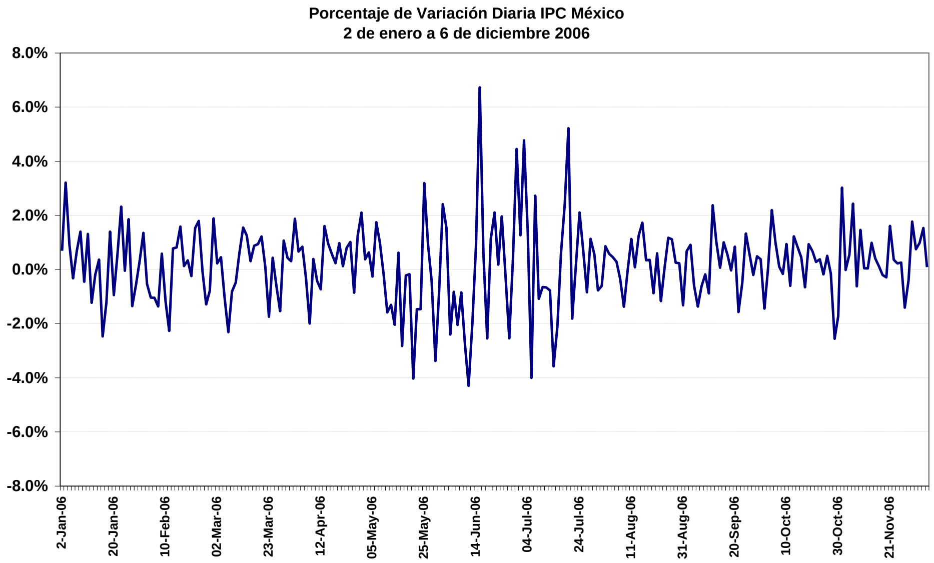
MODELOS ARCH Y GARCH

1 Introducción.

Los mercados financieros y muchas variables económicas, suelen presentar períodos de gran variabilidad seguidos de períodos de relativa calma. Esta condición de comportamiento de la varianza, observada desde mucho tiempo atrás, identificada como volatilidad, fue motivo de diversos análisis. Los modelos actuales tienen su origen en el trabajo publicado por R.F. Engle en 1982 en la revista Econométrica *"Autorregresive Condicio-nal Heterocedasticity with Estimates of de U.K. Inflation"* de donde procede el término ARCH, que significa Modelos Autorregresivos Condicionalmente Heteroscedástico y a partir de este han surgido diversas variaciones como GARCH, EARCH, IARCH, etc.

En la siguiente gráfica se presentan los cambios porcentuales diarios en los puntajes de cierre ajustado del Indice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (ROI), correspondientes al período del 2 de enero de 2006 al 6 de diciembre del mismo año. Destaca el período del 15 de mayo al 13 de julio por su elevada volatilidad con un mínimo de decremento porcentual con -4.3% el 12 de junio y el máximo 3 días después con un incremento porcentual respecto al día anterior de 6.73% el 15 de junio. Desde luego este período de gran volatilidad se explica por la crisis política en torno período de elecciones presidenciales.

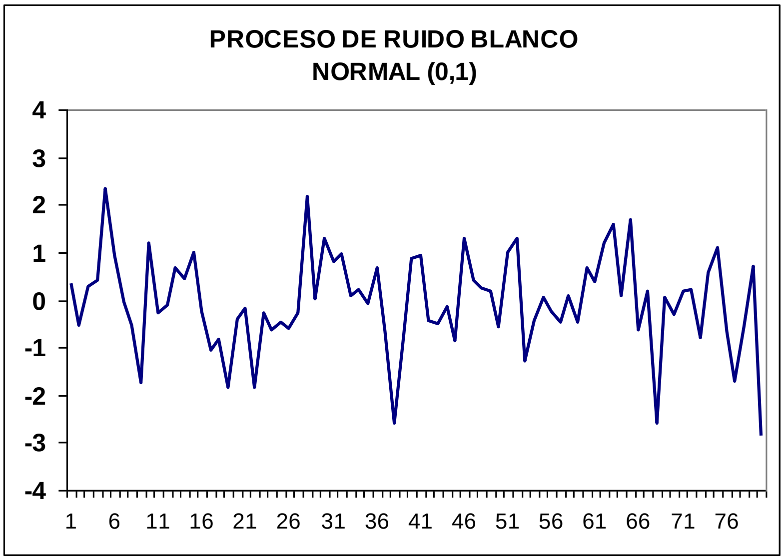


Los modelos autorregresivos, integrados y de medias móviles (ARIMA) que se han revisado suponen para su caracterización un proceso estacionario en el cual la media y la varianza permanecen constantes y las covarianzas entre observaciones separadas por períodos del mismo tamaño son iguales. En situación de volatilidad tomar decisiones en función de la varianza de largo plazo resulta inapropiado y entonces los modelos tradicionales no funcionan. Las suposiciones de un modelo ARIMA se resumen en las siguientes expresiones.

$$E(Z_t) = E(Z_{t+k}) = \mu$$
$$V(Z_t) = V(Z_{t+k}) = \sigma_Z^2$$
$$Cov(Z_t, Z_{t+k}) = Cov(Z_{t+m}, Z_{t+k+m})$$

Como caso particular de este tipo de procesos se puede mencionar el de "ruido blanco".

$$E(a_t) = E(a_{t+k}) = 0$$
$$V(a_t) = V(a_{t+k}) = \sigma^2$$
$$Cov(a_t, a_{t+k}) = Cov(a_{t+m}, a_{t+k+m}) = 0$$



Se argumenta que el hecho de que el ruido blanco no presente dependencia lineal no excluye la presencia de dependencia no lineal y puede existir por ejemplo dependencia recíproca, cuadrática, polinomial, exponencial, etc.

