

CAPÍTULO 7

Pruebas de Incorrelación

7.1. Introducción

En el Modelo de Componentes Estructurales aditivo la serie se asume que se puede descomponer en las componentes de tendencia, estacionalidad y error:

$$Y_t = T_t + S_t + \varepsilon_t, \quad t \in \mathbb{Z}.$$

La componente que interesa analizar en este capítulo es ε_t . La serie estimada $\hat{\varepsilon}_t$, denominada “residuo estructural”, resulta de sustraer de Y_t las estimaciones de las dos primeras componentes, es decir

$$\hat{\varepsilon}_t = Y_t - \hat{T}_t - \hat{S}_t.$$

Para calcular los pronósticos \hat{Y}_{T+j} , $j = 1, \dots, h$, se debe analizar la serie $\hat{\varepsilon}_t$ para detectar si tiene una dinámica autocorrelacionada ó, por el contrario, es solamente ruido aleatorio. En el primer caso es posible pronosticar los residuos estructurales e incorporar éstos a los pronósticos con la tendencia y la componente estacional. En el segundo caso el pronóstico de un ruido aleatorio es cero. Para clarificar la idea de ruido aleatorio se introduce la siguiente definición básica.

Definición 7.1.1 (Ruido Blanco). Una serie $(\varepsilon_t, t \in \mathbb{Z})$ se dice que es Ruido Blanco si cumple

$$\mathbb{E}(\varepsilon_t) \equiv 0, \quad (7.1a)$$

$$Var(\varepsilon_t) \equiv \sigma^2, \quad (7.1b)$$

$$\forall k \neq 0, Cov(\varepsilon_t, \varepsilon_{t+k}) = 0. \quad (7.1c)$$

Se denota $\varepsilon_t \sim RB(0, \sigma^2)$.

Si además cumple que $\varepsilon_t \sim N(0, \sigma^2)$ se dice que ε_t es Ruido Blanco Gaussiano y se denota $\varepsilon_t \sim RBG(0, \sigma^2)$. Nótese que en este caso la sucesión debe ser iid, $\varepsilon_t \sim iid N(0, \sigma^2)$, ya que una sucesión incorrelacionada que se distribuya Normal debe ser iid. En este caso también se cumple $\mathbb{E}(\varepsilon_{t+k}|\varepsilon_t) = \mathbb{E}(\varepsilon_{t+k}) = 0$, luego, en una sucesión $\varepsilon_t \sim iid N(0, \sigma^2)$ los pronósticos a k pasos son cero.

Obsérvese que si no se cumple $\forall k \neq 0, Cov(\varepsilon_t, \varepsilon_{t+k}) = 0$ entonces se tendría $\exists k \neq 0, Corr(\varepsilon_t, \varepsilon_{t+k}) \neq 0$. Lo cual implica que los errores estarán autocorrelacionados, y por tanto, podrían ser pronosticados. Por ejemplo, con $T_t = \beta_0 + \beta_1 t$,

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t, \quad t = 1, \dots, T, \quad \varepsilon_t \sim RBG(0, \sigma^2)$$

$$\hat{Y}_{T+k} = \hat{\beta}_0 + \hat{\beta}_1(T+k) + \hat{\varepsilon}_{T+k}, \quad k = 1, 2, \dots, h,$$

Verificar las condiciones (7.1a), (7.1b) y (7.1c) se denomina: “Validación de los supuestos de los errores”. La condición (7.1a) se cumple automáticamente si ε_t es el residuo de la regresión $Y_t = T_t + S_t + \varepsilon_t$. Las otras dos condiciones se comprueban aplicando pruebas de hipótesis adecuadas. El resto del capítulo se dedicará a exponer pruebas para (7.1c), denominadas “pruebas de incorrelación”.

Para la condición de varianza constante, (7.1b), las pruebas disponibles en la literatura, por ejemplo, White, Breusch-Pagan y Goldfeld-Quandt, se denominan “pruebas de heterocedasticidad”. Asumen que la hipótesis alterna equivale a suponer que $Var(\varepsilon_t)$ es una función de una combinación de las variables explicativas, por ejemplo, de t, t^2 , las variables indicadoras en las columnas de la matriz It , ó de los valores estimados \hat{Y}_t , que es una combinación particular de las anteriores.

Esta forma de la hipótesis alterna no es útil si se asume el modelo de componentes, ya que lo que se busca es precisamente eliminar el efecto de tales variables en los residuos estructurales. Rechazar la hipótesis nula de varianza constante llevaría a

tratar de corregir el efecto de heterocedasticidad y a modificar el modelo de manera sustancial. Por lo que optamos por no realizar la validación de este supuesto en esta etapa del modelo y concentrarnos en las pruebas de incorrelación.

7.2. Series Estacionarias en Covarianza

Para determinar si $\widehat{\varepsilon}_t, t = 1, \dots, T$ presenta autocorrelación o no, se introducen dos pruebas de hipótesis y la función de autocovarianza. El concepto de serie estacionaria es básico en la teoría de series de tiempo. Es diferente del concepto de serie estacional.

Al considerar una serie de tiempo Y_t se asumió que el índice t toma valores enteros, negativos y positivos, $t \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, escribiendo $(Y_t, t \in \mathbb{Z})$. Los valores observados de Y_t forman un conjunto finito que se denomina una *trayectoria muestral*, denotada $\{Y_1, Y_2, \dots, Y_T\}$.

Definición 7.2.1 (Serie de Tiempo Estacionaria en Covarianza). *Una serie $(Y_t, t \in \mathbb{Z})$ se dice estacionaria en covarianza si cumple dos condiciones.*

$$\mathbb{E}(Y_t) \equiv \mu, \quad (7.2a)$$

$$Cov(Y_{t_1}, Y_{t_2}) = R(t_2 - t_1). \quad (7.2b)$$

donde $Cov(Y_{t_1}, Y_{t_2}) = E((Y_{t_1} - \mu)(Y_{t_2} - \mu))$ y $R(t)$ es una función par, $R(-t) = R(t), \forall t$.

Es decir, la covarianza entre Y_{t_1} y Y_{t_2} depende únicamente de la distancia entre los tiempo t_2 y t_1 , y no de si $t_1 < t_2$ ó $t_1 > t_2$. Es decir,

$$Cov(Y_{t_1}, Y_{t_2}) = R(t_2 - t_1) = R(t_1 - t_2) = Cov(Y_{t_2}, Y_{t_1}).$$

Por ejemplo,

$$Cov(Y_{-4}, Y_3) = R(3 - (-4)) = R(7) = Cov(Y_0, Y_7),$$

$$Cov(Y_3, Y_4) = R(4 - 3) = R(1) = Cov(Y_0, Y_1).$$

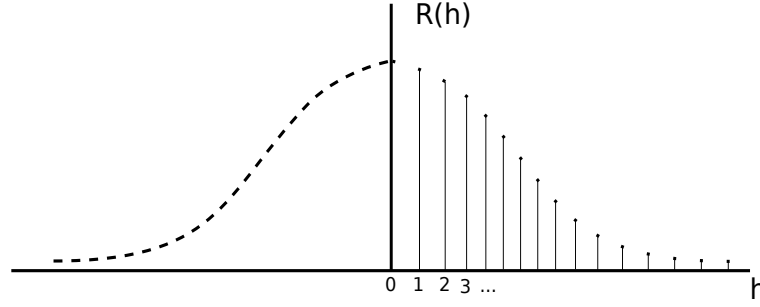


Figura 7.1: Ejemplo Función de Autocovarianza

Nótese que por ser funciones pares las autocovarianzas solamente se grafican en el semieje positivo. Las siguientes identidades son inmediatas.

$$\begin{aligned}
 Cov(Y_t, Y_{t+h}) &= E((Y_t - \mu)(Y_{t+h} - \mu)) \\
 &= E(Y_t Y_{t+h}) - \mu E(Y_t) - \mu E(Y_{t+h}) + \mu^2 \\
 &= E(Y_t Y_{t+h}) - \mu^2 - \mu^2 + \mu^2 \\
 &= E(Y_t Y_{t+h}) - \mu^2.
 \end{aligned}$$

Algunos ejemplos de funciones de autocovarianza para series estacionarias en covarianza se muestran a continuación.

