

**LA FUNCIÓN DE CORRELACIÓN CRUZADA
EN SERIES NO ESTACIONARIAS:
IDENTIFICACIÓN, TENDENCIAS
DETERMINÍSTICAS Y RAÍCES UNITARIAS**

ELKIN CASTAÑO VÉLEZ

**DIRECTOR: JORGE MARTÍNEZ COLLANTES, PH.D.
PROFESOR UNIVERSIDAD NACIONAL DE COLOMBIA**

**TESIS PRESENTADA PARA OPTAR EL TÍTULO DE M.S. EN
ESTADÍSTICA EN LA ESCUELA DE ESTADÍSTICA DE LA
UNIVERSIDAD NACIONAL DE COLOMBIA.**

UNIVERSIDAD DE COLOMBIA-SEDE MEDELLÍN

**NOVIEMBRE
2005**

AGRADECIMIENTOS

La realización de esta tesis demandó gran cantidad de trabajo y dedicación y en ella participaron directa o indirectamente muchas personas. Debo agradecer a mis padres, que fueron mi guía, a mi esposa Claudia María y a mis hijos Lina María, Andrés Esteban, Mariantonia y a mi nieta María Camila por su paciencia durante el tiempo que duró la elaboración de este trabajo. Así mismo agradezco a mi amigo Sergio Yáñez, quien siempre estuvo muy interesado en su realización. Finalmente, debo agradecer toda la dedicación de mi amigo y director Jorge Martínez para que este trabajo llegara a feliz término.

RESUMEN

La función de correlación cruzada muestral (FCCM) ha sido empleada para estudiar la fortaleza y la dirección de la relación lineal entre dos procesos estocásticos conjuntamente estacionarios. Rosales (2004) y Castaño y Rosales (2005) muestran que dicha función, calculada entre el proceso estacionario y los residuales de un modelo preliminar estimado, puede ser empleada como un diagnóstico adicional en la identificación del verdadero modelo $ARMA(p,q)$ asociado al proceso. El propósito de este trabajo es emplear la información contenida en la FCCM para i) Extender estos resultados al caso de procesos no estacionarios homogéneos, ii) Formular un diagnóstico gráfico para discriminar entre un modelo de paseo aleatorio con deriva y uno con tendencia determinística, iii) Construir una prueba de la hipótesis nula de que el proceso es generado por un paseo aleatorio con deriva (el proceso posee una raíz unitaria), contra la hipótesis alternativa de que el proceso es generado por un modelo de tendencia determinística (el proceso no posee una raíz unitaria), suponiendo que el término de error del modelo no está autocorrelacionado y, iv) Obtener una generalización de la prueba anterior, suponiendo que el término de error está autocorrelacionado. Los resultados obtenidos permiten concluir que:

- La FCCM calculada entre el proceso estacionario y los residuales de un modelo preliminar estimado puede ser empleada como un diagnóstico adicional en la identificación del verdadero modelo $ARIMA(p,d,q)$ asociado al proceso. Como una consecuencia de este resultado, es posible construir un diagnóstico gráfico para discriminar adecuadamente entre un modelo de paseo aleatorio con deriva y uno con tendencia determinística lineal.
- Bajo el supuesto de normalidad y no autocorrelación en el término de error del modelo, tanto la prueba exacta como la asintótica, tienen una alta potencia contra la alternativa, aun en muestras pequeñas (mayor que 0.9 para $n=30$). Además, la prueba es consistente y asintóticamente insesgada.

- Bajo el supuesto de normalidad y de autocorrelación en el término de error del modelo, los ensayos realizados para modelos MA(1) y AR(1) muestran que la prueba generalizada asintótica mantiene, en general, una alta potencia excepto en los casos de una raíz próxima a la unidad en el polinomio AR. El tamaño de la prueba asintótica se mantiene cerca de su respectivo nivel de significancia nominal, excepto cuando i) existe una raíz próxima a la unidad en el polinomio MA o ii) la magnitud del parámetro ϕ es pequeña.

CONTENIDO

1. Introducción	1
2. Empleo de la FCC en la identificación de series estacionarias y no Estacionarias	3
2.1 Conceptos Preliminares	3
2.1 Caracterización de un Proceso ARMA(p,q) por Medio de la FCC	5
2.3 Aplicación de la Proposición 2 a la Identificación de un Modelo ARMA(p,q)	19
2.4 La FCC en la Identificación de Series de Tiempo no Estacionarias	19
3. Empleo de la FCC en la discriminación de un paseo aleatorio	24
3.1 Conceptos Preliminares	24
3.2 La FCC del Modelo de Paseo Aleatorio	26
3.3 La FCC del Modelo de Tendencia Determinística Lineal más Ruido	28
3.4 Discriminación entre un Proceso de Paseo Aleatorio con Deriva y un Proceso de Tendencia Lineal Determinística	31
4. Prueba de la hipótesis de que el proceso generador de datos es un paseo aleatorio	36
4.1 Introducción	36
4.2. Derivación de un Contraste de Paseo Aleatorio basado en la FCCM	40
4.2.1 La Distribución simulada de la estadística TA	41
4.2.2 Potencia del contraste contra diferentes alternativas	46
4.2.2.1 Potencia del contraste contra la alternativa de tendencia lineal determinística	46
4.2.2.2 Potencia del contraste contra la alternativa de un cambio estructural	48
4.2.3 La Distribución asintótica de la estadística TA	50

4.2.3.1	Derivación del tamaño del contraste en el caso asintótico .	51
4.3	Contraste de las Hipótesis de Tendencia Aleatoria contra	
	Tendencia Determinística	55
4.3.1	La Distribución simulada de la estadística <i>TMIN</i>	57
4.3.2	Potencia de la prueba <i>TMIN</i>	61
4.3.3	Distribución Asintótica de la estadística <i>TMIN</i>	62
4.3.3.1	Tamaño de la prueba asintótica	64
4.3.3.2.	Potencia de la prueba asintótica	65
4.4	Generalización de la Prueba	66
4.4.1	Tamaño de la prueba asintótica	71
4.4.1.1	Tamaño de la prueba para procesos MA(1)	71
4.4.1.1.1	Para procesos MA(1) con θ positivo	71
4.4.1.1.2	Para procesos MA(1) con θ negativo	74
4.4.1.2	Tamaño de la prueba para el modelos AR(1)	76
4.4.1.2.1	Para procesos AR(1) con ϕ positivo	76
4.4.1.2.2	Para procesos AR(1) con ϕ negativo	79
4.4.2	Potencia de la prueba asintótica	81
4.4.2.1	Potencia de la prueba para modelos MA(1)	81
4.4.2.1.1	Para procesos MA(1) con θ positivo	81
4.4.2.1.2	Para modelos MA(1) con θ negativo	83
4.4.2.2	Potencia de la prueba para modelos AR(1)	85
4.4.2.2.1	Para modelos AR(1) con ϕ positivo	85
4.4.2.2.2	Para modelos AR(1) con ϕ negativo	88
5.	Conclusiones	91
Bibliografía	97

LISTA DE GRÁFICAS

Gráfica 2.1 Modelo $Z_t = (1 + 0.7B^4)a_t$	14
Gráfica 2.2 Modelo $Z_t = (1 + 0.4B^4)b_t$	15
Gráfica 2.3 Modelo $Z_t = 1/(1 - 0.5B^4)b_t$	15
Gráfica 2.4 Modelo $Z_t = 1/(1 - 0.4B)b_t$	16
Gráfica 2.5. Modelo $Z_t = (1 - 0.7B - 0.1B^2)/(1 - 0.4B + 0.2B^2)b_t$	17
Gráfica 2.6 Modelo $(1 - B)(1 - B^{12})Z_t = 1/(1 - 0.6B)(1 + 0.5B^{12})b_t$	17
Gráfica 2.7 Modelo $(1 - B)Z_t = -0.51 + b_t$	18
Gráfica 2.8 Modelo $Z_t = (1 - 0.3B^2)b_t$	18
Gráfica 2.9 Logaritmo natural del total de pasajeros	20
Gráfica 2.10 FCCM del modelo preliminar	21
Gráfica 2.11 FCCM del modelo final	22
Gráfica 2.12 FCCM para la serie diferenciada sólo con $(1 - B^{12})$	23
Gráfica 3.1 La FCC entre a_t y ∇Z_t de un paseo aleatorio (con y sin deriva)	28
Gráfica 3.2 La FCC entre a_t y ∇Z_t de un proceso de tendencia determinística	30
Gráfica 3.3 Realización de un proceso de paseo aleatorio con deriva	31
Gráfica 3.4 Realización de un proceso de tendencia lineal determinística	31
Gráfica 3.5 FCCM muestral para el modelo $\nabla Z_t = \theta_0 + a_t$	32
Gráfica 3.6 FCCM muestral para el modelo $\nabla Z_t = \beta_1 + a_t - a_{t-1}$	33
Gráfica 3.7 FCC bajo $\nabla Z_t = \beta_1 + a_t - a_{t-1}$	35
Gráfica 4.1 Distribución de TA , $n=30$	43
Gráfica 4.2 Distribución de TA , $n=50$	43
Gráfica 4.3 Distribución de TA , $n=100$	44
Gráfica 4.4 Distribución de TA , $n=5000$	44

Gráfica 4.5 Diagnóstico gráfico para el ejemplo 4	46
Gráfica 4.6 FCCM bajo H_0 entre \hat{a}_t y ∇Z_t , cuando ∇Z_t posee autocorrelación . .	67
Gráfica 4.7 Una realización del proceso ARIMA(1,1,1) del ejemplo 4.6	69
Gráfica 4.8 Una realización del proceso ARIMA(1,1,1) del ejemplo 4.7	70
Gráfica 4.9 Modelo bajo H_0 para diferentes valores de $\theta > 0$ y $n=50$	73
Gráfica 4.10 Tamaño de la prueba para $\alpha = 0.05$ y diferentes valores de n y $\theta > 0$.	74
Gráfica 4.11 Tamaño de la prueba para $\alpha = 0.05$ y diferentes valores de n y $\theta > 0$.	76
Gráfica 4.12 Modelo bajo H_0 para diferentes valores de ϕ y $n=50$	78
Gráfica 4.13 Tamaño de la prueba para $\alpha = 0.05$ y diferentes valores de $\phi > 0$	79
Gráfica 4.14 Tamaño de la prueba para $\alpha = 0.05$ y diferentes valores de $\phi < 0$	81
Gráfica 4.15 Potencia de la prueba para $\alpha = 0.05$ y diferentes valores de $\theta > 0$	83
Gráfica 4.16 Potencia de la prueba para $\alpha = 0.05$ y diferentes valores de $\theta > 0$	85
Gráfica 4.17 Modelo bajo H_1 para diferentes valores de ϕ y $n=60$	87
Gráfica 4.18 Potencia de la prueba para $\alpha = 0.05$ y diferentes valores de $\phi > 0$	88
Gráfica 4.19 Potencia de la prueba para $\alpha = 0.05$ y diferentes valores de $\phi > 0$	90
Gráfica 5.1 Tamaño de la prueba para modelos MA(1)	94
Gráfica 5.2 Tamaño de la prueba para modelos AR(1)	95
Gráfica 5.3 Potencia de la prueba para modelos MA(1)	95
Gráfica 5.4 Potencia de la prueba para modelos AR(1)	96

LISTA DE TABLAS

Tabla 2.1 Estimación del modelo preliminar	20
Tabla 4.1 Percentiles superiores para la distribución de la estadística TA bajo H_0 . . .	42
Tabla 4.2 Percentiles superiores suavizados para la distribución de la estadística TA bajo H_0	45
Tabla 4.3 Potencia del contraste contra la alternativa de una tendencia lineal Determinística	47
Tabla 4.4 Potencia del contraste contra la alternativa de un cambio estructural	49
Tabla 4.5 Tamaño del contraste bajo la distribución asintótica de TA	52
Tabla 4.6 Tamaño del contraste bajo la distribución asintótica de TA	54
Tabla 4.7 Percentiles superiores para la distribución de la estadística $TMIN$ bajo H_0 .	58
Tabla 4.8 Percentiles superiores suavizados para la distribución de la estadística $TMIN$ bajo H_0	58
Tabla 4.9 Tabla para la $P[TMIN = TA]$ bajo H_0	59
Tabla 4.10 Verdaderos niveles de significancia de la prueba $TMIN$	60
Tabla 4.11 Potencia de la prueba $TMIN$	61
Tabla 4.12 Tamaño de la prueba bajo la distribución asintótica de $TMIN$	64
Tabla 4.13 Potencia de la prueba asintótica	65
Tabla 4.14 Efectos de Errores correlacionados sobre el tamaño de la prueba	66
Tabla 4.15 Tamaño de la prueba para procesos MA(1) con $\theta > 0$	72
Tabla 4.16 Tamaño de la prueba para procesos MA(1) con $\theta < 0$	75
Tabla 4.17 Tamaño de la prueba para procesos AR(1) con $\phi > 0$	77
Tabla 4.18 Tamaño de la prueba para procesos AR(1) con $\phi < 0$	80
Tabla 4.19 Potencia de la prueba para procesos MA(1) con $\theta > 0$	82
Tabla 4.20 Potencia de la prueba para procesos MA(1) con $\theta < 0$	84
Tabla 4.21 Potencia de la prueba para procesos AR(1) con $\phi > 0$	86

Tabla 4.22 Potencia de la prueba para procesos AR(1) con $\phi < 0$ 89

CAPÍTULO 1

INTRODUCCIÓN

En la literatura de series de tiempo, para la identificación de un modelo autorregresivo y de medias móviles estacionario e invertible, $ARMA(p,q)$, generalmente se hace uso de las funciones de autocorrelación y autocorrelación parcial muestrales (denotadas por sus siglas en inglés como SACF y SPACF, respectivamente) para tratar de obtener una especificación adecuada de los órdenes p y q del proceso (Box y Jenkins, 1976). Para procesos puros $AR(p)$ y $MA(q)$ estas funciones generalmente suelen proporcionar suficiente información para la identificación correcta. Sin embargo, cuando el proceso es mixto, es decir, posee componentes autorregresivas y de medias móviles, la información proporcionada por la SACF y la SPACF, generalmente es insuficiente para lograr la identificación correcta de p y q . Debido a esto, se sugiere usar otras herramientas adicionales, tales como la función de autocorrelación extendida muestral (ESACF) propuesta por Tsay y Tiao (1984), el método de la más pequeña correlación canónica (SCAN) propuesta también por Tsay y Tiao (1985), el método de “la esquina” propuesto por Beguin, Gouriéroux y Monfort (1980) y los criterios de información tales como el AIC propuesto por Akaike (1974), el SC propuesto por Schwartz (1978), entre otros.

Ahora bien, la función de correlación cruzada muestral (FCCM) ha sido empleada para estudiar la fortaleza y la dirección de la relación entre dos procesos estocásticos conjuntamente estacionarios. Rosales (2004) y Castaño y Rosales (2005) muestran que dicha función, calculada entre el proceso estacionario y los residuales de un modelo preliminar estimado, puede ser empleada como un diagnóstico adicional en la identificación del verdadero modelo $ARMA(p,q)$ asociado al proceso.

El objetivo de este trabajo es extender estos resultados a modelos ARIMA estacionales y no estacionales y estudiar las características de la FCCM entre un proceso de paseo

aleatorio con sus errores y la FCCM de un proceso de tendencia determinística con sus errores, para elaborar un diagnóstico gráfico que ayude a discriminar entre los dos procesos, y construir pruebas de hipótesis que permitan detectar la presencia de una raíz unitaria en un proceso con tendencia lineal.

El plan de este documento es el siguiente: En el segundo capítulo se revisan los conceptos preliminares empleados en el trabajo, se presenta la caracterización de un proceso ARMA(p,q) por medio de la FCC, y se prueba una proposición que permite hacer operativos los resultados obtenidos en la caracterización anterior; finalmente los resultados son extendidos al caso no estacionario.

En el tercer capítulo se caracterizan las FCC para los procesos estacionarizados de paseo aleatorio con deriva y de tendencia lineal determinística y se propone un diagnóstico gráfico que ayude a discriminarlos.

