Q(15) 23.3     0.0772
## Ljung-Box Test     R      Q(20) 30.2     0.0666
## Ljung-Box Test     R^2  Q(10)  6.12     0.805
## Ljung-Box Test     R^2  Q(15) 14.8     0.465
## Ljung-Box Test     R^2  Q(20) 15.8     0.726
## LM Arch Test       R      TR^2  7.58     0.817
##
## Information Criterion Statistics:
##   AIC   BIC   SIC  HQIC
## -4.94 -4.91 -4.94 -4.93

m1.arch <- garchFit(~ arma(1,1) + garch(7,0), data = oex.rtn, trace=F)
summary(m1.arch)

##
## Title:
## GARCH Modelling
##
## Call:
## garchFit(formula = ~arma(1, 1) + garch(7, 0), data = oex.rtn,
##         trace = F)
##
## Mean and Variance Equation:
## data ~ arma(1, 1) + garch(7, 0)
## <environment: 0x7f89ae129030>
## [data = oex.rtn]
##
## Conditional Distribution:
## norm
##
```

```

## Coefficient(s):
##      mu      ar1      ma1      omega      alpha1
## 0.00408753 -0.60541229 0.52564106 0.00014286 0.23207776
##      alpha2      alpha3      alpha4      alpha5      alpha6
## 0.11058415 0.15819657 0.09482683 0.04190564 0.06808314
##      alpha7
## 0.06041564
##
## Std. Errors:
## based on Hessian
##
## Error Analysis:
##      Estimate Std. Error t value Pr(>|t|)
## mu      4.088e-03 7.992e-04 5.114 3.15e-07 ***
## ar1     -6.054e-01 1.997e-01 -3.031 0.002437 **
## ma1      5.256e-01 2.145e-01 2.450 0.014279 *
## omega    1.429e-04 1.625e-05 8.789 < 2e-16 ***
## alpha1   2.321e-01 3.371e-02 6.885 5.79e-12 ***
## alpha2   1.106e-01 3.077e-02 3.594 0.000326 ***
## alpha3   1.582e-01 3.498e-02 4.522 6.13e-06 ***
## alpha4   9.483e-02 3.324e-02 2.853 0.004337 **
## alpha5   4.191e-02 2.992e-02 1.401 0.161331
## alpha6   6.808e-02 2.777e-02 2.451 0.014230 *
## alpha7   6.042e-02 2.342e-02 2.579 0.009897 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
## 4733      normalized: 2.48
##
## Description:
## Mon Nov 25 14:08:04 2019 by user:
##
##
## Standardised Residuals Tests:
##
##      Statistic p-Value
## Jarque-Bera Test R Chi^2 182 0
## Shapiro-Wilk Test R W 0.985 4.74e-13
## Ljung-Box Test R Q(10) 6.54 0.768
## Ljung-Box Test R Q(15) 11.5 0.717
## Ljung-Box Test R Q(20) 19.2 0.511
## Ljung-Box Test R^2 Q(10) 6.45 0.776
## Ljung-Box Test R^2 Q(15) 16.6 0.343
## Ljung-Box Test R^2 Q(20) 17.8 0.604
## LM Arch Test R TR^2 7.5 0.823

```

```
##
## Information Criterion Statistics:
##   AIC   BIC   SIC  HQIC
## -4.95 -4.91 -4.95 -4.93
```

Se estima que el retorno promedio diario estimado en el primer modelo es

$$\hat{\mu} = .00253794 = .253794\%;$$

en el segundo modelo es igual a

$$\hat{\mu} = .00408753 / (1 + .60541229) = .002546 = .2546\%.$$

GARCH(1,1)

Los modelos de volatilidad que corresponden a las ecuación de media *ARMA*(0,0), *ARMA*(1,1), respectivamente, se dan a continuación. Se usarán innovaciones gaussianas, por el momento.

```
m0.garch <- garchFit(~ 1 + garch(1,1), data = oex.rtn, trace=F)
summary(m0.garch)
```

```
##
## Title:
##   GARCH Modelling
##
## Call:
##   garchFit(formula = ~1 + garch(1, 1), data = oex.rtn, trace = F)
##
## Mean and Variance Equation:
##   data ~ 1 + garch(1, 1)
## <environment: 0x7f89aa6c4638>
##   [data = oex.rtn]
##
## Conditional Distribution:
##   norm
##
## Coefficient(s):
##           mu           omega        alpha1        beta1
## 2.3739e-03  1.9594e-05  1.6539e-01  8.0393e-01
##
## Std. Errors:
##   based on Hessian
##
## Error Analysis:
##           Estimate Std. Error  t value Pr(>|t|)
## mu       2.374e-03   3.991e-04   5.948 2.71e-09 ***
```



```

## omega 1.959e-05 5.526e-06 3.546 0.000392 ***
## alpha1 1.654e-01 2.452e-02 6.746 1.52e-11 ***
## beta1 8.039e-01 2.919e-02 27.544 < 2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
## 4729 normalized: 2.48
##
## Description:
## Mon Nov 25 14:08:04 2019 by user:
##
##
## Standardised Residuals Tests:
##
## Statistic p-Value
## Jarque-Bera Test R Chi^2 211 0
## Shapiro-Wilk Test R W 0.985 4.72e-13
## Ljung-Box Test R Q(10) 18.7 0.0436
## Ljung-Box Test R Q(15) 24.2 0.0617
## Ljung-Box Test R Q(20) 32 0.0436
## Ljung-Box Test R^2 Q(10) 15.2 0.126
## Ljung-Box Test R^2 Q(15) 20 0.171
## Ljung-Box Test R^2 Q(20) 21.5 0.368
## LM Arch Test R TR^2 18.2 0.11
##
## Information Criterion Statistics:
## AIC BIC SIC HQIC
## -4.95 -4.94 -4.95 -4.95

m1.garch <- garchFit(~ arma(1,1) + garch(1,1), data = oex.rtn, trace=F)
summary(m1.garch)

##
## Title:
## GARCH Modelling
##
## Call:
## garchFit(formula = ~arma(1, 1) + garch(1, 1), data = oex.rtn,
## trace = F)
##
## Mean and Variance Equation:
## data ~ arma(1, 1) + garch(1, 1)
## <environment: 0x7f89af05d5d0>
## [data = oex.rtn]
##

```