$$R(h) = \sigma^2 e^{-\alpha|h|}, \quad \alpha > 0, \quad (7.3a)$$

$$R(h) = \sigma^2 e^{-\alpha|h|} \cos(\beta h), \quad \alpha > 0, \beta \in \mathbb{R}, \quad (7.3b)$$

$$R(h) = \sigma^2 e^{-\alpha|h|} \left(\cos(\beta h) + \frac{\alpha}{\beta} \sin(\beta|h|) \right), \quad \alpha > 0, \beta \in \mathbb{R}, \quad (7.3c)$$

$$R(h) = 2\alpha\sigma^2(1 - 2\alpha h^2)e^{-\alpha h^2}, \quad \alpha > 0, \quad (7.3d)$$

$$R(h) = \sigma^2 e^{-\alpha^2 h^2}, \quad \alpha > 0, \quad (7.3e)$$

$$R(h) = \sigma^2 e^{-\alpha|h|} \left(1 + \alpha|h| + \frac{1}{3}\alpha^2 h^2 \right), \quad \alpha > 0, \quad (7.3f)$$

Algunas propiedades de las series estacionarias en covarianza

Definición 7.2.2. *Un componente determinístico es una serie Z_t que tiene la propiedad de que si se conoce su historia hasta un tiempo t_0 , dada por $Z_s, s \leq t_0$, entonces se puede conocer Z_t para $t > t_0$.*

Por ejemplo, si $A \sim N(\mu, \sigma^2)$ entonces $Z_t = A \cos(\lambda t), t \geq 0$ es un componente determinístico.

Algunas propiedades permiten formarse una idea más clara de como es una serie estacionaria en covarianza.

1. Una serie estacionaria en covarianza con trayectorias monótonas (crecientes ó decrecientes) solamente puede ser una constante. Es decir,

$$\begin{aligned} Y_t \leq Y_{t+1}, \forall t \in \mathbb{Z} &\Rightarrow Y_t \equiv Y_0, \\ Y_t \geq Y_{t+1}, \forall t \in \mathbb{Z} &\Rightarrow Y_t \equiv Y_0. \end{aligned}$$

2. Una serie estacionaria en covarianza toma valores en un intervalo $[a, b]$ infinitas veces ó nunca. Es decir, si

$$\begin{aligned} A &= (\#\{t : Y_t \in [a, b]\} = 0), \\ B &= (\#\{t : Y_t \in [a, b]\} = \infty), \end{aligned}$$

entonces $\mathbb{P}(A \cup B) = 1$ y $\mathbb{P}(A \cap B) = 0$.

3. En vista de las dos propiedades anteriores, una serie estacionaria en covarianza fluctúa alrededor de la media $\mathbb{E}(Y_t) = \mu$. Y se mantiene dentro de un intervalo finito. Se dice entonces que presenta reversión en la media.
4. La función de autocovarianza de una serie estacionaria en covarianza sin componentes determinísticas, $R(h)$, tiende a cero cuando h tiende a infinito:

$$\lim_{h \rightarrow \infty} R(h) = 0.$$

5. Como este resultado es de la forma $p \Rightarrow q$, se cumple la contra-recíproca, es decir, $no(q) \Rightarrow no(p)$. Por lo tanto, si la función de autocorrelación $R(h)$ no tiende a cero la serie no puede ser estacionaria en covarianza.

En el modelo de componentes $Y_t = T_t + S_t + \epsilon_t$, lo que se espera es que los residuos ϵ_t sean una serie estacionaria en covarianza, incluyendo un ruido blanco. Un instrumento útil para examinar la posibilidad de que una serie sea estacionaria es la función de autocorrelación.

Definición 7.2.3 (Función de Autocorrelación). *La función de autocorrelación teórica de una serie estacionaria en covarianza (Y_t , $t \in \mathbb{Z}$) se define como*

$$\rho(h) := \text{Corr}(Y_t, Y_{t+h}) = \frac{R(h)}{R(0)}, \quad h = 0, 1, 2, \dots$$

La función de autocorrelación se indicará con las siglas “fac” (acf en inglés). ,

Como $|\rho(h)| \leq 1$ se cumple entonces que $\left| \frac{R(h)}{R(0)} \right| \leq 1$, es decir $|R(h)| \leq |R(0)| = |\sigma^2| = \sigma^2$, con $R(0) = \text{Var}(Y_t) = \sigma^2 > 0$, entonces $|R(h)| \leq \sigma^2 < \infty$, $\forall h$.

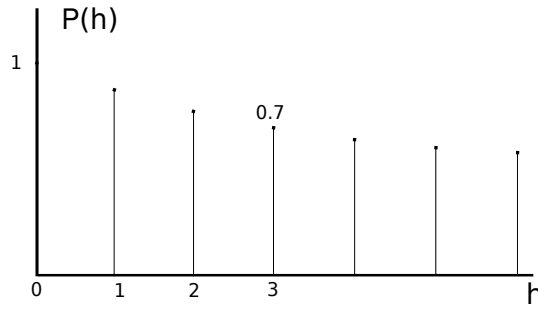


Figura 7.2: Ejemplo Función de Autocorrelación.

Los valores de $\rho(h)$ indican qué grado de asociación lineal existe entre dos valores de la serie Y_t separados una distancia en el tiempo de h . Por ejemplo, $\rho(3) = 0.7$ se interpreta como que Y_t y Y_{t+3} tienen una correlación de 0.7.

Si se observan las gráficas de $\rho(k)$ en las figuras (7.2), éstas parecen tender a cero, luego se puede aceptar que hay evidencia de que las series correspondientes son estacionarias en covarianza.

7.3. Estimación de las funciones de Autocovarianza y Autocorrelación.

A partir de la expresión teórica para $\rho(k)$

$$\rho(k) = \frac{E((Y_t - \mu)(Y_{t+k} - \mu))}{E(Y_t - \mu)^2} \quad (7.4)$$

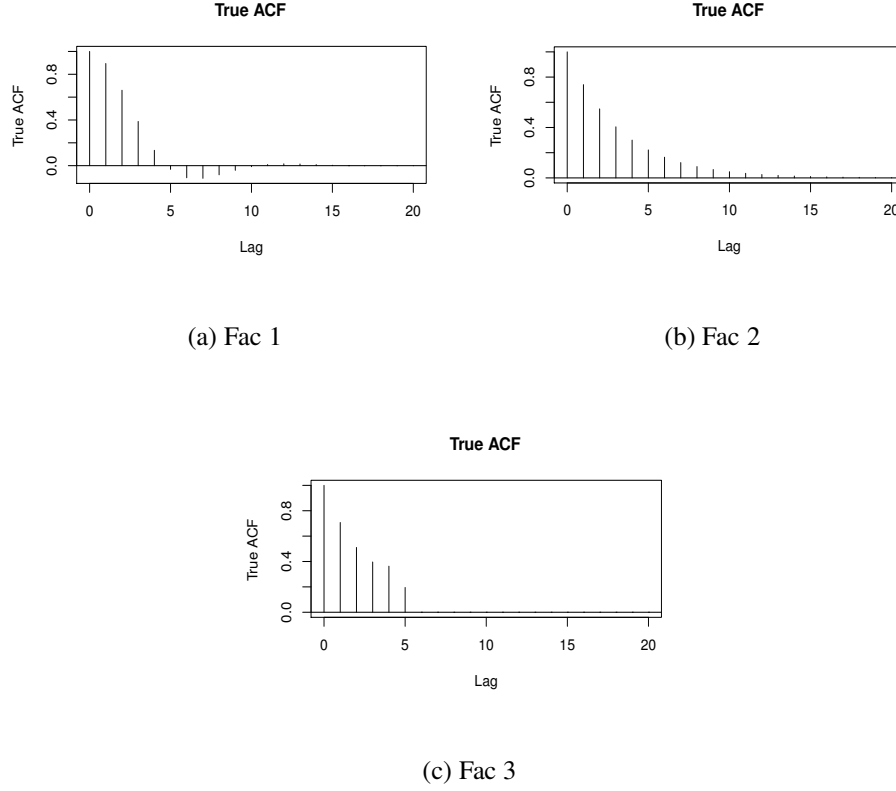


Figura 7.3: Ejemplos de Funciones de Autocorrelación de un Proceso Estacionario

se puede obtener el estimador (de momentos), $\hat{\rho}(k)$, basado en la trayectoria $\{Y_1, Y_2, \dots, Y_T\}$

$$\hat{\rho}(k) = \frac{\sum_{j=k+1}^T (Y_j - \bar{Y})(Y_{j-k} - \bar{Y})}{\sum_{j=1}^T (Y_j - \bar{Y})^2}, \quad (7.5)$$

donde $\bar{Y} = \frac{1}{T} \sum_{j=1}^T Y_j$ es la media de la trayectoria. Una cuestión práctica es decidir el rango m de los valores de $k = 1, 2, \dots, m$. Una posibilidad es $m = T/4$.