La primera parte del capítulo 4 presenta una revisión de algunos de los procedimientos empleados para la prueba de la hipótesis nula $H_0 : Z_t = \beta_0 + Z_{t-1} + a_t$ contra la hipótesis alternativa $H_1 : Z_t = \beta_0 + \beta_1 t + a_t$. A continuación, basados en la información contenida en la FCC del proceso bajo H_0 , se construye i) una prueba para contrastar la hipótesis H_0 contra la alternativa no informativa $H_1 : Z_t$ no es un paseo aleatorio con deriva, y se obtiene la distribución exacta y asintótica del estadístico de la prueba bajo H_0 , para investigar las propiedades de la prueba; ii) una prueba para contrastar las hipótesis $H_0 : Z_t = \beta_0 + Z_{t-1} + a_t$ contra $H_1 : Z_t = \beta_0 + \beta_1 t + a_t$, cuando a_t es ruido blanco normal, se obtiene distribución exacta y asintótica del estadístico de la prueba bajo H_0 para investigar las propiedades del contraste y, iii) una prueba generalizada asintótica para el caso donde el término de error del modelo es un proceso ARMA(p,q) estacionario e invertible.

Finalmente, en el capítulo 5, se presentan las conclusiones.

CAPÍTULO 2

EMPLEO DE LA FCC EN LA IDENTIFICACIÓN DE SERIES ESTACIONARIAS Y NO ESTACIONARIAS

En este capítulo se mostrará que todo modelo ARMA(p,q), estacionario e invertible, puede ser caracterizado en términos de la FCC entre el proceso y su término de error aleatorio. Adicionalmente se presenta un procedimiento que permite hacer operativos los resultados obtenidos.

2.1 Conceptos Preliminares: La Función de Correlación Cruzada.

Considere dos procesos conjuntamente estacionarios x_t y y_t , para $t = 0, \pm 1, \pm 2, \dots$. La covarianza cruzada de orden k entre x_t y y_t está definida como (Ver por ejemplo, Box, Jenkins, 1976; Wei, 1990):

$$\gamma_{xy}(k) = E[(x_t - \mu_x)(y_{t+k} - \mu_y)]$$

para $k=0, \pm 1, \pm 2, \dots$. Como función de k , $\gamma_{xy}(k)$ es llamada la función de covarianza cruzada entre x_t y y_t .

La estandarización de $\gamma_{xy}(k)$ produce la función de correlación cruzada (FCC)

$$\rho_{xy}(k) = \gamma_{xy}(k) / (\sigma_x \sigma_y)$$

para $k=0, \pm 1, \pm 2, \dots$, donde σ_x y σ_y son las desviaciones estándar de los procesos x_t y y_t .

La FCC mide no solamente la fortaleza de la relación sino también su dirección. Esta última propiedad es útil para identificar variables causales. Por esta razón es importante

examinar la FCC tanto para los valores positivos de k , como para los negativos. Para valores negativos de k , la FCC describe la influencia lineal de los valores pasados de y_t sobre x_t . Para valores positivos de k , la FCC indica la influencia lineal de los valores pasados de x_t sobre y_t . El gráfico de la FCC contra k , llamado correlograma cruzado, es útil para visualizar estas relaciones.

Dada una realización de n períodos para el proceso estacionario bivalente x_t, y_t , la FCC es estimada con la función de correlación cruzada muestral (FCCM):

$$\hat{\rho}_{xy}(k) = \frac{\hat{\gamma}_{xy}(k)}{\hat{\sigma}_x \hat{\sigma}_y}$$

para $k=0, \pm 1, \pm 2, \dots$, donde

$$\begin{aligned} \hat{\gamma}_{xy}(k) &= \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(y_{t+k} - \bar{y}), \quad \text{si } k \geq 0 \\ &= \frac{1}{n} \sum_{t=1-k}^n (x_t - \bar{x})(y_{t+k} - \bar{y}), \quad \text{si } k < 0 \end{aligned}$$

y donde $\hat{\sigma}_x = [\hat{\gamma}_{xx}(0)]^{1/2}$, $\hat{\sigma}_y = [\hat{\gamma}_{yy}(0)]^{1/2}$, \bar{x} y \bar{y} son las desviaciones estándar y las medias muestrales de las series x_t y y_t , respectivamente.

Bajo los supuestos de normalidad, que la serie x_t es ruido blanco y que las series x_t y y_t son incorrelacionadas, Bartlett (1955) probó que

$$\text{Var}[\hat{\rho}_{xy}(k)] \approx (n-k)^{-1}$$

Por tanto, cuando la serie x_t es ruido blanco y hay normalidad, podemos contrastar la hipótesis que las dos series tienen correlación cruzada nula comparando $\hat{\rho}_{xy}(k)$ con su error estándar aproximado $\sqrt{(n-k)^{-1}}$.

2.2 Caracterización de un proceso ARMA(p,q) por medio de la FCC

Proposición 1. Suponga que el proceso $\{Z_t\}$ sigue un modelo ARMA(p,q) estacionario e invertible bajo las condiciones de regularidad, y con proceso de ruido blanco $\{a_t\}$ de media cero y varianza σ_a^2 . Sea $\rho_{aZ}(k)$ la función de correlación cruzada entre el ruido blanco a_t y Z_t . Entonces $\rho_{aZ}(k)$ es tal que:

- i) Si $k \geq 0$, en general $\rho_{aZ}(k) \neq 0$
- ii) si $k < 0$, $\rho_{aZ}(k) = 0$
- iii) La forma de $\rho_{aZ}(k)$, para $k \geq 0$, permite caracterizar el modelo ARMA(p,q).

Demostración:

- i) Por definición

$$\rho_{aZ}(k) = \frac{\gamma_{aZ}(k)}{[\text{Var}(a_t)\text{Var}(Z_t)]^{1/2}}$$

Como Z_t es estacionario, este proceso puede escribir como

$$Z_t = \psi(B)a_t$$

siendo $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \psi_3 B^3 + \dots$, donde $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. Entonces,

$$\gamma_{aZ}(k) = E[a_t Z_{t+k}] = E[a_t \psi(B)a_{t+k}] = \sum_{j=0}^{\infty} \psi_j E[a_t a_{t+k-j}] = \psi_k \sigma_a^2 \neq 0, \text{ para al menos}$$

un $k \geq 0$, y por tanto, $\rho_{aZ}(k) \neq 0$ para al menos un $k \geq 0$

ii) Si $k < 0$, entonces

$$\gamma_{aZ}(k) = E[a_t Z_{t+k}] = E[a_t \psi(B) a_{t+k}] = \sum_{j=0}^{\infty} \psi_j E[a_t a_{t+k-j}] = 0$$

ya que no existirán subíndices comunes en $E[a_t a_{t+k-j}]$. Por tanto, $\rho_{aZ}(k) = 0$ si $k < 0$

iii) La derivación de los resultados teóricos se encuentran en (Rosales, 2004), los cuales se pueden resumir de la siguiente manera:

- El proceso es un MA(q): Si el lado derecho se anula a partir de $k=q+1$.
- El proceso es un AR(p): Dependiendo de las raíces de $\phi(B)=0$, el lado derecho es una combinación de decrecimientos exponenciales y/o mezcla de ondas sinusoidales amortiguadas
- El proceso es un ARMA(p,q): Dependiendo de las raíces de $\phi(B)=0$ y de $\theta(B)=0$, el lado derecho es una combinación de decrecimientos exponenciales y/o mezcla de ondas sinusoidales amortiguadas, a partir de $k=q$. Para valores de k menores o iguales a q no hay un patrón definido.
- En todos los casos anteriores $\rho_{aZ}(k) = 0$ si $k < 0$.

Es importante observar que a pesar de que la proposición anterior permite caracterizar el comportamiento de un modelo ARMA(p,q) en términos de su FCC, en la práctica su empleo requiere del conocimiento de una realización del verdadero ruido blanco del modelo, el cual en general, no es observable.

La siguiente proposición justifica el uso de la FCC entre el proceso y los residuales de un modelo preliminar ajustado a los datos, como una herramienta de diagnóstico adicional en la identificación del verdadero modelo ARMA(p,q) que generó los datos.

Proposición 2. Suponga que el proceso $\{Z_t\}$ sigue un modelo ARMA(p_1, q_1) estacionario e invertible bajo las condiciones de regularidad, con proceso de ruido blanco $\{a_t\}$ de media cero y varianza σ_a^2 . Sea b_t el proceso resultante de filtrar a Z_t usando un modelo ARMA(p_2, q_2) y sea que $\rho_{bZ}(k)$ la función de correlación cruzada entre el proceso filtrado, b_t , y Z_t . Si $\rho_{aZ}(k)$ es la función de correlación cruzada entre el ruido blanco a_t y Z_t , entonces la función de correlación cruzada $\rho_{bZ}(k)$, es tal que:

i) Si $k \geq 0$, $\rho_{bZ}(k)$ contiene el patrón de comportamiento de $\rho_{aZ}(k)$. Es decir, $\rho_{bZ}(k)$ “conserva la memoria” del verdadero proceso ARMA(p_1, q_1).

ii) Si $k < 0$, $\rho_{bZ}(k)$ es, en general, una función no nula.

Demostración:

i) Si el verdadero proceso de Z_t es un ARMA(p_1, q_1), entonces (sin perder generalidad haciendo la constante igual a cero)

$$Z_t = \frac{\theta_1(B)}{\phi_1(B)} a_t$$

donde:

$$\phi_1(B) = 1 - \phi_{11}B - \phi_{12}B^2 - \dots - \phi_{1p_1}B^{p_1} \text{ y } \theta_1(B) = 1 - \theta_{11}B - \theta_{12}B^2 - \dots - \theta_{1q_1}B^{q_1}$$

y donde a_t es un proceso de ruido blanco de media cero y varianza constante σ_a^2 .

Como Z_t es estacionario, lo podemos escribir en términos del proceso ruido blanco como

$$Z_t = \psi_1(B)a_t \tag{1}$$

donde

$$\psi_1(B) = \frac{\theta_1(B)}{\phi_1(B)} = 1 + \psi_{11}B + \psi_{12}B^2 + \psi_{13}B^3 + \dots = \sum_{j=0}^{\infty} \psi_{1j}B^j$$

con $\sum_{j=0}^{\infty} |\psi_{1j}| < \infty$, y $\psi_{10} = 1$.

Ahora bien, suponga que Z_t es filtrado usando un modelo ARMA(p_2, q_2), entonces

$$\frac{\phi_2(B)}{\theta_2(B)} Z_t = b_t$$

o,

$$Z_t = \frac{\theta_2(B)}{\phi_2(B)} b_t$$

donde:

$$\phi_2(B) = 1 - \phi_{21}B - \phi_{22}B^2 - \dots - \phi_{2p_2}B^{p_2} \text{ y}$$

$$\theta_2(B) = 1 - \theta_{21}B - \theta_{22}B^2 - \dots - \theta_{2q_2}B^{q_2}$$

Como Z_t es estacionario, también lo podemos escribir en términos del proceso de medias móviles como

$$Z_t = \psi_2(B) b_t \tag{2}$$

donde

$$\psi_2(B) = \frac{\theta_2(B)}{\phi_2(B)} = 1 + \psi_{21}B + \psi_{22}B^2 + \psi_{23}B^3 + \dots = \sum_{j=0}^{\infty} \psi_{2j}B^j$$

con $\sum_{j=0}^{\infty} |\psi_{2j}| < \infty$ y $\psi_{20} = 1$.

Observe que b_t no es necesariamente un proceso de ruido blanco, ya que de (1) y (2),

$$b_t = \frac{\psi_1(B)}{\psi_2(B)} a_t \quad (3)$$

donde, en general, $\frac{\psi_1(B)}{\psi_2(B)} \neq 1$.

Consideremos ahora la FCC entre b_t y Z_t ,

$$\rho_{bZ}(k) = \frac{\gamma_{bZ}(k)}{[\text{Var}(b_t)\text{Var}(Z_{t+k})]^{1/2}} \quad (4)$$

para $k=0, \pm 1, \pm 2, \dots$, donde

$$\gamma_{bZ}(k) = E(b_t Z_{t+k})$$

para $k=0, \pm 1, \pm 2, \dots$, es la función de covarianza cruzada entre b_t y Z_t .

De (3) se obtiene que,

$$b_t = \frac{\psi_1(B)}{\psi_2(B)} a_t = \psi_3(B) a_t = (1 + \psi_{31}B + \psi_{32}B^2 + \psi_{33}B^3 + \dots) a_t = \left(\sum_{j=0}^{\infty} \psi_{3j} B^j \right) a_t$$

donde, debido a la estacionaridad de Z_t , $\sum_{j=0}^{\infty} |\psi_{3j}| < \infty$, con $\psi_{30} = 1$.

Entonces, puesto que $\sum_{j=0}^{\infty} |\psi_{3j}| < \infty$,

$$\begin{aligned} \gamma_{bZ}(k) &= E[b_t Z_{t+k}] = E[Z_{t+k} \psi_3(B) a_t] = E[Z_{t+k} (1 + \sum_{j=1}^{\infty} \psi_{3j} B^j) a_t] \\ &= E[a_t Z_{t+k}] + \sum_{j=1}^{\infty} \psi_{3j} E[a_{t-j} Z_{t+k}] \\ &= \gamma_{aZ}(k) + \sum_{j=1}^{\infty} \psi_{3j} \gamma_{aZ}(k+j) \end{aligned} \quad (5)$$

Además,

$$\text{Var}(b_t) = \text{Var}(\psi_3(B)a_t) = \sigma_a^2 \sum_{j=0}^{\infty} \psi_{3j}^2 \quad (6)$$

Reemplazando (5) y (6) en (4), obtenemos,

$$\begin{aligned} \rho_{bz}(k) &= \frac{\gamma_{bz}(k)}{[\text{Var}(b_t)\text{Var}(Z_{t+k})]^{1/2}} = \frac{\gamma_{az}(k) + \sum_{j=1}^{\infty} \psi_{3j} \gamma_{az}(k+j)}{[\sigma_a^2 \sum_{j=0}^{\infty} \psi_{3j}^2 \text{Var}(Z_{t+k})]^{1/2}} \\ &= \frac{\gamma_{az}(k) + \sum_{j=1}^{\infty} \psi_{3j} \gamma_{az}(k+j)}{[\sum_{j=0}^{\infty} \psi_{3j}^2]^{1/2} [\text{Var}(a_t)\text{Var}(Z_t)]^{1/2}} = \frac{\gamma_{az}(k)}{[\sum_{j=0}^{\infty} \psi_{3j}^2]^{1/2} [\text{Var}(a_t)\text{Var}(Z_t)]^{1/2}} \\ &\quad + \frac{\sum_{j=1}^{\infty} \psi_{3j} \gamma_{az}(k+j)}{[\sum_{j=0}^{\infty} \psi_{3j}^2]^{1/2} [\text{Var}(a_t)\text{Var}(Z_t)]^{1/2}} \\ &= c \frac{\gamma_{az}(k)}{[\text{Var}(a_t)\text{Var}(Z_t)]^{1/2}} + c \sum_{j=1}^{\infty} \psi_{3j} \frac{\gamma_{az}(k+j)}{[\text{Var}(a_t)\text{Var}(Z_t)]^{1/2}} \end{aligned}$$

Por tanto,

$$\rho_{bz}(k) = c\rho_{az}(k) + c \sum_{j=1}^{\infty} \psi_{3j} \rho_{az}(k+j) \quad (7)$$

donde $c = \frac{1}{[\sum_{j=0}^{\infty} \psi_{3j}^2]^{1/2}}$ y $0 < c \leq 1$.

