

Series de Tiempo

VI Semestre

Grupo: B

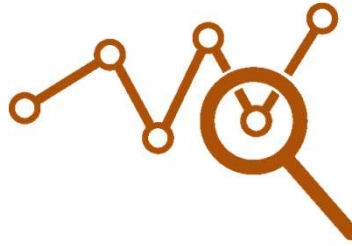
Mtr. Alcides Ramos Calcina

FINESI

CONSTRUCCIÓN DE MODELOS ARIMA

2°. Estimación y 3°. Validación

2) ESTIMACIÓN



- Una vez identificados el/los procesos estocásticos que han podido generar la serie temporal Y_t , la siguiente etapa consiste en estimar los parámetros desconocidos de dichos modelos:

$$\beta = (\delta, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$$

2) ESTIMACIÓN



- Los modelos los podemos agrupar en dos:

a) El modelo **autoregresivo** de orden p y el modelo **integrado** autoregresivo de orden $(p, d, 0)$ que se pueden expresar de la siguiente forma:

$$\text{AR}(p): \quad a_t = \phi(L)Y_t - \delta$$

$$\text{ARI}(p, d, 0): \quad a_t = \phi(L)\Delta^d Y_t - \delta$$

b) Los modelos de media móvil de orden q , autorregresivo de media móvil de orden (p, q) , integrado de media móvil de orden $(0, d, q)$

$$\text{MA}(q): \quad a_t = \theta(L)^{-1} (Y_t - \delta)$$

$$\text{IMA}(p, d, 0): \quad a_t = \theta(L)^{-1} (\phi(L)Y_t - \delta)$$

$$\text{ARMA}(p, q): \quad a_t = \theta(L)^{-1} (\Delta^d Y_t - \delta)$$

$$\text{ARIMA}(p, d, q): \quad a_t = \theta(L)^{-1} (\phi(L)\Delta^d Y_t - \delta)$$

Ejemplo 1



Estimación de los modelos propuestos:

- Modelo 1: $AR(3) \cong ARIMA(3, 0, 0)$: $Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + a_t$ con $Z_t = \log(Y_t)$

```
mod1 <- Arima(T.Yt, order = c(3,0,0), include.constant = T)
coeftest(mod1)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)
ar1	1.129439	0.091124	12.3945	< 2.2e-16 ***
ar2	-0.329780	0.137599	-2.3967	0.016544 *
ar3	-0.267812	0.090935	-2.9451	0.003229 **
intercept	6.675427	0.112465	59.3556	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Parámetros estimados del modelo:

$ARIMA(3, 0, 0)$: $Z_t = 6.67543 + 1.12944Z_{t-1} - 0.32978Z_{t-2} - 0.26781Z_{t-3} + a_t$

Ejemplo 1



- Modelo 2: ARIMA(0, 0, 2): $Z_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}$

```
mod2 <- Arima(T.Yt, order = c(0,0,2),include.constant = T)
coeftest(mod2)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)
ma1	1.060744	0.081001	13.0954	< 2.2e-16 ***
ma2	0.532477	0.082251	6.4738	9.556e-11 ***
intercept	6.669663	0.170361	39.1501	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Parámetros estimados del modelo:

ARIMA(0, 0, 2): $Z_t = 6.66966 + a_t + 1.060746a_{t-1} + 0.53248a_{t-2}$

Ejemplo 1



- Modelo 3: ARIMA(2, 0, 1): $Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t + \theta_1 a_{t-1}$

```
mod3 <- Arima(T.Yt, order = c(2,0,1),include.constant = T)
coeftest(mod3)
```

z test of coefficients:

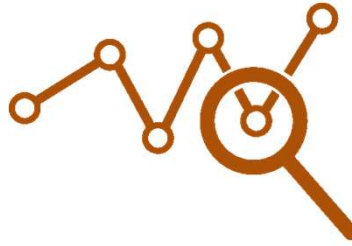
	Estimate	Std. Error	z value	Pr(> z)
ar1	1.484217	0.064353	23.0639	< 2.2e-16 ***
ar2	-0.818950	0.057206	-14.3159	< 2.2e-16 ***
ma1	-0.336677	0.097577	-3.4504	0.0005598 ***
intercept	6.675372	0.104597	63.8202	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Parámetros estimados del modelo:

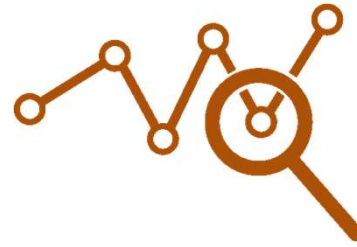
ARIMA(2, 0, 1): $Z_t = 6.67537 + 1.48422Z_{t-1} - 0.81895Z_{t-2} + a_t - 0.33668_1 a_{t-1}$

3) VALIDACIÓN



Se tiene en cuenta:

- a) Si las estimaciones de los coeficientes del modelo son significativas y cumplen las condiciones de **estacionariedad** e **invertibilidad** que deben satisfacer los parámetros del modelo.
- b) Si los residuos del modelo tienen un comportamiento similar a las innovaciones, es decir, si son **ruido blanco**.



3.1. Análisis de los coeficientes estimados

- Realizar los contrastes habituales de **significación individual** de los coeficientes AR y MA:

$$\beta = (\delta, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$$

- En el caso más general de un ARMA(p,q) con constante se plantean los siguientes contrastes:

$$\begin{array}{lll} \left\{ \begin{array}{l} H_0 : \delta = 0 \\ H_1 : \delta \neq 0 \end{array} \right. & \left\{ \begin{array}{l} H_0 : \phi_i = 0 \\ H_1 : \phi_i \neq 0 \end{array} \right. & \left\{ \begin{array}{l} H_0 : \theta_i = 0 \\ H_1 : \theta_i \neq 0 \end{array} \right.$$

a) Significación de los coeficientes

El estadístico es:

$$t_{\beta_i} = \frac{\hat{\beta}_i - 0}{\hat{\sigma}_{\hat{\beta}_i}} \sim t_{\left(\frac{\alpha}{2}, T-p-q-\delta\right)} \approx N(0,1)$$

Si el valor de probabilidad asociado al estadístico “t-Student”, si $p < \alpha$, entonces se rechaza la hipótesis de nulidad H_0 .

Ejemplo 1



- Modelo 1: ARIMA(3, 0, 0):

```
coeftest(mod1)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
ar1	1.129439	0.091124	12.3945	< 2.2e-16	***
ar2	-0.329780	0.137599	-2.3967	0.016544	*
ar3	-0.267812	0.090935	-2.9451	0.003229	**
intercept	6.675427	0.112465	59.3556	< 2.2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- AR(1) : $\phi_1 = 1.129439$ → $p = 0.00000 < 0.01$, altamente significativo.
- AR(2) : $\phi_2 = -0.329780$ → $p = 0.01654 < 0.05$, significativo.
- AR(3) : $\phi_3 = -0.267812$ → $p = 0.00323 < 0.01$, altamente significativo.

Ejemplo 1



- Modelo 2: ARIMA(0, 0, 1):

```
coeftest(mod2)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)
ma1	1.060744	0.081001	13.0954	< 2.2e-16 ***
ma2	0.532477	0.082251	6.4738	9.556e-11 ***
intercept	6.669663	0.170361	39.1501	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- MA(1) : $\theta_1 = 1.060744$ → $p = 0.00000 < 0.01$, altamente significativo.
- MA(2) : $\theta_2 = 0.532477$ → $p = 0.00000 < 0.01$, altamente significativo.

Ejemplo 1



- Modelo 3: ARIMA(2, 0, 1):

```
coeftest(mod3)
```

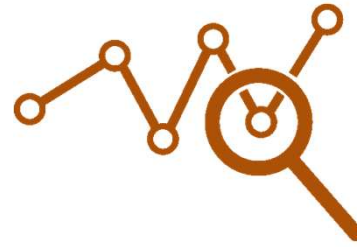
z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
ar1	1.484217	0.064353	23.0639	< 2.2e-16	***
ar2	-0.818950	0.057206	-14.3159	< 2.2e-16	***
ma1	-0.336677	0.097577	-3.4504	0.0005598	***
intercept	6.675372	0.104597	63.8202	< 2.2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Por el momento, se recomendaría utilizar los modelos 2 y 3 por tener coeficientes altamente significativos. Continuaremos con las demás pruebas.

- AR(1) : $\phi_1 = 1.484217$ $\rightarrow p = 0.00000 < 0.01$, altamente significativo.
- AR(2) : $\phi_2 = -0.818950$ $\rightarrow p = 0.00000 < 0.01$, altamente significativo.
- MA(1) : $\theta_1 = -0.336677$ $\rightarrow p = 0.0005598 < 0.01$, altamente significativo.



3.1. Análisis de los coeficientes estimados

b) Examen de la matriz de correlaciones entre los coeficientes

El objetivo es detectar si existe un problema de **multicolinealidad**.