2 Modelos para Varianza

Los modelos referidos tratan de modelar la esperanza condicional y la varianza condicional simultáneamente, para tomar en consideración las altas y bajas en la volatilidad. La esperanza condicional se refiere hacia el futuro pero sujeta a la información acumulada hasta el tiempo t y la no condicional hace caso omiso de ello.

Una forma de modelar la varianza es la introducción de una variable independiente Xt en el modelo que ayude a predecir la volatilidad. Por ejemplo el modelo:

$$Z_{t+1} = a_t X_t$$

Si Xt = M constante, entonces la varianza de Zt es constante, la varianza del proceso resulta constante

$$V(Z_t) = M^2 \sigma^2$$

Pero si Xt no es constante, entonces se tiene una varianza que cambia en cada período y la varianza condicional de Zt depende de los valores de Xt.

$$V(Z_t) = X_{t-1}^2 \sigma^2$$

Si Xt presenta aurtocorrelación positiva, entonces valores elevados de Xt tenderán a ser seguidos por valores elevados de Xt y como consecuencia intervalos de varianza elevada y baja. Desde luego un problema fundamental es cómo determinar cual es la variable independiente Xt que conviene y si está disponible.

En lugar de buscar una variable externa, Engle demuestra que es posible modelar simultáneamente la media y la varianza de las series. Hay que considerar que los pronósticos condicionados son superiores a los no condicionados. Suponga un modelo estacionario AR(1) del tipo $Z_t = \alpha_0 + \alpha_1 Z_{t-1} + a_t$. Si se pretende pronosticar Z_{t+1} su esperanza condicional dado Z_t es la combinación lineal:

$$E\left(Z_{t+1} / Z_t\right) = \alpha_0 + \alpha_1 Z_t$$

La varianza condicional de esta esperanza es la varianza del error.

$$V\left(Z_{t+1} / Z_t\right) = E\left[\left(Z_{t+1} - E(Z_{t+1} / Z_t)\right)^2\right]$$

$$= E\left[(Z_{t+1} - \alpha_0 - \alpha_1 Z_t)^2\right]$$

$$= E\left[\left(\alpha_0 + \alpha_1 Z_t + a_t - \alpha_0 - \alpha_1 Z_t\right)^2\right]$$

$$= E\left(a_t^2\right) = \sigma^2$$

Si se utiliza la esperanza y varianza no condicional entonces

$$E(Z_t) = E(\alpha_0 + \alpha_1 Z_{t-1} + a_t)$$

$$E(Z_t) = \alpha_0 + \alpha_1 E(Z_{t-1}) + E(a_t) = \alpha_0 + \alpha_1 E(Z_{t-1})$$
$$E(Z_t) - \alpha_1 E(Z_{t-1}) = \alpha_0$$

$$E(Z_t) = \frac{\alpha_o}{1-\alpha_1}$$

De manera análoga se obtiene la varianza no condicional

$$V(Z_t) = V(\alpha_0 + \alpha_1 Z_{t-1} + a_t)$$

$$V(Z_t) = \alpha_0 + V(\alpha_1 Z_{t-1}) + V(a_t)$$

$$V(Z_t) - \alpha_1^2 V(Z_{t-1}) = \sigma^2$$

Por coincidir las varianzas no condicionales de Z_t y Z_{t-1} se concluye:

$$V(Z_t) = \frac{\sigma^2}{1-\alpha_1^{\;2}}$$

Como el denominador de esta última expresión es menor que 1, entonces la varianza de la serie es mayor a la varianza del error.

3 Retorno de la Inversión (ROI)

Los activos financieros son activos intangibles. Son instrumentos financieros emitidos por entidades económicas empresas o gobierno, que otorgan a su comprador el derecho a recibir ingresos futuros generados por los activos reales del emisor. Se pueden considerar como certificados de deuda, es decir una promesa de pago por un determinado monto y plazo. Los montos y plazos pueden ser variables, ejemplo de ello son las acciones de las empresas que cotizan en las bolsas de valores, por las cuales los tenedores de las acciones pueden recibir dividendos en función de las utilidades o pérdidas, pues en este caso no hay certeza.

Consideremos que Pt, t=1,2,3,….. es el precio de una acción al tiempo t. Su tenencia en el período t-1 al t producirá un retorno simple definido como sigue. Los retornos son razones carentes de unidades que se pueden interpretar como porcentajes de utilidades o pérdidas según el caso.

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1$$

4 Definición de Volatilidad

La volatilidad se define como la varianza condicional de la serie subyacente. En las series financieras se pretende modelar la volatilidad de los retornos de inversión (ROI). Las series de retornos son en general estacionarias, pero su varianza suele presentar períodos de incremento, seguidas por períodos de menor varianza, son heteroscedásticas. Esta característica hace inadecuados los modelos usuales de series de tiempo (ARIMA) que suponen varianza constante. **En general se tiene interés para predicciones de corto plazo. Resultan de**

mayor interés para los inversionistas las predicciones de medias y varianzas condicionales que se tienen en pronósticos de corto plazo. Estas predicciones se utilizan regularmente en la administración de riesgos y análisis de portafolios de inversión.

Un buen modelo para la volatilidad debe tener la capacidad de pronosticarla; por lo tanto debe modelar sus características.

Estos pronósticos y estimaciones son utilizados en diversas actividades financieras: manejo de riesgo, selección de portafolios, posiciones cortas y largas en la tenencia de un activo, entre otras.

Un buen modelo para la volatilidad de los retornos debe reflejar las siguientes características:

Aglomeración de la volatilidad.- La volatilidad tiene tendencia aparecer agrupada por períodos, es decir, que la volatilidad puede ser alta durante un período y baja durante otro. A grandes cambio en la volatilidad siguen cambios grandes; a pequeños cambios siguen pequeños cambios de volatilidad. Este comportamiento ha sido reportado en varios estudios.