La ecuación (7) muestra la relación general entre $\rho_{bz}(k)$ y $\rho_{az}(k)$. La función $\rho_{bz}(k)$ contiene el patrón de comportamiento de $\rho_{az}(k)$ a través de $c\rho_{az}(k)$. Sin embargo, el comportamiento de $\rho_{az}(k)$ puede ser distorsionado por el término

$c \sum_{j=1}^{\infty} \psi_{3j} \rho_{aZ}(k+j)$, donde los coeficientes ψ_{3j} se obtienen igualando los coeficientes de las mismas potencias de B a ambos lados de la ecuación,

$$\psi_2(B)\psi_3(B) = \psi_1(B)$$

En general, la solución para ψ_{3j} es,

$$\begin{aligned} \psi_{3j} &= -\psi_{21}\psi_{3,j-1} - \psi_{22}\psi_{3,j-2} - \dots - \psi_{2,j-1}\psi_{31} - \psi_{2j} + \psi_{1,j} \\ \psi_{3j} &= (\psi_{1,j} - \psi_{2,j}) - \sum_{i=1}^{j-1} \psi_{2,i}\psi_{3,j-i} \end{aligned} \quad (8)$$

para $j=1,2,3,\dots$

Esta solución indica que,

a) Si $\psi_{1j} = \psi_{2j}$, para $j \geq 0$, es decir el modelo por el cual se filtra a Z_t es el verdadero modelo generador de los datos de Z_t , entonces, usando (8), $\psi_{3j} = 0$ para $j > 0$, y entonces $c=1$. Por tanto, como habría de esperarse,

$$\rho_{bZ}(k) = \rho_{aZ}(k)$$

para todo $k=0 \pm 1, \pm 2, \dots$

b) Si $\psi_{1j} \approx \psi_{2j}$, para $j > 0$, es decir el modelo por el cual se filtra a Z_t es un modelo aproximado al modelo generador de datos, entonces $\psi_{3j} \approx 0$, para $j > 0$, y entonces,

$$\rho_{bZ}(k) \approx \rho_{aZ}(k)$$

para $k \geq 0$. Es decir, $\rho_{bZ}(k)$ presenta aproximadamente el patrón de comportamiento de $\rho_{aZ}(k)$. Por ejemplo, suponga que el proceso Z_t es generado por $Z_t = a_t - \theta_1 a_{t-1}$ y

que el modelo para filtrar a Z_t es $Z_t = b_t - \delta_1 b_{t-1} - \delta_2 b_{t-2}$. Empleando la ecuación (7), y la caracterización de un MA(1) por medio de la FCC, si $k \geq 1$

$$\rho_{bZ}(k) = c\rho_{aZ}(k)$$

ya que $\rho_{aZ}(k+j) = 0$ para $j=1,2,3 \dots$. En este caso, $\rho_{bZ}(k)$ es solamente una contracción de $\rho_{aZ}(k)$ y $\rho_{bZ}(k)$ tiene la misma apariencia que $\rho_{aZ}(k)$.

Si $k = 0$,

$$\rho_{bZ}(0) = c\rho_{aZ}(0) + c\psi_{31}\rho_{aZ}(1)$$

donde, de la ecuación 8, $\psi_{31} = \psi_{11} - \psi_{21} = \theta_1 - \delta_1$. Por tanto, mientras más cerca estén θ_1 y δ_1 menor distorsión habrá. Resumiendo, si $k \geq 0$, observamos que $\rho_{bZ}(k) \approx \rho_{aZ}(k)$.

c) Si $\psi_{1j} \neq \psi_{2j}$ y $k=0,1,2,3,\dots$, entonces $\rho_{bZ}(k)$ aun presenta aproximadamente el patrón de comportamiento de $\rho_{aZ}(k)$ aunque el término $c \sum_{j=1}^{\infty} \psi_{3j} \rho_{aZ}(k+j)$ puede llegar a ser grande y en algunos casos puede distorsionar el comportamiento de $\rho_{aZ}(k)$.

Es posible establecer una cota para $c \sum_{j=1}^{\infty} \psi_{3j} \rho_{aZ}(k+j)$:

De la ecuación (7)

$$c \sum_{j=1}^{\infty} \psi_{3j} \rho_{aZ}(k+j) = \rho_{bZ}(k) - c\rho_{aZ}(k)$$

$$\left| c \sum_{j=1}^{\infty} \psi_{3j} \rho_{aZ}(k+j) \right| = \left| \rho_{bZ}(k) - c\rho_{aZ}(k) \right| \leq \left| \rho_{bZ}(k) \right| + \left| c\rho_{aZ}(k) \right| \leq 1 + c$$

Por tanto, $c \sum_{j=1}^{\infty} \psi_{3j} \rho_{aZ}(k+j)$ está acotado inferiormente por $-(1+c)$ y superiormente por $1+c$.

Nota: Sin embargo, cuando el verdadero proceso es un MA(q), entonces $\rho_{aZ}(k)=0$ si $k>q$, y en este caso, de (7):

$$\rho_{bZ}(k) = 0 \text{ si } k>q$$

lo que indica que para esta clase de procesos la función de autocorrelación cruzada entre b_t y Z_t desaparece a partir del verdadero orden q y el empleo de cualquier modelo para filtrar a Z_t no distorsiona el tipo de proceso ni su orden q.

Parte ii): Considere $k<0$ en la ecuación (7),

$$\rho_{bZ}(k) = c\rho_{aZ}(k) + c \sum_{j=1}^{\infty} \psi_{3j} \rho_{aZ}(k+j)$$

Entonces, puesto que para el verdadero proceso $\rho_{aZ}(k)=0$ si $k<0$,

$$\rho_{bZ}(k) = c \sum_{j=1}^{\infty} \psi_{3j} \rho_{aZ}(k+j) \neq 0$$

donde, en general, $\rho_{aZ}(k+j) \neq 0$ si $k+j \geq 0$. Por tanto, $\rho_{bZ}(k)$ presentará valores no nulos si $k<0$.

El siguiente ejemplo ilustra los resultados de la proposición 2.

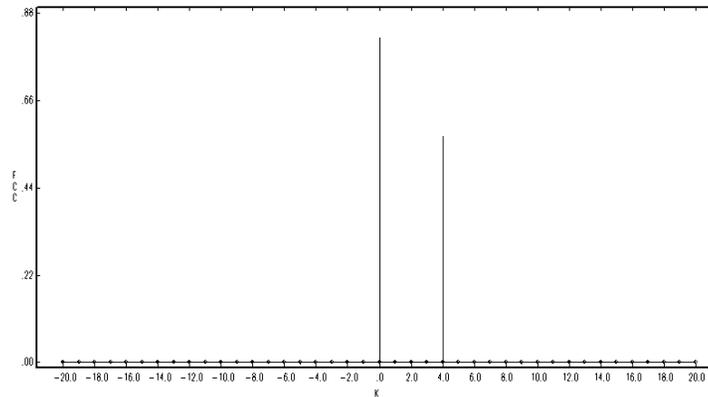
Ejemplo 2.1. Considere el modelo

$$Z_t = (1 + 0.7B^4)a_t \quad (9)$$

donde $\{a_t\}$ es un proceso de ruido blanco $N(0,1)$. A continuación se presentan los correlogramas cruzados entre Z_t y la serie Z_t filtrada por diferentes modelos ARMA.

i) Z_t es filtrada usando el modelo $Z_t = (1 + 0.7B^4)a_t$, es decir usamos el verdadero proceso generador de datos.

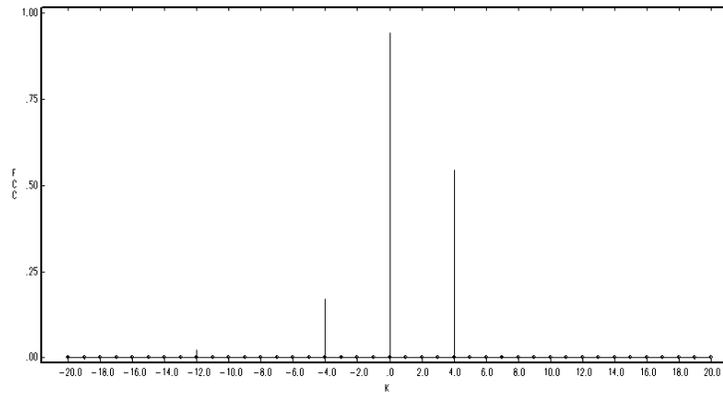
Gráfica 2.1 Modelo $Z_t = (1 + 0.7B^4)a_t$.



Observando la Gráfica 2.1, se concluye que el modelo que generó los datos es un MA(4) con restricciones de nulidad para los tres primeros coeficientes, el cual es el verdadero modelo. Esta información se encuentra representada en la parte derecha del correlograma, la cual es no nula únicamente para el rezago $k=4$, y en la nulidad de la parte izquierda de dicha función.

ii) Z_t es filtrada usando un modelo un poco diferente al verdadero, $Z_t = (1 + 0.4B^4)b_t$

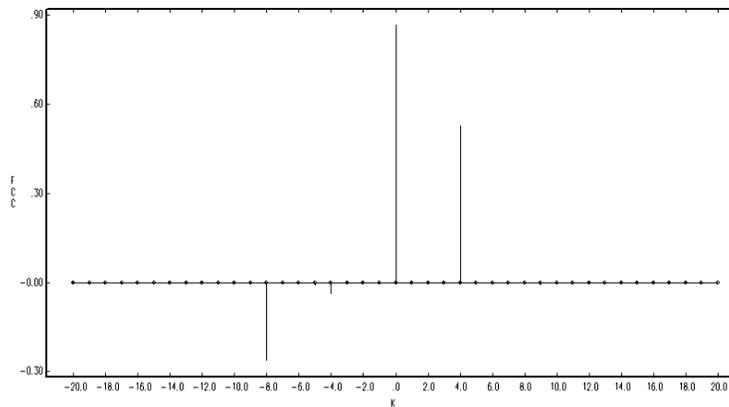
Gráfica 2.2 Modelo $Z_t = (1 + 0.4B^4)b_t$



La Gráfica 2.2, indica que el modelo que generó los datos es un MA(4) con restricciones de nulidad para los tres primeros coeficientes, el cual es el verdadero modelo. Sin embargo, el lado izquierdo no nulo señala que existe alguna especificación errónea del modelo, la cual, está relacionada con la magnitud del coeficiente del operador B^4 .

iii) Z_t es filtrada usando el modelo $Z_t = 1/(1 - 0.5B^4)b_t$

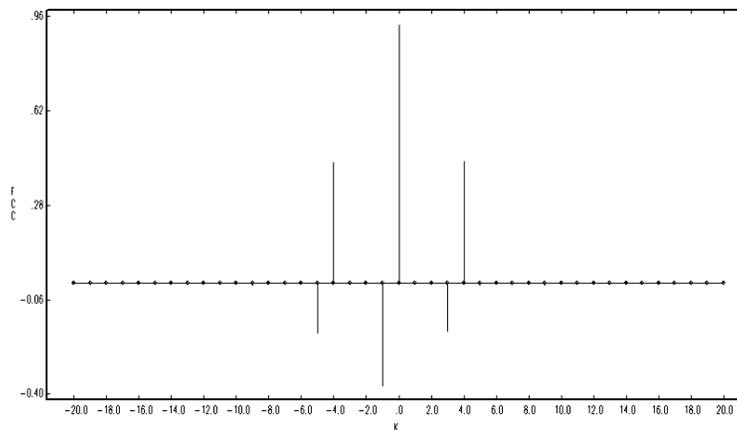
Gráfica 2.3 Modelo $Z_t = 1/(1 - 0.5B^4)b_t$



Aunque el modelo utilizado es un AR(4), la parte derecha del la FCC conserva la memoria señalando claramente un MA(4) como proceso generador. Sin embargo, la parte izquierda no nula indica que el AR(4) no es el modelo verdadero.

iv) Z_t es filtrada usando el modelo $Z_t = 1/(1 - 0.4B)b_t$

Gráfica 2.4 Modelo $Z_t = 1/(1 - 0.4B)b_t$

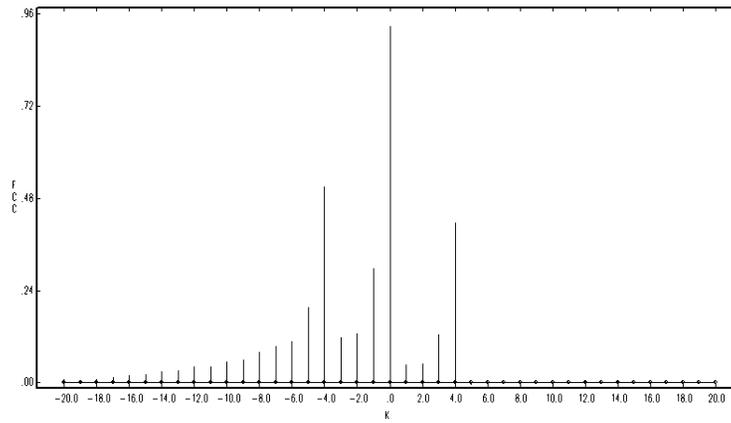


La parte derecha del correlograma conserva la memoria del MA(4), aunque con alguna distorsión, y la parte izquierda no nula señala que el modelo AR(1) no es el verdadero.

Los siguientes casos presentan resultados similares en el sentido de que la parte derecha del correlograma ($k > 0$) siempre conserva, en mayor o menor grado, la memoria del verdadero MA(4), y la parte izquierda ($k < 0$) no nula, señala lo inadecuado del modelo empleado (frente al verdadero modelo) para filtrar la serie.

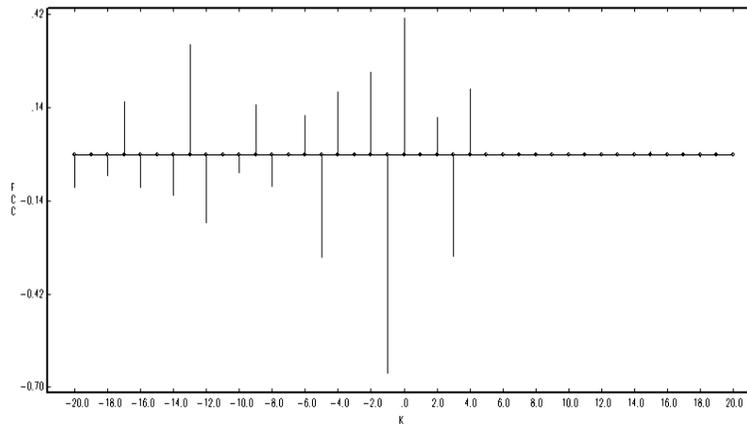
v) Z_t es filtrada usando el modelo $Z_t = (1 - 0.7B - 0.1B^2)/(1 - 0.4B + 0.2B^2)b_t$

Gráfica 2.5 Modelo $Z_t = (1 - 0.7B - 0.1B^2)/(1 - 0.4B + 0.2B^2)b_t$.



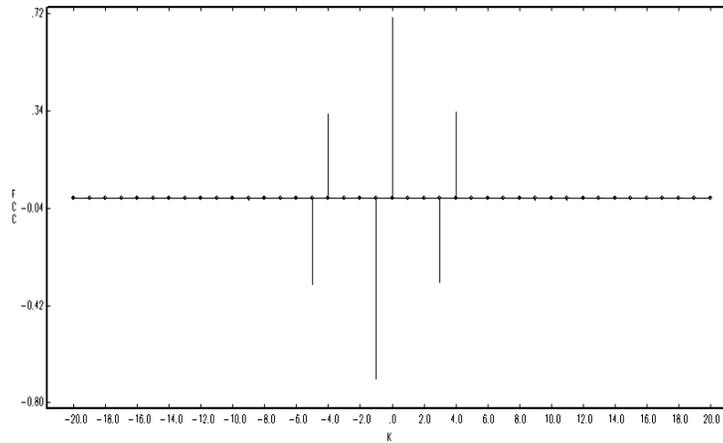
vi) Z_t es filtrada usando el modelo $(1 - B)(1 - B^{12})Z_t = 1/(1 - 0.6B)(1 + 0.5B^{12})b_t$

Gráfica 2.6 Modelo $(1 - B)(1 - B^{12})Z_t = 1/(1 - 0.6B)(1 + 0.5B^{12})b_t$.