También se utiliza

$$\hat{\rho}(k) = \frac{\frac{1}{T-k+1} \sum_{t=k+1}^T (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\frac{1}{T-1} \sum_{t=1}^T (Y_t - \bar{Y})^2}. \quad (7.6)$$

El cálculo de la fac estimada se realiza en el lenguaje R mediante la función `acf(x)`. La opción `ci.type = "ma"` calcula las “bandas de Bartlett”, que se definen más adelante.

```
acf(x, lag.max = m, ci.type = "ma",
type = c("correlation", "covariance", "partial"))
#---opcion para eliminar la autocorrelacion en cero
require(TSA)
TSA::acf(x, lag.max = m, ci.type = "ma",
drop.lag.0 = TRUE)
```

Pregunta: Cómo debería ser la Función de Autocorrelación para un Proceso Ruido Blanco?.

Si $(\varepsilon_t, t \in \mathbb{Z})$ es ruido blanco, como $E(\varepsilon_t) = 0$, se cumple que la media de ε_t es constante y como $Cov(\varepsilon_t, \varepsilon_{t+k}) = \sigma^2$, para $k = 0$, entonces la función de autocovarianza cumple:

$$R(k) = \begin{cases} \sigma^2, & \text{si } k = 0 \\ 0, & \text{si } k \geq 1. \end{cases}$$

La función de autocorrelación teórica $\rho(k)$ (fac) de un ruido blanco es cero excepto en $k = 0$. Sin embargo, las autocorrelaciones muestrales $\hat{\rho}(k)$ de un ruido blanco tienen la propiedad siguiente.

Proposición 7.3.1 (Teorema de Bartlett). *Si $(\varepsilon_t, t \in \mathbb{Z})$ es ruido blanco entonces las autocorrelaciones muestrales $\hat{\rho}(k) = \widehat{Cov}(\varepsilon_t, \varepsilon_{t+k})$, $k = 1, 2, \dots, m$, donde $1 < m < T$ es un entero arbitrario, con base en una muestra $\{\varepsilon_1, \dots, \varepsilon_T\}$, cumplen*

- (i) $\hat{\rho}(k)$ son independientes.
- (ii) $\hat{\rho}(k) \stackrel{a}{\sim} N(0, 1/T)$ para T grande.

donde $\stackrel{a}{\sim}$ significa: se distribuye aproximadamente.

Es decir, si T es grande y ε_t es ruido blanco entonces todas las $\hat{\rho}(k)$ se distribuyen Normales de media cero y varianza $\frac{1}{T}$. Por propiedad de la Normal que dice que para $X \sim N(0, \sigma^2)$ entonces

$$P(-2\sigma \leq X \leq 2\sigma) \approx 0.95,$$

las autocorrelaciones $\hat{\rho}(k)$ deben estar en el intervalo $[-2/\sqrt{T}, 2/\sqrt{T}]$, con una probabilidad de 0.95 aproximadamente. El valor de m se puede tomar, como se mencionó, como $T/4$.

Definición 7.3.1. La región en \mathbb{R}^2 dada por $\{(x, y) \in [0, m] \times [-2/\sqrt{T}, 2/\sqrt{T}]\}$ se denomina “bandas de Bartlett”.

Se puede realizar un diagnóstico gráfico, con base en $\hat{\rho}(k)$, la fac estimada de los residuos estructurales estimados $\hat{\varepsilon}_t$, con base en la banda de Bartlett, para decidir si los residuos son ruido blanco. Si se observa

$$\forall k \leq m, \hat{\rho}(k) \in [-2/\sqrt{T}, 2/\sqrt{T}] \quad (7.7)$$

entonces hay evidencia de que los residuos estructurales son ruido blanco y por tanto se puede pasar a realizar pronósticos de la serie con la parte estructural. Si se observan algunos k para los cuales $\hat{\rho}(k) \notin [-2/\sqrt{T}, 2/\sqrt{T}]$, entonces se tiene evidencia de que el residuo estructural no es ruido blanco y tiene una dinámica autocorrelacionada que se puede modelar, por ejemplo mediante un proceso ARMA, con el fin de mejorar los pronósticos estructurales.

Como ilustración, al observar la Figura (7.4(a)), se puede concluir que hay evidencia de que Y_t es ruido blanco. En la Figura (7.4(b)) no se cumple que $\forall k, \hat{\rho}(k) \in [-2/\sqrt{T}, 2/\sqrt{T}]$, por tanto, hay evidencia de que Y_t no es ruido blanco.

Ejemplo 7.3.1. En ⁽¹⁾: “El Índice General de la Bolsa de Valores de Colombia o IGBC era el índice bursátil más representativo de Colombia creado por la Bolsa de Valores de Colombia (BVC). Este índice fue reemplazado por el nuevo índice COLCAP (en nov-2013)”. El código siguiente calcula la fac de la serie $y = \log(\text{igbc})$ entre 2009-02-02 y 2010-12-30, y de su primera diferencia $r = \text{diff}(\log(y), 1, 1)$. El resultado se muestra en la Figura (7.5).

¹https://es.wikipedia.org/wiki/Indice_General_de_la_Bolsa_de_Valores_de_Colombia

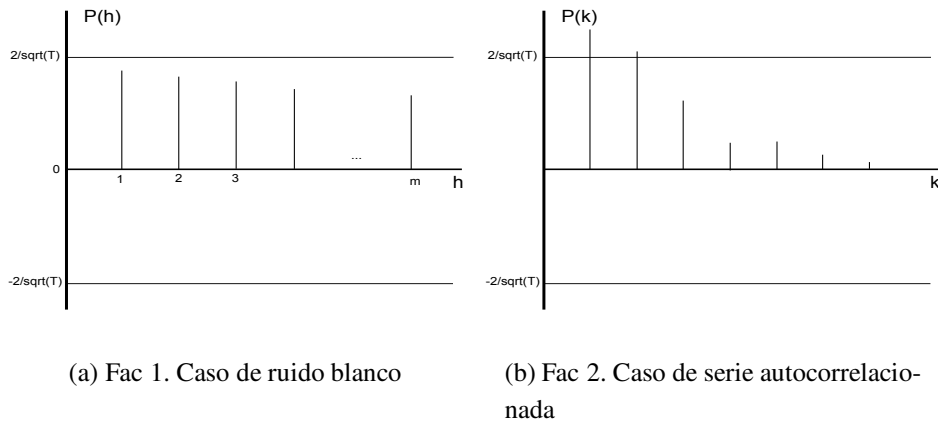


Figura 7.4: Ejemplos de Fac y Bandas de Bartlett

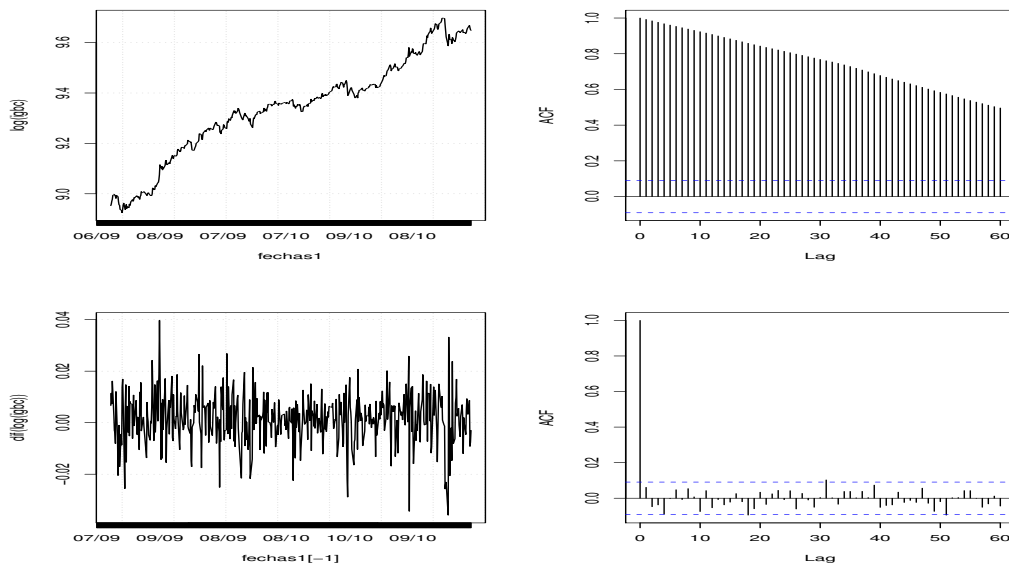


Figura 7.5: $\log(\text{igbc})$ y $\text{diff}(\log(\text{igbc}))$, a izquierda, con sus fac muestrales, a derecha