```

## Conditional Distribution:
## norm
##
## Coefficient(s):
##      mu      ar1      ma1      omega      alpha1
## 3.8825e-03 -6.3095e-01  5.4764e-01  1.9659e-05  1.6362e-01
##      beta1
## 8.0488e-01
##
## Std. Errors:
## based on Hessian
##
## Error Analysis:
##      Estimate Std. Error t value Pr(>|t|)
## mu      3.882e-03  7.603e-04   5.106 3.29e-07 ***
## ar1     -6.310e-01  1.834e-01  -3.439 0.000583 ***
## ma1      5.476e-01  1.984e-01   2.760 0.005782 **
## omega    1.966e-05  5.529e-06   3.556 0.000377 ***
## alpha1   1.636e-01  2.401e-02   6.813 9.55e-12 ***
## beta1    8.049e-01  2.893e-02  27.825 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log Likelihood:
## 4739      normalized:  2.48
##
## Description:
## Mon Nov 25 14:08:04 2019 by user:
##
##
## Standardised Residuals Tests:
##
##      Jarque-Bera Test  R      Chi^2  231      0
##      Shapiro-Wilk Test  R      W      0.984    9.27e-14
##      Ljung-Box Test    R      Q(10)  6.88     0.736
##      Ljung-Box Test    R      Q(15)  11.3     0.732
##      Ljung-Box Test    R      Q(20)  19.8     0.47
##      Ljung-Box Test    R^2    Q(10)  14       0.174
##      Ljung-Box Test    R^2    Q(15)  19.5     0.192
##      Ljung-Box Test    R^2    Q(20)  21.3     0.377
##      LM Arch Test      R      TR^2   16.6     0.165
##
## Information Criterion Statistics:
##      AIC      BIC      SIC      HQIC
## -4.96 -4.94 -4.96 -4.95

```

El retorno promedio diario estimado en el primer modelo es

$$\hat{\mu} = .002374 = .2374\%;$$

en el segundo modelo es igual a

$$\hat{\mu} = .0038825/(1 + .63095) = .002381 = .2381\%.$$

Diagnóstico

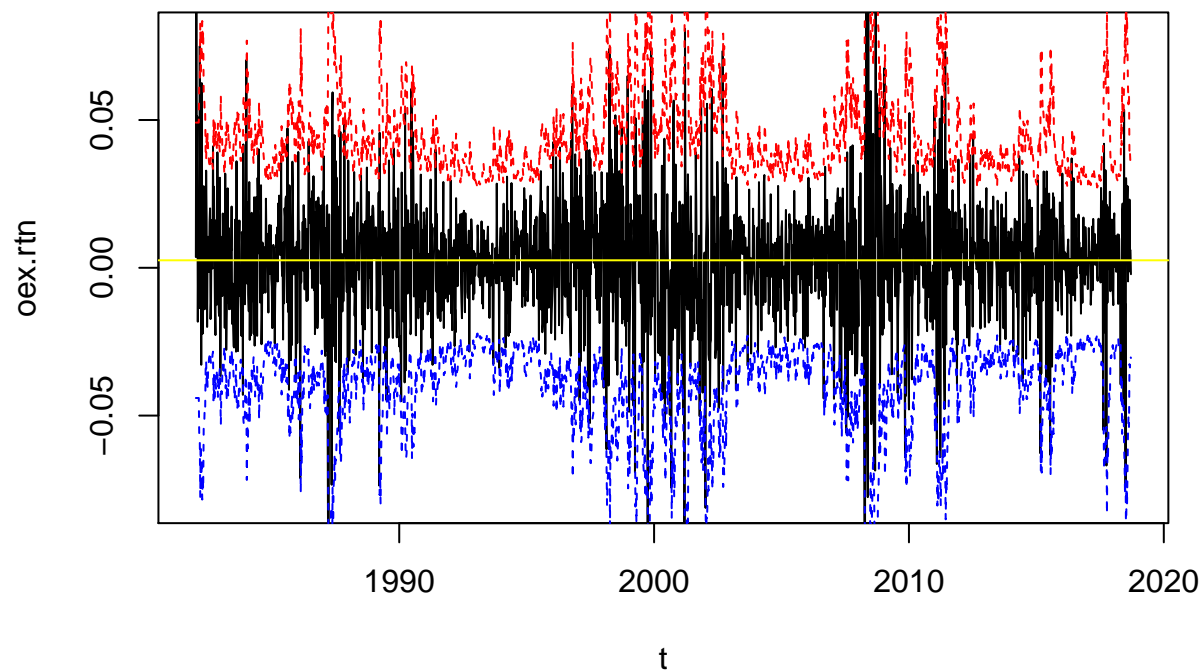
- El único modelo que no pasa uno de los diagnósticos es el $GARCH(1, 1)$, con ecuación de la media constante, `m0.garch`.
- Los criterios de información AIC, BIC son mejores (menores) en el modelo `m1.garch`.
- En ningún caso, las innovaciones son gaussianas.

Conclusión Usando el principio de parsimonia, concluimos que el $GARCH(1, 1)$ con ecuación de la media $ARMA(1, 1)$ es el mejor modelo de volatilidad.

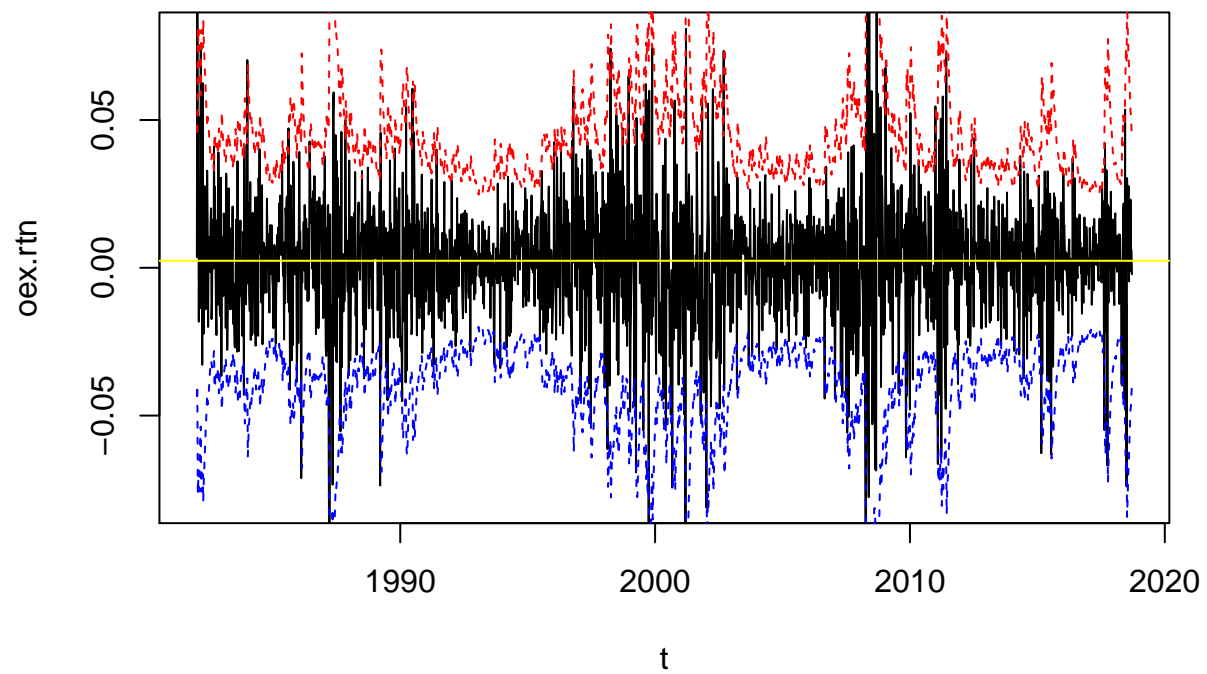
Gráficos con intervalos predictivos

A continuación se presentarán los intervalos predictivos de la volatilidad; por lo general, los modelos $GARCH(1, 1)$ siguen más la tendencia que los modelos $ARCH(7)$.

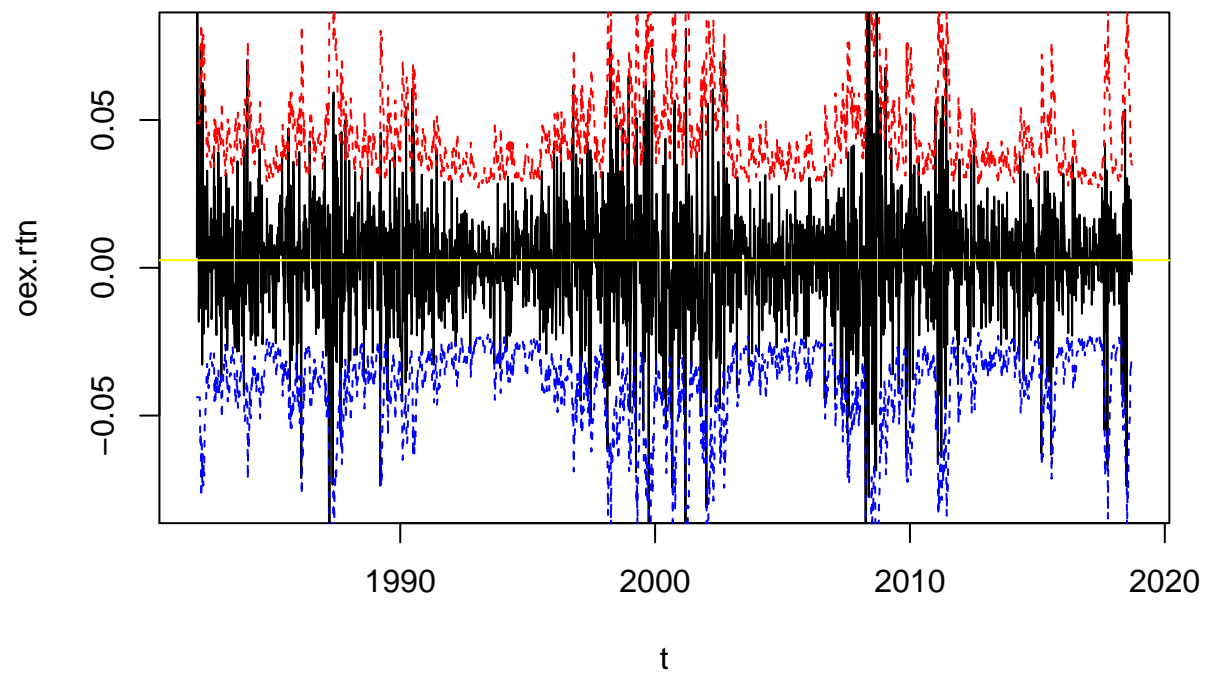
```
v0.arch <- volatility(m0.arch)
v1.arch <- volatility(m1.arch)
v0.garch <- volatility(m0.garch)
v1.garch <- volatility(m1.garch)
upper.v0.arch <- .0025379 + 2*v0.arch
lower.v0.arch <- .0025379 - 2*v0.arch
upper.v1.arch <- .002546 + 2*v1.arch
lower.v1.arch <- .002546 - 2*v1.arch
upper.v0.garch <- .002374 + 2*v0.garch
lower.v0.garch <- .002374 - 2*v0.garch
upper.v1.garch <- .002381 + 2*v1.garch
lower.v1.garch <- .002381 - 2*v1.garch
t <- c(1:1909)/52+1982
plot(t,oex.rtn, type = "l", ylim = c(-.08,.08))
lines(t,upper.v0.arch, lty=2, col = "red")
lines(t,lower.v0.arch, lty=2, col = "blue")
abline(h=c(.0025379), col = "yellow")
```



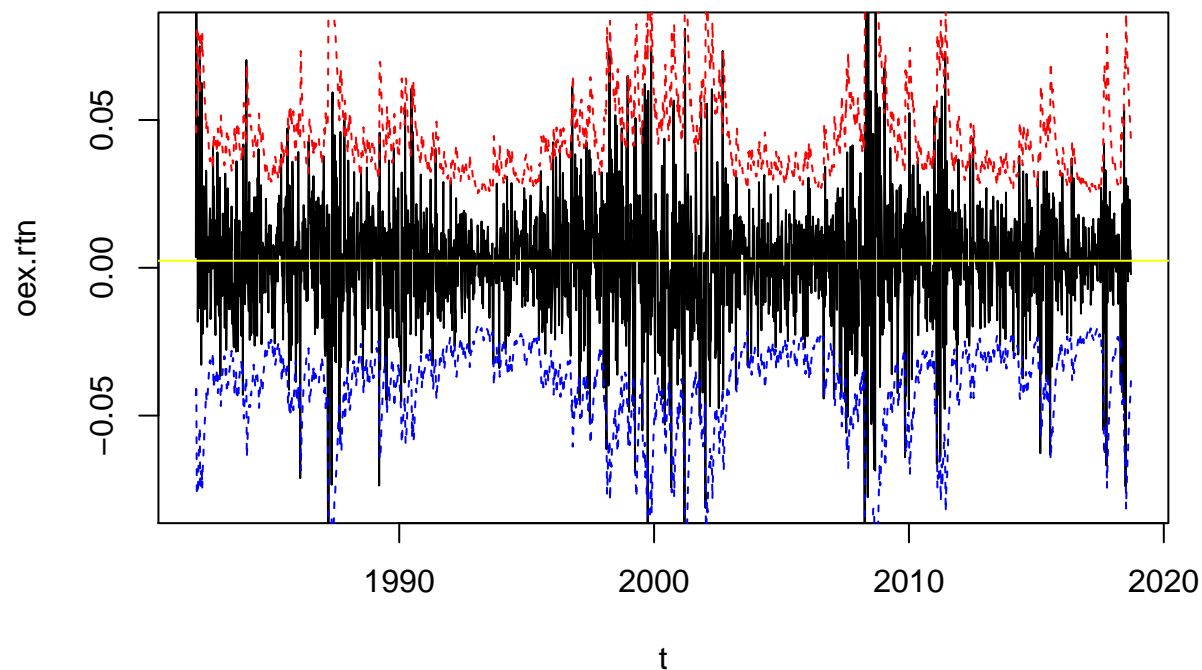
```
plot(t,oex.rtn, type = "l", ylim = c(-.08,.08))
lines(t,upper.v0.garch, lty=2, col = "red")
lines(t,lower.v0.garch, lty=2, col = "blue")
abline(h=c(.002374), col = "yellow")
```