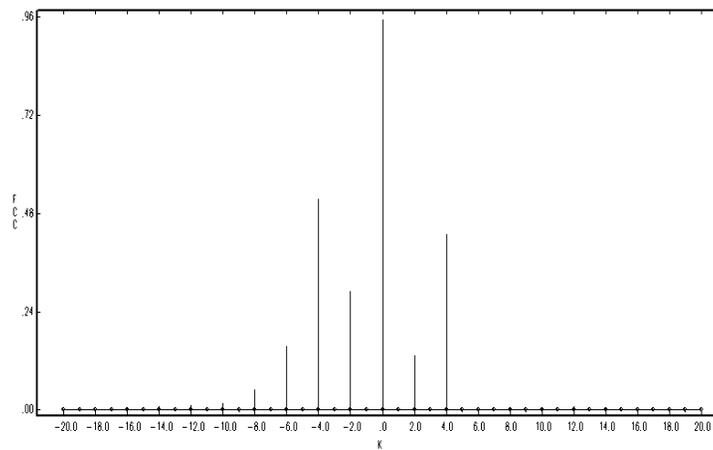
vii) Z_t es filtrada usando el modelo $(1 - B)Z_t = -0.51 + b_t$

Gráfica 2.7 Modelo $(1 - B)Z_t = -0.51 + b_t$.



viii) Z_t es filtrada usando el modelo $Z_t = (1 - 0.3B^2)b_t$

Gráfica 2.8 Modelo $Z_t = (1 - 0.3B^2)b_t$.



2.3 Aplicación de la Proposición 2 a la Identificación de un Modelo ARMA(p,q)

La Proposición 2 puede ser empleada para construir un procedimiento que ayude a la identificación de un modelo ARMA(p,q). A continuación se presentan los pasos que se deben seguir.

- i) Identifique un modelo preliminar, estímelo y calcule la FCCM entre los residuales del modelo estimado y la serie Z_t , $\hat{\rho}_{bZ}(k)$.
- ii) Para $k \geq 0$, el patrón de comportamiento de $\hat{\rho}_{bZ}(k)$, indicará el modelo adecuado para Z_t , y permitirá revisar si el modelo preliminar es correcto.
- iii) Para $k < 0$, el comportamiento no nulo de $\hat{\rho}_{bZ}(k)$, señalará que el modelo preliminar no es adecuado.
- iv) Cuando para $k < 0$, $\hat{\rho}_{bZ}(k)$ es nula, la evidencia sobre el modelo incorrecto se presentará en el comportamiento $\hat{\rho}_{bZ}(k)$ para $k \geq 0$.

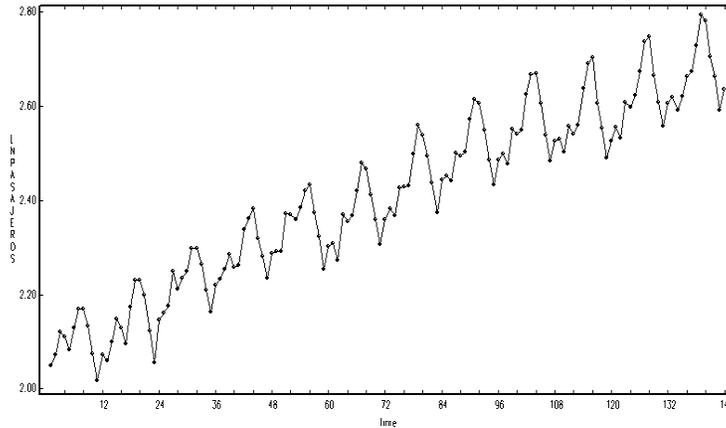
Castaño y Rosales (2005) ilustran el proceso de identificación en modelos ARMA.

2.4 La FCC en la Identificación de Series de Tiempo no Estacionarias.

Claramente, cuando las series son no estacionarias homogéneas todos los resultados obtenidos se cumplen para los procesos apropiadamente diferenciados. El siguiente ejemplo ilustra el procedimiento de identificación.

Ejemplo 2.2. Considere la serie mensual del total de pasajeros en vuelos internacionales (Serie G de Box y Jenkins, 1976), estabilizada en varianza usando el logaritmo natural, y cuyo gráfico aparece a continuación.

Gráfica 2.9. Logaritmo natural del total de pasajeros.



La serie presenta tendencia y estacionalidad, con período estacional $s=12$, y necesita ser diferenciada para estacionarizarla. Ahora bien, los operadores que pueden estacionarizar la serie son $(1-B^{12})$ o $(1-B)(1-B^{12})$.

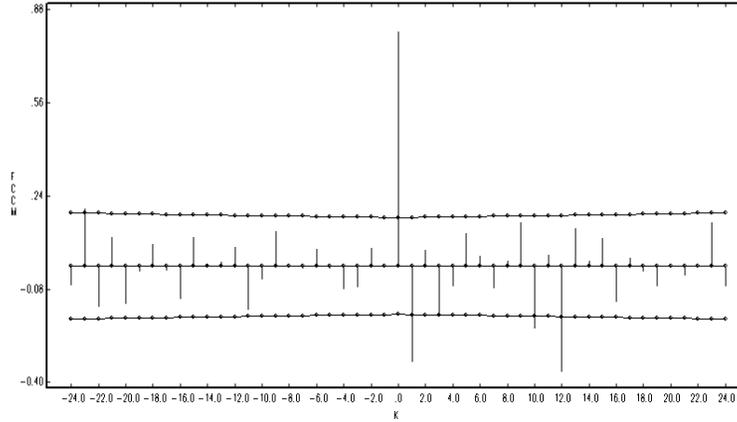
Partiendo de la aplicación del operador $(1-B)(1-B^{12})$ a la serie y de un modelo estacional preliminar $ARMA(1,1) \times (1,1)$ a la serie diferenciada, empleando el paquete estadístico SCA (Scientific Computing Associates Corp.), se obtuvo la siguiente estimación:

Tabla 1. Estimación de modelo preliminar

	LPASAJ	RANDOM		ORIGINAL	$(1-B)(1-B^{12})$			
PARAMETER LABEL	VARIABLE NAME	NUM./ DENOM.	FACTOR	ORDER	CONS- TRAJNT	VALUE	STD ERROR	T VALUE
1	LPASAJ	MA	1	1	NONE	.4445	.1762	2.52
2	LPASAJ	MA	2	12	NONE	.2147	.1258	1.71
3	LPASAJ	D-AR	1	1	NONE	-.0221	.1954	-.11
4	LPASAJ	D-AR	2	12	NONE	-.3439	.1274	-2.70

Claramente, de la tabla anterior, observamos que el modelo estimado parece no ser adecuado. La FCCM entre Z_t y los residuales \hat{a}_t de este modelo es:

Gráfica 2.10. FCCM del modelo preliminar



Esta gráfica indica que la selección del modelo preliminar es incorrecta, pues no parecen existir componentes autorregresivos en el modelo (se manifestarían por medio de mezclas de decrecimientos exponenciales y/o ondas sinusoidales en la parte derecha de la función). Ahora bien, fácilmente se puede probar que la FCC entre a_t y Z_t para el modelo estacionario e invertible

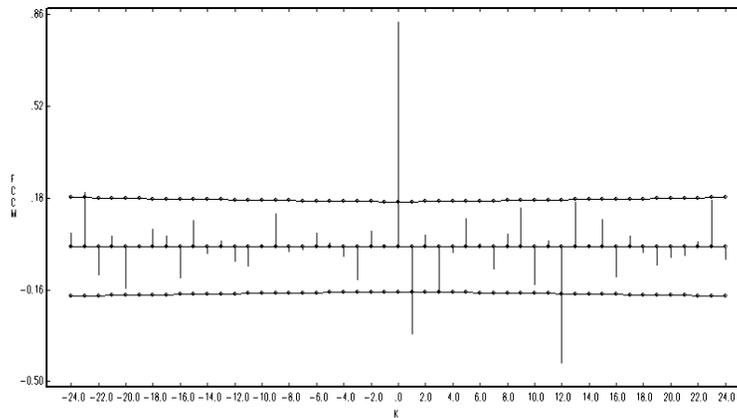
$$Z_t = (1 - \theta_1 B)(1 - \Theta_1 B^{12})a_t \quad (10)$$

es

$$\begin{aligned} \rho_{a\Delta Z}(k) &= 0 && \text{si } k < 0 \\ &= \frac{-\theta_1}{(1 + \theta_1^2 + \Theta_1^2 + \theta_1^2 \Theta_1^2)} && \text{si } k = 1 \\ &= 0 && \text{si } 1 < k < 12 \\ &= \frac{-\Theta_1}{(1 + \theta_1^2 + \Theta_1^2 + \theta_1^2 \Theta_1^2)} && \text{si } k = 12 \\ &= \frac{\theta_1 \Theta_1}{(1 + \theta_1^2 + \Theta_1^2 + \theta_1^2 \Theta_1^2)} && \text{si } k = 13 \\ &= 0 && \text{si } k > 13 \end{aligned}$$

Este patrón teórico es consistente con el patrón exhibido en la Gráfica 2.10, y señala la posibilidad de un modelo MA(1) para la parte no estacional de la serie y un MA(1) para la parte estacional. Esta identificación corresponde al modelo tradicionalmente identificado para esta serie. La estimación de este modelo produce la siguiente FCCM,

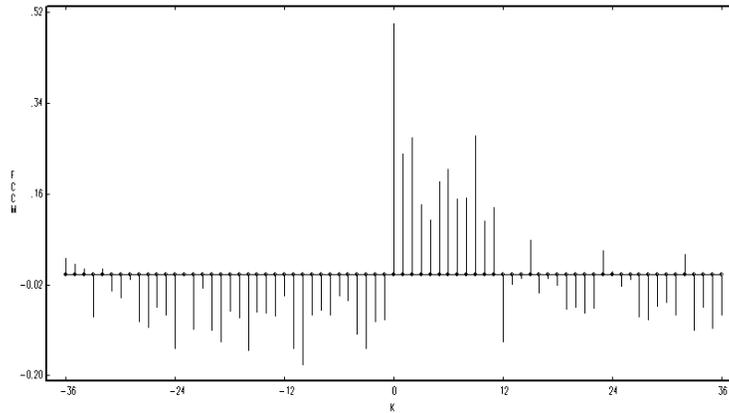
Gráfica 2.11 FCCM del modelo final



Como antes, la parte derecha identifica el modelo estimado, y la parte izquierda parece ser nula, indicando que la especificación parece correcta y consistente con el modelo (10).

Ahora bien, la identificación incorrecta de los órdenes de diferenciación produce el siguiente correlograma cruzado para el modelo preliminar.

Gráfica 2.12 FCCM para la serie diferenciada sólo con $(1-B^{12})$



En este caso es aparente la incorrecta especificación del modelo y la necesidad de diferenciar usando el operador completo $(1-B)(1-B^{12})$.

Finalmente, es importante anotar que este diagnóstico de especificación puede ser obtenido fácilmente, pues la mayoría de programas para el análisis de Series de Tiempo disponen de una opción para calcular la FCCM.

A continuación se presentan algunos resultados de interés para algunas series no estacionarias.

CAPÍTULO 3

EMPLEO DE LA FCC EN LA DISCRIMINACIÓN DE UN PASEO ALEATORIO

3.1 Conceptos preliminares

Considere los procesos

$$Z_t = \phi_1 Z_{t-1} + a_t, \quad |\phi_1| < 1 \quad (10)$$

$$Z_t = Z_{t-1} + a_t \quad (11)$$

donde a_t es ruido blanco $N(0, \sigma_a^2)$.

Los dos procesos son AR(1), y (11) es un caso especial de (10) cuando $\phi_1 = 1$.

Sin embargo mientras que (10) es un proceso estacionario, (11) es un proceso no estacionario homogéneo y sus propiedades estadísticas son muy diferentes: para (10) la media y la varianza son constantes y las autocovarianzas y autocorrelaciones sólo dependen de k , el número de períodos comprendidos entre Z_t y Z_{t+k} , y para (11) dichas funciones dependen de t .

Ahora bien, la función de autocorrelación (ACF) del proceso estacionario (10) es de la forma

$$\rho_k = \phi_1^k, \quad (12)$$

Dado que $|\phi_1| < 1$, ρ_k decae geométricamente cuando k crece, y a medida que $\phi_1 \rightarrow 1$, el decaimiento es cada vez más lento. En el límite, es decir, cuando $\phi_1 = 1$, el

proceso (11) es llamado un *proceso de paseo aleatorio*. Dadas las características de la ACF en (12), el proceso de paseo aleatorio puede ser caracterizado por grandes autocorrelaciones que decrecen muy lentamente.

El proceso (11) también puede ser escrito en la forma:

$$Z_t = \mu_t + a_t$$

donde $\mu_t = Z_{t-1}$ y donde μ_t es llamada la tendencia del proceso. Dado que Z_{t-1} es aleatorio, se dice que el proceso Z_t posee una *tendencia aleatoria*.

Si agregamos una constante a los procesos (10) y (11), obtenemos

$$Z_t = \theta_0 + \phi_1 Z_{t-1} + a_t, \quad |\phi_1| < 1 \quad (13)$$

$$Z_t = \theta_0 + Z_{t-1} + a_t \quad (14)$$

cuyas soluciones son, respectivamente

$$Z_t = \phi_1^t Z_0 + \theta_0 \sum_{j=0}^{t-1} \phi_1^j + \sum_{j=0}^{t-1} \phi_1^j a_{t-j} \quad (15)$$

$$Z_t = Z_0 + \theta_0 t + \sum_{j=1}^{t-1} a_{t-j} \quad (16)$$

donde Z_0 es un valor inicial fijo. Las soluciones difieren bastante: mientras que (16) además de contener una tendencia aleatoria, también contiene una tendencia determinística lineal¹, la evolución de (15) es la típica de un proceso estacionario.

¹ Para $t > k$, el proceso también puede ser escrito $Z_t = Z_k + (t-k)\theta_0 + \sum_{j=k+1}^t a_j, \quad t > k$

Puesto que (11) y (14) no son procesos estacionarios, la teoría asintótica tradicional no puede ser aplicada. Una manera fácil de analizar esta clase de procesos es volviéndolos estacionarios por medio de diferenciación. Diferenciando una vez los procesos (11) y (14) obtenemos, respectivamente

$$\begin{aligned}\nabla Z_t &= a_t \\ \nabla Z_t &= \theta_0 + a_t\end{aligned}$$

donde $\nabla Z_t = Z_t - Z_{t-1}$. Estos procesos pertenecen a la clase de procesos *estacionarios en diferencias* (Difference Stationary, DS). También son denominados procesos de raíz unitaria, dado que la solución B a la ecuación $1 - \phi_1 B = 0$ está dada por $B=1$, lo que implica que $\phi_1 = 1$.