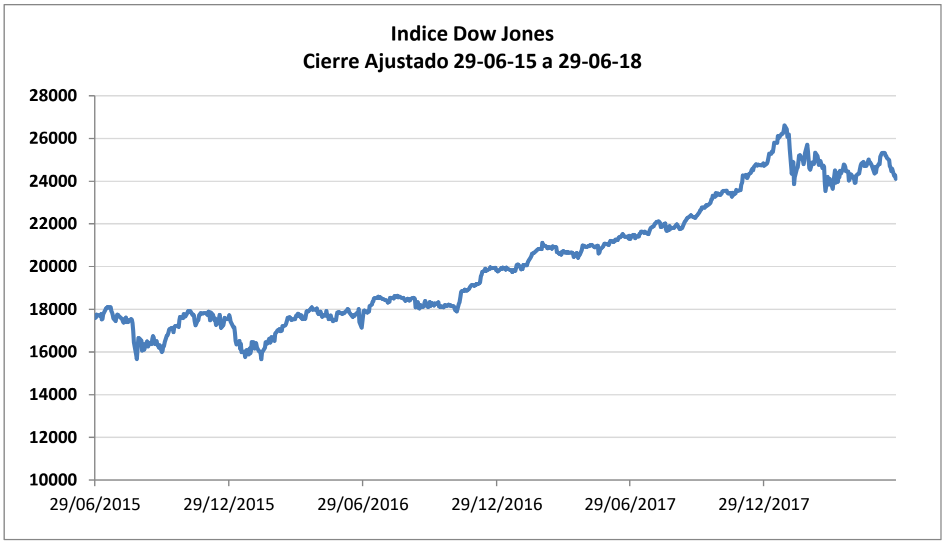
Reversión a la media. La aglomeración de la volatilidad implica que a un período de alta volatilidad, eventualmente, seguirá otro de volatilidad normal, y a un período de baja volatilidad seguirá otro de alta volatilidad la cual está retorna eventualmente. Los pronósticos a largo plazo convergerán todos al nivel normal de la volatilidad, sin importar cuando fueron hechos.

La volatilidad es asimétrica. La volatilidad se comporta diferente frente a innovaciones positivas y negativas. No reacciona de la misma manera frente a una gran alza en el precio que frente a una caída en su precio.

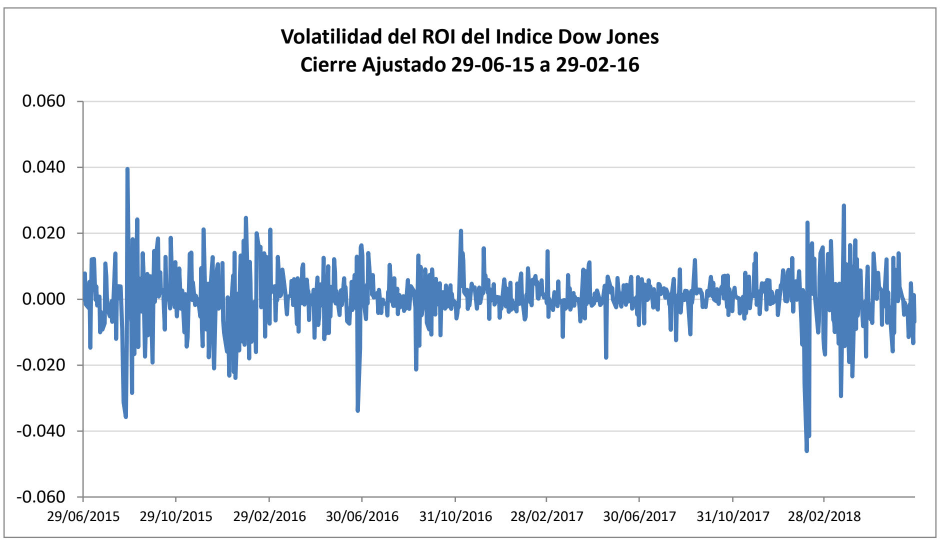
Influencia de variables exógenas. No se considera que los precios de los activos financieros evolucionan independientemente de los mercados alrededor de éstos y, por ello, se espera que existan variables que contengan información relevante para la información de la serie. Un ejemplo sería al reportar el PIB en USA, como influye en las bolsas de inversión como el Dow Jones

Distribución de probabilidad. La distribución de probabilidad de los retornos tiene colas pesadas y, en general, exceso de curtosis.

Como ejemplo, la siguiente serie corresponde al valor del Índice de la Bolsa de New York, Dow Jones del 29-06-15 al 29-06-18



Los retornos de inversión diarios (ROI), calculados para la misma serie del Dow Jones se presentan a continuación. Observe los periodos de alta volatilidad al inicio y final de la serie. Estos último, consideramos, motivados por la "Guerra Comercial" iniciada en 2018 por Estados Unidos.