- Si la correlación entre dos coeficientes estimados esté próxima a uno (ejemplo: **0.9**) entonces existe un **problema grave de multicolinealidad**.
- Si esto ocurre, los coeficientes estimados serían de carácter **muy inestables** con lo que podrían cambiar mucho de una realización a otra.
- Para evitar este problema puede ser conveniente eliminar algún parámetro del modelo.
- El coeficiente de correlación entre dos coeficientes se define:

$$\rho_{\phi_i \phi_j} = \frac{Cov(\hat{\phi}_i \hat{\phi}_j)}{\sqrt{Var(\hat{\phi}_i) Var(\hat{\phi}_j)}} = \frac{\sigma_{\hat{\phi}_i \hat{\phi}_j}}{\sigma_{\hat{\phi}_i} \sigma_{\hat{\phi}_j}}$$

la información requerida para calcular estos coeficientes se encuentra en la estimación del modelo.

Ejemplo 1



La matriz de correlaciones de los coeficientes para los dos modelos estimado es el siguiente:

- Modelo 1: ARIMA(3, 0, 0)

```
vcov(mod1)
```

	ar1	ar2	ar3	intercept
ar1	8.303634e-03	-0.011020379	5.861802e-03	3.168035e-05
ar2	-1.102038e-02	0.018933416	-1.107504e-02	-1.781600e-05
ar3	5.861802e-03	-0.011075038	8.269135e-03	-2.285447e-05
intercept	3.168035e-05	-0.000017816	-2.285447e-05	1.264838e-02

- Modelo 2: ARIMA(0, 0, 2)

```
vcov(mod2)
```

	ma1	ma2	intercept
ma1	0.0065611558	-1.330452e-04	6.865070e-05
ma2	-0.0001330452	6.765233e-03	-3.627862e-05
intercept	0.0000686507	-3.627862e-05	2.902301e-02

Ejemplo 1



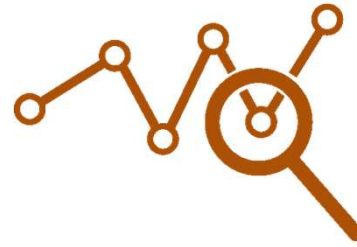
La matriz de correlaciones de los coeficientes para los dos modelos estimado es el siguiente:

- Modelo 3: ARIMA(2, 0, 1)

```
vcov(mod3)
```

	ar1	ar2	ma1	intercept
ar1	4.141245e-03	-3.098085e-03	-3.669202e-03	2.200616e-05
ar2	-3.098085e-03	3.272484e-03	2.414807e-03	-3.251553e-05
ma1	-3.669202e-03	2.414807e-03	9.521275e-03	2.925611e-05
intercept	2.200616e-05	-3.251553e-05	2.925611e-05	1.094043e-02

Se observa claramente que ningún coeficiente esta próximo ni cercano a 0.9, por tanto, podemos indicar que estamos **no hay problema de multicolinealidad** en los modelos propuestos.



3.1. Análisis de los coeficientes estimados

c) Condición de Convergencia e Invertibilidad

Se tiene dos casos:

- **Condición de convergencia.** Este paso permite conocer si la suma de todos los coeficientes de la parte autorregresiva del modelo ARIMA es menor a 1.

De ser el caso, se cumple la primera **condición de estacionariedad** en sentido débil: que el valor esperado de la serie dada la filtración sea igual a la media y que sea tanto constante como finita.

- **Condición de raíces invertibles.** El teorema de invertibilidad establece que si una serie tiene raíces invertibles, un proceso $AR(p)$ se puede convertir a uno $MA(\infty)$ y viceversa (convertir un proceso $MA(q)$ a uno $AR(\infty)$).

Ejemplo 1

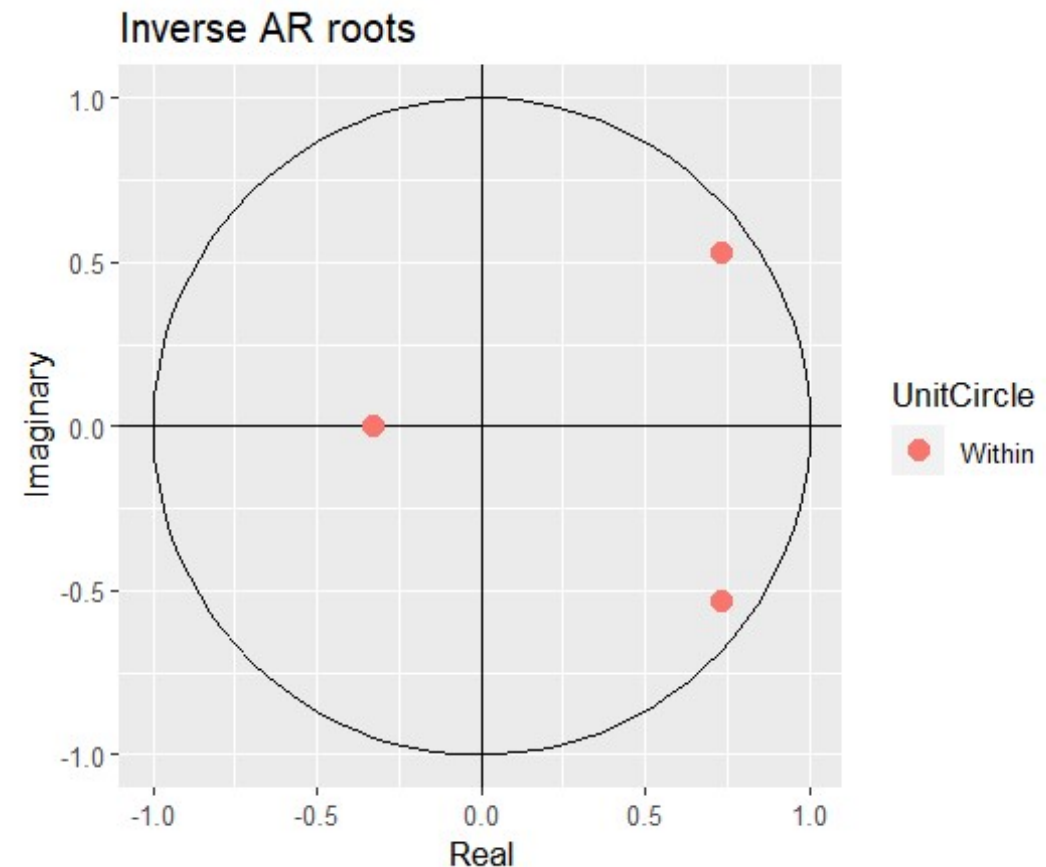


NOTA: Es importante indicar que, para graficar las raíces inversas es necesario estimar los parámetros del modelo haciendo uso de la función **Arima()**.

```
mod1 <- Arima(T.Yt, order = c(3,0,0), include.constant = T)
autoplot(mod1)
```

- Modelo 1: ARIMA(3, 0, 0)

En la figura de raíces inversas de AR, se observa en el gráfico las raíces características se encuentran dentro del círculo, es decir que **cumplen con la condición de estacionariedad** para la parte autorregresiva.



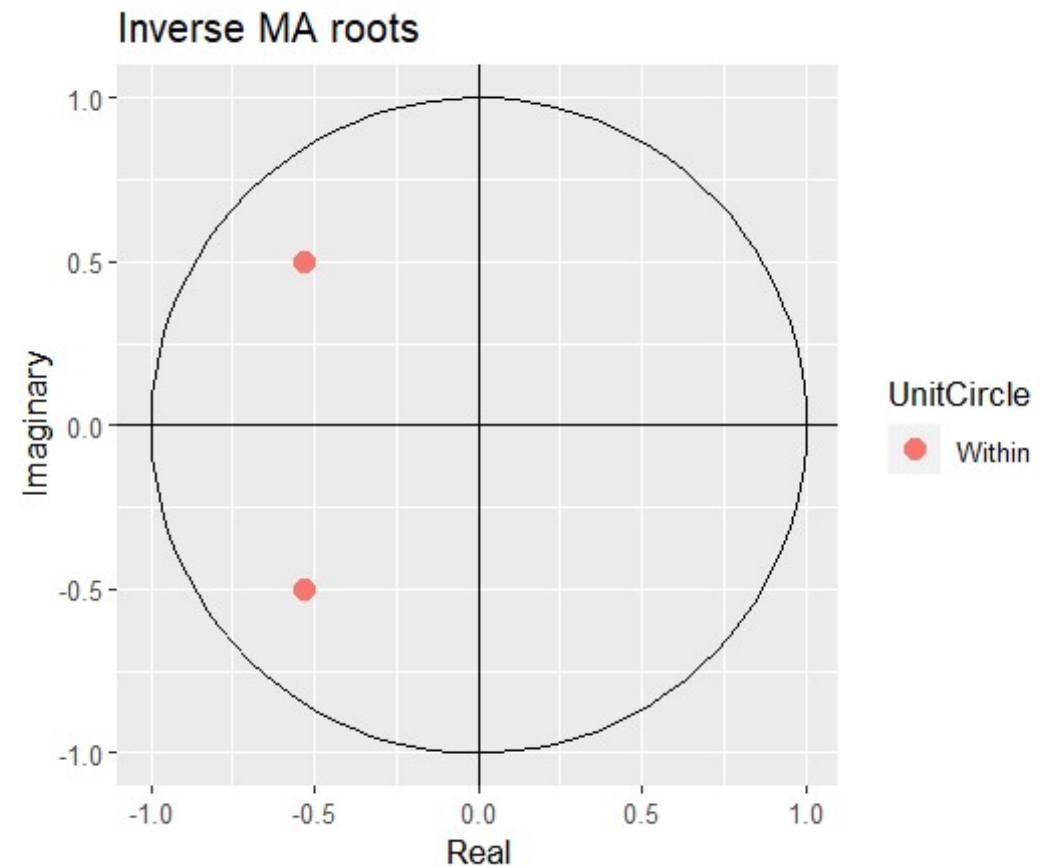
Ejemplo 1



- Modelo 2: ARIMA(0, 0, 1)

```
mod2 <- Arima(T.Yt, order = c(0,0,2), include.constant = T)
autoplot(mod2)
```

En la figura de raíces inversas de MA, se observa en el gráfico las raíces características se encuentran dentro del círculo, es decir que **cumplen con la condición de invertibilidad** para la parte de media móvil.