3.2 La FCC del Modelo de Paseo Aleatorio.

El modelo (14)

$$Z_t = \theta_0 + Z_{t-1} + a_t$$

es también llamado un proceso de paseo aleatorio con deriva (con constante). Como vimos la solución de esta ecuación en diferencias es

$$Z_t = Z_0 + \theta_0 t + \sum_{j=1}^t a_j, \quad t \geq 1$$

donde Z_0 es el valor inicial fijo. De la expresión anterior, fácilmente obtenemos que

$$Var(Z_t) = t\sigma_a^2$$

La covarianza cruzada de orden k entre a_t y Z_t es

$$\begin{aligned}
\gamma_{aZ}(t, k) &= E(a_t Z_{t+k}) = E[a_t (Z_0 + \theta_0(t+k) + \sum_{j=1}^{t+k} a_j)] = E[a_t (\sum_{j=1}^{t+k} a_j)] \\
&= \sigma_a^2, \text{ si } k=0, 1, 2, \dots \\
&= 0 \quad \text{si } k < 0
\end{aligned}$$

Por tanto la correlación cruzada de orden k es

$$\begin{aligned}
\rho_{aZ}(t, t+k) &= \gamma_{aZ}(t, t+k) / [\text{var}(a_t) \text{var}(Z_{t+k})]^{1/2} \\
&= \sigma_a^2 / [\sigma_a^2(t+k)\sigma_a^2]^{1/2} \\
&= 1/(t+k)^{1/2} \quad \text{si } k = 0, 1, 2, \dots \\
&= 0 \quad \text{si } k < 0
\end{aligned}$$

Para t grande, los coeficientes $1/(t+k)^{1/2}$ serán pequeños y similares. En consecuencia, en el lado derecho de la FCC aparecerá un patrón de correlaciones cruzadas pequeñas que decrecen muy lentamente. El lado izquierdo de la función es nulo. Observe que para el cálculo de esta FCCM no podemos utilizar la estimación tradicional de la FCC, puesto que Z_t no es estacionaria. Sin embargo, podemos obtener la FCC del modelo estacionarizado

$$(1-B)Z_t = \theta_0 + a_t$$

o,

$$\nabla Z_t = \theta_0 + a_t \quad (17)$$

y caracterizar el comportamiento del paseo aleatorio a través de su FCC. Su derivación aparece a continuación. Para el modelo (17),

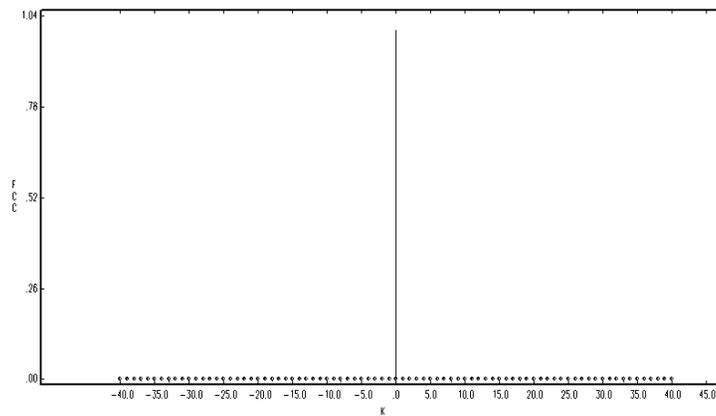
$$\begin{aligned}
\gamma_{a\nabla Z}(k) &= E[a_t \nabla Z_{t+k}] = E[a_t (\theta_0 + a_{t+k})] \\
&= 0 \quad \text{si } k < 0 \\
&= \sigma_a^2 \quad \text{si } k = 0 \\
&= 0 \quad \text{si } k > 0
\end{aligned}$$

además, $Var(\nabla Z_{t+k}) = Var(\theta_0 + a_{t+k}) = \sigma_a^2$. Por tanto, la FCC es de la forma

$$\begin{aligned} \rho_{a\nabla Z}(k) &= 1 \quad \text{si } k=0 \\ &= 0 \quad \text{si } k \neq 0 \end{aligned} \quad (18)$$

Como era de esperar, hay una correlación cruzada contemporánea positiva y perfecta entre a_t y ∇Z_t , y para $k \neq 0$, la FCC es nula. Gráficamente,

Gráfica 3.1 La FCC entre a_t y ∇Z_t de un paseo aleatorio (con o sin deriva)



3.3 La FCC del Modelo de Tendencia Determinística Lineal más Ruido.

Un proceso con tendencia determinística lineal (Trend Stationary, TS) toma la forma,

$$Z_t = \beta_0 + \beta_1 t + a_t$$

donde a_t es ruido blanco (en general, a_t puede ser un proceso estacionario ARMA(p,q)).

Para este modelo, la covarianza cruzada de orden k entre a_t y Z_t , es

$$\begin{aligned}
\gamma_{aZ}(k) &= E[a_t Z_{t+k}] = E[a_t(\beta_0 + \beta_1(t+k) + a_{t+k})] \\
&= \sigma_a^2 \quad \text{si } k = 0 \\
&= 0 \quad \text{si } k \neq 0
\end{aligned}$$

y,

$$\text{Var}(Z_t) = \sigma_a^2$$

Por tanto la correlación cruzada de orden k es

$$\begin{aligned}
\rho_{aZ}(k) &= \frac{\gamma_{aZ}(k)}{[\text{Var}(a_t)\text{Var}(Z_{t+k})]^{1/2}} \\
&= 1 \quad \text{si } k = 0 \\
&= 0 \quad \text{si } k \neq 0
\end{aligned}$$

Esta es la misma FCC entre a_t y ΔZ_t en el modelo de paseo aleatorio.

Ahora bien, el modelo de tendencia determinística lineal no es estacionario, pero sí lo es su primera diferencia

$$\nabla Z_t = \beta_1 + a_t - a_{t-1}$$

la cual corresponde a un modelo IMA(1,1) estrictamente no invertible. La covarianza cruzada entre a_t y ∇Z_t de orden k, es

$$\begin{aligned}
\gamma_{a\nabla Z}(k) &= E[a_t \nabla Z_{t+k}] = E[a_t(\beta_1 + a_{t+k} - a_{t+k-1})] \\
&= \sigma_a^2 \quad \text{si } k = 0 \\
&= -\sigma_a^2 \quad \text{si } k = 1 \\
&= 0 \quad \text{si } k < 0 \text{ o } k > 1
\end{aligned}$$

La varianza de ∇Z_t , es

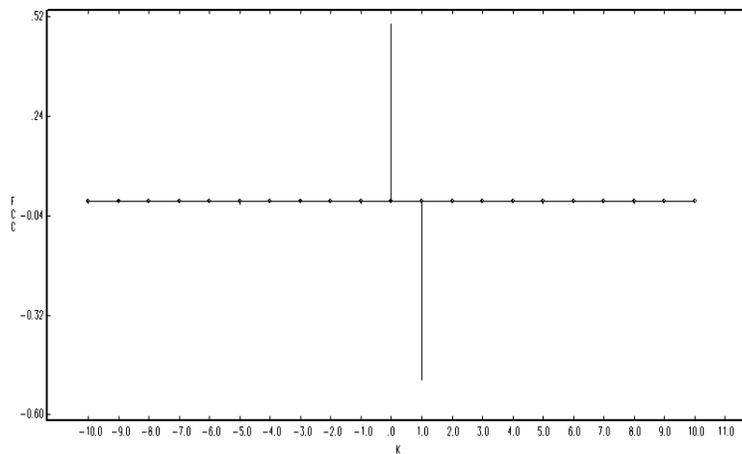
$$\text{Var}[\nabla Z_t] = \text{Var}[\beta_1 + a_t - a_{t-1}] = 2\sigma_a^2$$

Entonces, la correlación cruzada de orden k es

$$\begin{aligned} \rho_{a\nabla Z}(k) &= \frac{\gamma_{a\nabla Z}(k)}{[\text{Var}(a_t)\text{Var}(\nabla Z_{t+k})]^{1/2}} = \frac{\gamma_{a\nabla Z}(k)}{[\sigma_a^2(2\sigma_a^2)]^{1/2}} \\ &= \frac{1}{\sqrt{2}} \quad \text{si } k = 0 \\ &= -\frac{1}{\sqrt{2}} \quad \text{si } k = 1 \\ &= 0 \quad \text{si } k < 0 \text{ o } k > 1 \end{aligned} \quad (19)$$

Por tanto, en un proceso de tendencia lineal determinística más ruido blanco, la primera diferencia genera un modelo IMA(1,1) no invertible con parámetro $\theta_1 = 1$, y la FCC entre a_t y ∇Z_t es nula para $k > 1$ y $k < 0$; para $k=0$ y $k=1$ los coeficientes son de la misma magnitud, pero de signo contrario. Gráficamente,

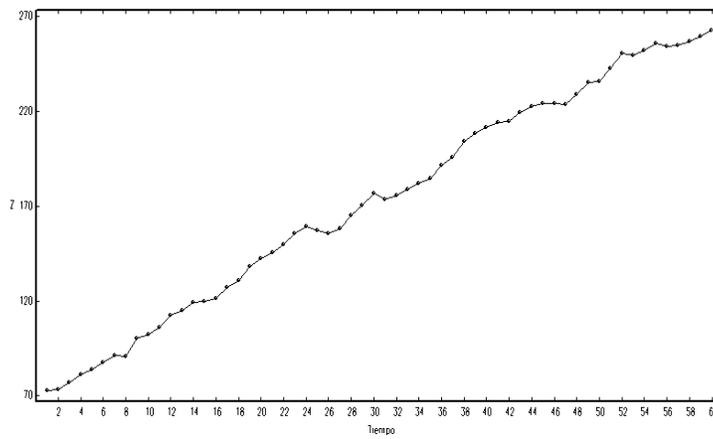
Gráfica 3.2 La FCC entre a_t y ∇Z_t de un proceso de tendencia determinística



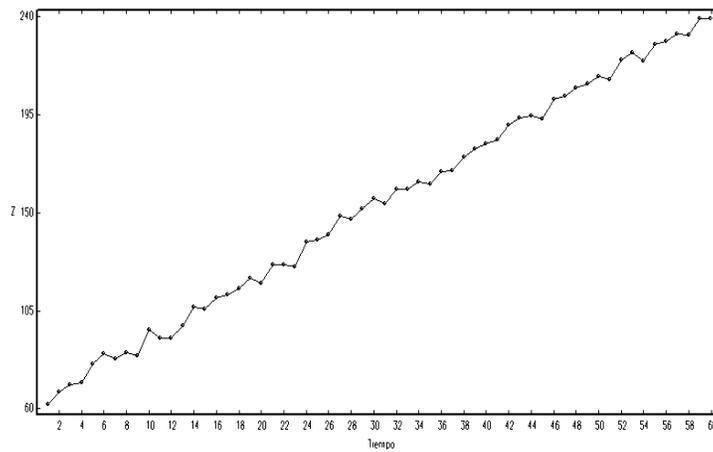
3.4 Discriminación entre un Proceso de Paseo Aleatorio con Deriva y un Proceso de Tendencia Lineal Determinística

No siempre es fácil distinguir entre un modelo de paseo aleatorio con deriva de un modelo con tendencia determinística. Considere los siguientes gráficas.

Gráfica 3.3 Realización de un proceso de paseo aleatorio con deriva



Gráfica 3.4 Realización de un proceso de tendencia lineal determinística

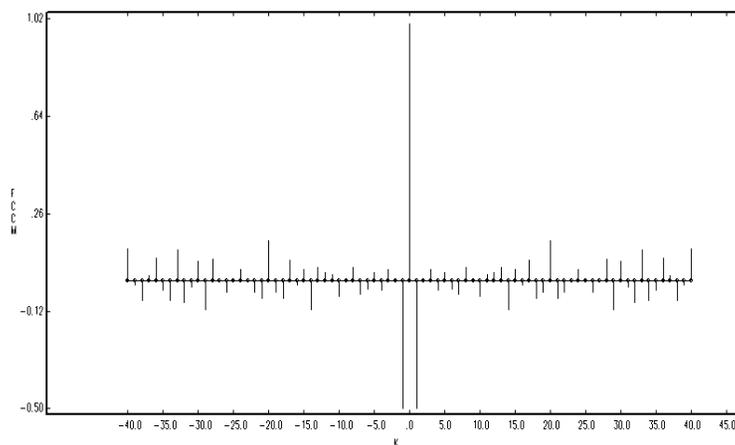


La observación de las realizaciones, en general, no proporciona elementos de juicio suficientes para la elección entre uno de los dos modelos. Sin embargo, usando los resultados obtenidos en la sección anterior, se puede elaborar un diagnóstico gráfico que permita discriminar entre ellos. El procedimiento sería el siguiente:

- i. Obtenga los residuales \hat{a}_t resultantes de la estimación del modelo $\nabla Z_t = \theta_0 + a_t$ y calcule la FCCM entre \hat{a}_t y ∇Z_t .
- ii. Estime el modelo $\nabla Z_t = \beta_1 + a_t - \theta_1 a_{t-1}$, bajo la restricción $\theta_1 = 1$, y calcule la FCCM entre los residuales \hat{a}_t y ∇Z_t .
- iii. Compare las FCCM obtenidas en i. y ii. con los patrones en (18) y (19).

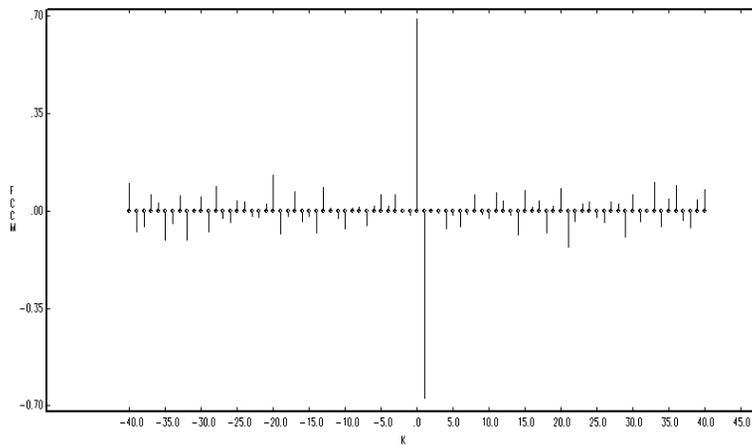
Ejemplo 3. Se simularon $n=200$ datos del proceso $Z_t = \beta_0 + \beta_1 t + a_t$, donde $\beta_0 = 1.0$, $\beta_1 = 0.3$ y a_t sigue una distribución $N(0,1)$. Aplicando el proceso anterior, el gráfico de la FCCM entre \hat{a}_t y ∇Z_t , donde \hat{a}_t son los residuales del modelo $\nabla Z_t = \theta_0 + a_t$ es el siguiente:

Gráfica 3.5 FCCM muestral para el modelo $\nabla Z_t = \theta_0 + a_t$



El siguiente gráfico presenta la FCCM entre \hat{a}_t y ∇Z_t , donde \hat{a}_t son los residuales del modelo $\nabla Z_t = \beta_1 + a_t - a_{t-1}$ (estimado como un IMA(1,1), bajo la restricción $\theta_1 = 1$):

Gráfica 3.6 FCCM muestral para el modelo $\nabla Z_t = \beta_1 + a_t - a_{t-1}$



La comparación de las FCCM obtenidas en los distintos modelos, muestran evidencia a favor de que el proceso generador de datos es un proceso de tendencia lineal determinística, donde los coeficientes en $k=0$ y 1 son los únicos diferentes de cero, con magnitudes iguales pero de signo contrario. (Observe que el primer gráfico sugiere que se estimó un modelo incorrecto, dado que su parte izquierda es no nula).