5 Modelo ARCH

El modelo más sencillo propuesto por Engle es un modelo ARCH(1) en el que interviene at "ruido blanco", con media 0 y varianza no condicional constante igual a σ^2 =1 y la media de Zt condicional es constante w igual a 0 y su varianza condicional es variable.

$$Z_t = a_t \sqrt{\alpha_0 + \alpha_1 Z_{t-1}^2}$$

Cuyos valores esperados por independencia de a_t es la siguiente:

$$E(Z_{t+1}) = E(a_t)E\left(\sqrt{\alpha_0 + \alpha_1 Z_{t-1}^2}\right) = 0 * E\left(\sqrt{\alpha_0 + \alpha_1 Z_{t-1}^2}\right) = 0$$

Para la varianza no condicional se procede a continuación

$$Z_t = a_t \sqrt{\alpha_0 + \alpha_1 Z_{t-1}^2}$$

Se eleva al cuadrado y puesto que la media de Zt es cero y $a_t \rightarrow N(0,1)$

$$Z_t^2 = a_t^2 \left(\alpha_0 + \alpha_1 Z_{t-1}^2\right)$$

$$V(Z_t) = E(Z_t^2) = E(a_t^2) E\left(\alpha_0 + \alpha_1 Z_{t-1}^2\right)$$

$$E(Z_t^2) = E(a_t^2)\left(\alpha_0 + \alpha_1 E\left(Z_{t-1}^2\right)\right)$$

$$E(Z_t^2) = \alpha_0 + \alpha_1 E\left(Z_{t-1}^2\right)$$

$$E(Z_t^2) - \alpha_1 E\left(Z_{t-1}^2\right) = \alpha_0$$

$$E(Z_t^2)(1 - \alpha_1) = \alpha_0$$

$$V(Z_t) = \frac{\alpha_0}{1 - \alpha_1}$$

Por su parte la esperanza condicional de Zt también es cero

$$E(Z_{t+1} / Z_t) = 0$$

Una vez dado el valor de Zt, éste se trata como constante y la varianza condicional adopta la siguiente forma:

$$V(Z_{t+1} / Z_t) = (\alpha_0 + \alpha_1 Z^2{}_t)\sigma^2$$

Donde **a_t** se recuerda que es ruido blanco N(0,1) y se tiene σ^2 =1 no correlacionado y **Z_t** es un proceso estacionario.

En términos más generales un modelo ARCH(q) tiene la forma:

$$Z_t = a_t \sigma_t$$

Donde $\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i Z_{t-i}^2$

$$Z_t = a_t \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i Z_{t-i}^2}$$

$$E(Z_t) = 0_t$$

La varianza incondicional de Zt tiene la siguiente fórmula.

$$V(Z_t) = \frac{\alpha_o}{1 - \sum_{i=1}^{q} \alpha_i}$$

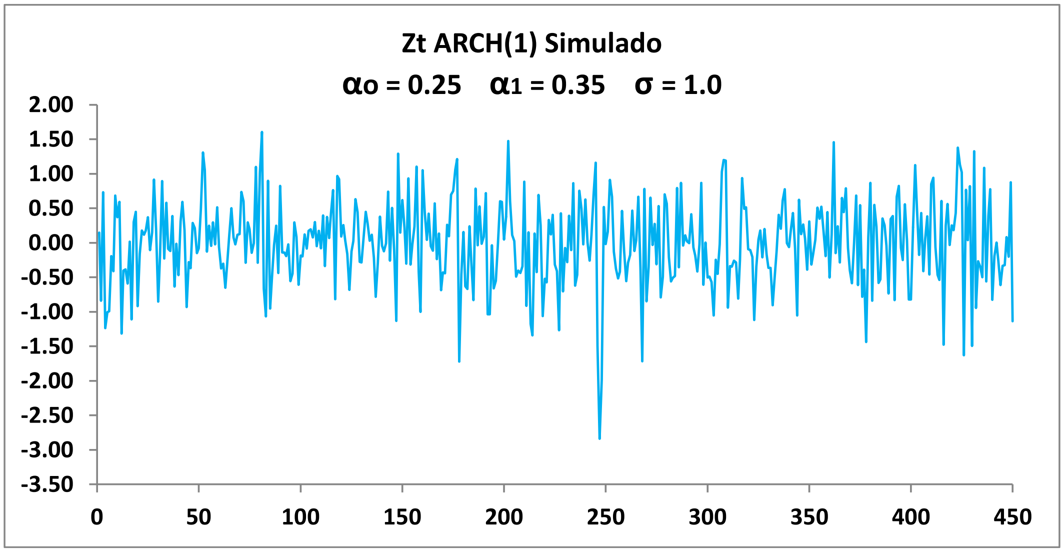
a_t tiene distribución normal para toda t con media 0 y varianza igual a $\sigma^2=1$ esto es N(0,1).

La suma de todos los parámetros α_t debe ser menor a la unidad para garantizar la que la serie sea estacionaria.

Si a_t se distribuye normal, entonces Zt se distribuye normal con varianza σ_t^2

SIMULACIÓN ARCH(1)

Como ejemplo se tiene una realización de 450 observaciones del modelo ARCH(1) con $\alpha_0 = 0.25$ y $\alpha_1 = 0.35$ y $\sigma = 1.0$ Los valores de los parámetros cumplen $\alpha_0 > 0$ y $0 < \alpha_1 < 1$ y su suma es menor a 1. La varianza teórica de V(Zt) es 0.03846 y su esperanza 0.



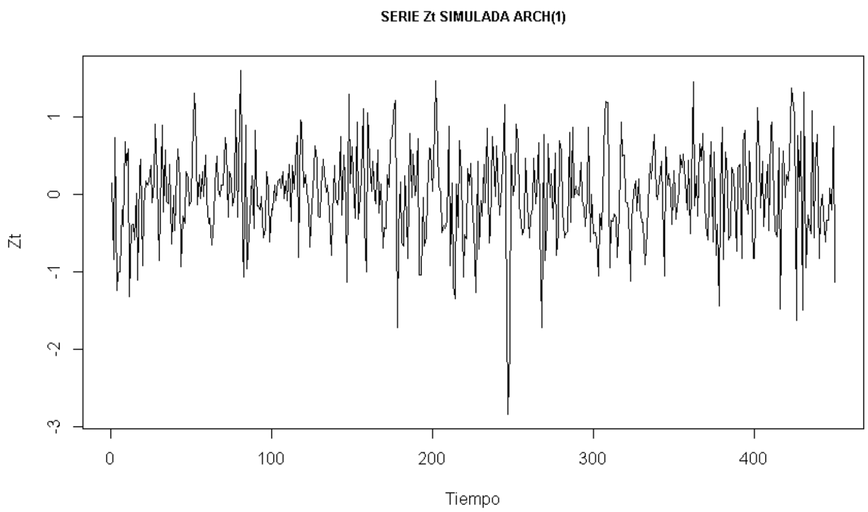
Los 450 datos de se consultan como anexo (SIMULA ARCH1 NOTAS V2.csv). La varianza empírica de Zt resultó S2=0.376132, cercana al valor teórico y su media empírica = -0.003828, cercana a cero.