Puede apreciarse que la fac de $\log(\text{igbc})$ aparece muy persistente, tendiendo a cero lentamente, lo cual se puede interpretar como que no es estacionaria en covarianza. En el capítulo 7 se analizará este hecho mediante pruebas de hipótesis. En cambio, la fac de la serie de las diferencias aparece con todos sus valores dentro de las bandas de Bartlett, por lo que puede concluirse que puede tratarse de un ruido blanco. El

valor $m = 60$ corresponde a 60 días, y se escogió arbitrariamente.

```
# cálculo de la fac
require(TSA)
t = seq(1,length(y))
r = diff(y,1,1)
par(mfrow=c(2,2))
plot(t,y,type='l')
TSA::acf(y, lag.max = 60, ci.type = "ma",
drop.lag.0 = TRUE)
plot(t[-1],r,type='l')
acf(r,60)
```

Ejemplo 7.3.2. Retomando el Ejemplo (3.7.1) en la pag. 69, para modelación de la serie de producción de cemento Portland, trimestral, en toneladas por mil, entre Q1 1956 y Q3 1994, en Australia. Para la serie se asumió un modelo de la forma

$$Y_t = \beta_0 + \beta_1 t + \sum_{j=1}^3 \delta_j I_t(t) + \varepsilon_t.$$

Es decir, un modelo lineal para la tendencia, más una componente estacional, de período $s = 4$. Al estimar los residuos estructurales $\hat{\varepsilon}_t$ se obtienen las gráficas siguientes.

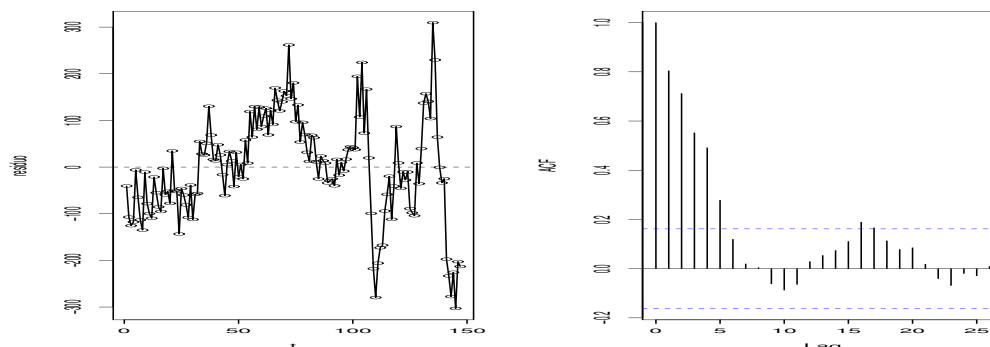


Figura 7.6: Fac de los Residuos Estructurales Serie Cementos

Es evidente que los residuos no son ruido blanco. En la gráfica de la fac muestral, en el panel derecho de la Figura 7.6, se puede apreciar cómo las primeras cinco

autocorrelaciones se salen de las Bandas de Bartlett. Además, en la gráfica del panel izquierdo, la gráfica de $\hat{\varepsilon}_t$ muestra períodos de incremento y luego de decrecimiento, que puede tomarse como evidencia de autocorrelación.

7.4. Pruebas de Incorrelación

Decidir si una serie es ruido blanco usando la función de autocorrelación estimada $\hat{\rho}(k)$ es equivalente a decidir si todas son estadísticamente cero, para $k \geq 1$. La gráfica de $\hat{\rho}(k)$ con las bandas de Bartlett dá un diagnóstico. Dos pruebas de hipótesis para chequear la incorrelación son:

1. Prueba Ljung-Box.
2. Prueba Durbin-Watson.

En las pruebas de hipótesis consideramos cuatro pasos:

1. Enunciado de las Hipótesis de la Prueba.
2. Estadístico de la Prueba.
3. Distribución del Estadístico bajo H_0 .
4. Decisión.

El Error Tipo I es rechazar la Hipótesis nula H_0 siendo ésta cierta. Se trata de un falso positivo porque en ocasiones la alterna H_1 es el resultado de interés. El nivel de significación α se define como

$$\mathbb{P}(\text{Error Tipo I}) = \alpha.$$

El Error Tipo II es no rechazar la nula H_0 siendo la alterna H_1 la cierta. Se trata de un falso negativo. La probabilidad de cometer este error es

$$\mathbb{P}(\text{Error Tipo I}) = \beta.$$

Además, la probabilidad $1 - \beta$, denominada la potencia de la prueba, se interpreta como la probabilidad de rechazar la nula correctamente, es decir, la probabilidad de rechazar H_0 cuando H_1 es cierta.

Las pruebas que tienen una baja potencia tienden a no rechazar la nula H_0 . No tienen capacidad para detectar H_1 . En ocasiones, la alterna H_1 puede estar muy cerca de la nula H_0 , y se requiere una prueba con alta potencia para detectarla.

7.4.1. Prueba Ljung-Box (LB)

A partir del Teorema de Bartlett, (7.3.1), si $\varepsilon_t \sim RB(0, \sigma^2)$, $t = 1, \dots, T$ y $\hat{\rho}(k) = \widehat{Corr}(\varepsilon_t, \varepsilon_{t+k})$, $k = 1, 2, \dots, m$, entonces, para T grande, se tiene

1. $\hat{\rho}(k) \stackrel{a}{\sim} N(0, 1/T)$.
2. $\hat{\rho}(k)$ son idenpendientes.

Luego $\sqrt{T}\hat{\rho}(k) \stackrel{a}{\sim} N(0, 1)$, y $(\sqrt{T}\hat{\rho}(k))^2 \stackrel{a}{\sim} \chi_1^2$, independientes, por lo tanto, por la propiedad reproductiva de la distribución χ_ν^2 ,

$$Q = T \sum_{k=1}^m \hat{\rho}^2(k) \stackrel{a}{\sim} \chi_m^2. \quad (7.8)$$

El estadístico de la prueba Ljung-Box se define a partir de este resultado.

1. Hipótesis de la Prueba

$$\begin{aligned} H_0 : \varepsilon_t &\sim RB(0, \sigma^2) \\ H_1 : no(H_0) \end{aligned}$$

2. Estadístico de Prueba

Ljung y Box modificaron el estadístico Q en (7.8) para mejorar la aproximación de Q a la χ_m^2 en muestras pequeñas, usando

$$Q_{LB} = T(T+2) \sum_{k=1}^m \frac{1}{T-k} \hat{\rho}^2(k). \quad (7.9)$$

3. Distribución del Estadístico de la Prueba bajo H_0

Si $H_0 : \varepsilon_t \sim RB(0, \sigma^2)$ es cierta se cumple $Q_{LB} \stackrel{a}{\sim} \chi_m^2$.

4. Decisión

- a) Colocando Q_{obs} el estadístico observado y *Valor p* $= P(\chi_m^2 \geq Q_{\text{obs}} | H_0 \text{ cierto})$, si *Valor p* < 0.05 , se rechaza H_0 . Si *Valor p* > 0.05 , no se rechaza H_0 .

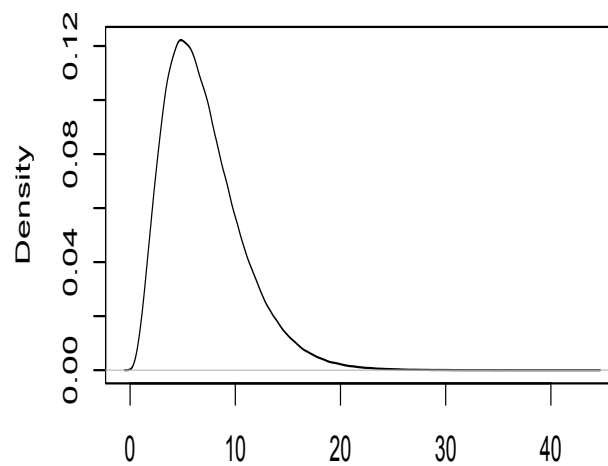


Figura 7.7: Densidad χ_m^2 .