Nota: También se puede derivar el diagnóstico de la siguiente manera. Suponga que ajustamos a Z_t un proceso de paseo aleatorio con deriva, es decir, $\nabla Z_t = \theta_0 + a_t$. Si el modelo supuesto es correcto, la FCC entre a_t y ∇Z_t es nula para todo $k \neq 0$ y toma el valor de 1 si $k=0$. Sin embargo, si el verdadero proceso asociado a Z_t es el proceso de tendencia lineal determinística $Z_t = \beta_0 + \beta_1 t + a_t$, entonces el proceso estacionarizado por medio de la primera diferencia es $\nabla Z_t = \beta_1 + a_t - a_{t-1}$. Esto implica que el modelo

ajustado está incorrectamente especificado y este hecho se debería reflejar en la FCC entre a_t y ∇Z_t . En este caso el modelo incorrecto es de la forma

$$\nabla Z_t = \theta_0 + e_t$$

donde e_t no es ruido blanco. Comparando con el modelo correcto,

$$e_t = a_t - a_{t-1} = \nabla a_t$$

De aquí que, para conocer la apariencia que tendrá la FCC entre el ruido y ∇Z_t , cuando a un modelo de tendencia determinística se ajusta un modelo de paseo aleatorio con deriva, debemos derivar la FCC entre e_t y $\nabla Z_t = \beta_1 + a_t - a_{t-1}$. La covarianza cruzada de orden k entre e_t y ∇Z_t es

$$\begin{aligned} \gamma_{e\nabla Z}(k) &= E[\nabla a_t \nabla Z_{t+k}] = E[(a_t - a_{t-1})(\beta_1 + a_{t+k} - a_{t+k-1})] \\ &= E[a_t a_{t+k}] - E[a_t a_{t+k-1}] - E[a_{t-1} a_{t+k}] + E[a_{t-1} a_{t+k-1}] \\ &= -\sigma_a^2 \text{ si } k = -1 \\ &= 2\sigma_a^2 \text{ si } k = 0 \\ &= -\sigma_a^2 \text{ si } k = 1 \\ &= 0 \text{ si } k > 1 \text{ o si } k < -1 \end{aligned}$$

La varianza de ∇Z_t bajo H_1 , es

$$\text{Var}[\nabla Z_t] = \text{Var}[\beta_1 + a_t - a_{t-1}] = 2\sigma_a^2$$

La varianza de e_t , es

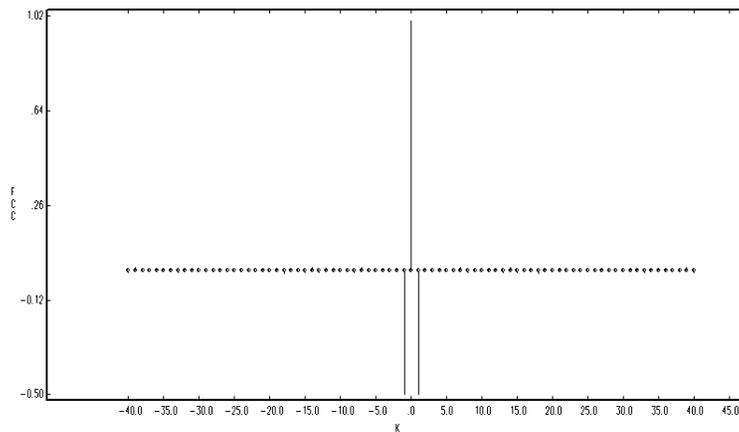
$$\text{Var}[e_t] = \text{Var}[a_t - a_{t-1}] = 2\sigma_a^2$$

Entonces la FCC es,

$$\begin{aligned}
\rho_{e\nabla Z}(k) &= \frac{\gamma_{a\nabla Z}(k)}{[\text{Var}(e_t)\text{Var}(\nabla Z_{t+k})]^{1/2}} = \frac{\gamma_{a\nabla Z}(k)}{[(2\sigma_a^2)(2\sigma_a^2)]^{1/2}} \\
&= -\frac{1}{2} \quad \text{si } k = -1 \\
&= 1 \quad \text{si } k = 0 \\
&= -\frac{1}{2} \quad \text{si } k = 1 \\
&= 0 \quad \text{si } k > 1 \text{ o si } k < -1
\end{aligned}$$

Gráficamente,

Gráfica 3.7 FCC bajo $\nabla Z_t = \beta_1 + a_t - a_{t-1}$



Observe que el primer gráfico (Gráfica 3.5) del ejemplo anterior recuerda claramente este patrón teórico.

Como conclusión, el correlograma cruzado entre \hat{a}_t y ∇Z_t puede ser empleado para elaborar un diagnóstico gráfico que permita discriminar entre un proceso de tendencia lineal determinística y un proceso de tendencia aleatoria con deriva.

CAPÍTULO 4

PRUEBA DE LA HIPÓTESIS DE QUE EL PROCESO GENERADOR DE DATOS ES UN PASEO ALEATORIO

4.1 Introducción.

La información contenida en la FCCM del modelo de paseo aleatorio también puede ser empleada para elaborar un contraste de la hipótesis $H_0 : Z_t = \beta_0 + Z_{t-1} + a_t$. En la literatura de series de tiempo, se encuentran diferentes enfoques para contrastar esta hipótesis. Los más empleados son:

- i. Contraste de raíz unitaria en el polinomio autorregresivo.
- ii. Contraste de raíz unitaria en el polinomio de medias móviles.
- iii. Contraste sobre la variabilidad de la componente de tendencia en series de tiempo estructurales o de componentes no observables.
- iv. Contraste basado en errores de predicción.

A continuación se hace una breve descripción de los procedimientos mencionados.

Caso i.: Nelson y Plosser (1982), proporcionan un contraste de la hipótesis

$$\begin{aligned} H_0 : Z_t &= \beta_0 + Z_{t-1} + a_t, \\ &\text{contra} \\ H_1 : Z_t &= \beta_0 + \beta_1 t + a_t \end{aligned}$$

El procedimiento se basa en probar la existencia de una raíz unitaria en polinomio autorregresivo del proceso. Su contraste usa el modelo

$$\begin{aligned} Z_t &= \beta_0 + \beta_1 t + a_t \\ a_t &= \rho a_{t-1} + u_t \end{aligned}$$

donde u_t es un proceso de ruido blanco $N(0, \sigma_u^2)$. El modelo puede escribirse como

$$Z_t = \beta_0 + \beta_1 t + \rho[Z_{t-1} - \beta_0 - \beta_1(t-1)] + u_t$$

o, en forma reducida

$$Z_t = \alpha + \delta t + \rho Z_{t-1} + u_t$$

o, restando a ambos lados de la ecuación Z_{t-1}

$$\nabla Z_t = \alpha + \delta t + (\rho - 1)Z_{t-1} + u_t$$

o, re-escribiendo

$$\nabla Z_t = \alpha + \delta t + \pi Z_{t-1} + u_t$$

Sobre este modelo, el contraste de H_0 es equivalente a contrastar H_0 : $\pi = 0$ y $\delta = 0$, el cual puede ser realizado usando el contraste Dickey y Fuller (1979) para una raíz unitaria. Otros contrastes de raíces unitarias, tales como el de Phillips y Perron (1988), Sargan y Bhargava (1983), Cochrane (1988), Hall (1989), Choi (1992) pueden ser usados para contrastar H_0 .

Sin embargo, para valores de ρ cercanos a 1 (π cercanos a cero) y muestras pequeñas, Schwert (1989) y De Jong et al. (1992b) señalan que la *potencia del contraste es baja* y aun menor en presencia de correlaciones entre las u_t . De Jong et al. (1992b) argumentan que, en general, las pruebas de raíces unitarias tienen baja potencia contra la alternativa de tendencia determinística y concluyen la necesidad de desarrollar nuevas pruebas que tengan potencias mayores.

Cochrane (1991) señala que cualquier contraste de la hipótesis $H_0: \theta = \theta_0$ tiene una baja potencia contra la hipótesis $H_1: \theta = \theta_0 - \varepsilon$, en muestras pequeñas y valores

arbitrariamente bajos de ε , pero que en muchos casos la diferencia entre θ y $\theta_0 - \varepsilon$ puede no ser importante. Sin embargo, en el caso de una raíz unitaria, la baja potencia tiene una importancia particular debido a la discontinuidad de la teoría de la distribución en la raíz unitaria y sus implicaciones en la interpretación del proceso bajo estudio.

Caso ii.: Considere el modelo

$$\begin{aligned} Z_t &= \mu_t + e_t \\ \nabla e_t &= (1 - \theta_1 B)a_t \end{aligned}$$

donde a_t es un proceso estacionario y μ_t es una componente determinística. Si $\theta_1 = 1$, es decir la componente MA tiene una raíz unitaria, el modelo es llamado un modelo *estrictamente* no invertible. En este caso (por sustitución repetida),

$$e_t = a_t + e_0 - a_0$$

y e_t será un proceso estacionario y por tanto el proceso Z_t tiene tendencia determinística. Si $|\theta_1| < 1$, el proceso es invertible y e_t tiene una raíz unitaria lo que implica que el proceso Z_t tiene tendencia aleatoria. Por tanto, el contraste de la hipótesis $\theta_1 = 1$ contra $|\theta_1| < 1$ es equivalente al contraste de la hipótesis nula de que Z_t tiene tendencia determinística contra Z_t tiene tendencia aleatoria.

Para contrastar esta hipótesis algunos autores tales como Arellano y Pantula (1990), Saikkonen y Luukkonen (1993), Tsay (1993), Tanaka (1993), Choi y Yu (1997) proponen distintos contrastes sobre la existencia de una raíz unitaria en la componente MA.

Caso iii.: Enfoque basado en el análisis de modelo de componentes no observables o modelos de series de tiempo estructurales. Esta aproximación está descrita en Harvey

(1989), Jones (1993) Kitagawa y Gersch (1996), West y Harrison (1989), Young (1984), Shephard y Harvey (1990) y Shephard (1993).

El modelo de series de tiempo más simple para el análisis de la tendencia es el modelo de componente de tendencia aleatoria y un término irregular aleatorio.

$$y_t = \mu_t + \varepsilon_t$$

$$\mu_t = \mu_{t-1} + \eta_t, \quad t = 1, 2, \dots, T$$

donde ε_t y η_t son ruidos blancos distribuidos $N(0, \sigma_\varepsilon^2)$ y $N(0, \sigma_\eta^2)$, respectivamente, y son mutuamente independientes. Este modelo es denominado modelo de *nivel local* y la tendencia sigue un paseo aleatorio. Cuando σ_η^2 es cero, el nivel es constante.

Un modelo más general, que permita que la componente de tendencia tenga una pendiente aleatoria, es el modelo de *tendencia local*

$$y_t = \mu_t + \varepsilon_t$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$

$$\beta_t = \beta_{t-1} + \omega_t, \quad t = 1, 2, \dots, T$$

donde ε_t , η_t y ω_t son ruidos blancos distribuidos $N(0, \sigma_\varepsilon^2)$, $N(0, \sigma_\eta^2)$ y $N(0, \sigma_\omega^2)$, respectivamente, y son mutuamente independientes. Si σ_η^2 y σ_ω^2 son iguales a cero, la tendencia es determinística, es decir,

$$\mu_t = \beta_0 + \beta t$$

Cuando solamente $\sigma_\omega^2 = 0$, la tendencia se reduce a un paseo aleatorio más deriva,

$$\mu_t = \mu_{t-1} + \beta + \eta_t$$

La decisión sobre el tipo de tendencia se basa en el contraste de la hipótesis de que σ_{η}^2 y σ_{ω}^2 son iguales a cero (Harvey, 1989; Harvey 2000).

Caso iv. Recientemente, Flores de Frutos y Jerez (2002) proponen otro enfoque diferente a los anteriores, que se basa en comparar los pronósticos de un modelo probablemente sobrediferenciado con aquellos obtenidos por el correspondiente modelo sin diferenciar. El estadístico del contraste tiene una distribución estándar $\chi^2(1)$ bajo la hipótesis nula de que el proceso sea invertible, es decir $|\theta_1| < 1$.

4.2 Derivación de un Contraste de Paseo Aleatorio basado en la FCCM

Un contraste alternativo a los anteriores, puede ser construido usando como base la información de la FCCM entre \hat{a}_t y ∇Z_t del modelo $Z_t = \beta_0 + Z_{t-1} + a_t$.

Proposición 3. Suponga que a_t es un proceso de ruido blanco $N(0, \sigma_a^2)$ y que β_0 es una constante. El contraste de la hipótesis

$$\begin{aligned} H_0 : Z_t &= \beta_0 + Z_{t-1} + a_t \text{ (} Z_t \text{ es un paseo aleatorio con deriva), contra} \\ H_1 : Z_t &\text{ no es un paseo aleatorio con deriva} \end{aligned}$$

puede ser llevado a cabo usando el estadístico

$$TA = n \sum_{\substack{j=-K \\ j \neq 0}}^K \hat{\rho}_{\hat{a}\nabla Z}^2(j)$$

donde $\hat{\rho}_{\hat{a}\nabla Z}(j)$ es el estimador de $\rho_{\hat{a}\nabla Z}(j)$.

Demostración:

Bajo H_0 y el supuesto de normalidad de a_t , se cumple que:

- El proceso ∇Z_t es estacionario
- El proceso ∇Z_t es un cambio de origen del proceso a_t , y por tanto la estructura de correlación cruzada entre a_t y ∇Z_t es la misma que determina la función de autocorrelación de a_t , es decir, $\rho_{a\nabla Z}(j) = \rho_{aa}(j) = \rho_a(j) = \rho_a(-j)$.
- La FCC entre a_t y ∇Z_t es

$$\rho_{a\nabla Z}(k) = 1 \quad \text{si } k = 0 \\ = 0 \quad \text{si } k \neq 0$$

- $Var[\hat{\rho}_{a\nabla Z}(j)] = Var[\hat{\rho}_a(j)] \approx \frac{1}{n}$, si $j \neq 0$

Empleando los resultados anteriores, cuando H_0 es cierto, el estadístico

$$TA = n \sum_{\substack{j=-K \\ j \neq 0}}^K \hat{\rho}_{a\nabla Z}^2(j)$$

TA tiende a tomar valores pequeños cercanos a cero. Esto implica que valores grandes de TA conducen al rechazo de H_0 . Por tanto, si $TA(\alpha, n, K)$ denota el percentil α -superior de la distribución nula de TA , se rechaza H_0 a un nivel de significancia α si $TA > TA(\alpha, n, K)$.

4.2.1 La Distribución simulada de la estadística TA

La distribución de TA para muestras finitas puede ser derivada por simulación. Para su obtención, es importante observar que ella es invariante al valor de β_0 y de σ_a^2 . La siguiente tabla presenta los percentiles superiores de la distribución de TA , para niveles de significación de $\alpha = 0.01, 0.025$ y 0.05 , valores de $K=5, 10$ y 15 y diferentes tamaños

muestrales. Se usaron 100000 simulaciones para cada n , K y α^2 . En el proceso de simulación se empleó el paquete SCA (Hudak y Liu, 1992).

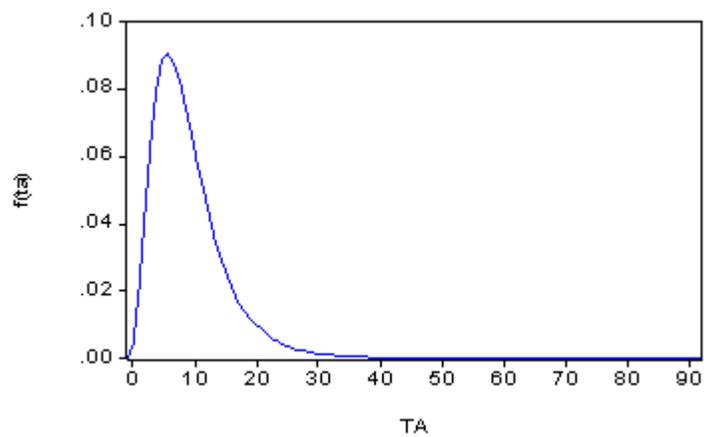
Tabla 4.1 Percentiles superiores para la distribución de TA bajo H_0

n	K=5			K=10			K=15		
	0.05	0.025	0.01	0.05	0.025	0.01	0.05	0.025	0.01
30	19.4	23.0	27.7	30.8	36.0	42.9	38.6	44.3	52.4
40	20.2	23.9	28.9	32.3	37.6	44.8	42.7	49.1	58.1
50	20.4	24.2	29.3	33.3	38.5	45.5	43.3	50.9	59.3
60	20.8	24.4	29.3	33.9	38.9	46.1	45.3	51.9	61.2
100	21.3	25.0	29.9	35.2	40.1	46.5	47.6	53.7	61.5
150	21.7	25.2	30.1	35.8	40.5	46.8	48.5	54.4	62.4
200	21.7	25.1	29.9	35.9	40.6	46.5	48.9	54.7	62.0
250	21.8	25.3	29.9	36.1	40.8	46.9	49.2	54.7	62.1
300	21.9	25.4	29.9	36.2	40.7	46.6	49.2	54.7	61.8
350	21.9	25.5	30.2	36.2	40.7	46.7	49.5	54.9	61.8
400	22.0	25.5	30.0	36.3	40.7	46.6	49.7	55.0	62.0
450	21.9	25.5	30.3	36.3	40.9	46.7	49.6	54.9	61.6
500	21.9	25.4	30.2	36.4	40.8	46.6	49.7	54.8	61.7
1000	22.1	25.6	30.0	36.4	40.8	46.6	49.6	54.8	61.3
5000	22.1	25.7	30.2	36.4	40.6	46.4	50.0	55.0	61.3

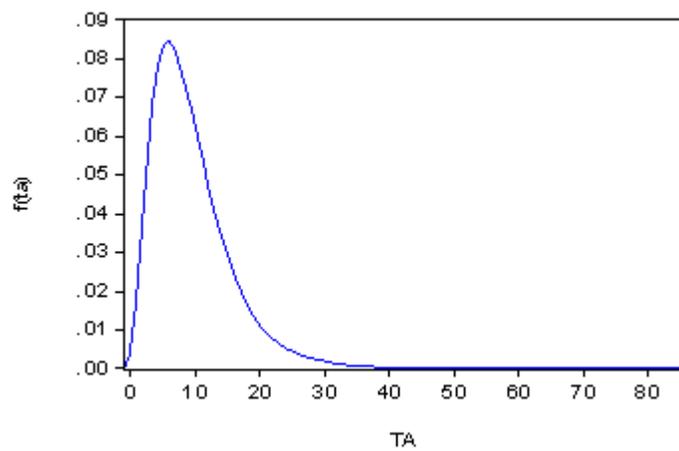
A continuación se presenta el histograma suavizado por medio del kernel de Epanechnikov, usando los 100000 datos y para valores de $n=30, 50, 100$ y 5000 . Para una descripción del procedimiento véase Silverman (1986), Härdle (1991).