La estimación de los parámetros α_i se lleva a efecto por máxima verosimilitud y es usual que se tome una solución de mínimos cuadrados ordinarios como solución inicial para proseguir con un procedimiento de optimización iterativo del tipo de Newton Rapson.

Estimación del Modelo con R

A partir de los 450 datos simulados se procede inicialmente a la lectura, verificación de la recuperación correcta de datos y su gráfica.

```
################################################################################
#           CARGA BIBLIOTECA tseries

library(tseries)

#  LEE LOS DATOS DE ARCHIVO SEPARADO POR COMAS  csv

ARCH1<-read.csv(file="C:/R LENGUAJE ESTADISTICO/EJERCICIOS DE PRUEBA/SIMULA ARCH1 NOTAS V2.csv",header=TRUE,dec=".",fill=TRUE)

ARCH1

plot(ARCH1$Zt,type="l",main="SERIE Zt SIMULADA ARCH(1)",cex.main=0.8,xlab="Tiempo",ylab="Zt")
```



A continuación se codifica el modelo a ajustar y se obtiene el resumen del reporte de resultados.

```
#           -------- AJUSTE DEL MODELO GARCH(0,1) -----------------

MODELO_ARCH<-garch(ARCH1$Zt,order=c(0,1))

summary(MODELO_ARCH)
```

R emite el siguiente resumen. Se emiten estadísticas con el mínimo, máximo, primer cuartil, mediana y tercer cuartil de la serie. Los valores puntuales de los coeficientes (0.2588 Y 0.30254) con sus errores estándares y las estadísticas t para probar la significancia de la hipótesis de nulidad de los parámetros con la probabilidad asociada, que resulta significativa en ambos estimadores.

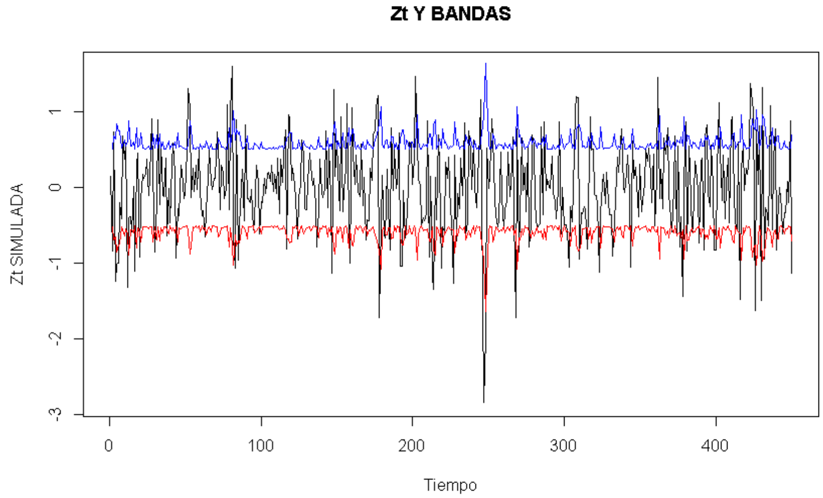
```
Model:
GARCH(0,1)

Residuals:
     Min       1Q   Median       3Q      Max 
-3.12947 -0.64987 -0.02408  0.72020  2.84201 
Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)    
a0   0.25880     0.02459   10.523  < 2e-16 ***
a1   0.30254     0.07589    3.987  6.7e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Se incluyen pruebas de hipótesis de normalidad y de autocorrelación de residuales de primer orden.

```
Diagnostic Tests:
        Jarque Bera Test
data:  Residuals
X-squared = 0.69425, df = 2, p-value = 0.7067

        Box-Ljung test
data:  Squared.Residuals
X-squared = 0.016369, df = 1, p-value = 0.8982
```

Como siguiente paso se asignan las predicciones del modelo y se grafica la serie original junto con las bandas de 95% de confianza para la varianza.

```
#          BANDAS CON AMBAS LINEAS EN UN SOLO ARCHIVO Y GRÁFICA
PRONOS<-predict(MODELO_ARCH)
PRONOS

plot(ARCH1$Zt,type="l",xlab="Tiempo",ylab="Zt SIMULADA",main="Zt Y BANDAS")

lines(PRONOS[,1],col="blue",type="l")