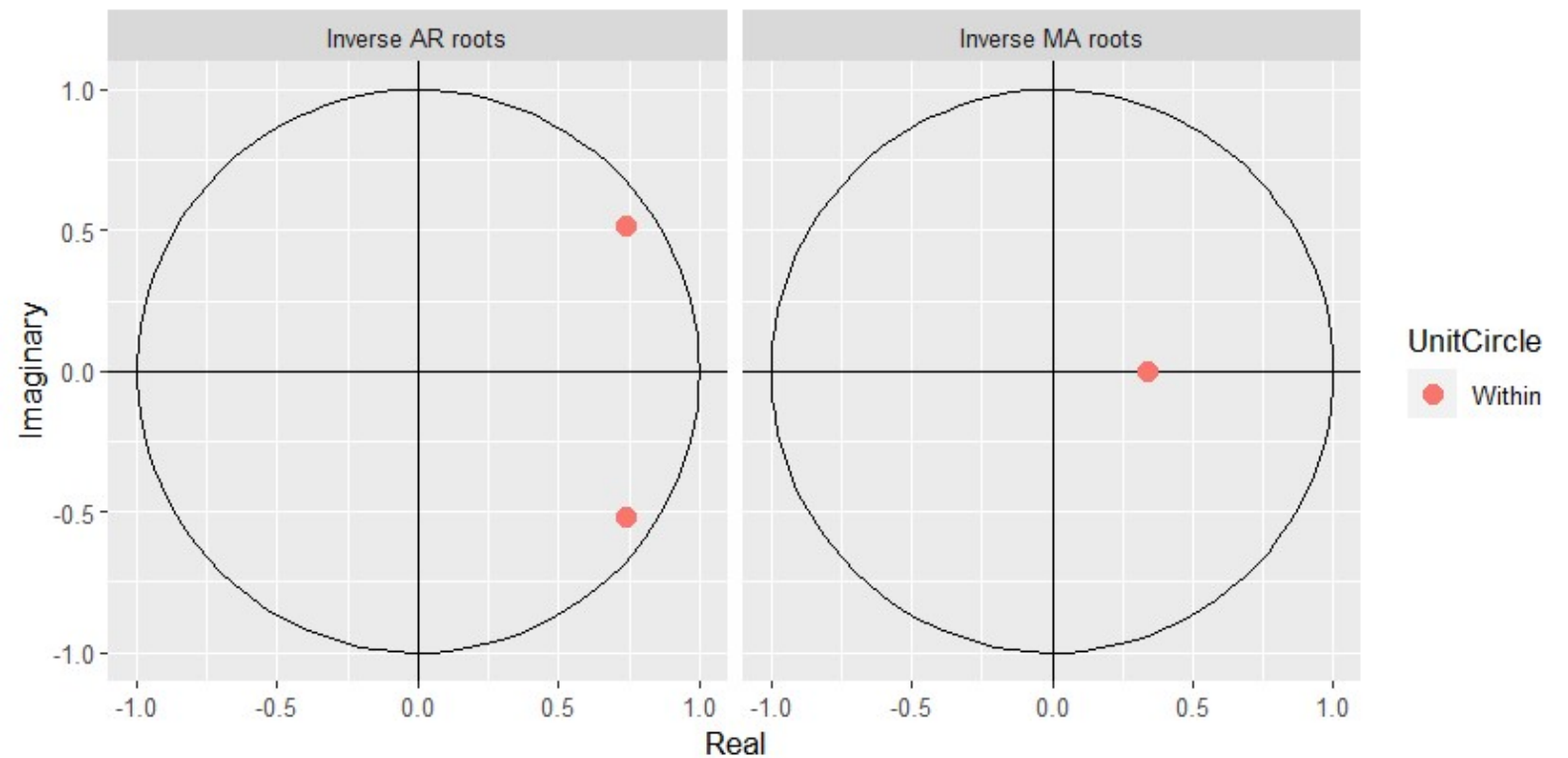
Ejemplo 1

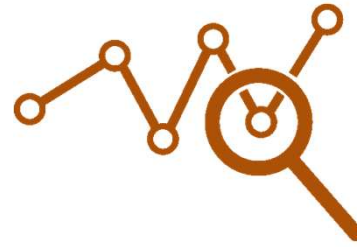


- Modelo 3: ARIMA(2, 0, 1)

```
mod3 <- Arima(T.Yt, order = c(2,0,1), include.constant = T)
autoplot(mod3)
```

Al estar los valores dentro de la circunferencia unitaria es un indicativo de que el modelo se ajusta correctamente. Tanto en su parte AR, como en su parte MA.

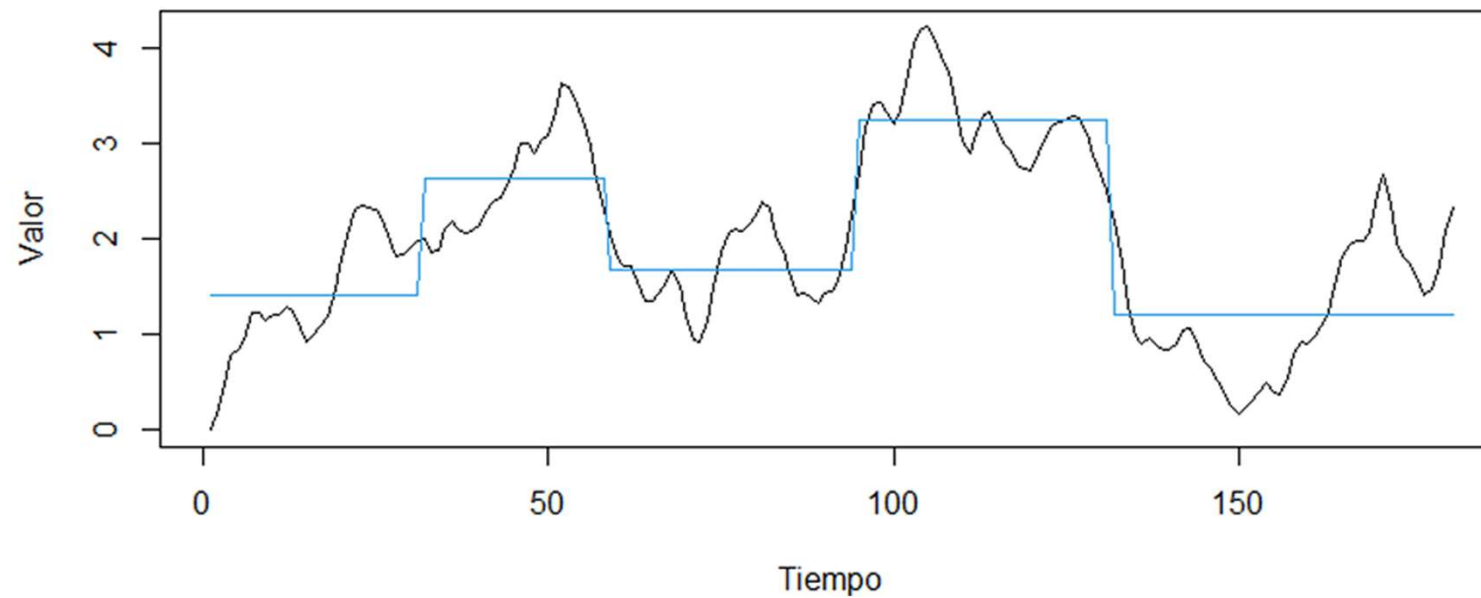


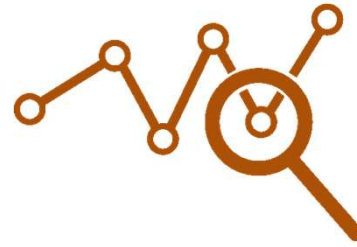


3.1. Análisis de los coeficientes estimados

d) Análisis de la estabilidad

- Consiste en un análisis de cambio estructural, esta consiste cuando hay un cambio instantáneo o permanente e inesperado en una variable recogida a lo largo del tiempo.
- De acuerdo al cambio estructural, si se estima un modelo con una serie sin tener en cuenta los cambios estructurales puede llevar a varias consecuencias.





3.1. Análisis de los coeficientes estimados

- Una prueba estadística para detectar la existencia de un cambio estructural es hacer el test de **Chow**.