² En cada una de las realizaciones simuladas se descartaron los primeros 50 datos con el objeto de eliminar la influencia del valor inicial.

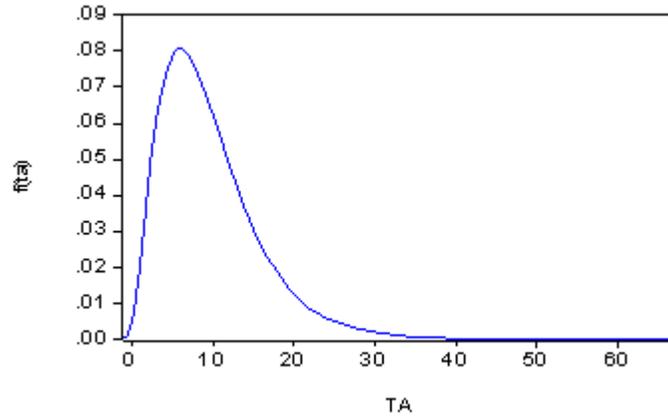
Gráfica 4.1 Distribución de TA , $n=30$, $K=5$



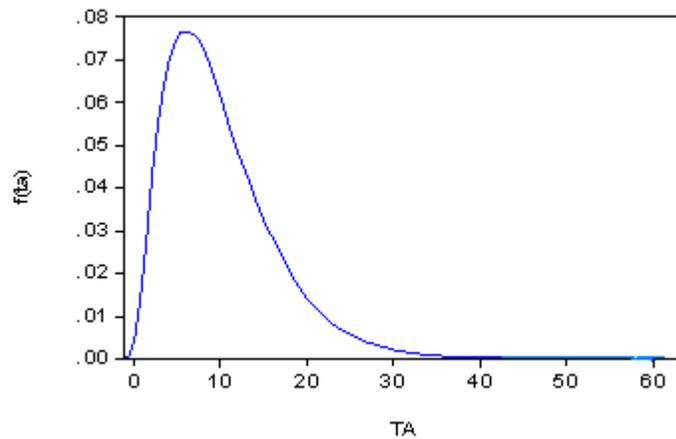
Gráfica 4.2 Distribución de TA , $n=50$, $K=5$



Gráfica 4.3 Distribución de TA , $n=100$, $K=5$



Gráfica 4.4 Distribución de TA , $n=5000$, $K=5$



Nota: En la Tabla 4.1 anterior, a pesar del elevado número de simulaciones realizadas el comportamiento de los percentiles no es monótono. Para lograrlo, los percentiles obtenidos para cada n , K y α fueron suavizados usando la técnica de suavizamiento de diagramas de dispersión denominada Lowess y descrita por Cleveland (1993, 1994). La siguiente tabla presenta los resultados.

Tabla 4.2 Percentiles superiores suavizados para la distribución de TA bajo H_0

N	K=5			K=10			K=15		
	0.05	0.025	0.01	0.05	0.025	0.01	0.05	0.025	0.01
30	19.4	23.1	27.8	30.8	36.0	43.1	38.6	44.9	52.4
40	19.9	23.6	28.4	32.1	37.3	44.2	42.7	47.9	56.5
50	20.4	24.0	29.0	33.2	38.4	45.3	43.4	50.9	59.6
60	20.9	24.5	29.7	33.9	38.9	46.3	45.3	52.3	61.1
100	21.3	24.8	29.7	35.1	39.7	46.3	47.6	53.7	61.6
150	21.6	25.0	29.9	35.6	40.4	46.5	48.3	54.2	61.8
200	21.7	25.2	29.9	35.9	40.6	46.7	48.9	54.7	62.0
250	21.8	25.3	30.0	36.1	40.7	46.7	49.1	54.7	62.0
300	21.9	25.4	30.0	36.2	40.7	46.7	49.3	54.8	61.9
350	21.9	25.4	30.1	36.2	40.7	46.7	49.5	54.9	61.9
400	22.0	25.5	30.2	36.3	40.8	46.6	49.6	54.9	61.8
450	22.0	25.5	30.3	36.3	40.8	46.6	49.7	54.9	61.8
500	22.0	25.5	30.3	36.4	40.8	46.6	49.7	54.9	61.6
1000	22.1	25.6	30.3	36.4	40.8	46.6	49.7	54.9	61.3
5000	22.1	25.7	30.3	36.4	40.6	46.4	50.0	55.0	61.3

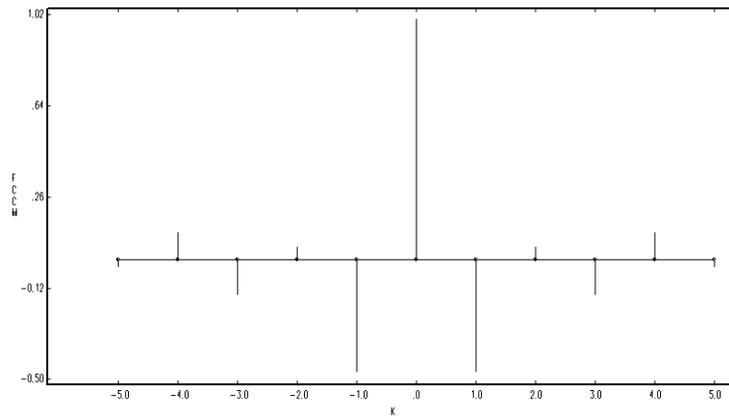
A continuación se presentan algunos ejemplos para ilustrar el desempeño del contraste en muestras pequeñas simuladas.

Ejemplo 4.1. Se simularon $n=40$ datos del proceso $Z_t = \beta_0 + Z_{t-1} + a_t$, donde $\beta_0 = 3.0$, y a_t sigue una distribución $N(0,1)$. Aplicando el procedimiento anterior, se obtuvo para $K=10$, que $TA=16.475$. Para un nivel $\alpha = 0.05$, de la Tabla 4.2 el valor crítico es $TA(\alpha, n, K)=32.1$ y no se rechaza H_0 . Se concluye que los datos presentan evidencia a favor de un proceso paseo aleatorio con deriva.

Ejemplo 4.2. Se simularon $n=40$ datos de un proceso no estacionario gaussiano Z_t . Aplicando el contraste anterior, para $K=5$ se obtuvo un valor observado de $TA=55.609$. Para un nivel $\alpha = 0.05$, de la Tabla 4.2 se obtuvo que $TA(\alpha, n, K)=19.9$. Como $TA > 19.9$, se concluye que los datos rechazan la hipótesis de que el proceso generador de

los datos es paseo aleatorio con deriva. ¿Cuál es el verdadero proceso? La gráfica de la FCCM entre \hat{a}_t y ∇Z_t es la siguiente,

Gráfica 4.5 Diagnóstico gráfico para el ejemplo 4.2



El diagnóstico gráfico presenta evidencia de que el modelo generador de los datos parece ser un proceso de tendencia determinística. En efecto, los 40 datos fueron generados por el proceso $Z_t = \beta_0 + \beta_1 t + a_t$, donde $\beta_0 = 2.0$, $\beta_1 = 1.5$ y a_t sigue una distribución $N(0,1)$.

4.2.2 Potencia del contraste contra diferentes alternativas

A continuación se presenta el comportamiento de la potencia del contraste frente a diferentes hipótesis alternativas.

4.2.2.1 Potencia del contraste contra la alternativa de tendencia lineal determinística

Uno de los problemas de interés en medición económica tiene que ver con la determinación del tipo de tendencia, entre aleatoria o determinística, de una serie de tiempo. La naturaleza de la permanencia de los ‘shocks’ macroeconómicos sobre la serie

depende del tipo tendencia que posea: mientras que el ‘shock’ solamente tiene una permanencia *temporal* en la serie cuando la tendencia es determinística, en una tendencia aleatoria dichos impactos son de naturaleza *permanente*.

Debido a la importancia de esta interpretación, a continuación se presenta un estudio de la potencia del contraste contra la alternativa de una tendencia lineal, para diferentes tamaños de muestra, valores de K y de α (en negrilla), usando la distribución de *TA*. Se usaron 20000 simulaciones. En este caso, la hipótesis alternativa es

$$H_1: Z_t = \beta_0 + \beta_1 t + a_t$$

Tabla 4.3 Potencia del contraste contra la alternativa de una tendencia lineal determinística

n	K=5			K=10			K=15		
	0.050	0.025	0.010	0.050	0.025	0.010	0.050	0.025	0.010
30	0.54	0.40	0.25	0.43	0.30	0.18	0.41	0.29	0.18
40	0.73	0.58	0.39	0.60	0.44	0.29	0.53	0.39	0.24
50	0.88	0.74	0.57	0.73	0.58	0.40	0.68	0.50	0.34
60	0.97	0.89	0.75	0.85	0.72	0.52	0.77	0.62	0.42
100	1.00	1.00	0.99	1.00	0.98	0.93	0.98	0.94	0.85
150	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99
200	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
250	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
300	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
350	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
400	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
450	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Los resultados muestran:

- Una gran capacidad del contraste para discriminar el modelo de paseo aleatorio con deriva del modelo de tendencia determinística lineal, aun en tamaños muestrales pequeños ($n > 40$).

- También se observa que a medida que disminuye el tamaño muestral, el contraste es más potente usando $K=5$.
- A partir de tamaños muestrales alrededor de 100, alcanza la máxima potencia de 1 y es independiente de K .
- El contraste clásico (sin constante ni tendencia) de Dickey-Fuller tuvo potencia cero contra la alternativa de tendencia lineal. Esto significa que dicha prueba es incapaz de distinguir entre una tendencia determinística lineal y la tendencia mixta aleatoria y determinística del modelo de paseo aleatorio con deriva. En otras palabras, ante la presencia de una tendencia determinística lineal en los datos (modelo Trend-Stationary), la prueba clásica de Dickey-Fuller siempre concluiría que el proceso generador de los datos es un modelo de tendencia aleatoria con deriva (modelo Difference-Stationary).
- Para cada valor de K y de α , la potencia tiende a 1 cuando $n \rightarrow \infty$. Es decir, para el caso estudiado, la prueba es consistente para dicha alternativa.
- La potencia de la prueba depende del valor de K . Para la elección del valor de K es importante tener en cuenta que la prueba está diseñada para datos no estacionales. Para datos de baja frecuencia, los resultados obtenidos apoyan la sugerencia de Ljung (1986) de escoger a K relativamente pequeño, por ejemplo alrededor de $K=5$. Para datos de alta frecuencia no estacionales, la elección de K depende de la frecuencia de medición.

4.2.2.2. Potencia del contraste contra la alternativa de un cambio estructural

En muchos casos, la ocurrencia de un cambio estructural en el proceso conduce al no rechazo de la presencia de una raíz unitaria (Perron, 1989). Con el fin de investigar la robustez de la prueba propuesta, la siguiente tabla presenta su potencia contra la alternativa

$$H_1: Z_t = \theta_0 + \delta N_t + a_t$$

donde N_t es una variable que indica un cambio de nivel situado en la mitad del tamaño de n , cuyo efecto es elevar el nivel del proceso en δ . Los valores empleados fueron $\delta = 1.5$, $\theta_0 = 2$ y $a_t \sim N(0, 1)$. Se usaron 20000 simulaciones.

Tabla 4.4 Potencia del contraste contra la alternativa de un cambio estructural

N	K=5			K=10			K=15			DF		
	0.05	0.025	0.01	0.05	0.025	0.01	0.05	0.025	0.01	0.05	0.025	0.01
30	0.50	0.36	0.22	0.41	0.29	0.17	0.39	0.28	0.17	0.00	0.00	0.00
40	0.68	0.53	0.35	0.56	0.40	0.26	0.49	0.36	0.22	0.02	0.00	0.00
50	0.85	0.72	0.53	0.70	0.56	0.38	0.65	0.47	0.31	0.08	0.01	0.00
60	0.94	0.86	0.70	0.82	0.69	0.49	0.74	0.59	0.40	0.21	0.03	0.00
100	1.00	1.00	0.99	0.99	0.98	0.92	0.97	0.93	0.82	0.93	0.61	0.12
150	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	0.87
200	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
250	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
300	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
350	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
400	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
450	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Los resultados muestran:

- Una buena capacidad para discriminar entre los dos procesos aun en muestras pequeñas, y lo hace mucho mejor que el contraste de Dickey-Fuller (columna DF en la tabla anterior).
- En muestras moderadas y grandes (entre 100 y 150), el contraste alcanza una potencia mayor de 0.90 y se comporta mejor que la prueba de DF.
- Para muestras pequeñas, el contraste es más potente para K=5.
- Para muestras grandes (mayores que 150) la potencia de la prueba es 1, y es independiente de K. En este caso comportamiento es similar al del contraste DF.

- Para cada valor de K y de α , la potencia tiende a 1 cuando $n \rightarrow \infty$. Es decir, la prueba es consistente para dicha alternativa.

4.2.3 La Distribución asintótica de la estadística TA

A continuación se deriva la distribución asintótica de TA .

Proposición 4: Bajo H_0 y el supuesto de normalidad de a_t , lo siguiente es cierto:

i) El estadístico TA puede ser escrito

$$TA = n \sum_{\substack{j=-K \\ j \neq 0}}^K \hat{\rho}_{a\nabla Z}^2(j) = n \sum_{\substack{j=-K \\ j \neq 0}}^K \hat{\rho}_a^2(j) = 2n \sum_{j=1}^K \hat{\rho}_a^2(j)$$

ii) $TA \xrightarrow{d} 2\chi^2(K)$

Demostración:

i) Bajo H_0 , el proceso ∇Z_t es un cambio de origen del proceso a_t , y por tanto la estructura de correlación cruzada entre a_t y ∇Z_t es la misma que determina la ACF de a_t , es decir, $\rho_{a\nabla Z}(j) = \rho_{aa}(j) = \rho_a(j) = \rho_a(-j)$. Entonces,

$$TA = n \sum_{\substack{j=-K \\ j \neq 0}}^K \hat{\rho}_{a\nabla Z}^2(j) = n \sum_{\substack{j=-K \\ j \neq 0}}^K \hat{\rho}_a^2(j) = 2n \sum_{j=1}^K \hat{\rho}_a^2(j)$$

ii) Bajo H_0 y normalidad en a_t , $Var[\hat{\rho}_{aVZ}(j)] = Var[\hat{\rho}_a(j)] \approx \frac{1}{n}$, si $j \neq 0$.

Entonces $n \sum_{j=1}^K \hat{\rho}_a^2(j)$ sigue una distribución asintótica $\chi^2(K)$ (Box y Pierce, 1970; Box y Jenkins, 1976; Brockwell y Davis, 1987). Por tanto,

$$TA \xrightarrow{d} 2\chi^2(K).$$

De acuerdo con este resultado, si $\chi^2(\alpha, K)$ denota el percentil α -superior de la distribución $\chi^2(K)$, para muestras grandes se rechaza H_0 a un nivel de significancia aproximado α si

$$TA \geq 2\chi^2(\alpha, K)$$

Ejemplo 4.3. En el ejemplo 4.1, para $K=10$ y un nivel de significancia aproximado de $\alpha=0.05$, el valor crítico $2\chi^2(\alpha, K)=2(18.307)=36.614$. Por tanto, se mantiene la conclusión obtenida antes.