Un Valor muy pequeño de Q_{obs} significa que *todas* las $\hat{\rho}^2(k)$ son pequeñas, luego no se rechaza H_0 .

- b) **Usando valores críticos:** Si $\chi_{m,\alpha}^2$ es valor crítico de χ_m^2 correspondiente a la probabilidad α , α nivel de significación, entonces si $Q_{\text{obs}} < \chi_{m,\alpha}^2$, no se rechaza H_0 y si $Q_{\text{obs}} > \chi_{m,\alpha}^2$ se rechaza H_0 . Es una prueba de cola derecha.

La prueba Ljung-Box se programa con la función `Box.test`. Por ejemplo

```
Box.test(rn, lag = 35 , type = "Ljung-Box")
X-squared = 379.43, df = 35, p-value < 2.2e-16,
```

por tanto, rechaza la hipótesis nula de que los residuos son Ruido Blanco.

Ejemplo 7.4.1. *Este es un ejemplo sobre la importancia de la hipótesis de Ruido Blanco en la prueba Ljung-Box. Suponga una sucesión periódica de período $p = 126$, dada por*

$$\sigma_t = 10(\sin(2\pi t/p) + 2), t = 1, 2, \dots$$

Y suponga una sucesión aleatoria $Z_t \sim iidN(0, 1)$. Se forma la sucesión $Y_t = \sigma_t Z_t$. Se tiene inmediatamente

$$\begin{aligned}\mathbb{E}(Y_t) &\equiv 0, \\ \text{Var}(Y_t) &= \text{Var}(\sigma_t Z_t) = \sigma_t^2, \\ \text{Corr}(Y_t, Y_{t+s}) &= \frac{\mathbb{E}(Y_t Y_{t+s})}{\sigma_t \sigma_{t+s}} = \frac{\mathbb{E}(Y_t) \mathbb{E}(Y_{t+s})}{\sigma_t \sigma_{t+s}} = 0, \quad t, s = 1, 2, \dots\end{aligned}$$

Entonces Y_t no es Ruido Blanco, es una serie incorrelacionada heterocedástica. En la Figura 7.8 se muestran trayectorias simuladas de Z_t y Y_t , y la fac de ésta. Qué resultado dá la prueba Ljung-Box?

```
#-----ruido  periódico
n = 350; p = 126; t = seq(1,n);
sn = 10*(sin(2*pi*t/p) + 2)
en = rnorm(n,0,1)
rn = sn*en
Box.test(rn, lag = 16 , type = "Ljung-Box")
data:  rn
X-squared = 28.407, df = 16, p-value = 0.02825
Box.test(rn, lag = 32 , type = "Ljung-Box")
Box-Ljung test
data:  rn
X-squared = 51.864, df = 32, p-value = 0.01462
```

Ejemplo 7.4.2. *Ejemplo con datos de la accion de Siemens AG. Ver la Figura 7.9. Calcular la prueba Ljung-Box para los rendimientos logarítmicos y calcular la fac para la serie nivel (precio) y la serie retorno. Se utiliza la función de R, `Box.test`.*

Código R 7.4.1.

```
#-----ejemplo accion Siemens bolsa Frankfurt
```

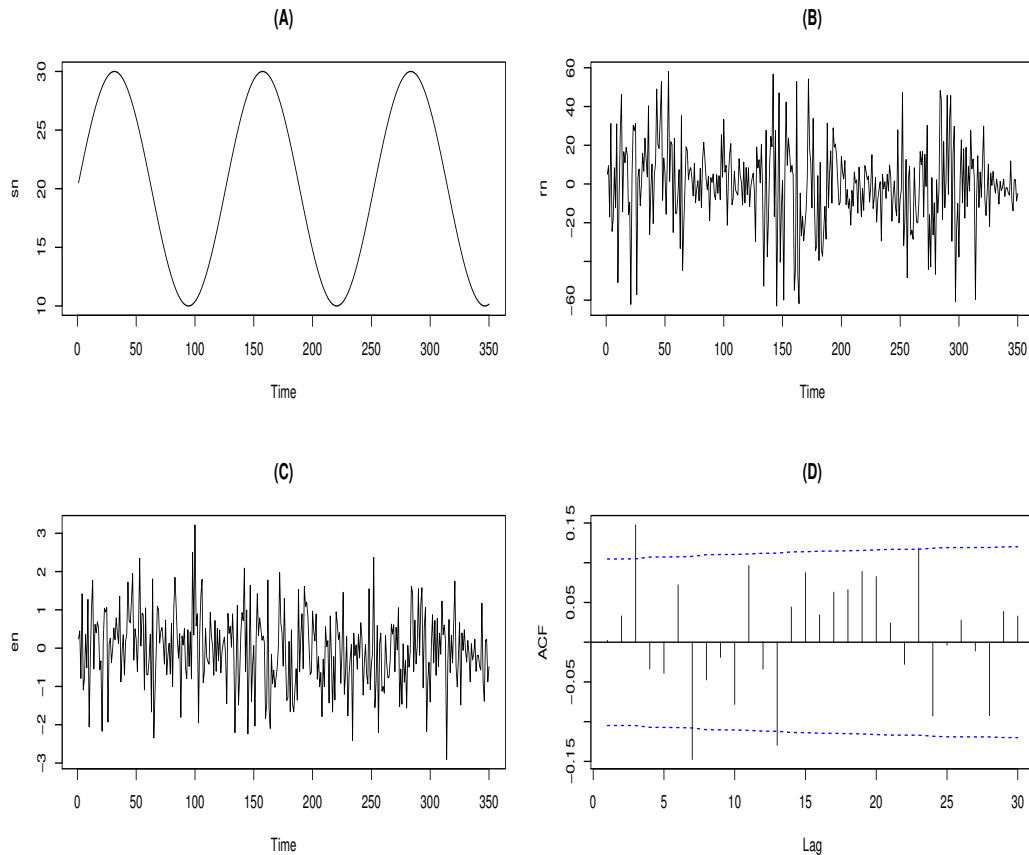


Figura 7.8: (A): σ_t , (B): Y_t , (C): Z_t , (D): fac de Y_t

```

https://es-us.finanzas.yahoo.com/q/hp?s=SIE.F+Precios+históricos
D = read.csv("siemens.SIE.F.csv", header=TRUE, stringsAsFactors=FALSE)
Xn = D$Close; Zn = diff(log(Xn),1,1);
require(TSA)
par(mfrow=c(2,2))
ts.plot(Xn)
TSA::acf(Xn,90,ci.type="ma",drop.lag.0 = TRUE)
ts.plot(Zn)
TSA::acf(Zn,30,ci.type="ma",drop.lag.0 = TRUE)
#-----pruebas Ljung-Box
Box.test(Zn, lag = 15 , type = "Ljung-Box")
X-squared = 30.983, df = 15, p-value = 0.008831