```
plot(t,oex.rtn, type = "l", ylim = c(-.08,.08))
lines(t,upper.v1.arch, lty=2, col = "red")
lines(t,lower.v1.arch, lty=2, col = "blue")
abline(h=c(.002546), col = "yellow")
```



```
plot(t,oex.rtn, type = "l", ylim = c(-.08,.08))  
lines(t,upper.v1.garch, lty=2, col = "red")  
lines(t,lower.v1.garch, lty=2, col = "blue")  
abline(h=c(.002381), col = "yellow")
```



Otros modelos

Modelo GARCH-M "GARCH in the mean"

Un modelo GARCH(1,1)-M es un modelo GARCH(1,1) en donde la ecuación de la media incluye un término de volatilidad:

$$\begin{aligned} r_t &= \mu + c\sigma_t^2 + u_t \\ u_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned}$$

El coeficiente c se puede interpretar como una *prima de riesgo*.

Si $c > 0$, entonces los retornos son positivamente correlacionados con la volatilidad pasada.

Determina si los retornos de Intel Corporation tienen una prima de riesgo.

La hipótesis estadística es

$$\begin{aligned} H_0 : c &= 0 \\ H_1 : c &\neq 0 \end{aligned}$$

```
library(rugarch)
```

```
spec.archm = ugarchspec(variance.model=list(model="sGARCH"),
                        mean.model=list(armaOrder=c(0,0),archm=TRUE)) ## <-- opción "archm=TRUE"
fit.archm = ugarchfit(data=intc.rtn,spec=spec.archm)
fit.archm
```

```
##
## *-----*
## *          GARCH Model Fit          *
## *-----*
##
## Conditional Variance Dynamics
## -----
## GARCH Model   : sGARCH(1,1)
## Mean Model    : ARFIMA(0,0,0)
## Distribution   : norm
##
## Optimal Parameters
## -----
##      Estimate  Std. Error  t value Pr(>|t|)
## mu      0.010447    0.015457  0.67585  0.49914
## archm   -0.023962    0.165007 -0.14522  0.88454
## omega    0.000070    0.000090  0.77538  0.43812
## alpha1   0.057199    0.016775  3.40988  0.00065
## beta1    0.936924    0.019774 47.38267  0.00000
##
## Robust Standard Errors:
##      Estimate  Std. Error  t value Pr(>|t|)
## mu      0.010447    0.014564  0.71731  0.473185
## archm   -0.023962    0.166171 -0.14420  0.885344
## omega    0.000070    0.000127  0.55137  0.581381
## alpha1   0.057199    0.020545  2.78406  0.005368
## beta1    0.936924    0.028719 32.62429  0.000000
##
## LogLikelihood : 409
##
## Information Criteria
## -----
##
## Akaike          -1.7295
## Bayes           -1.6851
## Shibata         -1.7298
## Hannan-Quinn   -1.7121
##
```



```

## Weighted Ljung-Box Test on Standardized Residuals
## -----
##               statistic p-value
## Lag[1]         0.08134  0.7755
## Lag[2*(p+q)+(p+q)-1][2]  0.19402  0.8579
## Lag[4*(p+q)+(p+q)-1][5]  1.64742  0.7036
## d.o.f=0
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## -----
##               statistic p-value
## Lag[1]         0.06424  0.7999
## Lag[2*(p+q)+(p+q)-1][5]  0.66850  0.9289
## Lag[4*(p+q)+(p+q)-1][9]  1.19218  0.9769
## d.o.f=2
##
## Weighted ARCH LM Tests
## -----
##               Statistic Shape Scale P-Value
## ARCH Lag[3]    0.07839 0.500 2.000  0.7795
## ARCH Lag[5]    0.27848 1.440 1.667  0.9459
## ARCH Lag[7]    0.36579 2.315 1.543  0.9889
##
## Nyblom stability test
## -----
## Joint Statistic:  0.933
## Individual Statistics:
## mu      0.1363
## archm   0.1467
## omega   0.2808
## alpha1  0.1886
## beta1   0.1739
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:      1.28 1.47 1.88
## Individual Statistic:  0.35 0.47 0.75
##
## Sign Bias Test
## -----
##               t-value   prob sig
## Sign Bias      0.1103 0.9123
## Negative Sign Bias 0.1650 0.8690
## Positive Sign Bias 0.5254 0.5996
## Joint Effect    0.5444 0.9090

```

```
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## -----
##   group statistic p-value(g-1)
## 1    20      15.87      0.6660
## 2    30      29.90      0.4193
## 3    40      38.27      0.5031
## 4    50      42.53      0.7313
##
##
## Elapsed time : 0.834
```

Se concluye que no existe evidencia de una prima de riesgo asociada con la volatilidad en los retornos de Intel.

Modelos EGARCH(1,1)

Un modelo EGARCH(1,1) se escribe como

$$\ln(\sigma_t^2) = \omega + g(\varepsilon_{t-1}) + \beta_1 \ln(\sigma_{t-1}^2),$$

donde

$$g(\varepsilon_{t-1}) = \alpha_1 \varepsilon_{t-1} + \gamma_1 [|\varepsilon_{t-1}| - E(|\varepsilon_{t-1}|)].$$

Nótese que

$$g(\varepsilon_{t-1}) = \begin{cases} (\alpha_1 - \gamma_1)\varepsilon_{t-1} - E(|\varepsilon_{t-1}|), & \varepsilon_{t-1} < 0 \\ (\alpha_1 + \gamma_1)\varepsilon_{t-1} - E(|\varepsilon_{t-1}|), & \varepsilon_{t-1} \geq 0 \end{cases}$$

Se determina que existe efecto de apalancamiento, si $\alpha_1 < 0$: innovaciones negativas tienen más impacto que innovaciones positivas.

Para probar un efecto de apalancamiento se prueba la hipótesis estadística

$$H_0 : \alpha_1 = 0$$

$$H_1 : \alpha_1 < 0$$

Problema Determinar si se observa un efecto de apalancamiento en los retornos de Intel Corporation.

```
spec.egarch = ugarchspec(variance.model=list(model="eGARCH"), ## <-- opción "eGARCH"
                        mean.model=list(armaOrder=c(0,0)))
fit.egarch = ugarchfit(data=intc.rtn,spec=spec.egarch)
fit.egarch
```

```

##
## *-----*
## *          GARCH Model Fit          *
## *-----*
##
## Conditional Variance Dynamics
## -----
## GARCH Model   : eGARCH(1,1)
## Mean Model    : ARFIMA(0,0,0)
## Distribution   : norm
##
## Optimal Parameters
## -----
##      Estimate  Std. Error  t value Pr(>|t|)
## mu      0.010749    0.004130   2.6028 0.009247
## omega  -0.022665    0.008636  -2.6244 0.008679
## alpha1  0.032029    0.017668   1.8128 0.069857
## beta1   0.994881    0.001819 546.9193 0.000000
## gamma1  0.107795    0.018451   5.8421 0.000000
##
## Robust Standard Errors:
##      Estimate  Std. Error  t value Pr(>|t|)
## mu      0.010749    0.004338   2.4778 0.013219
## omega  -0.022665    0.014031  -1.6154 0.106232
## alpha1  0.032029    0.026975   1.1873 0.235091
## beta1   0.994881    0.003022 329.1622 0.000000
## gamma1  0.107795    0.018062   5.9680 0.000000
##
## LogLikelihood : 411
##
## Information Criteria
## -----
##
## Akaike      -1.7369
## Bayes       -1.6926
## Shibata     -1.7372
## Hannan-Quinn -1.7195
##
## Weighted Ljung-Box Test on Standardized Residuals
## -----
##
##              statistic p-value
## Lag[1]              0.04017  0.8412
## Lag[2*(p+q)+(p+q)-1] [2]  0.08324  0.9316
## Lag[4*(p+q)+(p+q)-1] [5]  2.10722  0.5933
## d.o.f=0

```

```

## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## -----
##
##              statistic p-value
## Lag[1]              0.9527  0.3290
## Lag[2*(p+q)+(p+q)-1][5]  1.0992  0.8368
## Lag[4*(p+q)+(p+q)-1][9]  2.0324  0.9011
## d.o.f=2
##
## Weighted ARCH LM Tests
## -----
##
##              Statistic Shape Scale P-Value
## ARCH Lag[3]  0.001577 0.500 2.000  0.9683
## ARCH Lag[5]  0.397468 1.440 1.667  0.9134
## ARCH Lag[7]  0.645447 2.315 1.543  0.9634
##
## Nyblom stability test
## -----
## Joint Statistic:  0.994
## Individual Statistics:
## mu      0.08271
## omega   0.19209
## alpha1  0.15712
## beta1   0.22414
## gamma1  0.02696
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:      1.28 1.47 1.88
## Individual Statistic:  0.35 0.47 0.75
##
## Sign Bias Test
## -----
##
##              t-value   prob sig
## Sign Bias      0.07594 0.9395
## Negative Sign Bias 0.68897 0.4912
## Positive Sign Bias 0.58188 0.5609
## Joint Effect    0.82339 0.8439
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## -----
##      group statistic p-value(g-1)
## 1      20      22.55      0.2577
## 2      30      29.12      0.4586

```

```
## 3      40      39.30      0.4566
## 4      50      39.75      0.8245
##
##
## Elapsed time : 0.208
```

Modelos EGARCH(1,1) con prima de riesgo

Se puede también considerar un modelo EGARCH(1,1), incluyendo una prima de riesgo en la ecuación de la media:

```
spec.egarch.m = ugarchspec(variance.model=list(model="eGARCH"), ## <-- opción "eGARCH"
mean.model=list(armaOrder=c(0,0), archm=T)) ## <-- prima de riesgo
fit.egarch.m = ugarchfit(data=intc.rtn,spec=spec.egarch.m)
fit.egarch.m
```