lines(PRONOS[,2],col="red",type="l")
```



6 Modelo GARCH

El modelo GARCH (Generalizad Autorreresive Condicional Heterocedasticity) y constituye una ampliación del modelo ARCH propuesta por Bollerslev en 1986. El modelo ARCH (q) puede complicarse en su estimación cuando se incluye un número elevado de retardos, demasiadas iteraciones y restricciones en los parámetros.

El modelo GARCH (p,q) incorpora otro grupo de parámetros se expresa de la manera siguiente:

$$Z_t = a_t \sigma_t$$

Donde

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i Z_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

Entonces el modelo ARCH(q) es un caso particular de GARCH (q,0), únicamente con parámetros del tipo α_i.

Las características de un modelo GARCH son las siguientes:

Las $\mathbf{a_t}$ se distribuyen con media 0 y varianza 1

Los parámetros α_0, α_i y $\beta_j \geq 0$ para i=1,2,….q y j=1,2,…..p. La suma de todos los parámetros debe ser menor a la unidad.

Un modelo GARCH (1,1) tiene los siguientes valores esperados y varianzas no condicionales y condicionales:

$$Z_t = a_t \sqrt{\alpha_0 + \alpha_1 Z_{t-1}^2 + \beta_1 \sigma_{t-1}^2}$$

$$Z_t^2 = a_t^2 \left(\alpha_0 + \alpha_1 Z_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \right)$$

$$E(Z_{t+1}) = 0$$

De manera análoga al modelo ARCH, se expresan la varianza no condicional y la esperanza y varianza condicionales.

$$V(Z_{t+1}) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

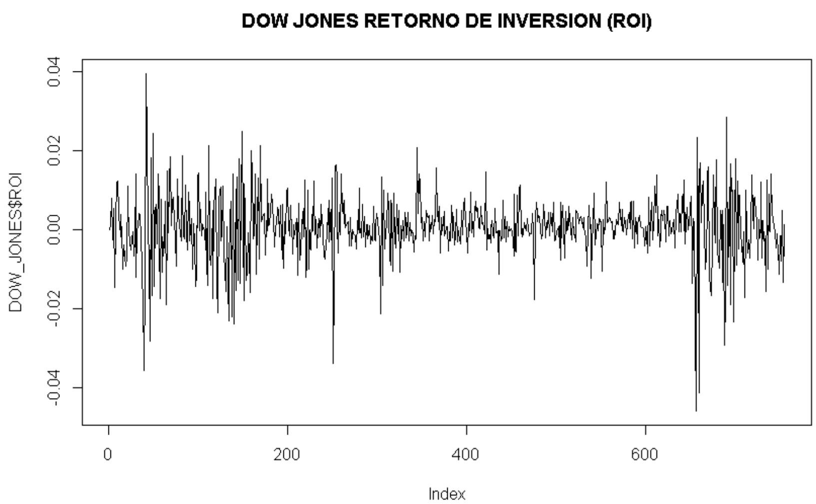
$$E(Z_{t+1} / Z_t) = 0$$

$$V(Z_{t+1} / Z_t) = \sigma_t^2$$

Ejemplo 2 Ajuste de modelo GARCH(1,1) a la serie de R0I del Dow Jones

Se tienen datos de valores del índice de la Bolsa de New York, Dow Jones en el archivo anexo (DOW JONES 2015-2018 DJI.XLSX). Se utiliza archivo formato CSV que contiene fecha y ROI.

```
##############################################################################
#
#           CARGA BIBLIOTECA tseries

library(tseries)

#  LEE LOS DATOS DE ARCHIVO SEPARADO POR COMAS  csv DEL ROI DOW JONES

DOW_JONES<-read.csv(file="C:/R LENGUAJE ESTADISTICO/EJERCICIOS DE
PRUEBA/DOW_JONES_ROI_29-06-15-8.csv",header=TRUE,dec=".",fill=TRUE)

DOW_JONES

plot(DOW_JONES$ROI,type="l",main="DOW JONES RETORNO DE INVERSION (ROI)")
```



A continuación se procede al ajuste del modelo. Se ensayaron varios modelos y el mejor resultó ser un GARCH(1,1)

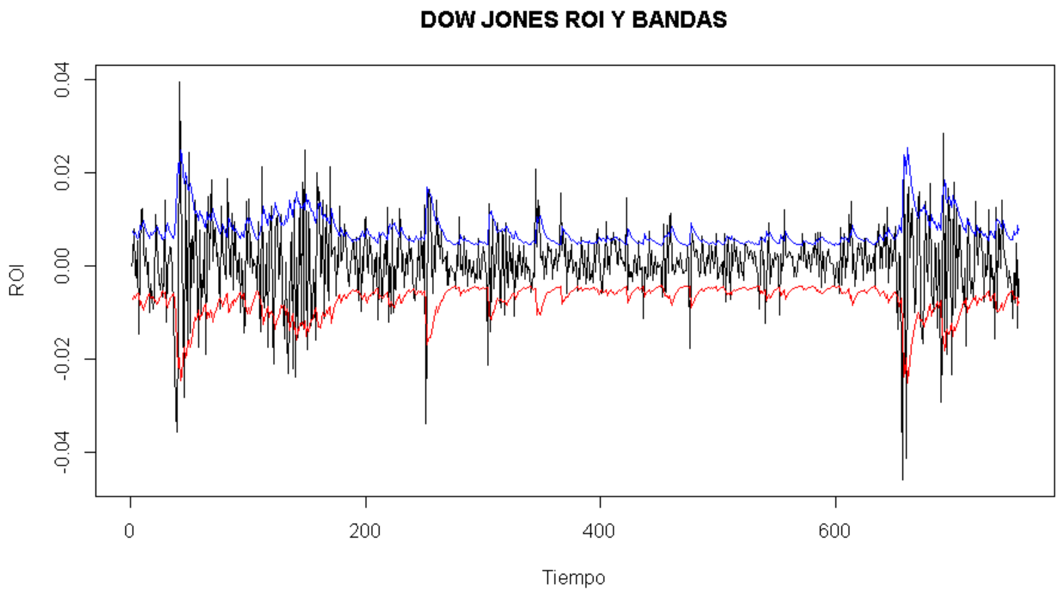
$$\sigma_t^2 = \alpha_0 + \alpha_i Z_{t-1}^2 + \beta_j \sigma_{t-1}^2$$

```
#           -------- AJUSTE DEL MODELO GARCH(1,1) -----------------

GARCH_DJ<-garch(DOW_JONES$ROI,order=c(1,1))

summary(GARCH_DJ)

Model:
GARCH(1,1)
```

```
Residuals:
     Min       1Q   Median       3Q      Max 
-4.60176 -0.41097  0.07209  0.64899  4.27774 