```

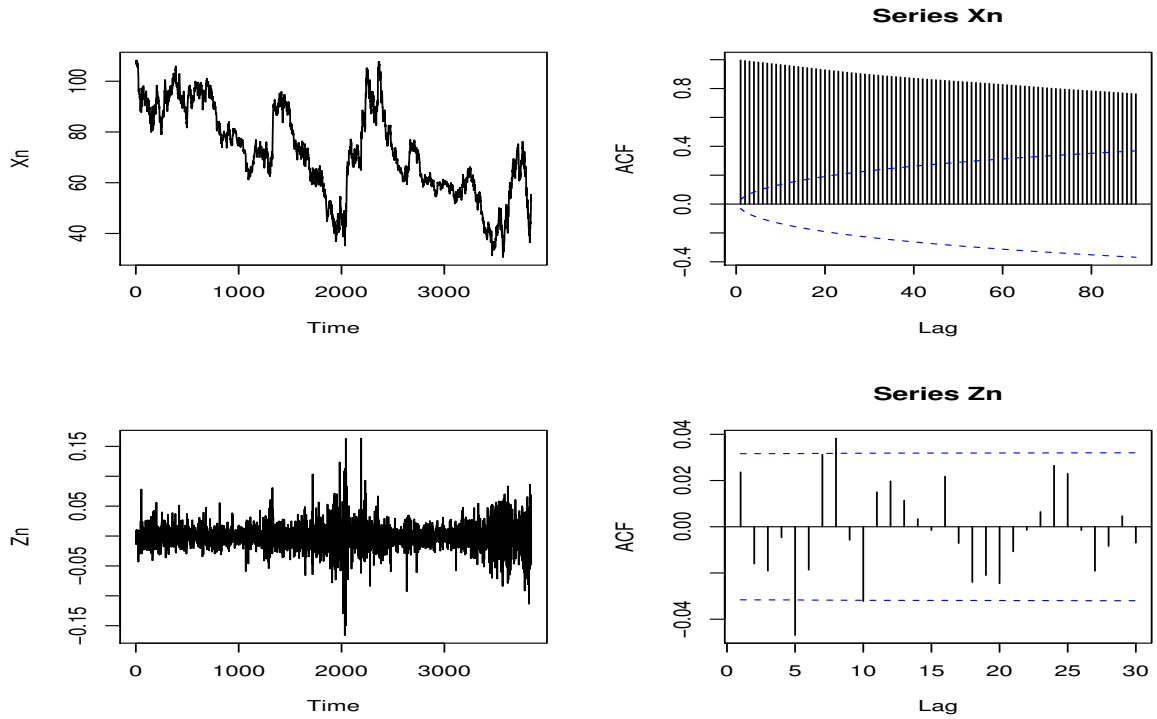



Figura 7.9: Compara fac de serie no estacionaria versus estacionaria.

7.4.2. Prueba Durbin-Watson (DW)

La prueba DW se basa en la definición del modelo Auto regresivo de orden uno, $AR(1)$.

Definición 7.4.1. Una serie $(Y_t, t \in \mathbb{Z})$ se dice que sigue un modelo $AR(1)$ con media μ si cumple

$$Y_t = \mu(1 - \phi) + \phi Y_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (7.10)$$

donde $|\phi| < 1$, $\mu \in \mathbb{R}$ y $\varepsilon_t \sim RB(0, \sigma^2)$. Se escribirá $Y_t \sim AR(1)$.

Propiedades. Si $Y_t \sim AR(1)$ entonces es un proceso estacionario en covarianza, tal que

- $E(Y_t) \equiv \mu$

- $Var(Y_t) \equiv \frac{\sigma^2}{1 - \phi}$
- $Cov(Y_t, Y_{t+k}) = R(k) = \frac{\sigma^2 \phi^k}{1 - \phi^2}, k = 0, 1, \dots$
- $\rho(k) = \phi^k, k = 0, 1, 2, \dots$

luego la fac teórica de un $AR(1)$ tiene las formas generales

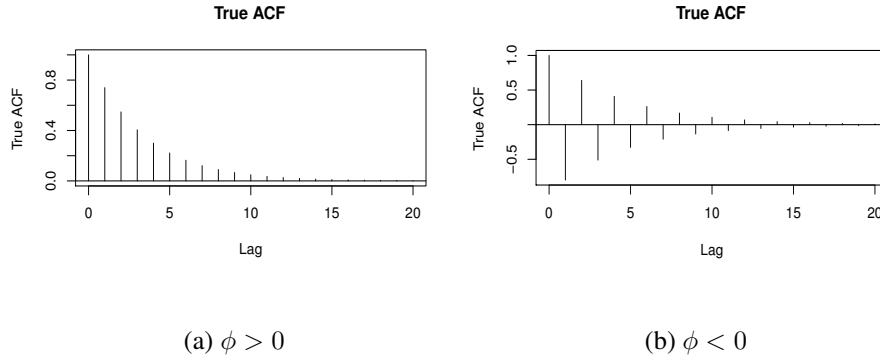


Figura 7.10: Fac teóricas de un $AR(1)$.

Ahora, si se tiene una serie $(Y_t, t = 1, \dots, T)$ y se quiere probar que es incorrelacionada, asumimos que $\mathbb{E}(Y_t) = 0$; en caso contrario, se transforma $Y_t^* = Y_t - \bar{Y}_t$.

La prueba DW se basa en asumir primero que $Y_t \sim AR(1)$, es decir, Y_t sigue el modelo (7.10), con $\mu = 0$.

1. Hipótesis de la Prueba

$$H_0 : \phi = 0$$

$$H_1 : \phi \neq 0.$$

Nótese que si H_0 es cierta, entonces $Y_t = \varepsilon_t$ y como $\varepsilon_t \sim RB(0, \sigma^2)$ y $Y_t \sim RB(0, \sigma^2)$.

2. Estadístico de Prueba

Definición 7.4.2. El estadístico DW se define como

$$d = \frac{\sum_{t=2}^T (Y_t - Y_{t-1})^2}{\sum_{t=1}^T Y_t^2}. \quad (7.11)$$

3. Distribución del estadístico bajo H_0 . La distribución de d bajo H_0 depende de la matriz de diseño X , para todos los modelos de regresión $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$. Por lo que debe ser calculada para cada X .

4. Decisión.

De (7.11), como $\sum_{t=2}^T (Y_t - Y_{t-1})^2 = \sum_{t=2}^T Y_t^2 + \sum_{t=2}^T Y_{t-1}^2 - 2 \sum_{t=2}^T Y_t Y_{t-1}$, se aproxima

$$\begin{aligned} d &= \frac{\sum_{t=2}^T Y_t^2 + \sum_{t=1}^{T-1} Y_t^2 - 2 \sum_{t=2}^T Y_t Y_{t-1}}{\sum_{t=1}^T Y_t^2} \\ &\approx \frac{\sum_{t=2}^T Y_t^2 + \sum_{t=2}^T Y_t^2 - 2 \sum_{t=2}^T Y_t Y_{t-1}}{\sum_{t=1}^T Y_t^2} \\ &= 2 - 2\hat{\rho}(1). \end{aligned}$$

es decir,

$$d \approx 2(1 - \hat{\rho}(1)). \quad (7.12)$$

A partir de (7.12) se puede concluir que, como $-1 < \hat{\rho}(1) < 1$, entonces $0 < d < 4$. Y si $\hat{\rho}(1) \approx 0$ entonces $d \approx 2$ y no rechaza la nula.

- $d = 2 \Leftrightarrow 1 - \hat{\rho}(1) = 1 \Leftrightarrow \hat{\rho}(1) = 0$
- $d < 2 \Leftrightarrow 1 - \hat{\rho}(1) < 1 \Leftrightarrow \hat{\rho}(1) > 0$
- $d > 2 \Leftrightarrow \hat{\rho}(1) < 0$

luego un $d < 2$ indica posible autocorrelación positiva en la serie y $d > 2$ posible autocorrelación negativa.

Colocando d_{obs} el estadístico observado y $\text{Valor } p = P(d \geq d_{\text{obs}} | H_0 \text{ cierto})$, si $\text{Valor } p < 0.05$, se rechaza H_0 . Si $\text{Valor } p > 0.05$, no se rechaza H_0 .

4. Prueba DW en R

- a) La función `dwtest` en la librería `lmtest`. Permite probar hipótesis alternas $H_1 : \phi > 0, \phi \neq 0, \phi < 0$, colocando `dwtest(yt ~ t + It)`. Calcula d y el Valor p .
- b) La función `durbin.watson` en la librería `car`. Utiliza bootstrap para calcular la distribución del estadístico d . Se programa `durbinWatsonTest(yt ~ t + It)`.