```
##
## *-----*
## *          GARCH Model Fit          *
## *-----*
##
## Conditional Variance Dynamics
## -----
## GARCH Model   : eGARCH(1,1)
## Mean Model    : ARFIMA(0,0,0)
## Distribution   : norm
##
## Optimal Parameters
## -----
##      Estimate Std. Error  t value Pr(>|t|)
## mu      0.013201   0.011675   1.13065 0.258201
## archm   -0.027115   0.129337  -0.20965 0.833942
## omega   -0.023282   0.006474  -3.59647 0.000323
## alpha1   0.032038   0.017466   1.83428 0.066612
## beta1    0.994776   0.001353 735.30502 0.000000
## gamma1   0.106108   0.013118   8.08898 0.000000
##
## Robust Standard Errors:
##      Estimate Std. Error  t value Pr(>|t|)
## mu      0.013201   0.012877   1.02512 0.305307
## archm   -0.027115   0.140425  -0.19309 0.846885
## omega   -0.023282   0.011059  -2.10529 0.035266
## alpha1   0.032038   0.025727   1.24531 0.213016
## beta1    0.994776   0.002275 437.24289 0.000000
```

```

## gamma1  0.106108    0.012289    8.63453 0.000000
##
## LogLikelihood : 411
##
## Information Criteria
## -----
##
## Akaike          -1.7327
## Bayes           -1.6795
## Shibata         -1.7330
## Hannan-Quinn   -1.7118
##
## Weighted Ljung-Box Test on Standardized Residuals
## -----
##
##                statistic p-value
## Lag[1]                0.04369  0.8344
## Lag[2*(p+q)+(p+q)-1] [2]  0.08878  0.9276
## Lag[4*(p+q)+(p+q)-1] [5]  2.14690  0.5841
## d.o.f=0
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## -----
##
##                statistic p-value
## Lag[1]                1.013  0.3142
## Lag[2*(p+q)+(p+q)-1] [5]  1.159  0.8229
## Lag[4*(p+q)+(p+q)-1] [9]  2.130  0.8890
## d.o.f=2
##
## Weighted ARCH LM Tests
## -----
##
##                Statistic Shape Scale P-Value
## ARCH Lag[3] 0.0005389 0.500 2.000  0.9815
## ARCH Lag[5] 0.4019011 1.440 1.667  0.9121
## ARCH Lag[7] 0.6714248 2.315 1.543  0.9603
##
## Nyblom stability test
## -----
## Joint Statistic:  1.06
## Individual Statistics:
## mu      0.08819
## archm   0.09257
## omega   0.18446
## alpha1  0.15763
## beta1   0.21556

```

```

## gamma1 0.02736
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:          1.49 1.68 2.12
## Individual Statistic:     0.35 0.47 0.75
##
## Sign Bias Test
## -----
##                t-value   prob sig
## Sign Bias      0.1369 0.8912
## Negative Sign Bias 0.6781 0.4980
## Positive Sign Bias 0.6193 0.5361
## Joint Effect    0.8681 0.8331
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## -----
##  group statistic p-value(g-1)
## 1      20      20.58      0.3604
## 2      30      31.57      0.3393
## 3      40      45.12      0.2314
## 4      50      35.25      0.9301
##
##
## Elapsed time : 0.52

```

Regresiones espurias; cointegración

Las series de tiempo que ocurren típicamente en Economía y Finanzas no son, por lo general, estacionarias (*e.g.*, PIB, precios, etc.). Las correspondientes series diferenciadas (una vez, o más) lo son. Ya vimos en la sección *Estacionariedad de un modelo AR(p)*, pruebas para reconocer series estacionarias; se usarán estas pruebas para determinar si combinaciones lineales de dos series es válida, o no. En el primer caso, se dirá que las series son *cointegradas*; en el segundo caso, la relación es una regresión espuria.

Regresiones espurias y estacionariedad

Una situación que se puede presentar, especialmente cuando no se conoce a fondo la naturaleza de los datos que se están manejando, es encontrar relaciones artificiales que son significativas estadísticamente. Véase el artículo clásico de Yule, [Y].

Una regresión espuria es un ejemplo de tales relaciones. El fuerte resultado estadístico es consecuencia de un efecto de tendencia o de no estacionariedad de los datos.

Considérese, por ejemplo, el caso de dos caminatas aleatorias independientes. Si se regresa una sobre la obra, entonces se encontrará una estimación que es estadísticamente significativa que, pero es completamente artificial. Presentamos un ejemplo de esta relación a continuación. El valor de la semilla es para reproducibilidad de los resultados.

Se definen primeramente las dos caminatas aleatorias, x_1 , y_1 :

```
set.seed(12321)
n <- 1000
x1 <- cumsum(sample(c(-1, 1), n, TRUE))
y1 <- cumsum(sample(c(-1, 1), n, TRUE))
```

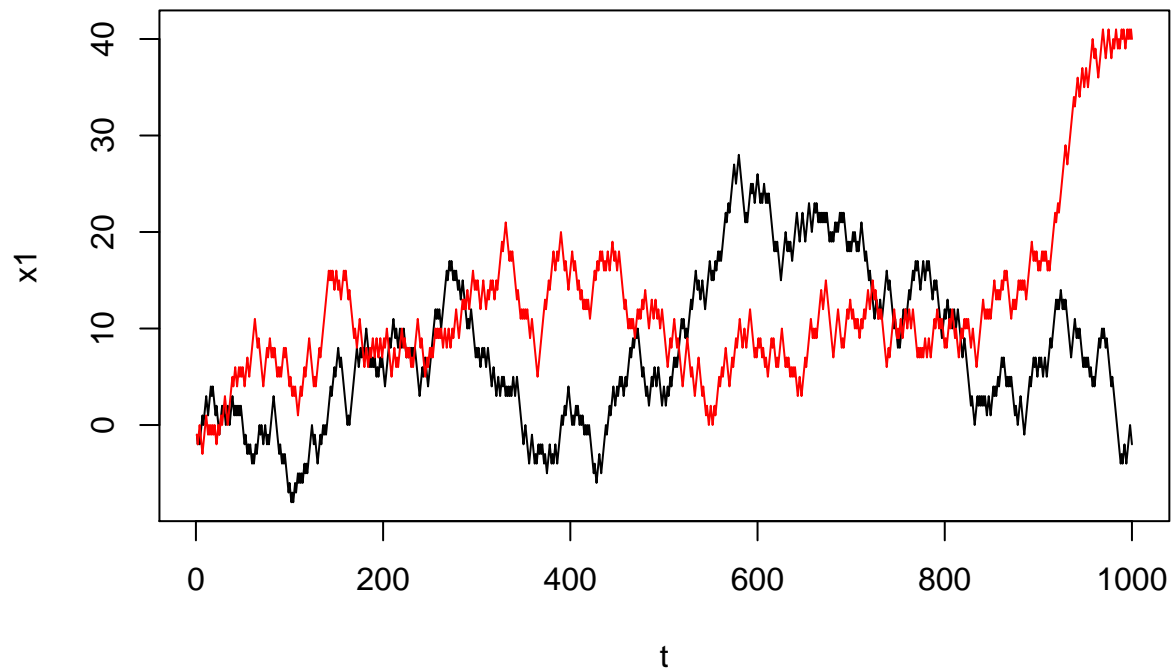
Se calcula un prueba de hipótesis de correlación entre x , y y se grafican:

```
cor.test(x1,y1)

##
## Pearson's product-moment correlation
##
## data: x1 and y1
## t = -4, df = 998, p-value = 1e-05
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.1995 -0.0779
## sample estimates:
## cor
## -0.139
```

La correlación es positiva, de manera altamente significativa. Las gráficas a continuación de x_1 , y_1 nos permite comprender esta situación.

```
plot(x1, type = "l", xlab = "t",
      ylim = c(min(min(x1), min(y1)), max(max(x1), max(y1))))
lines(y1, col="red")
```

Después, regresamos y sobre x y desplegamos el resultado

```
lm1 <- lm(y1~x1)
lm1.summary <- summary(lm1)
lm1.summary
```

```
##
## Call:
## lm(formula = y1 ~ x1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.96  -4.46  -1.34   2.04  29.22
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  13.1061    0.3667   35.74  <2e-16 ***
## x1          -0.1472    0.0332   -4.44   1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.32 on 998 degrees of freedom
```

```
## Multiple R-squared:  0.0194, Adjusted R-squared:  0.0184
## F-statistic: 19.7 on 1 and 998 DF,  p-value: 9.95e-06
```

El modelo es significativo, así como el intercepto y la pendiente; el valor de $R^2 = 0.019$.

Una señal de atención para una regresión espuria es la presencia de una raíz unitaria en los residuos:

```
library(tseries)
adf.test(as.ts(lm1$residuals))
```

```
##
## Augmented Dickey-Fuller Test
##
## data:  as.ts(lm1$residuals)
## Dickey-Fuller = -1, Lag order = 9, p-value = 0.9
## alternative hypothesis: stationary
```

```
pp.test(as.ts(lm1$residuals))
```

```
##
## Phillips-Perron Unit Root Test
##
## data:  as.ts(lm1$residuals)
## Dickey-Fuller Z(alpha) = -8, Truncation lag parameter = 7, p-value
## = 0.7
## alternative hypothesis: stationary
```

```
kpss.test(as.ts(lm1$residuals))
```

```
## Warning in kpss.test(as.ts(lm1$residuals)): p-value smaller than printed p-
## value
```

```
##
## KPSS Test for Level Stationarity
##
## data:  as.ts(lm1$residuals)
## KPSS Level = 4, Truncation lag parameter = 7, p-value = 0.01
```

Por otro lado, la regresión de las series diferenciadas no tiene ninguna información importante:

```
x1.diff <- diff(x1)
y1.diff <- diff(y1)
lm1.diff <- lm(y1.diff ~ x1.diff)
summary(lm1.diff)
```