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)    
a0 4.526e-06   7.438e-07    6.085 1.17e-09 ***
a1 2.062e-01   2.506e-02    8.227 2.22e-16 ***
b1 7.260e-01   3.087e-02   23.517  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals 
X-squared = 208.68, df = 2, p-value < 2.2e-16

        Box-Ljung test

data:  Squared.Residuals 
X-squared = 2.344, df = 1, p-value = 0.1258
```

A continuación se obtienen las bandas de confianza

```
#          SE GRAFICA LA SERIE ORIGINAL DE ROY Y AMBAS BANDAS

PRONOS<-predict(GARCH_DJ)

PRONOS

plot(DOW_JONES$ROI,type="l",xlab="Tiempo",ylab="ROI",main="DOW JONES ROI Y BANDAS")

lines(PRONOS[,1],col="blue",type="l")

lines(PRONOS[,2],col="red",type="l")
```



7 Prueba de Jarque Bera

La prueba Jarque Bera es una prueba de bondad de ajuste específica para la normalidad. Se apoya en los valores de asimetría y curtosis de la Normal. Como se sabe la normal es simétrica y por tanto su coeficiente de asimetría es cero y en el caso específico de una Normal (0,1), su curtosis es cercana a 3. La prueba combina amas estadísticas y en la medida que se alejan del valor asociado a una normal, la hipótesis de normalidad tenderá a rechazarse.

Ho: Los datos proceden de una distribución normal
Ha: Los datos No proceden de una distribución normal

La estadística Jarque Bera se distribuye asintóticamente como una Ji cuadrada con 2 grados de libertad.

$$JB = \frac{n}{6}\left(A^2 + \frac{1}{4}(C-3)^2 \right)$$

$$A = \frac{\frac{1}{n}\sum\left(X_i - \bar{X}\right)^3}{\left(\frac{1}{n}\sum\left(X_i - \bar{X}\right)^2\right)^{3/2}}$$

$$C = \frac{\frac{1}{n}\sum\left(X_i - \bar{X}\right)^4}{\left(\frac{1}{n}\sum\left(X_i - \bar{X}\right)^2\right)^{2}}$$

8 Prueba Box Ljung

Permite probar la hipótesis de que un conjunto de autocorrelaciones son iguales a cero. La hipótesis nula plantea de la siguiente forma:

Ho: Los datos se distribuyen de forma independiente o en forma equivalente el conjunto de correlaciones de la población de donde proviene la muestra **son todas cero.**

H0: $\rho_t = 0$ para toda t

H1: $\rho_t \neq 0$ para alguna t. Los datos no se distribuyen en forma independiente.

La estadística de prueba:

$$Q(k) = n(n+2)\sum_{t=1}^{k} \frac{r_k^2}{(n-k)}$$

La Q(k) se distribuye Ji cuadrada con k grados de libertad. Cuando se aplica a los residuales de un modelo ARIMA, a los grados de libertad se les debe restar el número de parámetros autorregresivos y de medias móviles asociados al modelo.

9 Criterio de información de Akaike (AIC)

Es una medida de la calidad relativa de un modelo estadístico, para una serie de datos. Como tal, el AIC proporciona un medio para la selección del modelo.

AIC maneja una compensación “*trade-off*” entre la bondad de ajuste del modelo y la complejidad del modelo. Se basa en la entropía de información: se ofrece una estimación relativa de la información perdida cuando se utiliza un modelo determinado para representar el proceso que genera los datos.

AIC no proporciona una prueba de un modelo en el sentido de probar una hipótesis nula, es decir AIC no puede decir nada acerca de la calidad del modelo en un sentido absoluto. Si todos los modelos candidatos encajan mal, AIC no dará ningún aviso de ello.

Criterio de información Bayesiano normalizado (BIC normalizado)

Una medida general del ajuste global del modelo que intenta tener en cuenta la complejidad del modelo. Es una medida basada en el error cuadrático promedio que incluye una penalización para el número de parámetros presentes en el modelo y la longitud de la serie.

La penalización elimina la ventaja de los modelos con mayor número de parámetros, haciendo que el estadístico sea fácil de comparar entre distintos modelos para la misma serie.

10. Estimación del ARIMA-GARCH

Para la estimación de estos modelos es recomendable:

1. **Verificar que la serie es estacionaria**
   Se analiza la serie de rendimientos, con prueba de raíces unitarias. Para la estimación de media condicional con la metodología de Box – Jenkis, esta es válida en series estacionarias. Se puede ver si la serie de rendimientos es estacionaria

Para confirmar esta sospecha realizaremos la prueba de raíces unitarias sobre la serie de retornos. Si en el resultado de la prueba obtenemos que el p- value < 0.05, indica rechazo de Ho en favor de H1. Se continuar con el análisis.

2. **Estimar un modelo de media ARIMA** para la serie de retorno, con el fin de eliminar cualquier dependencia lineal en la serie.- Se busca encontrar un modelo ARMA, ARIMA o ARIMA estacional que ajuste a la serie de datos analizada.

   Realizando los siguientes procesos:

   2.1 Identificación.- se analiza la serie de retornos y sus autocorrelaciones simples y parciales, para determinar los valores más apropiados para los parámetros (p,d,q). De aquí se puede encontrar varios modelos que ajustan a la serie, en la parte de validación se elegirá al mejor de ellos, con sus coeficientes y el criterio de información AIC.

   Se debe estar consiente que la serie de retornos no hay tendencia, es estacional y tiene varianza diferente no constante por periodos. De aquí que el valor d es igual a cero y para determinar el valor de los para metros p y q, analizando las autocorrelaciones simples y parciales de la serie.