Test de Chow

- Este contraste consiste en comprar si las últimas observaciones muestrales disponibles presente cambio respecto a las anteriores.
- Se plantea la siguiente hipótesis:

$$H_0 : \begin{pmatrix} \beta_0^1 \\ \beta_1^1 \\ \vdots \\ \beta_k^1 \end{pmatrix} = \begin{pmatrix} \beta_0^2 \\ \beta_1^2 \\ \vdots \\ \beta_k^2 \end{pmatrix} \qquad H_1 : \begin{pmatrix} \beta_0^1 \\ \beta_1^1 \\ \vdots \\ \beta_k^1 \end{pmatrix} \neq \begin{pmatrix} \beta_0^2 \\ \beta_1^2 \\ \vdots \\ \beta_k^2 \end{pmatrix}$$

- Se rechaza H_0 , si la probabilidad asociada al estadístico F (p-valor.) es menor al nivel de significancia, es decir: $p < \alpha$.

Ejemplo 1



- La tercera parte de la muestra es 37 observaciones y la mitad de la muestra es 55 observaciones, entonces vamos a considerar 46 observaciones y aplicamos la prueba de Chow para cada uno de los modelos, en R utilizaremos la librería **strucchange** y las funciones **Fstats** y **sctest**.

- **Modelo 1**

```
library(strucchange) # Generar un modelo respecto a una constante
Chow_mod1 <- Fstats(mod1$fitted ~ 1, from = 0.43)
sctest(Chow_mod1)
```

```
supF test
```

```
data: Chow_mod1
sup.F = 0.16992, p-value = 1
```

- **Modelo 2**

```
Chow_mod2 <- Fstats(mod2$fitted ~ 1, from = 0.43)
sctest(Chow_mod2)
```

```
supF test
```

```
data: Chow_mod2
sup.F = 0.099595, p-value = 1
```

Ejemplo 1



- **Modelo 3**

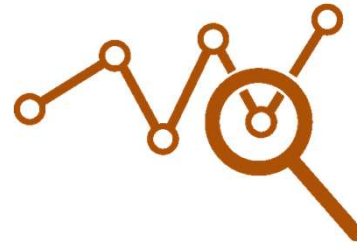
```
Chow_mod3 <- Fstats(mod2$fitted ~ 1, from = 0.3)
sctest(Chow_mod3)
```

supF test

```
data: Chow_mod3
sup.F = 0.17598, p-value = 1
```

En las tres pruebas se acepta la hipótesis nula ($p > \alpha = 0.05$), es decir, **existe estabilidad** de coeficientes..





3.2. Análisis de los residuos

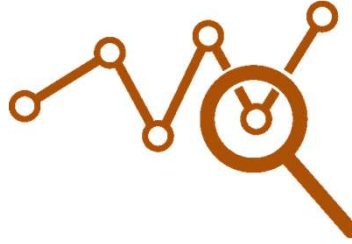
- Suponga un modelo ARMA(p, q) elegido para la serie estacionaria $Z_t = \log(Y_t)$ (por simplicidad, suponemos que $E(Z_t) = 0$)

$$\phi_p(L)Z_t = \theta_q(L)a_t$$

es correcto, entonces $\hat{a}_t = \frac{\phi_p(L)}{\theta_q(L)}Z_t$ es un proceso ruido blanco. Donde \hat{a}_t son estimaciones de a_t

- El análisis de residuos consiste en una serie de contrastes de diagnóstico con el objetivo de determinar si los residuos replican el comportamiento de un ruido blanco.
- Es decir, un proceso ruido blanco con **media cero**, **varianza constante** y las **autocorrelaciones nulas**.

3.2. Análisis de los residuos



a) Media es igual a cero

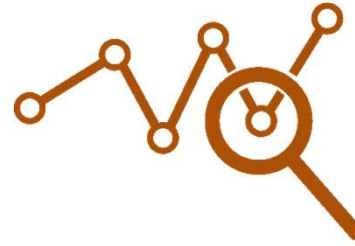
- Se realiza un **análisis gráfico**, representando los residuos a lo largo del tiempo y observando si los valores **oscilan alrededor de cero**.
- Por tanto, como método confirmatorio se puede llevar a cabo el siguiente contraste estadístico de hipótesis:

$$H_0: E(a_t) = 0$$

$$H_1: E(a_t) \neq 0$$

- Si la probabilidad asociada al estadístico de contraste t -Student es menor al nivel de significancia, es decir: $p < \alpha$, se rechaza la hipótesis nula H_0 .

3.2. Análisis de los residuos



b) Homocedasticidad o varianza constante

- Si en el gráfico de los residuos, la dispersión de los mismos es constante, concluiremos que la varianza de y permanece constante.
- Además de las gráficas, también se ha aplicado la prueba de **Breusch-Pagan** para estudiar la homocedasticidad definiendo como hipótesis:

H_0 : *Los residuos tienen varianza constante (homocedasticidad)*

H_1 : *Los residuos no tienen varianza constante (heterocedasticidad)*

- Si: $p < \alpha$, se rechaza la hipótesis nula H_0 .

Ejemplo 1



Media igual a cero

- **Modelo 1**

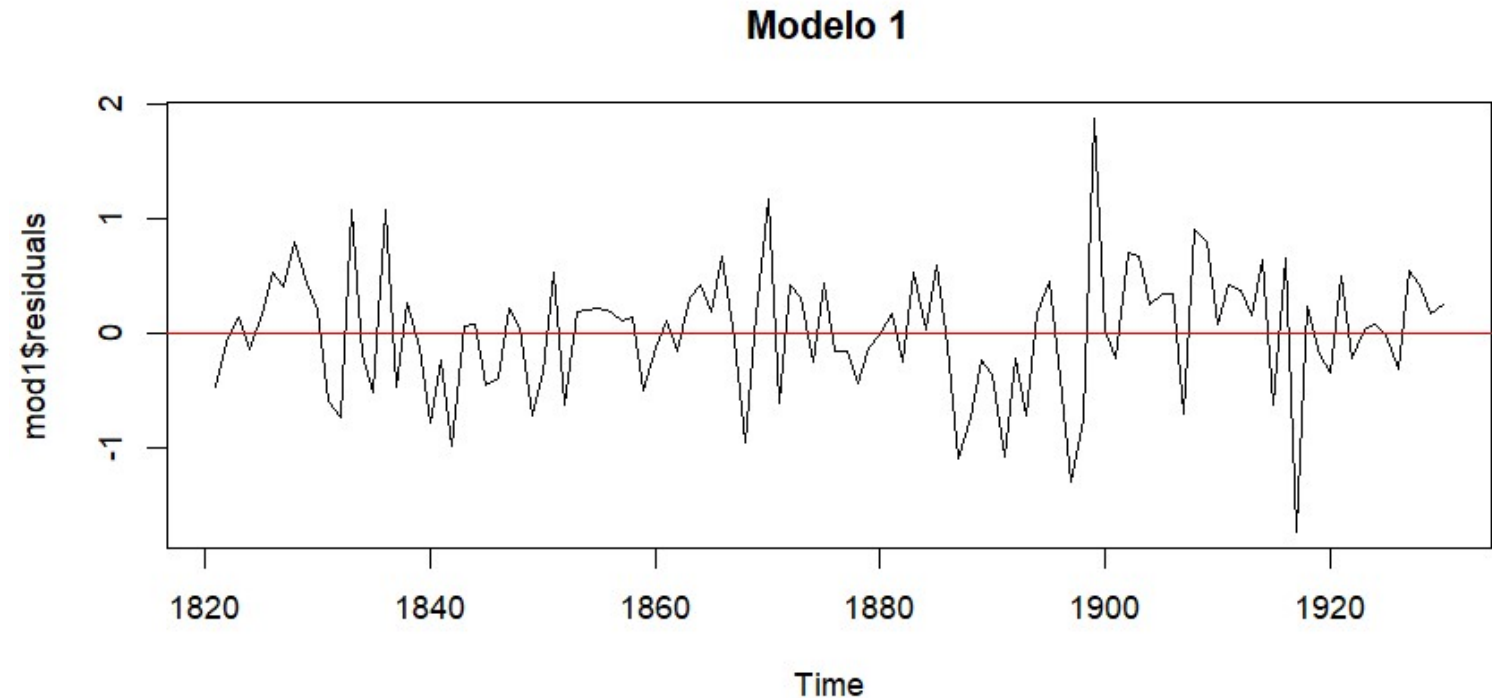
```
plot(mod1$residuals)
abline(h = 0, col = "red")
```

Prueba t

```
t.test(mod1$residuals, mu = 0)
```

One Sample t-test

```
data: mod1$residuals
t = -0.003873, df = 109, p-value = 0.9969
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.1043491  0.1039421
sample estimates:
 mean of x
-0.0002035144
```



Parece indicar que un buen número de residuales están en torno a la media igual a cero.

Para confirmar lo mencionado se realiza la prueba t:
Como $p = 0.9969 > \alpha = 0.05$, se acepta H_0 , es decir la media es igual a cero.