```
##
## Call:
## lm(formula = y1.diff ~ x1.diff)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.096 -0.986  0.904  0.904  1.014
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0410     0.0316   1.30   0.195
## x1.diff       -0.0550     0.0316  -1.74   0.082 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.999 on 997 degrees of freedom
## Multiple R-squared:  0.00303,    Adjusted R-squared:  0.00203
## F-statistic: 3.03 on 1 and 997 DF,  p-value: 0.082
```

```
adf.test(as.ts(lm1.diff$residuals))
```

```
## Warning in adf.test(as.ts(lm1.diff$residuals)): p-value smaller than
## printed p-value
```

```
##
## Augmented Dickey-Fuller Test
##
## data:  as.ts(lm1.diff$residuals)
## Dickey-Fuller = -11, Lag order = 9, p-value = 0.01
## alternative hypothesis: stationary
```

Los residuos son estacionarios:

```
adf.test(as.ts(lm1.diff$residuals))
```

```
## Warning in adf.test(as.ts(lm1.diff$residuals)): p-value smaller than
## printed p-value
```

```
##
## Augmented Dickey-Fuller Test
##
## data:  as.ts(lm1.diff$residuals)
## Dickey-Fuller = -11, Lag order = 9, p-value = 0.01
## alternative hypothesis: stationary
```

```
pp.test(as.ts(lm1.diff$residuals))
```

```
## Warning in pp.test(as.ts(lm1.diff$residuals)): p-value smaller than printed
## p-value
```

```
##
## Phillips-Perron Unit Root Test
##
```

```
## data:  as.ts(lm1.diff$residuals)
## Dickey-Fuller Z(alpha) = -926, Truncation lag parameter = 7,
## p-value = 0.01
## alternative hypothesis: stationary

kpss.test(as.ts(lm1.diff$residuals))

## Warning in kpss.test(as.ts(lm1.diff$residuals)): p-value greater than
## printed p-value

##
## KPSS Test for Level Stationarity
##
## data:  as.ts(lm1.diff$residuals)
## KPSS Level = 0.2, Truncation lag parameter = 7, p-value = 0.1
```

Modelos VAR (Vectoriales Autoregresivos)

La modelación VAR consiste en la estimación y análisis de dos, o más, series de tiempo en conjunto. Por simplicidad de exposición, se considerarán dos series de tiempo (dos ecuaciones).

Un modelo $VAR(p)$ para las series de tiempo x_t, y_t es

$$\begin{aligned}x_t &= \beta_{10} + \beta_{11}x_{t-1} + \cdots + \beta_{1p}x_{t-p} + \gamma_{11}y_{t-1} + \cdots + \gamma_{1p}y_{t-p} + u_{1,t} \\y_t &= \beta_{20} + \beta_{21}x_{t-1} + \cdots + \beta_{2p}x_{t-p} + \gamma_{21}y_{t-1} + \cdots + \gamma_{2p}y_{t-p} + u_{2,t}\end{aligned}$$

La matriz de covarianza Σ de los términos de error, $u_{1,t}, u_{2,t}$ se supone nonsingular ($\det(\Sigma) \neq 0$). Los errores pueden presentar correlación instantánea (endogeneidad):

$$\text{cov}(u_{1,t}, u_{2,t}) \neq 0$$

Supón

$$\begin{aligned}E[u_{i,t}] &= 0, \quad i = 1, 2; \quad t = 1, \dots, T; \\var(u_{i,t}) &= \sigma_i^2, \quad i = 1, 2; \quad t = 1, \dots, T; \\E[u_{i,t} u_{j,t}] &= \sigma_{ij}, \quad i = 1, 2; \\E[u_{i,t} u_{j,s}] &= 0, \quad i = 1, 2; t \neq s.\end{aligned}$$

El método de mínimos cuadrados ordinarios permite la estimación eficiente de los coeficientes del modelo.

Ventajas y desventajas de la modelación VAR

Ventajas

- La modelación VAR considera todas las variables como endógenas, por lo que no se necesita determinar variables instrumentales.
- La estimación de los modelos VAR es sencilla.
- Es fácil el pronóstico del modelo estimado.

Desventajas

- Cada variable en cada ecuación del modelo debe tener el mismo número de rezagos; por ejemplo, en un modelo VAR con dos ecuaciones y 10 rezagos, se necesitan estimar 42 parámetros.
- Los modelos VAR no están respaldados por la teoría económico-financiera, por lo que no son modelo prescriptivos.

Después de la estimación

- Pronóstico del modelo estimado
- Diagnóstico
 - Evidencia de autocorrelación serial residua
 - Evidencia de efecto arch residuo
 - normalidad de los errores
 - estabilidad estructural
- Causalidad de Granger: determinar si una variable lidera las otras.
- Función de impulso respuesta (IRF, *Impulse-Response Function*): cuánto un impulso unitario de una variable afecta a si misma y a las otras variables en el modelo.
- Decomposición de la varianza del error de pronóstico (FEVD, *Forecast Error Variance Decomposition*): cuánto un choque en una variable afecta la varianza del error de pronóstico de si misma y de las otras variables en el modelo.

Además, se pueden también determinar posibles puntos de quiebre para cada serie (cambio estructural) y para el modelo.

Estimación de un modelo VAR

En la estimación VAR se usará la librería `vars`:

A propósitos ilustrativos, se estimarán modelos $VAR(p)$ en donde la variable x_t corresponde a la serie de tiempo de S&P500 y la variable y_t a la de MorganStanley.

Bajaremos los datos de Yahoo Finance, usando la librería `quantmod`; se usarán fechas de inicios diferentes, para poder explicar sucesivamente cómo usar la función `window()`.

```
library(vars)
library(tseries)
library(quantmod)
getSymbols("MS",
           from = "2000-01-01",
           periodicity = "weekly")
```

```
## [1] "MS"
```

```
getSymbols("^GSPC",
           from = "1999-01-01",
           periodicity = "weekly")
```

```
## [1] "^GSPC"
```

Después definimos las series de tiempo asociadas (precios al cierre) y usamos la función `window()` para que ambas series de tiempo empiecen en la misma fecha:

```
ms.ts <- ts(MS$MS.Close, frequency = 52, start = c(2000,01,01),
            end = c(2019,11,1))
sp.ts <- ts(GSPC$GSPC.Close, frequency = 52, start = c(1999,01,01),
            end = c(2019,11,1))
sp.ts <- window(sp.ts, start = c(2000, 1))
```

Analizaremos los retornos de las series, por lo que

```
ms.rtn <- diff(log(ms.ts))
sp.rtn <- diff(log(sp.ts))
```

Las funciones y métodos de la librería `vars` manejan dataframes, por lo que se necesita crear un dataframe de los retornos:

```
spms.rtn <- data.frame(sp.rtn, ms.rtn)
```

En este dataframe, la primera variable corresponde a los retornos de Standard&Poor's. La estimación no se ve afectada por el orden de las variables, pero el análisis de decomposición de la varianza del error de pronóstico (FEVD) podría verse afectado. Podría ser necesario, entonces, considerar diferentes órdenes de las variables en el análisis FEVD.

Determinación del orden y estimación

La determinación del orden del modelo $VAR(p)$ se realiza especificando el orden máximo a usar:

```
VARselect(spms.rtn, lag.max = 12, type = "const")
```

```
## $selection
## AIC(n)  HQ(n)  SC(n) FPE(n)
```

```
##      7      6      2      7
##
## $criteria
##           1           2           3           4           5           6
## AIC(n) -1.34e+01 -1.35e+01 -1.35e+01 -1.35e+01 -1.35e+01 -1.35e+01
## HQ(n)  -1.34e+01 -1.35e+01 -1.34e+01 -1.34e+01 -1.34e+01 -1.35e+01
## SC(n)  -1.34e+01 -1.34e+01 -1.34e+01 -1.34e+01 -1.34e+01 -1.34e+01
## FPE(n)  1.45e-06  1.41e-06  1.42e-06  1.42e-06  1.40e-06  1.36e-06
##           7           8           9          10          11          12
## AIC(n) -1.35e+01 -1.35e+01 -1.35e+01 -1.35e+01 -1.35e+01 -1.35e+01
## HQ(n)  -1.35e+01 -1.34e+01 -1.34e+01 -1.34e+01 -1.34e+01 -1.34e+01
## SC(n)  -1.34e+01 -1.33e+01 -1.33e+01 -1.33e+01 -1.33e+01 -1.32e+01
## FPE(n)  1.35e-06  1.36e-06  1.37e-06  1.37e-06  1.37e-06  1.38e-06
```

Elegimos una estimación sin tendencia, `type = "const"`, ya que los datos son estacionarios.

La elección del orden se realiza con base en uno de los criterios de información. Para empezar, seleccionamos $p = 2$, usando el criterio de Schwarz bayesiano (SC) (más parsimonioso).

Pasamos a la estimación del modelo elegido