Ejemplo 7.4.3. Ejemplos de resultados con las dos funciones anteriores para residuos de un modelo $Y_t = \beta_0 + \beta_1 t + \varepsilon_t$.

```
require(car)
durbinWatsonTest(lm(yt ~ t), simulate=TRUE, reps=1000)
#-----resultado
lag Autocorrelation D-W Statistic p-value
  1          0.1353341          1.729167          0
Alternative hypothesis: rho != 0
#-----
require(lmtest)
dwtest(yt ~ t)
#-----resultado
data: yt ~ t
DW = 1.7292, p-value = 4.473e-05
alternative hypothesis: true autocorrelation
is greater than 0
#-----fac residuos estimados
require(TSA)
acf(et, 90, ci.type='ma', drop.lag.0=TRUE)
t = seq(1, length(et))
```

Ejemplo 7.4.4. Retomando el Ejemplo (7.3.2) en la pag. 141, para modelación de la serie de producción de cemento Portland, trimestral se ajustó un modelo de la forma

$$Y_t = \beta_0 + \beta_1 t + \sum_{j=1}^3 \delta_j I_t(t) + \varepsilon_t.$$

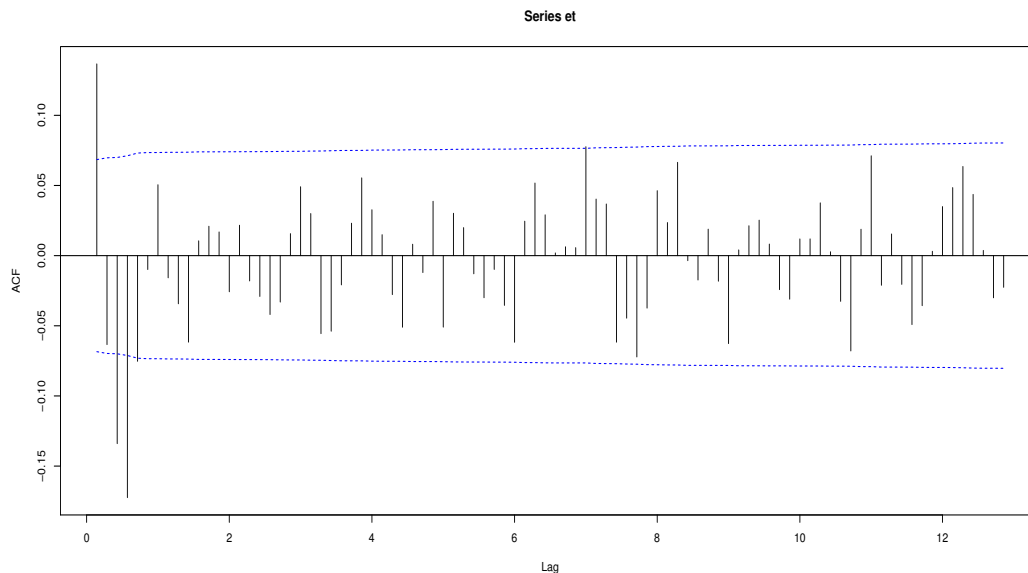


Figura 7.11: La fac de residuos $Y_t = \beta_0 + \beta_1 t + \varepsilon_t$

Al estimar los residuos estructurales $\hat{\varepsilon}_t$ se observó que posiblemente muestran una dinámica autocorrelacionada. Se puede confirmar este hecho con las pruebas Ljung-Box y Durbin-Watson (con las dos funciones de R mencionadas). Los resultados de R se muestran a continuación del código siguiente.

Es claro que los residuos estructurales presentan autocorrelación. Es decir, el supuesto de incorrelación no es válido.

Código R 7.4.2.

```
# ejemplo de pruebas LB, DW
# para la serie de produccion de cemento
library(car)
library(forecast)
library(lmtest)
#----
E = read.table(cementq.dat, header = TRUE)
attach(E)
y = ts(y,frequency=4,start=c(1956,1),end=c(1994,3))
#----modelo con tendencia lineal y estacionalidad
```

```

#----con variables indicadoras estacionales
t = seq(1,length(y))
It = seasonaldummy(y)
mod1 = lm(y ~ t + It)
summary(mod1)
et = residuals(mod1)
# pruebas Ljung-Box
Box.test(et, lag = 25, type = "Box-Pierce")
# resultado -----
Box-Ljung test
X-squared = 293.9028, df = 25, p-value < 2.2e-16

# pruebas Durbin-Watson
#-----con dwtest
dwtest(mod1) # de la libreria lmtest
Durbin-Watson test
DW = 0.3646, p-value < 2.2e-16
alternative hypothesis: true autocorrelation is greater than 0

durbinWatsonTest(mod1) # de la libreria car
#-----con durbin.watson
lag Autocorrelation D-W Statistic p-value
1 0.8043803 0.3645806 0
Alternative hypothesis: rho != 0
-----

```

7.4.3. Prueba Durbin-Watson Generalizada

Las pruebas Ljung-Box y Durbin-Watson coincidirían en su decisión siempre que la auto correlación de orden 1, $\hat{\rho}(1)$, sea dominante, es decir, sea la mayor de todas.

Pero la prueba DW puede no rechazar la hipótesis nula de incorrelación, siendo falsa, en el caso en que las primeras autocorrelaciones no sean significativas pero las de rezagos mayores sí.

Para incorporar esta posibilidad, Vinod [1973] introdujo la prueba Durbin-Watson Generalizada (DWG). Consiste en una serie de pruebas en secuencia, para $j = 1, 2, \dots$, asumiendo en cada j un modelo de la forma $Y_t = \phi_j Y_{t-j} + \eta_t$. La prueba para el paso j es

$$H_0 : \phi_j = 0$$

$$H_1 : \phi_j \neq 0.$$

dado que en las pruebas anteriores $j - 1, \dots, 1$, no se rechazó la hipótesis nula. El estadístico Durbin-Watson generalizado para el paso j se define como

$$d_j = \frac{\sum_{t=j+1}^T (Y_t - Y_{t-j})^2}{\sum_{t=1}^T Y_t^2}. \quad (7.13)$$

La prueba DWG se encuentra implementada en la librería `car`. Por ejemplo, colocando

```
durbinWatsonTest(mod,max.lag=10)
```

se obtienen los valores de los estadísticos d_j , para $j = 1, \dots, 10$, los correspondientes valores de $\hat{\rho}(j)$ y los valores p . Esta secuencia de pruebas permite determinar el orden de la menor auto correlación significativa.

Ejemplo 7.4.5. *Un caso de aplicación de la prueba GDW, muestra que no rechaza la nula de incorrelación en los rezagos $k = 1, 2, 3, 4$, pero a partir del rezago $k = 5$ rechaza la nula. Luego, se concluye que los residuos no son ruido blanco.*

Código R 7.4.3.

```
# ejemplo de prueba DWG
library(car)
library(forecast)
library(lmtest)
#---modelo con tendencia lineal y estacionalidad
#---con variables indicadoras estacionales
t = seq(1,length(y))
It = seasonaldummy(y)
```

```

mod1 = lm(y ~ t + It)
summary(mod1)
r = residuals(mod1)
# prueba Durbin-Watson Generalizada
# de la libreria car
durbinWatsonTest(mod1,max.lag=10)

```

Tabla 7.1: Resultados de la Prueba DW Generalizada

	Lag	rho.hat	dw	Valor.p
1	1.00	0.10	1.71	0.21
2	2.00	0.07	1.74	0.19
3	3.00	-0.01	1.89	0.93
4	4.00	-0.06	1.96	0.53
5	5.00	0.80	0.39	0.00
6	6.00	0.71	0.55	0.00
7	7.00	0.60	0.76	0.00
8	8.00	0.53	0.89	0.00
9	9.00	0.34	1.26	0.00
10	10.00	0.20	1.53	0.01

7.5. Alternativas cuando los Residuos Estructurales muestran Autocorrelación

En el caso de detectarse autocorrelación en los residuos estructurales el procedimiento consistirá en buscar un modelo tipo ARMA, autorregresivo de media móvil, que se tratará en el Capítulo 8. Pero es posible utilizar otras alternativas.

7.5.1. Incluir rezagos de la variable dependiente en el modelo

Una alternativa consiste en incluir en el modelo una ó varias componentes autorregresivas, es decir, aumentar el número de variables explicativas incluyendo rezagos

de la variables dependiente. Supongamos que el modelo inicial es un modelo con tendencia lineal y componente estacional dado por

$$Y_t = \beta_0 + \beta_1 t + \sum_{j=1}^{s-1} \delta_j I_t(t) + \varepsilon_t. \quad (7.14)$$

Y se incluyen en el modelo, por ejemplo, dos rezagos de la variable dependiente

$$Y_t = \beta_0 + \beta_1 t + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \sum_{j=1}^{s-1} \delta_j I_t(t) + \varepsilon_t. \quad (7.15)$$

Con este cambio se trata de “capturar” la autocorrelación en los residuos estructurales estimados $\hat{\varepsilon}_t$. En la librería `dynlm` se implementa este modelo mediante la función `dynlm`. La forma de aplicarla es con base en operadores de rezago. Por ejemplo, para el modelo (7.15) se coloca

```
require(dynlm)
mod2 = dynlm(y ~ t + It + L(y,1) + L(y,2))
summary(mod2)
```

Sin embargo, se sabe que cuando se agregan rezagos de la variable dependiente, el estadístico Durbin-Watson se sesga hacia 2 y no rechaza la nula de ruido blanco, rebaja la potencia.