Ejemplo 1



- **Modelo 2**

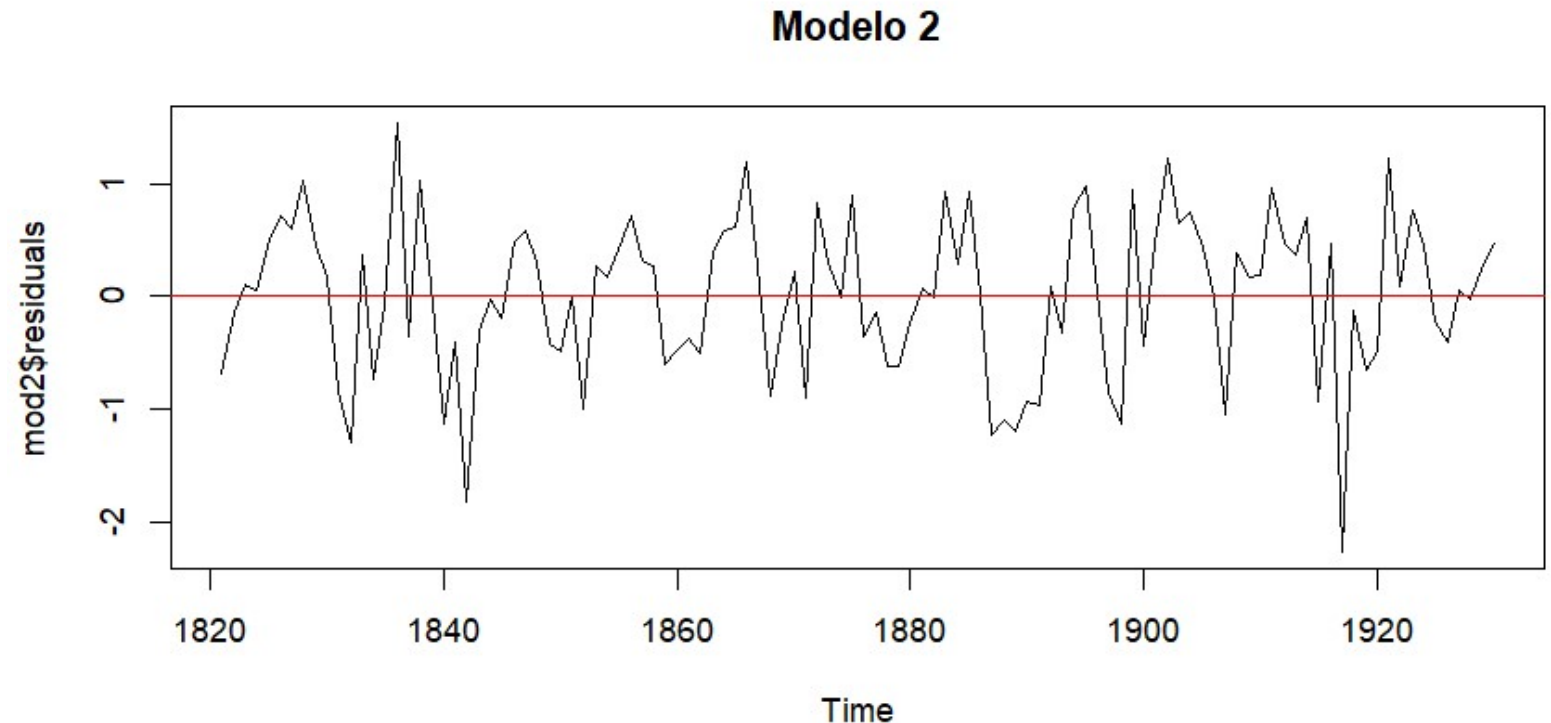
```
plot(mod2$residuals)
abline(h = 0, col = "red")
```

Prueba t

```
t.test(mod2$residuals, mu = 0)
```

One Sample t-test

```
data: mod2$residuals
t = 0.055737, df = 109, p-value = 0.9557
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.1280626  0.1354738
sample estimates:
 mean of x
0.003705576
```



También indica que un buen número de residuales están en torno a la media igual a cero.

Para confirmar lo mencionado se realiza la prueba t:
Como $p = 0.9557 > \alpha = 0.05$, se acepta H_0 , es decir la media es igual a cero.

Ejemplo 1



- **Modelo 3**

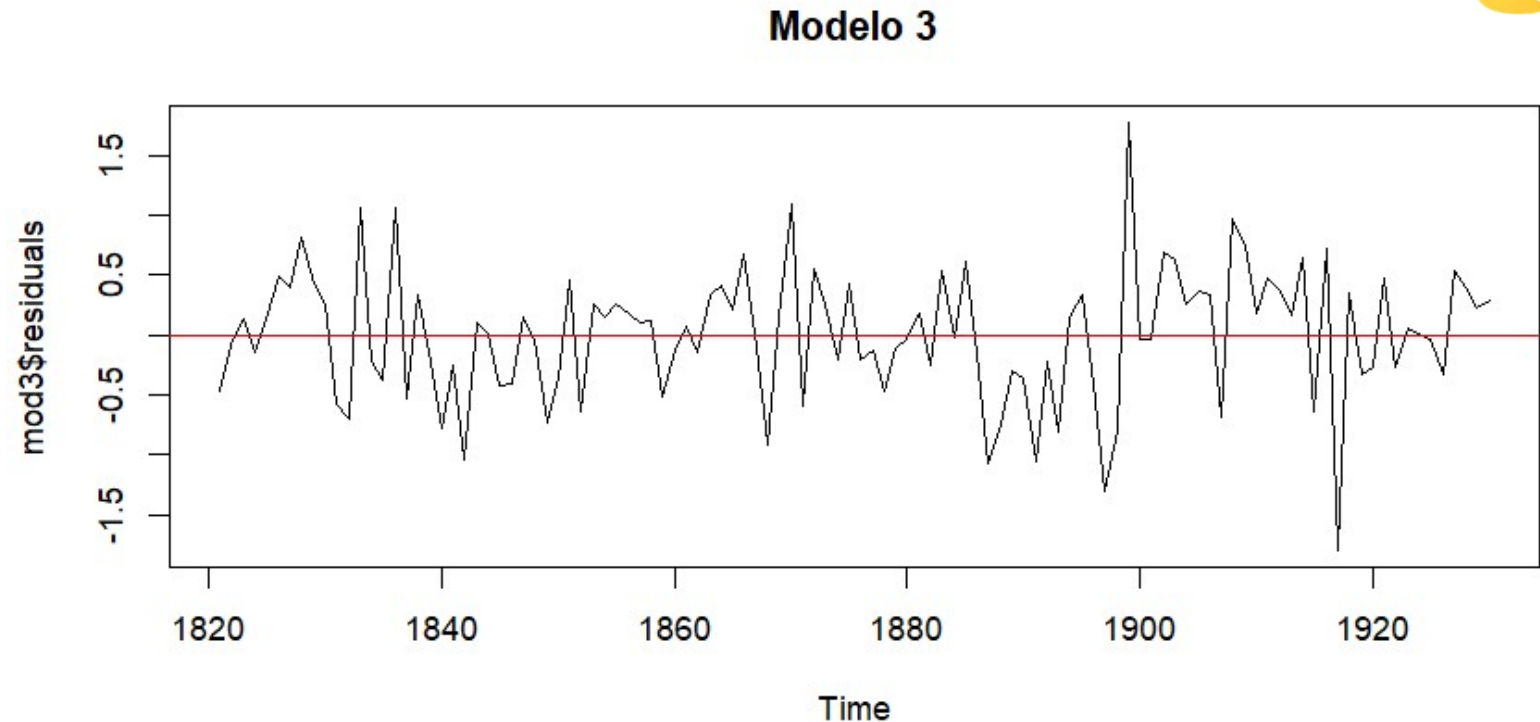
```
plot(mod3$residuals)
abline(h = 0, col = "red")
```

Prueba t

```
t.test(mod3$residuals, mu = 0)
```

One Sample t-test

```
data: mod3$residuals
t = -0.0086125, df = 109, p-value = 0.9931
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.1046138  0.1037085
sample estimates:
 mean of x
-0.0004526259
```



Del mismo modo un buen número de residuales están en torno a la media igual a cero.

Para confirmar lo mencionado se realiza la prueba t:
Como $p = 0.9931 > \alpha = 0.05$, se acepta H_0 , es decir la media es igual a cero.