```
spms.var1 <- VAR(spms.rtn, p=2, type = "const")
summary(spms.var1)
```

```
##
## VAR Estimation Results:
## =====
## Endogenous variables: GSPC.Close, MS.Close
## Deterministic variables: const
## Sample size: 996
## Log Likelihood: 3884.715
## Roots of the characteristic polynomial:
## 0.395 0.395 0.255 0.255
## Call:
## VAR(y = spms.rtn, p = 2, type = "const")
##
##
## Estimation results for equation GSPC.Close:
## =====
## GSPC.Close = GSPC.Close.l1 + MS.Close.l1 + GSPC.Close.l2 + MS.Close.l2 + const
##
##           Estimate Std. Error t value Pr(>|t|)
## GSPC.Close.l1 -0.275788  0.040872  -6.75  2.6e-11 ***
## MS.Close.l1    0.087468  0.014540   6.02  2.5e-09 ***
## GSPC.Close.l2  0.017822  0.039742   0.45  0.6539
## MS.Close.l2    0.047328  0.014653   3.23  0.0013 **
## const          0.000884  0.000744   1.19  0.2354
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.0234 on 991 degrees of freedom
## Multiple R-Squared: 0.0578, Adjusted R-squared: 0.0539
## F-statistic: 15.2 on 4 and 991 DF, p-value: 4.68e-12
##
##
## Estimation results for equation MS.Close:
## =====
## MS.Close = GSPC.Close.l1 + MS.Close.l1 + GSPC.Close.l2 + MS.Close.l2 + const
##
##              Estimate Std. Error t value Pr(>|t|)
## GSPC.Close.l1 -0.117769   0.115858  -1.02   0.310
## MS.Close.l1   -0.192404   0.041216  -4.67 3.5e-06 ***
## GSPC.Close.l2 -0.192788   0.112653  -1.71  0.087 .
## MS.Close.l2   0.058740   0.041535   1.41  0.158
## const         -0.000329   0.002110  -0.16  0.876
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.0665 on 991 degrees of freedom
## Multiple R-Squared: 0.0519, Adjusted R-squared: 0.0481
## F-statistic: 13.6 on 4 and 991 DF, p-value: 9.03e-11
##
##
##
## Covariance matrix of residuals:
##              GSPC.Close MS.Close
## GSPC.Close    0.00055  0.00100
## MS.Close      0.00100  0.00442
##
## Correlation matrix of residuals:
##              GSPC.Close MS.Close
## GSPC.Close    1.000    0.645
## MS.Close      0.645    1.000

```

La matriz de correlación da una medida de la interacción instantánea entre las dos variables.

Diagnóstico

El diagnóstico del modelo estimado considera las siguientes pruebas:

- Pruebas de ruido blanco sobre los residuales:
 - Prueba portmanteau(Ljung-Box; Box-Pierce)
 - Prueba de Breusch-Godfrey de correlación serial
- Prueba de efectos ARCH
- Prueba de normalidad
- Prueba de estabilidad

Pruebas de ruido blanco Se considera que los residuales cumplen el supuesto de ruido blanco si no se rechaza la hipótesis nula

$$H_0 : \tau_1 = 0, \dots, \tau_s = 0$$

$$H_1 : \tau_1 \neq 0, \dots, \tau_s \neq 0$$

con $s > 1$.

Prueba portmanteau El estadístico de prueba es una χ_m^2 , con m el número de grados de libertad del modelo. Para muestras pequeñas, es disponible también un estadístico corregido.

```
serial.test(spms.var1, lags.pt = 16, type = "PT.asymptotic")
```

```
##
##  Portmanteau Test (asymptotic)
##
## data:  Residuals of VAR object spms.var1
## Chi-squared = 133, df = 56, p-value = 3e-08
```

```
serial.test(spms.var1, lags.pt = 16, type = "PT.adjusted")
```

```
##
##  Portmanteau Test (adjusted)
##
## data:  Residuals of VAR object spms.var1
## Chi-squared = 134, df = 56, p-value = 2e-08
```

Existe evidencia de autocorrelación serial (los residuales no son ruido blanco).

Prueba de Breusch-Godfrey Es una prueba ML, con estadístico de prueba es una χ_n^2 , n número de grados de libertad. También hay un estadístico ajustado para muestra pequeña (prueba F de Edgerton-Shukur):

```
serial.test(spms.var1, lags.bg = 15, type = "BG")
```

```
##
##  Breusch-Godfrey LM test
##
## data:  Residuals of VAR object spms.var1
## Chi-squared = 136, df = 60, p-value = 8e-08
```

```
serial.test(spms.var1, lags.bg = 15, type = "ES")
```

```
##  
## Edgerton-Shukur F test  
##  
## data: Residuals of VAR object spms.var1  
## F statistic = 2, df1 = 60, df2 = 1920, p-value = 5e-08
```

Después de las cuatro pruebas, concluimos que existe fuerte evidencia de autocorrelación serial. Para eliminarla, se aumentará el orden del modelo VAR. Se usará $p = 6$, como indica el criterio de información de Hannan-Quinn; después, pasaremos al diagnóstico de autocorrelación serial.

```
spms.var6 <- VAR(spms.rtn, p=6, type = "const")
```

Prueba portmanteau El estadístico de prueba es una χ_m^2 , con m el número de grados de libertad del modelo. Para muestras pequeñas, es disponible también un estadístico corregido.

```
serial.test(spms.var6, lags.pt = 16, type = "PT.asymptotic")
```

```
##  
## Portmanteau Test (asymptotic)  
##  
## data: Residuals of VAR object spms.var6  
## Chi-squared = 66, df = 40, p-value = 0.006
```

```
serial.test(spms.var6, lags.pt = 16, type = "PT.adjusted")
```

```
##  
## Portmanteau Test (adjusted)  
##  
## data: Residuals of VAR object spms.var6  
## Chi-squared = 67, df = 40, p-value = 0.005
```

Existe evidencia de autocorrelación serial (los residuales no son ruido blanco).

Prueba de Breusch-Godfrey Es una prueba ML, con estadístico de prueba es una χ_n^2 , n número de grados de libertad. También hay un estadístico ajustado para muestra pequeña (prueba F de Edgerton-Shukur):

```
serial.test(spms.var6, lags.bg = 15, type = "BG")
```

```
##  
## Breusch-Godfrey LM test  
##  
## data: Residuals of VAR object spms.var6  
## Chi-squared = 115, df = 60, p-value = 2e-05
```

```
serial.test(spms.var6, lags.bg = 15, type = "ES")
```

```
##  
##  Edgerton-Shukur F test  
##  
## data:  Residuals of VAR object spms.var6  
## F statistic = 2, df1 = 60, df2 = 1896, p-value = 2e-05
```

Todavía se encuentra evidencia de autocorrelación serial; incrementando sucesivamente el orden del modelo VAR, se encuentra que el modelo $VAR(15)$ es adecuado. Se presenta a continuación el diagnóstico de ruido blanco y, después, las otras pruebas del diagnóstico.

```
spms.var15 <- VAR(spms.rtn, p=15, type = "const")
```

Prueba portmanteau El estadístico de prueba es una χ_m^2 , con m el número de grados de libertad del modelo. Para muestras pequeñas, es disponible también un estadístico corregido.

```
serial.test(spms.var15, lags.pt = 20, type = "PT.asymptotic")
```

```
##  
##  Portmanteau Test (asymptotic)  
##  
## data:  Residuals of VAR object spms.var15  
## Chi-squared = 28, df = 20, p-value = 0.1
```

```
serial.test(spms.var15, lags.pt = 20, type = "PT.adjusted")
```

```
##  
##  Portmanteau Test (adjusted)  
##  
## data:  Residuals of VAR object spms.var15  
## Chi-squared = 29, df = 20, p-value = 0.09
```

Existe evidencia de autocorrelación serial (los residuales no son ruido blanco).

Prueba de Breusch-Godfrey Es una prueba ML, con estadístico de prueba es una χ_n^2 , n número de grados de libertad. También hay un estadístico ajustado para muestra pequeña (prueba F de Edgerton-Shukur):

```
serial.test(spms.var15, lags.bg = 15, type = "BG")
```

```
##  
##  Breusch-Godfrey LM test  
##  
## data:  Residuals of VAR object spms.var15  
## Chi-squared = 96, df = 60, p-value = 0.002
```

```
serial.test(spms.var15, lags.bg = 15, type = "ES")
```

```
##
```

```
## Edgerton-Shukur F test
##
## data: Residuals of VAR object spms.var15
## F statistic = 2, df1 = 60, df2 = 1842, p-value = 0.004
```

Consideramos que el orden $p = 15$ es adecuado, aún si las pruebas de Breusch-Godfrey indican cierta autocorrelación serial residua.

Prueba de normalidad

Realizaremos la prueba de normalidad para cada ecuación y para el modelo vectorial.

```
normality.test(spms.var15, multivariate.only = F)
```

```
## $GSPC.Close
##
## JB-Test (univariate)
##
## data: Residual of GSPC.Close equation
## Chi-squared = 3794, df = 2, p-value <2e-16
##
##
## $MS.Close
##
## JB-Test (univariate)
##
## data: Residual of MS.Close equation
## Chi-squared = 67581, df = 2, p-value <2e-16
##
##
## $JB
##
## JB-Test (multivariate)
##
## data: Residuals of VAR object spms.var15
## Chi-squared = 14804, df = 4, p-value <2e-16
##
##
## $Skewness
##
## Skewness only (multivariate)
##
## data: Residuals of VAR object spms.var15
## Chi-squared = 409, df = 2, p-value <2e-16
##
##
## $Kurtosis
```

```
##
## Kurtosis only (multivariate)
##
## data: Residuals of VAR object spms.var15
## Chi-squared = 14395, df = 2, p-value <2e-16
```

Existe evidencia que los errores no son normales, aunque el tamaño muestral permite el uso de normalidad asintótica,

Efecto arch residuo La prueba de heterocedasticidad se realiza usando la función *arch.test()*

```
arch.test(spms.var15,
          lags.single = 10, lags.multi = 10, multivariate.only = F)
```

```
## $GSPC.Close
##
## ARCH test (univariate)
##
## data: Residual of GSPC.Close equation
## Chi-squared = 72, df = 10, p-value = 2e-11
##
##
## $MS.Close
##
## ARCH test (univariate)
##
## data: Residual of MS.Close equation
## Chi-squared = 132, df = 10, p-value <2e-16
##
##
##
## ARCH (multivariate)
##
## data: Residuals of VAR object spms.var15
## Chi-squared = 935, df = 90, p-value <2e-16
```

Existe evidencia que el modelo es heterocedástico. Esta observación será importante en la prueba de causalidad de Granger.

Causalidad según Granger

Considérese el modelo $VAR(p)$

$$x_t = \beta_{10} + \beta_{11}x_{t-1} + \cdots + \beta_{1p}x_{t-p} + \gamma_{11}y_{t-1} + \cdots + \gamma_{1p}y_{t-p} + u_{1,t}$$

$$y_t = \beta_{20} + \beta_{21}x_{t-1} + \cdots + \beta_{2p}x_{t-p} + \gamma_{21}y_{t-1} + \cdots + \gamma_{2p}y_{t-p} + u_{2,t}$$

Definición Se dice que la serie y_t causa (según Granger) la variable x_t si los rezagos de y_t explican el valor actual de x_t .

Hay una definición semejante para x_t que causa y_t .

Para probar si y_t causa x_t , se plantea la siguiente prueba de hipótesis

$$H_0 : \gamma_{11} = 0, \dots \gamma_{1p} = 0$$

$$H_1 : \gamma_{11} \neq 0, \dots \gamma_{1p} \neq 0$$

La hipótesis nula es equivalente a decir que los rezagos de y_t no explican el valor actual de x_t .

Si se rechaza la hipótesis nula, entonces se dice que y_t causa x_t .

Se prueban hipótesis estadísticas semejantes para determinar si

- los rezagos de la variable x_t explican el valor actual de y_t ;
- los rezagos de la variable y_t (resp., x_t) explican el valor actual de la misma variable y_t (resp., x_t).

En la prueba de causalidad de Granger se puede usar el estadístico F cuando las series son estacionarias.

Comentarios

- Vale la pena subrayar que la causalidad según Granger no implica que una variable es causa de la otra. Analizando activos financieros, se puede decir, más bien, que un activo “lidera” (o anticipa) el otro en los cambios de sus valores.
- Si x_t causa y_t , y viceversa, entonces se dice que existe “feedback” bidireccional.
- Si, por ejemplo, x_t causa y_t , pero y_t no causa x_t , se dice que x_t es fuertemente exógena (en la ecuación para y_t).
- Si ninguna variable causa la otra, se dice que x_t , y_t son independientes.

Pasaremos ahora a determinar si alguna variable explica la otra.

Empezamos con determinar si los retornos de Morgan&Stanley explican los retornos de Standard&Poor's; a fines ilustrativos, usaremos dos variantes de errores robustos a heterocedasticidad.