Ejemplo 7.5.1. *Retomando el Ejemplo (7.4.4), para modelación de la serie de producción de cemento Portland, trimestral. Se puede preguntar si el modelo (7.15) logra capturar la estructura autoregresiva, detectada anteriormente. El programa siguiente muestra los pasos en la estimación de los modelos (7.14) y (7.15).*

Código R 7.5.1.

```
# ejemplo de pruebas LB, DW y DWG
# para la serie de produccion de cemento
# ademas ejemplo con rezagos con libreria dynlm
library(car)
library(forecast)
library(lmtest)
# modelo incorporando terminos rezagados de la serie
```

```
# utilizando la libreria dynlm
require(dynlm)
mod2 = dynlm(y ~ t + It + L(y,1)+ L(y,2))
summary(mod2)
```

Los resultados de la estimación con el modelo (7.15) se muestran en la Tabla 7.2 siguiente.

Tabla 7.2: Resultados del Modelo estructural con dos rezagos 1,2 de Y_t

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	100.1147	35.5231	2.82	0.0055
t	0.8695	0.4035	2.15	0.0329
ItQ1	-159.4054	15.5466	-10.25	0.0000
ItQ2	74.9556	21.9855	3.41	0.0009
ItQ3	65.1675	17.0202	3.83	0.0002
L(yi, 1)	0.6505	0.0828	7.86	0.0000
L(yi, 2)	0.2180	0.0839	2.60	0.0103

La pruebas LB rechaza la nula de ruido blanco, pero la DW no.

```
r2 = residuals(mod2)
# pruebas Ljung-Box
Box.test(r2, lag = 15, type = "Ljung-Box")
Box-Ljung test
data:  r2
X-squared = 50.508, df = 15, p-value = 9.944e-06

# pruebas Durbin-Watson

dwtest(mod2)
Durbin-Watson test
data:  mod2
DW = 1.9078, p-value = 0.2669
alternative hypothesis: true autocorrelation is greater than 0
```

```

durbinWatsonTest(mod2)
lag Autocorrelation D-W Statistic p-value
  1      0.01205702      1.907794    0.546
Alternative hypothesis: rho != 0
durbinWatsonTest(mod2,max.lag=10)

```

La prueba DWG con 10 rezagos no rechazan la nula para las auto correlaciones de órdenes 1,2,3. Pero la de orden 4 sí es significativa, y se rechaza la hipótesis de ruido blanco para este modelo. Ver la Tabla 7.3. La estrategia de incluir variables rezagadas no produce un resultado concluyente en este caso particular.

Tabla 7.3: Resultados Prueba DWG para el modelo con rezagos 1,2

	r	dw	p
1	0.0231	1.9484	0.6980
2	0.0793	1.8313	0.2720
3	-0.0762	2.1134	0.4100
4	0.3807	1.1955	0.0000
5	-0.0747	2.0854	0.4140
6	-0.2402	2.3958	0.0040
7	-0.1780	2.2050	0.0840
8	0.0815	1.6810	0.0920
9	-0.1215	2.0790	0.2360
10	-0.1959	2.2123	0.0480
11	-0.1247	2.0237	0.2900
12	0.1753	1.4218	0.0000
13	-0.0313	1.7453	0.5680
14	-0.0868	1.8542	0.7960
15	-0.0424	1.7558	0.6980

Como se anotó, la prueba DW se afecta por el termino $\phi_1 Y_{t-1}$, y el estadístico se sesga hacia 2, es decir, la prueba tiende a aceptar H_0 . La prueba tiene baja potencia para detectar la hipótesis alterna H_a .

7.5.2. Transformación Cochran-Orcutt

Una es la transformación Cochran-Orcutt, que permite re-estimar los parámetros cuando existe una autocorrelación de tipo AR(1) en los residuos.

Si la prueba DW detecta la alterna $\phi = \rho(1) \neq 0$, entonces se procede a transformar el modelo para la serie Y_t . Escribiendo el modelo estructural como un modelo de regresión lineal en (4.1), $Y_t = \underline{X}_t' \underline{\beta} + \varepsilon_t$, donde la prueba DW asume $\varepsilon_t = \phi \varepsilon_{t-1} + \eta_t$ y η_t es ruido blanco, la transformación Cochran-Orcutt se define como

$$Y_t - \phi Y_{t-1} = \alpha(1 - \hat{\rho}(1)) + \beta(\underline{X}_t - \hat{\rho}(1)\underline{X}_{t-1}) + \eta_t. \quad (7.16)$$

De esta manera, los residuos η_t son ruido blanco. Esta transformación depende de una autocorrelación en los residuos tipo AR(1). No sería de esperar que funcionara en casos de estructuras de autocorrelación más complejas, como será el caso de la mayoría de las series de tiempo a considerar.

Ejemplo 7.5.2. *Retomando el Ejemplo (7.4.4), para modelación de la serie de producción de cemento Portland, trimestral, al aplicar la transformación Cochran-Orcutt mediante la función `cochrane.orcutt` de la librería `orcutt`, se obtiene que se rechaza la nula de ruido blanco. La función `cochrane.orcutt`, desafortunadamente, no es compatible con las funciones para Durbin-Watson y DWG.*

```
# transformacion Cochran-Orcutt
require(orcutt)
mod2 = cochrane.orcutt(mod1)
r = mod2$Cochrane.Orcutt$residuals
Box.test(r, lag = 25, type = "Ljung-Box")
# pruebas Durbin-Watson
dwtest(mod2) # no funciona!
durbinWatsonTest(mod2) # no funciona!
```

7.5.3. Pruebas de Heterocedasticidad

Ejemplo 7.5.3. *Retomando el Ejemplo (7.3.2) en la pag. 141, para modelación de la serie de producción de cemento Portland, trimestral.*

Al estimar los residuos estructurales $\hat{\varepsilon}_t$ es evidente que muestran heterocedasticidad, como se aprecia a partir de la Figura 7.12.

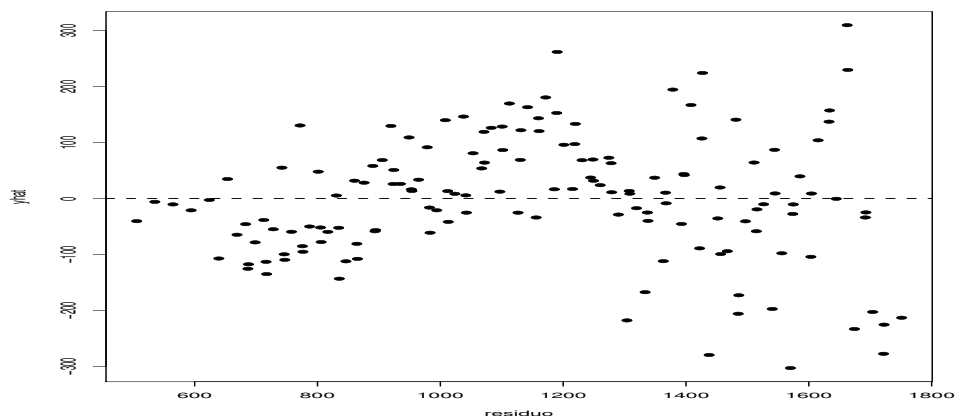


Figura 7.12: Y estimada versus residuales estructurales

Se pueden aplicar las pruebas de heterocedasticidad, Breusch-Pagan, Prueba de Scores para Varianza no constante y Goldfeld-Quandt. Los resultados se muestran a continuación.

studentized Breusch-Pagan test

BP = 19.9866, df = 4, p-value = 0.0005025

Non-constant Variance Score Test

Variance formula: ~ fitted.values

Chisquare = 23.03938 Df = 1 p = 1.587171e-06

Goldfeld-Quandt test

GQ = 6.8213, df1 = 69, df2 = 68, p-value = 7.005e-14

Sin embargo, la estrategia consiste en buscar un modelo tipo ARMA(p,q) para los residuos estructurales y no una transformación como la que se sugeriría a partir de los resultados de estas pruebas. Por lo que la verificación de la homocedasticidad de la varianza de los residuos estructurales no se incluirá en las pruebas de validación de los residuos.