Ejemplo 1

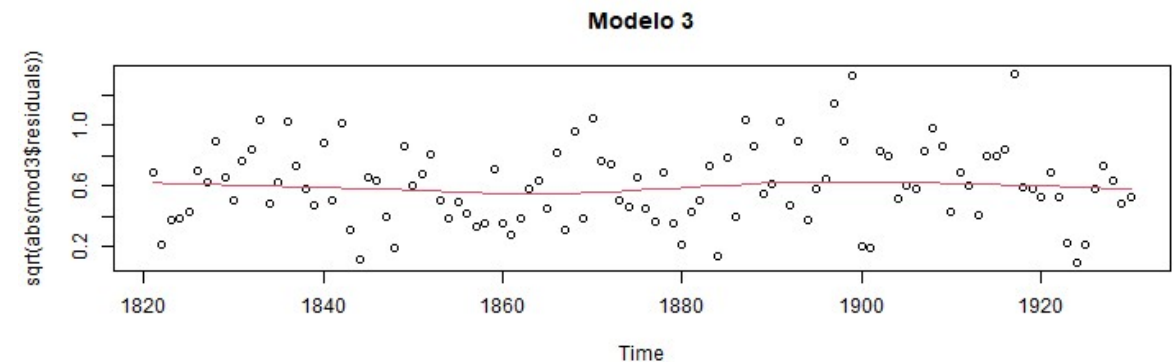
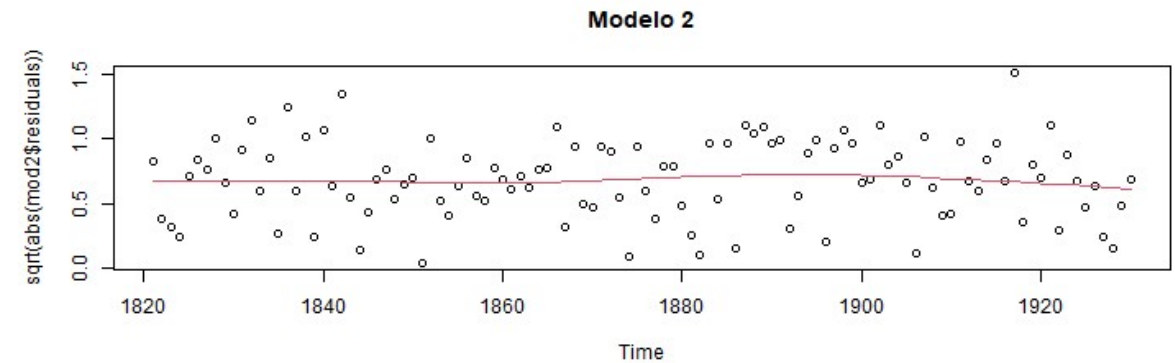
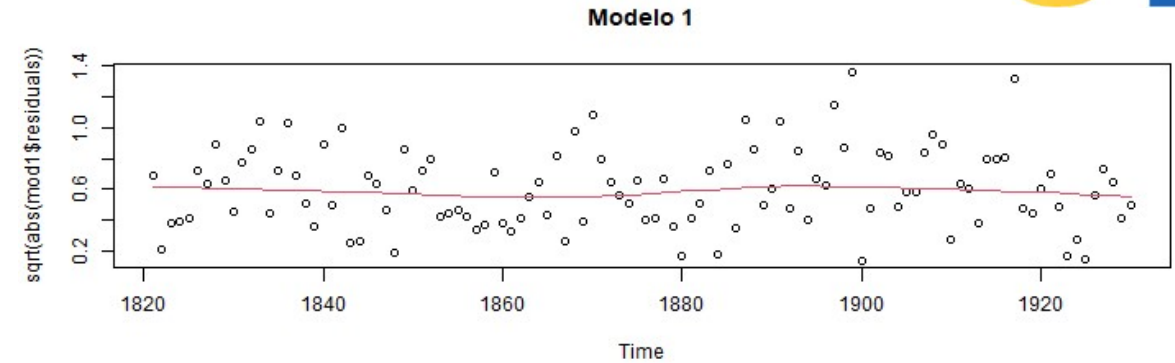


Homocedasticidad de los residuales (Varianza constante de los residuos)

Para analizar la homocedasticidad de los residuales realizaremos el diagrama de dispersión de la **raíz cuadrada de los residuales absolutos**.

```
par(mfrow = c(3,1))
scatter.smooth(sqrt(abs(mod1$residuals)), 1,
pars=list(col=2), main = "Modelo 1")
scatter.smooth(sqrt(abs(mod2$residuals)), 1,
pars=list(col=2), main = "Modelo 2")
scatter.smooth(sqrt(abs(mod3$residuals)), 1,
pars=list(col=2), main = "Modelo 3")
```

Se observa que los datos parecen no presentar una **variabilidad considerable**, por tanto, será necesario realizar la prueba de **Breusch-Pagan** para determinar finalmente si las varianzas constantes para los modelos.



Ejemplo 1



Prueba de Breusch - Pagan

Modelo 1:

```
obs=get(mod1$series)
bptest(resid(mod1)~I(obs-resid(mod1)))
      studentized Breusch-Pagan test
```

```
data:  resid(modelo1) ~ I(obs - resid(modelo1))
BP = 6.9701, df = 1, p-value = 0.008288
```

Modelo 2:

```
obs=get(mod2$series)
bptest(resid(mod2)~I(obs-resid(mod2)))
      studentized Breusch-Pagan test
```

```
data:  resid(modelo2) ~ I(obs - resid(modelo1))
BP = 0.82692, df = 1, p-value = 0.3632
```

Modelo 3:

```
obs=get(mod3$series)
bptest(resid(mod3)~I(obs-resid(mod3)))
      studentized Breusch-Pagan test
```

```
data:  resid(modelo2) ~ I(obs - resid(modelo1))
BP = 6.3902, df = 1, p-value = 0.01148
```

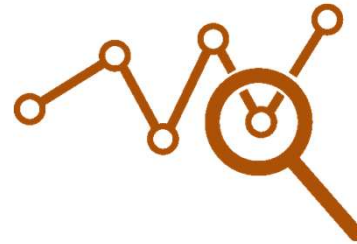
El valor de probabilidad (p-valor) asociado al estadístico BP asume un valor de 0.008288 y 0.01148 para el modelo 1 y modelo 3 respectivamente y estos son menores $\alpha = 0.05$, entonces se rechaza H_0 . Por tanto, estos residuos **no tienen varianza constante**.

En cambio, en el modelo 2, el p-valor = $0.3334 > 0.05$, con esto podemos decir que los residuales de este modelo **son constantes**.

Finalmente, podemos indicar que solo en modelo 2 estaría cumpliendo este supuesto.



Para solucionar este problema, aparecen modelos nombrados como modelos con **heterocedasticidad condicional autoregresiva (ARCH)**, los cuales se analizarán en los siguientes capítulos.



3.2. Análisis de los residuos

c) Ausencia de correlación serial

Si los residuos se comportaran como un ruido blanco, los coeficientes de la FAS y FAP muestrales deben ser prácticamente nulos para todos los retardos. Para comprobarlo, se pueden llevar a cabo:

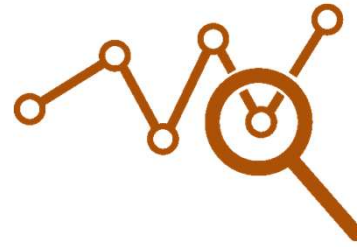
- **Contrastes de significatividad individual** sobre los coeficientes de autocorrelación:

$$H_0: \rho_k(a_t) = 0 \text{ (} a_t \text{ es ruido blanco)}$$

$$H_1: \rho_k(a_t) \neq 0 \text{ (} a_t \text{ no es ruido blanco)}$$

$$\left. \begin{array}{l} H_0: \rho_k(a_t) = 0 \text{ (} a_t \text{ es ruido blanco)} \\ H_1: \rho_k(a_t) \neq 0 \text{ (} a_t \text{ no es ruido blanco)} \end{array} \right\} \rho_k(a_t) \sim N\left(0, \frac{1}{T}\right)$$

Si: $p < \alpha$, se rechaza la hipótesis nula H_0 .



3.2. Análisis de los residuos

- Contrastes global

La prueba de Ljung-Box

- Es un tipo de prueba estadística de si un grupo cualquiera de autocorrelaciones de una serie de tiempo son diferentes de cero.
- Es conocida como la prueba Q de Ljung-Box, y está estrechamente relacionada con la prueba de Box-Pierce (1970).

La hipótesis es:

$H_0: \rho_1 = \rho_2 = \rho_3 = \dots = \rho_k = 0$ (residuos independientes)

$H_1: \rho_1 \neq \rho_2 \neq \rho_3 \neq \dots \neq \rho_k \neq 0$ (residuos no independientes)

Estadístico de prueba: $Q_{BL} = n(n+2) \sum_{i=1}^k \frac{\hat{\rho}_i^2}{n-i} \approx \chi_{k,1-\alpha}^2$

Se rechaza H_0 , si $Q_{BL} > \chi_{k,1-\alpha}^2$, también si: $p < \alpha$, se rechaza la hipótesis nula H_0 .