```
causality(spms.var15, cause = "MS.Close", vcov.=vcovHC(spms.var15))
```

```
## $Granger
##
## Granger causality H0: MS.Close do not Granger-cause GSPC.Close
##
## data: VAR object spms.var15
## F-Test = 2, df1 = 15, df2 = 1904, p-value = 0.04
##
##
```

```
## $Instant
##
## H0: No instantaneous causality between: MS.Close and GSPC.Close
##
## data: VAR object spms.var15
## Chi-squared = 291, df = 1, p-value <2e-16
causality(spms.var15, cause = "MS.Close", vcov.=vcovHC(spms.var15))
```

```
## $Granger
##
## Granger causality H0: MS.Close do not Granger-cause GSPC.Close
##
## data: VAR object spms.var15
## F-Test = 2, df1 = 15, df2 = 1904, p-value = 0.04
##
##
## $Instant
##
## H0: No instantaneous causality between: MS.Close and GSPC.Close
##
## data: VAR object spms.var15
## Chi-squared = 291, df = 1, p-value <2e-16
```

Podemos rechazar la hipótesis nula, con ambas variantes, y concluir que sí, los retornos de Morgan&Stanley explican los retornos de Standard&Poor's.

Probamos ahora si los retornos de Standard&Poor's explican los retornos de Morgan&Stanley:

```
causality(spms.var15, cause = "GSPC.Close", vcov.=vcovHC(spms.var15))

## $Granger
##
## Granger causality H0: GSPC.Close do not Granger-cause MS.Close
##
## data: VAR object spms.var15
## F-Test = 0.2, df1 = 15, df2 = 1904, p-value = 1
##
##
## $Instant
##
## H0: No instantaneous causality between: GSPC.Close and MS.Close
##
## data: VAR object spms.var15
## Chi-squared = 291, df = 1, p-value <2e-16
```

```
causality(spms.var15, cause = "GSPC.Close", vcov.=vcovHC(spms.var15))
```

```
## $Granger
##
##   Granger causality H0: GSPC.Close do not Granger-cause MS.Close
##
## data:   VAR object spms.var15
## F-Test = 0.2, df1 = 15, df2 = 1904, p-value = 1
##
##
## $Instant
##
##   H0: No instantaneous causality between: GSPC.Close and MS.Close
##
## data:   VAR object spms.var15
## Chi-squared = 291, df = 1, p-value <2e-16
```

En este caso, no hay evidencia que los retornos de Standard&Poor's expliquen los retornos de Morgan&Stanley. Concluimos que los retornos de Morgan&Stanley son fuertemente exógenos.

Respuesta a impulso unitario

En la estimación de un modelo $VAR(p)$, la causalidad de Granger permite determinar si una variable depende de los rezagos de las variables en el modelo, pero no cuantifica esta relación.

La respuesta a impulso unitario permite la cuantificación de esta relación mediante la respuesta del modelo $VAR(p)$ a choques (unitarios) aplicados a cada variable del modelo.

En R, la función `irf()` permite la estimación de la respuesta al impulso unitario. Se puede pasar un parámetro booleano adicional, `ortho`. Cuando `ortho = TRUE`, entonces la estimación usa la decomposición de Cholesky de la matriz de covarianza; en este caso, el orden de las variables es importante.

Usando el parámetro `runs = n` se calcula un intervalo de confianza de la respuesta al impulso, usando una distribución de probabilidad empírica.

Empezamos con el impulso aplicado a la variable *MS.Close*:

```
irf(spms.var15, impulse = "MS.Close",
    response = "GSPC.Close", n.ahead = 5, ortho = FALSE,
    runs = 1000)
```

```
##
## Impulse response coefficients
## $MS.Close
##       GSPC.Close
## [1,]      0.00000
```



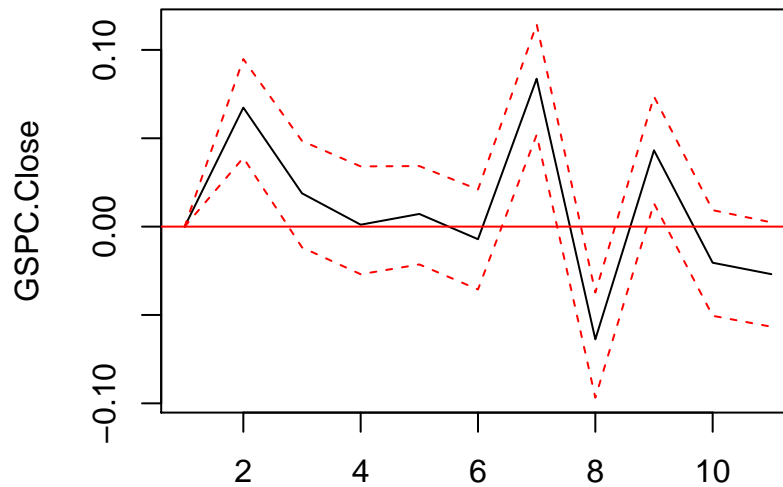
```

## [2,]    0.06736
## [3,]    0.01884
## [4,]    0.00108
## [5,]    0.00717
## [6,]   -0.00713
##
##
## Lower Band, CI= 0.95
## $MS.Close
##      GSPC.Close
## [1,]    0.0000
## [2,]    0.0363
## [3,]   -0.0110
## [4,]   -0.0285
## [5,]   -0.0220
## [6,]   -0.0383
##
##
## Upper Band, CI= 0.95
## $MS.Close
##      GSPC.Close
## [1,]    0.0000
## [2,]    0.0997
## [3,]    0.0478
## [4,]    0.0301
## [5,]    0.0379
## [6,]    0.0223

plot(irf(spms.var15, impulse = "MS.Close",
        response = "GSPC.Close", n.ahead = 10, ortho = FALSE,
        runs = 1000))

```

Impulse Response from MS.Close



95 % Bootstrap CI, 1000 runs

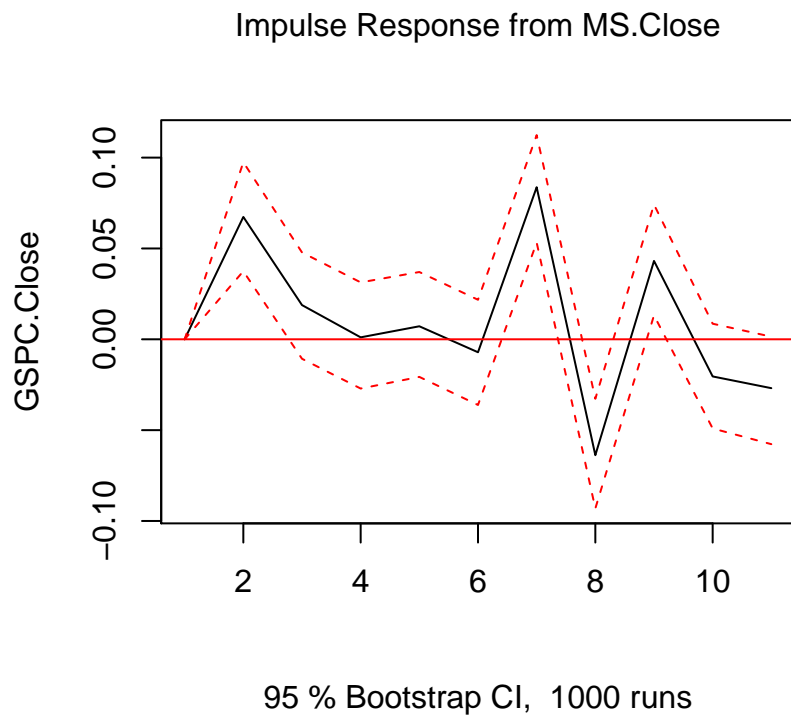
El impulso aplicado a la variable *GSPC.Close*:

```
irf(spms.var15, impulse = "MS.Close",
    response = "GSPC.Close", n.ahead = 5, ortho = FALSE,
    runs = 1000)
```

```
##
## Impulse response coefficients
## $MS.Close
##      GSPC.Close
## [1,] 0.00000
## [2,] 0.06736
## [3,] 0.01884
## [4,] 0.00108
## [5,] 0.00717
## [6,] -0.00713
##
##
## Lower Band, CI= 0.95
## $MS.Close
##      GSPC.Close
## [1,] 0.00000
```