   2.2 Estimación.- se determina los posibles modelos que pueden estimar la serie de retornos, verificaremos la significancia de sus coeficientes y elegiremos el mejor de ellos mediante el criterio de información AIC.

   2.3 Validación.- se ven los residuos del modelo seleccionado para asegurar este cumpla con el supuesto de ruido blanco, mediante el empleo de la serie de residuos, curva de normalidad, autocorrelaciones y pruebas de hipótesis.

3. Análisis de la Serie de Residuos del Modelo ARIMA, efectos GARCH

Para determinar si existen efectos GARCH en la serie de residuos del modelo ARIMA, se realiza las pruebas de hipótesis: Ljung-Box y Multiplicador de Lagrange para efectos GARCH. El resultado de estas pruebas, nos indicará si es necesario realizar la estimación de un modelo GARCH.

Sí en la prueba de hipótesis Box Ljung, obtenemos un p-value menor a 0.05. Se rechaza la hipótesis nula, es decir los errores al cuadrado están correlacionados entre sí. Por tanto es necesario realizar un estimación GARCH para esta serie de tiempo, esto nos permite tener un buen estimador de la serie de retornos, ya que el modelo ARIMA(p,0,q) por sí solo no lo es. En la prueba de hipótesis Multiplicador de Lagrange para efectos GARCH, obtenemos un p-value menor 0.05. Se rechaza la hipótesis nula, es decir los errores tiene efectos GARCH significativos, que se pueden estimar mediante un Modelo GARCH.

En R la función es "ArchTest()"

3.1 Utilizar los residuos estandarizados y los residuos estandarizados al cuadrado del modelo de media para probar los efectos de GARCH.

En la validación se revisan los residuos del modelo elegido para asegurar este cumpla con el supuesto de ruido blanco, mediante el empleo de la serie de residuos, curva de normalidad, autocorrelociones simples y parciales.

Para determinar si existen efectos GARCH en la serie de residuos del modelo ARIMA, se revisan las pruebas de hipótesis: Ljung-Box y Multiplicador de Lagrange para efectos GARCH.

Instrucción en R. Box.test(residuals(ModeloArimaElegido)^2,type = 'Ljung-Box',lag = 12

Si en la prueba de hipótesis Box Ljung, se obtiene un p-value < 0.05 se rechaza la hipótesis nula, es decir que existe evidencia que los errores al cuadrado están correlacionados entre sí.

Instrucción en R. ArchTest(residuals(ModeloArimaElegido),lag = 5)

En la prueba de hipótesis Multiplicador de Lagrange para efectos GARCH, si se obtiene un p-value < 0.05. Se rechaza la hipótesis nula, es decir los errores tiene efectos GARCH significativos, que necesitan ser estimados mediante un Modelo GARCH.

Ambas prueba, confirman la necesidad de estimar un modelo GARCH para serie de retornos, ya que el modelo ARIMA(p,d,q) no cumple con la condición requerida de homocedasticidad de los errores, para ser un buen estimador de media condicional.

3.2 Estimar un modelo de volatilidad si los efectos GARCH son estadísticamente significativos.

Para estimar un modelo de varianza condicional GARCH, se debe identificar su orden, estimar los coeficientes del modelo y realizar las pruebas de validación pertinentes. Sin embargo, en la mayor parte de casos basta un modelo GARCH(1,1) para ajustar el comportamiento de la varianza condicional de una serie de retornos financieros.

Para Identificar un modelo se analizan las autocorrelociones simples y parciales de los residuos y de los residuos al cuadrado.

Se procede con la Estimación para la varianza condicional de la serie de retornos, verificando la significancia de los coeficientes y elegir el mejor de los propuestos con el criterio de información AIC ó BIC.

$$\sigma_t^2 = \alpha_0 + \alpha_i Z_{t-1}^2 + \beta_j \sigma_{t-1}^2$$

3.3 Comprobar el modelo ajustado

Se valida el ajuste del modelo estimado, utilizando la serie de residuos, el gráfico de normalidad asociado y correlogramas.

Si observamos en la serie de residuos un comportamiento aleatorio, y en el gráfico de normalidad vemos que los valores se aproximan a la línea de normalidad. Y en los correlogramas de los errores al cuadrado, se ve las autocorrelaciones simples y parciales son significativas, entonces se puede concluir que el Modelo Garch proporciona un buen ajuste de la Serie de Residuos del Modelo ARIMA(p,d,q), estimado anteriormente.

Como resultado de la estimación se obtiene el Modelo ARIMA(p.d.q) - GRACH(p,q) es el mejor estimador de la Serie de Retornos en el periodo analizado.

Ejemplo suponga un Modelo: ARIMA(1,0,1) - GRACH(1,1), se tendría:

Modelo ARIMA(1,0,1)

	Estimate	Std. Error	z value	Pr(>|z|)
ar1	-0.969230	0.023723	-40.856	< 2e-16 ***
ma1	0.940202	0.031987	29.393	< 2e-16 ***
intercept	0.425248	0.160777	2.645	0.00817 **

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

Modelo GARCH(1,1)

	Estimate	Std. Error	z value	Pr(>|z|)
a0	0.944583	0.378338	2.4967	0.0125367 *
a1	0.066551	0.017924	3.7130	0.0002049 ***
b1	0.877930	0.037007	23.7235	< 2.2e-16 ***

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

$$r_t = -0.9692 * r_{t-1} + 0.4252 + 0.9402 * a_{t-1}$$

Donde:

$\boldsymbol{rt}$: Es el rendimiento de una acción en el tiempo t.

$a_t = \sqrt{\sigma_{ta_t}}$

$\sigma_t^2 = 0.9445 + 0.0665 * r_{t-1}^2 + 0.8779 * \sigma_{t-1}^2$

a$\boldsymbol{t}$: Es un Ruido Blanco de variables independientes

Gráficamente queda:

Retornos Vs Ajuste de Retornos de un Modelo ARIMA(1,0,1)-GARCH(1,1)