Ejemplo 1



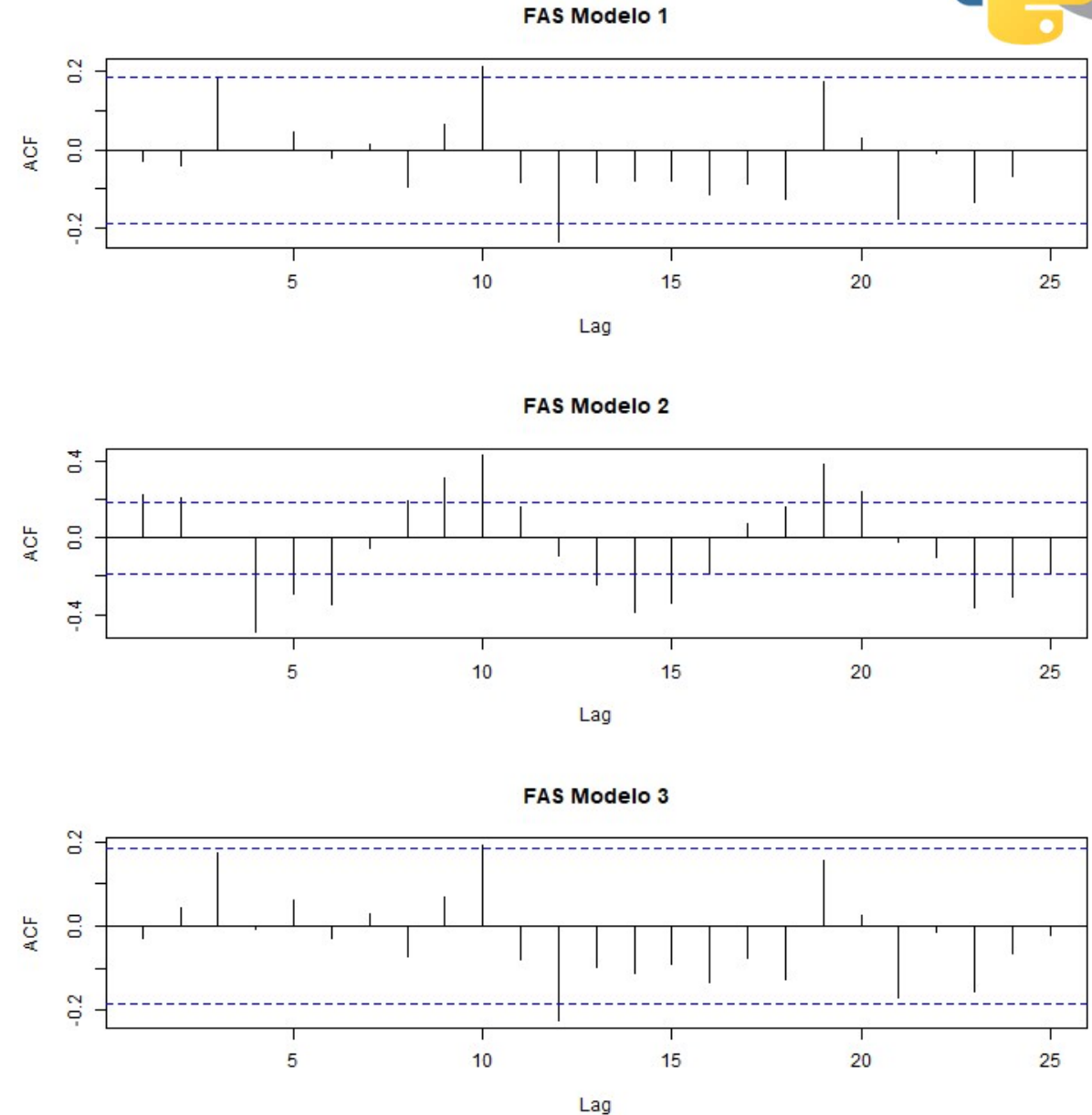
Correlograma de los residuos

```
resid_m1 <- as.vector(mod1$residuals)
resid_m2 <- as.vector(mod2$residuals)
resid_m3 <- as.vector(mod3$residuals)
```

```
par(mfrow = c(3,1))
FAS_e.m1 <- acf(resid_m1, lag.max = 25,
main="FAS Modelo 1", level = 0.95)
FAS_e.m2 <- acf(resid_m2, lag.max = 25,
main="FAS Modelo 2", level = 0.95)
FAS_e.m3 <- acf(resid_m3, lag.max = 25,
main="FAS Modelo 3", level = 0.95)
```

Se observa que casi la totalidad de los coeficientes del FAS para los modelos 1 y 3 se encuentran dentro de las bandas de no significación, sobre todo los de los primeros retardos. Por el contrario, el modelo 2 presenta muchos coeficientes que superan los límites de confianza indicando problemas de autocorrelación entre residuales..

Por tanto, tenemos **altos indicios** de que los residuos de los modelos 1 y 3 sean **ruido blanco**.



Ejemplo 1



Prueba de Ljung - Box

```
Box.test(resid_m1,type = "Ljung-Box")
```

Box-Ljung test

```
data: resid_m1  
X-squared = 0.096786, df = 1, p-value = 0.7557
```

```
Box.test(resid_m2,type = "Ljung-Box")
```

Box-Ljung test

```
data: resid_m2  
X-squared = 5.4867, df = 1, p-value = 0.01916
```

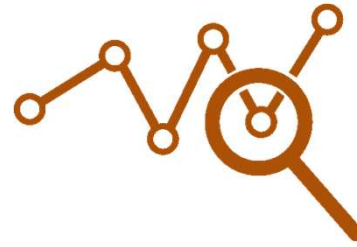
```
Box.test(resid_m3,type = "Ljung-Box")
```

Box-Ljung test

```
data: resid_m3  
X-squared = 0.10067, df = 1, p-value = 0.751
```

Confirmamos lo mencionado en los correlogramas, en los modelos 1 y 3 se acepta la hipótesis nula ($p\text{-valor} > 0.05$) de que los coeficientes de autocorrelación son cero; es decir, **los residuos son independientes o están incorrelacionados**.

Y en el modelo 2 se rechaza la hipótesis nula ($p\text{-valor} < 0.05$) indicándonos que los residuos están correlacionados.



3.2. Análisis de los residuos

d) Contraste de normalidad

- Para examinar el supuesto de normalidad del ruido del proceso, existen diversas pruebas tanto de tipo gráfico, como de tipo analítico.
- El estadístico más utilizado es el de Jarque-Bera, una prueba que tiene propiedades óptimas de potencia asintótica, es la referida por Jarque y Bera, (1980) planteando la siguiente hipótesis:

H_0 : a_t se aproxima a una distribución normal

H_1 : a_t no se aproxima a una distribución normal

El estadístico es:

$$JB = \frac{N-k}{6} \left(A_s^2 + \frac{(K-3)^2}{4} \right) \approx \chi_{(2)g.l.}^2$$

donde: A_s es el coeficiente de asimetría y K es el coeficiente de curtosis.

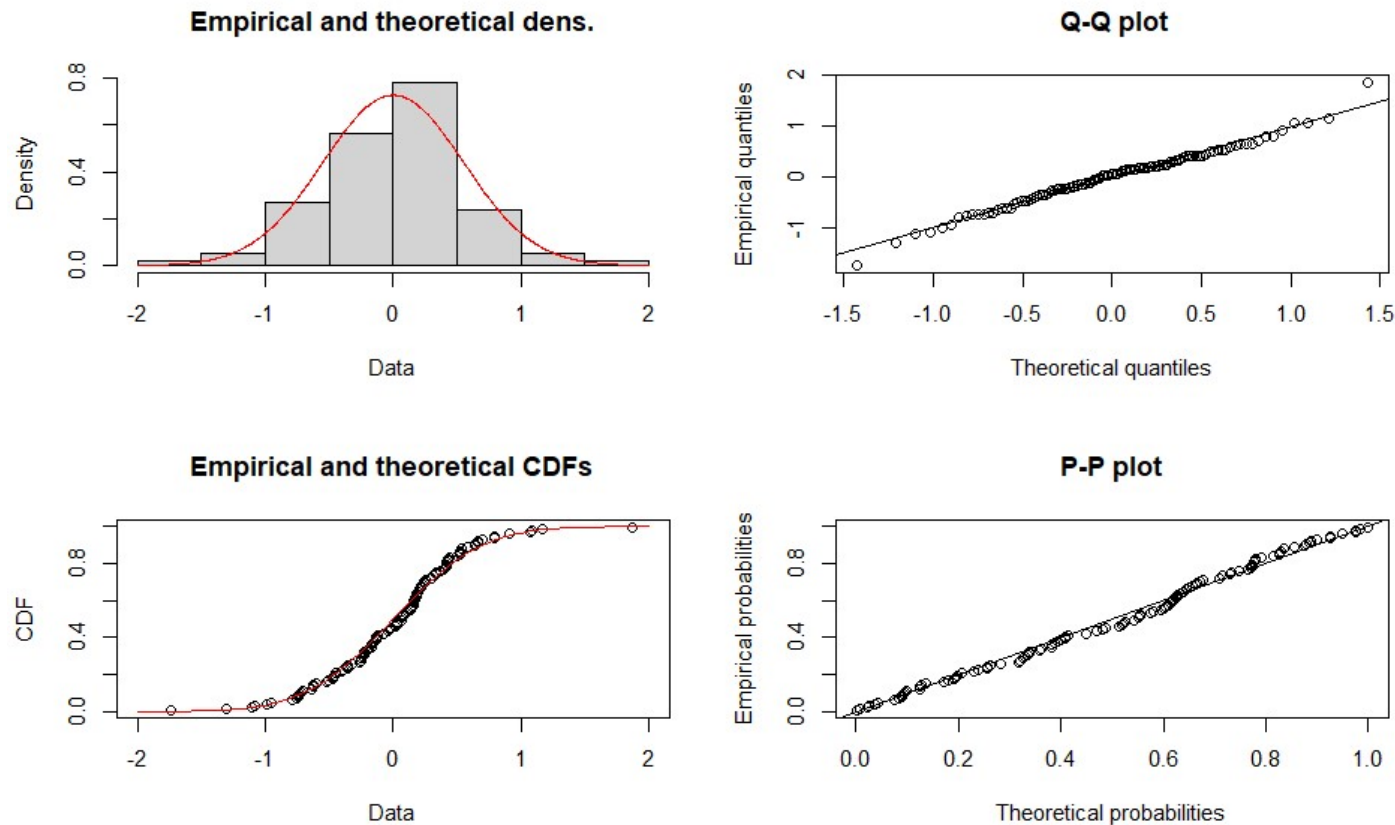
Se rechaza la hipótesis nula de normalidad al nivel de significación del 5% si $JB > 5.99$. También, si: $p < \alpha$, se rechaza la hipótesis nula H_0 .