```
## [2,]      0.0410
## [3,]     -0.0078
## [4,]     -0.0271
## [5,]     -0.0213
## [6,]     -0.0355
##
##
## Upper Band, CI= 0.95
## $MS.Close
##      GSPC.Close
## [1,]      0.0000
## [2,]      0.0959
## [3,]      0.0454
## [4,]      0.0330
## [5,]      0.0338
## [6,]      0.0224
```

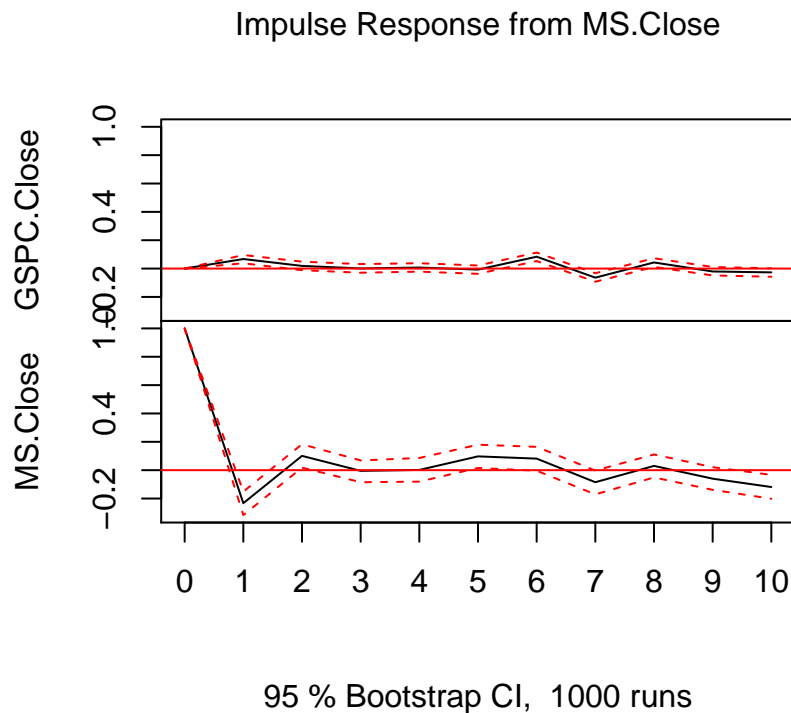
```
plot(irf(spms.var15, impulse = "MS.Close",
         response = "GSPC.Close", n.ahead = 10, ortho = FALSE,
         runs = 1000))
```



El impacto de los retornos de Morgan&Stanley sobre Standard&Poor's no es tan pronunciado.

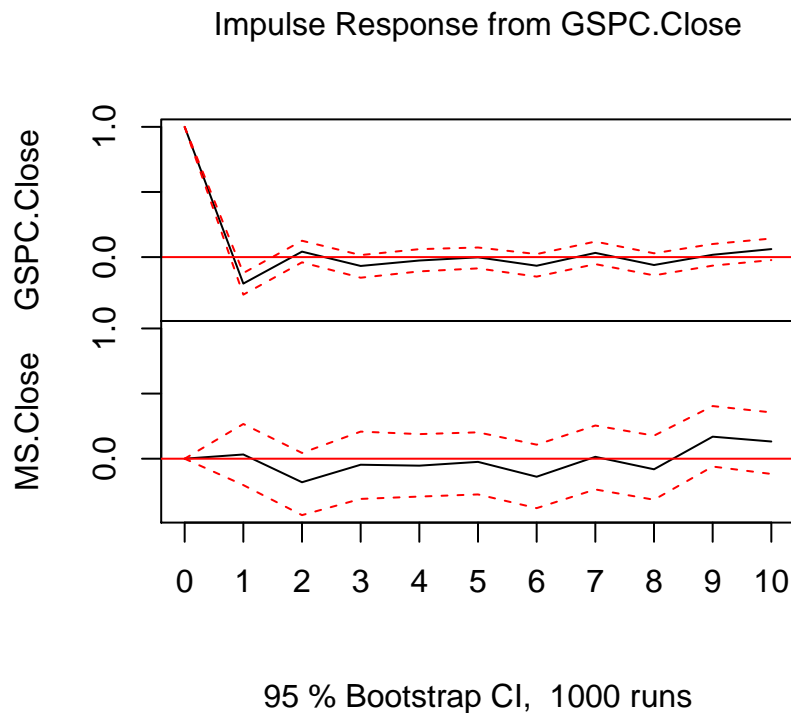
También se puede analizar la respuesta al impulso unitario que la variable *MS.Close* tiene sobre si misma y sobre la otra variable.

```
plot(irf(spms.var15, impulse = "MS.Close", n.ahead = 10, ortho = FALSE,
        runs = 1000))
```



Y la respuesta al impulso que la variable *GSPC.Close* tiene sobre si misma y sobre la otra variable.

```
plot(irf(spms.var15, impulse = "GSPC.Close", n.ahead = 10, ortho = FALSE,
        runs = 1000))
```



Descomposición del error del error de pronóstico

Usando la descomposición de Cholesky en la matriz de covarianza, podemos también determinar el porcentaje que tiene una variable sobre el error de pronóstico, sobre la misma variable y sobre la otra.

En R, usaremos la función `fevd()`.

```
fevd(spms.var15, n.ahead = 10)
```

```
## $GSPC.Close
##      GSPC.Close MS.Close
## [1,]      1.000  0.0000
## [2,]      0.979  0.0206
## [3,]      0.978  0.0220
## [4,]      0.978  0.0219
## [5,]      0.978  0.0222
## [6,]      0.978  0.0224
## [7,]      0.948  0.0518
## [8,]      0.932  0.0677
## [9,]      0.925  0.0749
```

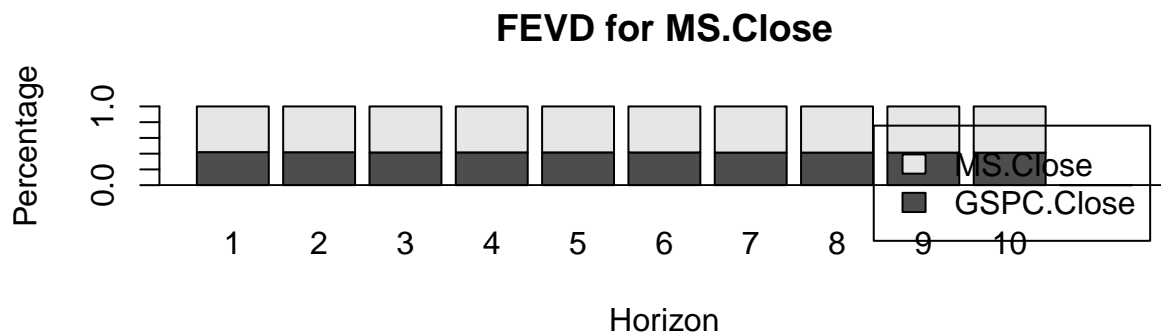
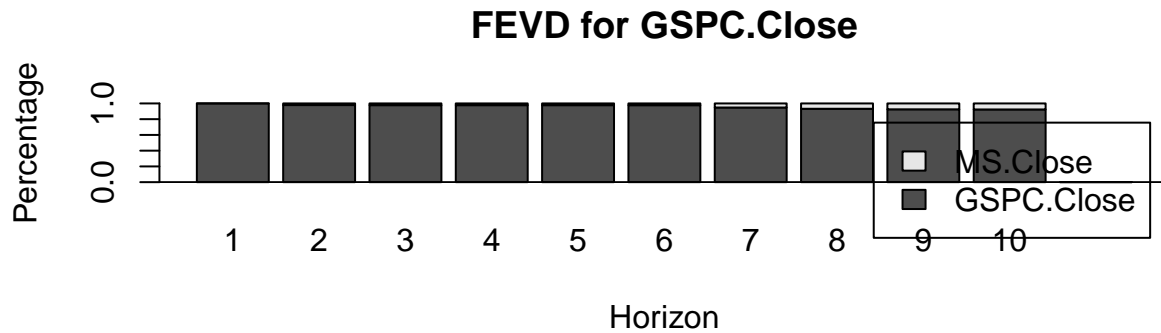
```
## [10,]      0.924    0.0765
##
## $MS.Close
##      GSPC.Close MS.Close
## [1,]      0.420    0.580
## [2,]      0.418    0.582
## [3,]      0.416    0.584
## [4,]      0.416    0.584
## [5,]      0.416    0.584
## [6,]      0.415    0.585
## [7,]      0.414    0.586
## [8,]      0.414    0.586
## [9,]      0.414    0.586
## [10,]     0.413    0.587
```

Podemos concluir que el 92.4% de la varianza del error de pronóstico de *GSPC.Close* es debido a la misma variable; el remanente a los retornos de Morgan&Stanley.

Por otro lado, el 41.3% de la varianza del error de pronóstico de *MS.Close* es debido a los retornos de Standard&Poor's, con el remanente debido a *MS.Close*.

La misma información se puede observar de manera gráfica:

```
plot(fevd(spms.var15, n.ahead = 10))
```



Bibliografía

- Br** Brooks C, *Introductory Econometrics for Finance*, Third edition, Cambridge University Press, 2014.
- BP** Box GEP, Pierce DA., "Distribution of residual autocorrelation in autoregressive-integrated moving average time series models." *J Am Stat Assoc.* 1970;65:1509–1526. doi: 10.1080/01621459.1970.10481180.
- KPSS** Kwiatkowski D, Phillips PCB, Schmidt P., Shin Y, "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root". *Journal of Econometrics.* 1992, 54:159–178.
- LB** Ljung GM, Box GEP, "On a measure of lack of fit in time series models." *Biometrika.* 1978; 65:297–303. doi: 10.1093/biomet/65.2.297.
- M** Mahdi E, "Portmanteau test statistics for seasonal serial correlation in time series models", *Springerplus.* 2016;5(1):1485. Published 2016 Sep 5. doi:10.1186/s40064-016-3167-4, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5011475/>.
- ShS** Shumway R.H., Stoffer, D.S., *Time Series and Its Applications with R Examples*, Third edition, Springer Text in Statistics, Springer, 2010.

- T** Tsay R.S., *An introduction to Analysis of Financial Data with R*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, New Jersey, 2013.
- Y** Yule GU, "Why do we Sometimes get Non-Sense Correlations between Time-Series?—A Study in Sampling and the Nature of Time-Series", Institute of the Royal Statistical Society, Vol. 89, No. 1 (January 1926), pp. 1-63