Ejemplo 1



Método gráfico - Modelo 1

```
library(fitdistrplus)
ajuste_m1<-fitdist(data = resid_m1, distr="norm")
plot(ajuste_m1)
```



Prueba de normalidad de Jarque-bera

```
JB_m1 <- jarque.bera.test(resid_m1)
JB_m1
```

Jarque Bera Test

```
data: resid_m1
X-squared = 4.7522, df = 2, p-value = 0.09291
```

En las figuras se observa que los residuales del modelo 1 presentan indicios de ajuste hacia la distribución normal.

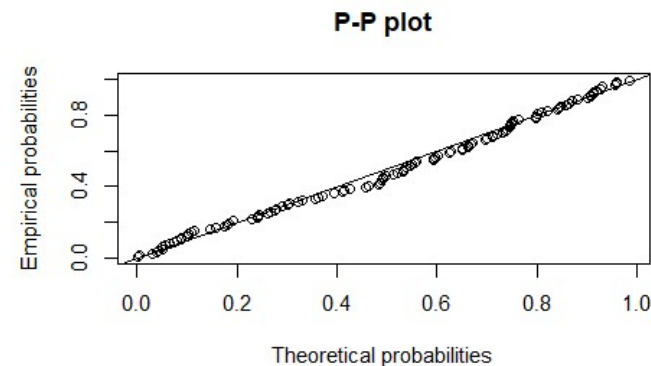
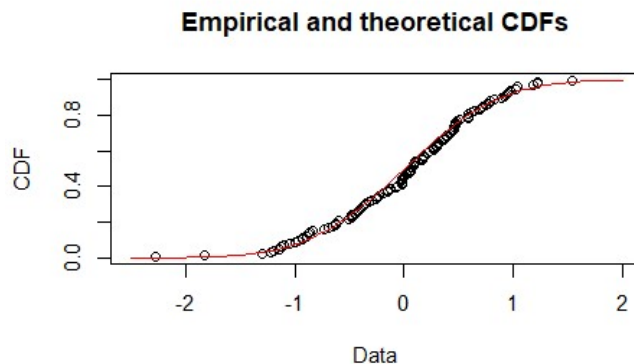
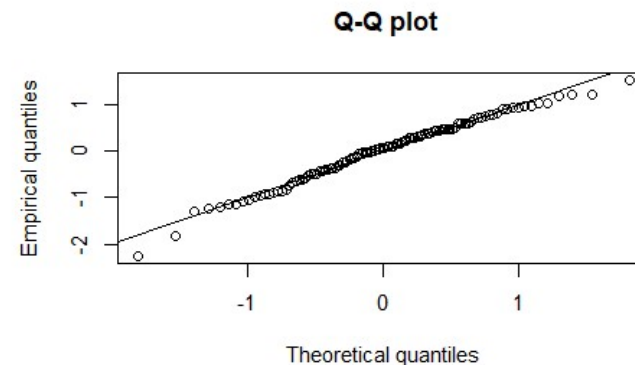
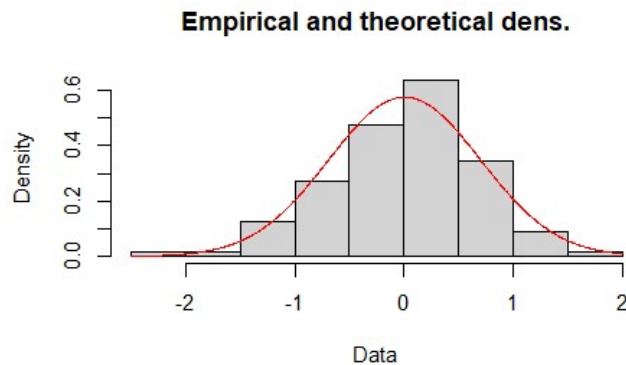
En la prueba JB, como $p = 0.09291 > 0.05$, se acepta H_0 , es decir, los residuos **se aproximan a una distribución normal**.

Ejemplo 1



Método gráfico - Modelo 2

```
ajuste_m2<-fitdistr(data = resid_m2, distr="norm")  
plot(ajuste_m2)
```



Prueba de normalidad de Jarque-bera

```
JB_m2 <- jarque.bera.test(resid_m2)  
JB_m2
```

Jarque Bera Test

```
data: resid_m2  
X-squared = 3.8952, df = 2, p-value = 0.1426
```

En las figuras se observa que los residuales del modelo 2 también presentan altos indicios de ajuste hacia la distribución normal.

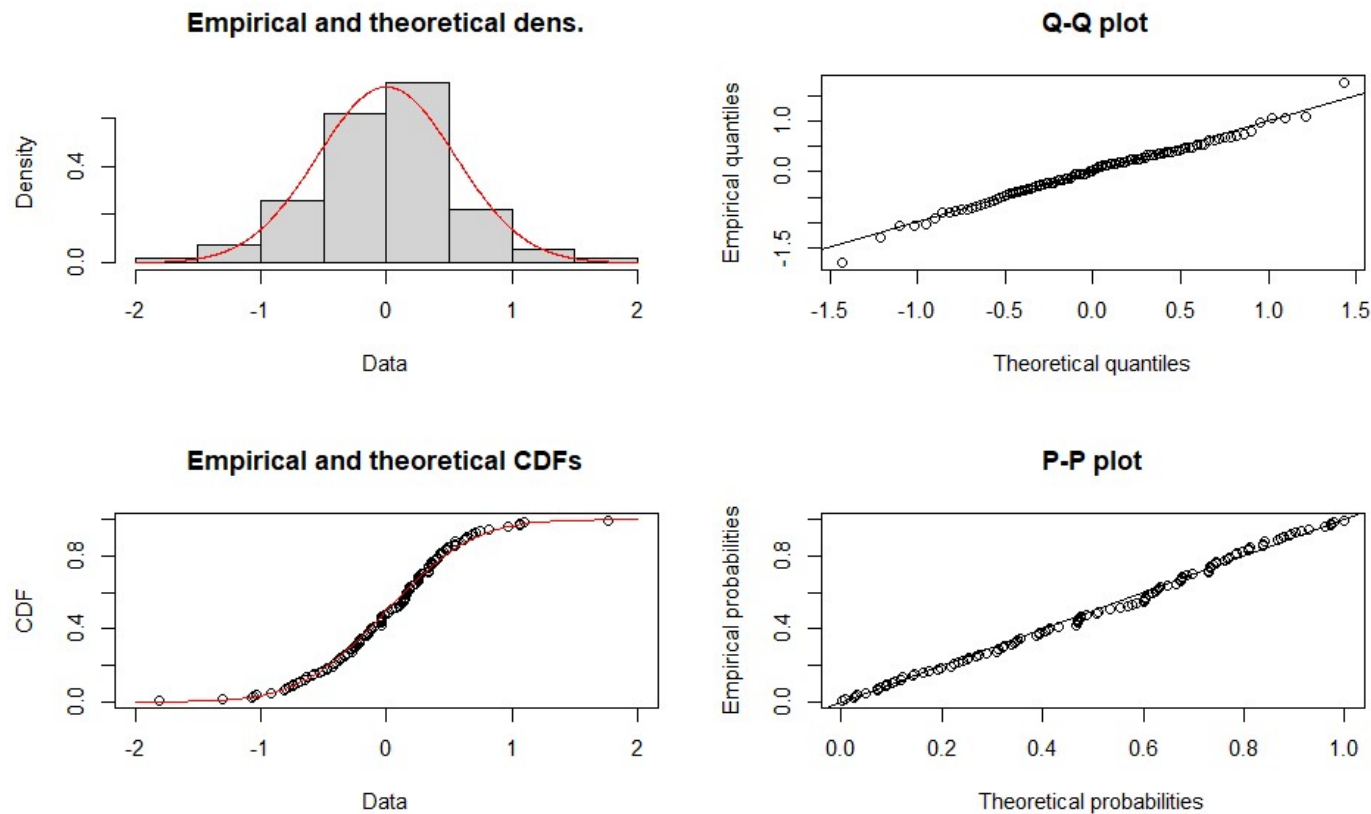
En la prueba JB, como $p = 0.1426 > 0.05$, se acepta H_0 , es decir, los residuos **se aproximan a una distribución normal**.

Ejemplo 1



Método gráfico - Modelo 3

```
ajuste_m3<-fitdist(data = resid_m3, distr="norm")  
plot(ajuste_m3)
```



Prueba de normalidad de Jarque-bera

```
JB_m3 <- jarque.bera.test(resid_m2)  
JB_m3
```

Jarque Bera Test

```
data: resid_m3  
X-squared = 3.8414, df = 2, p-value = 0.1465
```

También en las figuras se observa que los residuales del modelo 3 presentan altos indicios de ajuste hacia la distribución normal.

En la prueba JB, como $p = 0.1465 < 0.05$, se acepta H_0 , es decir, los residuos **se aproximan a una distribución normal**.

Los tres modelos propuestos satisfacen los contrastes de diagnóstico por lo que recogen la estructura de la serie Pieles de Lince.





The background of the slide features a warm, orange-toned aerial view of a city, likely New York City, with a river visible. Overlaid on this are several abstract elements: a network diagram in the top left with nodes in pink, yellow, and blue connected by lines; a molecular model in the top right with blue and yellow spheres; and two thick, dark diagonal lines in the corners. A large, dark brown circle with a dashed border is centered on the slide, containing the text.

GRACIAS

<https://aulavirtual2.unap.edu.pe/>