Ejemplo 4.4. En el ejemplo 4.2, para $K=5$ y un nivel de significancia aproximado de $\alpha=0.05$, el valor crítico $2\chi^2(\alpha, K)=2(11.071)=22.142$. Por tanto, se mantiene la conclusión obtenida antes de rechazar H_0 .

4.2.3.1 Derivación del tamaño del contraste en el caso asintótico

La siguiente tabla presenta el tamaño del contraste usando la distribución asintótica de TA , para diferentes valores de K , de n y para $\alpha=0.01, 0.025$ y 0.05 . Se usaron 20000 simulaciones.

Tabla 4.5 Tamaño del contraste bajo la distribución asintótica de TA

n	K=5			K=10			K=15		
	0.050	0.025	0.010	0.050	0.025	0.010	0.050	0.025	0.010
30	0.031	0.014	0.007	0.021	0.013	0.008	0.012	0.009	0.005
40	0.035	0.020	0.009	0.029	0.016	0.008	0.022	0.014	0.008
50	0.040	0.019	0.008	0.032	0.017	0.008	0.027	0.017	0.008
60	0.038	0.019	0.008	0.032	0.017	0.009	0.028	0.015	0.008
100	0.040	0.022	0.008	0.040	0.022	0.011	0.038	0.021	0.011
150	0.047	0.022	0.009	0.043	0.023	0.011	0.045	0.025	0.012
200	0.047	0.023	0.009	0.042	0.022	0.009	0.041	0.023	0.010
250	0.044	0.022	0.009	0.049	0.025	0.009	0.044	0.025	0.011
300	0.048	0.025	0.010	0.047	0.027	0.012	0.048	0.025	0.012
350	0.047	0.024	0.010	0.047	0.023	0.009	0.046	0.023	0.009
400	0.046	0.023	0.010	0.050	0.025	0.010	0.049	0.026	0.010
450	0.048	0.023	0.009	0.045	0.023	0.010	0.045	0.024	0.010
500	0.048	0.022	0.009	0.050	0.026	0.011	0.049	0.026	0.011
1000	0.050	0.023	0.009	0.050	0.024	0.011	0.049	0.026	0.011

De la tabla anterior se observa que

- Para cada valor de K , el tamaño de la prueba tiende a su nivel nominal α a medida que $n \rightarrow \infty$.
- Para tamaños muestrales pequeños, el nivel nominal de α es mayor que el nivel observado. Esto implica que, en muestras pequeñas, la prueba tiende a no rechazar H_0 menos veces de lo esperado: el valor de TA parece ser más pequeño de lo esperado bajo H_0 .
- Para n pequeño y $K=5$, el tamaño de la prueba está más cerca al valor nominal del nivel de significancia α .

Notas:

1) El contraste de H_0 usando el estadístico TA es equivalente al contraste basado en el estadístico de Box-Pierce(1970)

$$Q = n \sum_{j=1}^K \hat{\rho}_a^2(j)$$

para contrastar la hipótesis de que un proceso estocástico es ruido blanco. Bajo H_0 , los percentiles superiores $Q(\alpha, n, K)$, de la distribución exacta de Q se derivan de la Tabla 4.1 como

$$Q(\alpha, n, K) = 0.5 TA(\alpha, n, K)$$

Para muestras grandes, puesto que bajo H_0 , $Q \xrightarrow{d} \chi^2(K)$, la prueba puede ser realizada usando los percentiles superiores $\chi^2(\alpha, K)$ de la distribución $\chi^2(K)$.

2) Ljung-Box (1978) han sugerido que la distribución chi-cuadrado no proporciona una aproximación suficientemente precisa a la distribución del estadístico Q (y por tanto de TA) bajo H_0 . Señalan que los valores de Q (y por tanto los de TA) tienden a ser más pequeños que los esperados bajo la distribución chi-cuadrado. Por tanto, proponen modificar el estadístico Q como

$$Q^* = n(n+2) \sum_{j=1}^K (n-j)^{-1} \hat{\rho}_a^2(j)$$

el cual, bajo H_0 tiene una distribución asintótica $\chi^2(K)$. La correspondiente corrección en TA será

$$TA^* = n(n+2) \sum_{\substack{j=1 \\ j \neq 0}}^K (n-j)^{-1} \hat{\rho}_{aVZ}^2(j)$$

Bajo H_0 , $TA^* \xrightarrow{d} 2\chi^2(K)$.

La siguiente tabla presenta el tamaño del contraste usando la distribución asintótica de TA^* , para diferentes valores de K , de n y de $\alpha = 0.01, 0.025$ y 0.05 . Se usaron 20000 simulaciones.

Tabla 4.6. Tamaño del contraste bajo la distribución asintótica de TA^*

n	K=5			K=10			K=15		
	0.050	0.025	0.010	0.050	0.025	0.010	0.050	0.025	0.010
30	0.061	0.035	0.015	0.074	0.045	0.026	0.083	0.057	0.033
40	0.058	0.032	0.017	0.070	0.044	0.024	0.083	0.054	0.033
50	0.058	0.032	0.015	0.066	0.039	0.021	0.077	0.050	0.029
60	0.052	0.028	0.014	0.063	0.035	0.018	0.075	0.047	0.025
100	0.052	0.027	0.012	0.060	0.034	0.016	0.065	0.041	0.021
150	0.052	0.027	0.012	0.057	0.031	0.015	0.063	0.038	0.019
200	0.052	0.026	0.011	0.051	0.027	0.012	0.055	0.032	0.016
250	0.049	0.024	0.010	0.055	0.031	0.013	0.058	0.030	0.015
300	0.051	0.027	0.011	0.054	0.031	0.014	0.059	0.032	0.016
350	0.050	0.025	0.011	0.054	0.028	0.010	0.055	0.029	0.011
400	0.049	0.025	0.011	0.055	0.028	0.012	0.057	0.030	0.013
450	0.051	0.025	0.010	0.049	0.025	0.011	0.051	0.028	0.013
500	0.048	0.023	0.009	0.053	0.029	0.013	0.056	0.030	0.013
1000	0.050	0.024	0.010	0.052	0.026	0.011	0.052	0.028	0.011

De la tabla anterior se observa que

- Para tamaños muestrales pequeños el valor nominal de α es menor que el nivel observado.
- Para cada valor de K , a medida que n crece el tamaño de la prueba tiende a su nivel de significancia nominal.
- Para n pequeño y $K=5$ está más cerca el valor nominal α del verdadero tamaño de la prueba.

3) Las debilidades del estadístico TA serán las mismas que las del estadístico Q . En particular su potencia contra un proceso AR(1) será cada vez más baja a medida que el coeficiente ϕ se acerca a 1.

4) Una desventaja de las pruebas basadas en estos estadísticos es que no son informativos en el sentido de que el rechazo de H_0 no proporciona información sobre la forma alternativa del modelo.

4.3 Contraste de la Hipótesis de Tendencia Aleatoria Contra Tendencia Determinística

El contraste elaborado en la sección anterior no es informativo en el sentido allí descrito. A continuación se presenta la derivación de una prueba de la hipótesis nula

$$H_0 : Z_t = \beta_0 + Z_{t-1} + a_t$$

contra la hipótesis alternativa

$$H_1 : Z_t = \beta_0 + \beta_1 t + a_t,$$

Es importante tener en cuenta que los métodos presentados a continuación parten del supuesto de que se quiere diferenciar entre un proceso de tendencia aleatoria con deriva y un proceso de tendencia determinística. Esto significa, que el gráfico de la serie exhibirá claramente una tendencia lineal y el objetivo es determinar su origen.

Proposición 5. Si a_t es ruido blanco normal, para contrastar la hipótesis nula $H_0 : Z_t = \beta_0 + Z_{t-1} + a_t$ contra la hipótesis alternativa $H_1 : Z_t = \beta_0 + \beta_1 t + a_t$, se puede emplear el siguiente procedimiento:

1) Suponga que H_0 es cierto y estime el modelo $\Delta Z_t = \beta_0 + a_t$. Obtenga la FCCM entre \hat{a}_t y ∇Z_t . Sean $\hat{\rho}_{\hat{a}\nabla Z}^{(0)}(j)$, $j = \pm 1, \pm 2, \dots$, los coeficientes estimados de dicha FCCM. Obtenga el estadístico

$$TA = n \sum_{\substack{j=-K \\ j \neq 0}}^K \left[\hat{\rho}_{\hat{a}\nabla Z}^{(0)}(j) \right]^2$$

2) Suponga que H_1 es cierto y estime el modelo $Z_t = \beta_0 + \beta_1 t + a_t$. Obtenga la FCCM entre \hat{a}_t y $\hat{Z}_t = Z_t - \hat{\beta}_0 - \hat{\beta}_1 t$, que es equivalente a obtener la función de autocorrelación muestral de \hat{a}_t . Sean $\hat{\rho}_{\hat{a}\hat{Z}}^{(1)}(j)$, $j = \pm 1, \pm 2, \dots$, los coeficientes estimados de dicha FCCM. Obtenga el estadístico

$$TD = n \sum_{\substack{j=-K \\ j \neq 0}}^K \left[\hat{\rho}_{\hat{a}\hat{Z}}^{(1)}(j) \right]^2$$

3) Defina el estadístico

$$TMIN = \text{Mínimo} \{ TA, TD \}$$

No rechace H_0 a un nivel de significancia α^* si

$$TMIN = TA \text{ y } TMIN \leq TMIN(\alpha, n, K)$$

Rechace H_0 si

$$TMIN = TD \text{ o } TMIN > TMIN(\alpha, n, K)$$

donde $TMIN(\alpha, n, K)$ es el percentil α -superior de la distribución nula exacta de $TMIN$ y donde $\alpha^* = 1 - (1 - \alpha)P[TMIN = TA | H_0]$

Demostración:

Bajo H_0 , por construcción TA estará cercano a cero. Entonces, si $TMIN = TA$ valores pequeños de $TMIN$ conducirán al no rechazo de H_0 . Por tanto si $TMIN(\alpha, n, K)$ es el percentil α -superior de la distribución nula de $TMIN$,

No rechace H_0 a un nivel de significancia α^* si

$$TMIN = TA \text{ y } TMIN \leq TMIN(\alpha, n, K)$$

Rechace H_0 si

$$TMIN = TD \text{ o } TMIN > TMIN(\alpha, n, K)$$

donde,

$$\begin{aligned} \alpha^* &= P[TMIN = TD \text{ o } TMIN > TMIN(\alpha, K, n) / H_0] \\ &= 1 - P[TMIN \leq TMIN(\alpha, K, n) / TMIN = TA, H_0] P[TMIN = TA / H_0] \\ &= 1 - (1 - \alpha) P[TMIN = TA / H_0] \end{aligned}$$

4.3.1 La Distribución simulada de la estadística $TMIN$

La distribución para $TMIN$ puede ser derivada por simulación. Para su obtención, es importante observar que la distribución es invariante al valor de β_0 y de σ_a^2 . La siguiente tabla presenta los percentiles superiores de la distribución de $TMIN$, para niveles de significación de $\alpha = 0.01, 0.025$ y 0.05 , valores de $K=5, 10$ y 15 y diferentes tamaños muestrales. Se usaron 100000 simulaciones para cada n, K y α . Los resultados fueron suavizados usando la técnica de suavizamiento de diagramas de dispersión denominada Lowess y descrita por Cleveland (1993, 1994).

Tabla 4.7 Percentiles superiores para la distribución de $TMIN$ bajo H_0

N	K=5			K=10			K=15		
	0.05	0.025	0.01	0.05	0.025	0.01	0.05	0.025	0.01
30	18.4	21.5	25.8	29.5	33.9	39.9	37.3	42.4	49.4
40	19.9	23.4	28.1	31.9	36.9	43.6	42.2	48.4	56.9
50	20.4	24.1	29.2	33.2	38.3	45.1	44.2	50.7	59.0
60	20.8	24.4	29.3	33.9	38.9	46.1	45.3	51.8	61.1
100	21.3	25.0	29.9	35.2	40.1	46.5	47.6	53.7	61.5
150	21.7	25.2	30.1	35.8	40.5	46.8	48.5	54.4	62.4
200	21.7	25.1	29.9	35.9	40.6	46.5	48.9	54.7	62.0
250	21.8	25.3	29.9	36.1	40.8	46.9	49.2	54.7	62.1
300	21.9	25.4	29.9	36.2	40.7	46.6	49.2	54.7	61.8
350	21.9	25.5	30.2	36.2	40.7	46.7	49.5	54.9	61.8
400	22.0	25.5	30.0	36.3	40.7	46.6	49.7	55.0	62.0
450	21.9	25.5	30.3	36.3	40.9	46.7	49.6	54.9	61.6
500	21.9	25.4	30.2	36.4	40.8	46.6	49.7	54.8	61.7
1000	22.1	25.6	30.0	36.4	40.8	46.6	49.6	54.8	61.3
5000	22.1	25.7	30.2	36.4	40.6	46.4	50.0	55.0	61.3

Tabla 4.8 Percentiles superiores suavizados para la distribución de $TMIN$ bajo H_0

n	K=5			K=10			K=15		
	0.05	0.025	0.01	0.05	0.025	0.01	0.05	0.025	0.01
30	18.4	21.5	26.3	29.5	33.9	40.4	37.3	42.5	49.4
40	18.8	22.8	27.5	31.5	35.6	43.0	40.1	48.0	54.7
50	19.2	24.1	28.6	33.0	37.2	44.9	42.8	50.2	59.0
60	19.6	24.4	29.8	33.9	38.9	46.2	45.6	51.8	61.1
100	21.3	25.0	29.8	35.2	39.9	46.3	46.8	53.7	61.3
150	21.5	25.1	29.9	35.6	40.2	46.6	47.8	54.3	62.0
200	21.7	25.2	29.9	35.9	40.6	46.7	48.8	54.6	62.2
250	21.8	25.3	30.0	36.1	40.7	46.7	49.1	54.7	62.0
300	21.9	25.4	30.0	36.2	40.7	46.7	49.3	54.8	61.9
350	21.9	25.5	30.1	36.2	40.7	46.7	49.4	54.9	61.9
400	21.9	25.5	30.1	36.3	40.8	46.7	49.5	54.9	61.8
450	21.9	25.5	30.2	36.3	40.8	46.6	49.6	54.9	61.8
500	21.9	25.5	30.2	36.4	40.8	46.6	49.7	54.9	61.6
1000	22.1	25.6	30.2	36.4	40.8	46.6	49.7	54.9	61.3
5000	22.1	25.7	30.2	36.4	40.6	46.4	50.0	55.0	61.3

Comparando la Tabla 4.1 y la Tabla 4.7, se observa que, a medida que n crece y bajo H_0 , la distribución de $TMIN$ converge rápidamente a la distribución de TA .

Obtención del nivel de significancia de la prueba:

Dada la regla de decisión de la prueba, su verdadero nivel de significancia es $\alpha^* = 1 - (1 - \alpha)P[TMIN = TA | H_0]$, donde $P[TMIN = TA]$ se calculó usando las 100000 simulaciones realizadas para el cálculo de la distribución exacta de $TMIN$ bajo H_0 , para los casos en los cuales $TMIN = TA$. La siguiente tabla presenta la probabilidad de que $TMIN = TA$, para diferentes valores de n y de K .

Tabla 4.9 Tabla para la $P[TMIN = TA]$ bajo H_0

n	K=5	K=10	K=15
30	0.95433	0.95552	0.96044
40	0.98984	0.98858	0.98826
50	0.99816	0.99732	0.99727
60	0.99972	0.99961	0.99931
100	1.00000	1.00000	1.00000
150	1.00000	1.00000	1.00000
200	1.00000	1.00000	1.00000
250	1.00000	1.00000	1.00000
300	1.00000	1.00000	1.00000
350	1.00000	1.00000	1.00000
400	1.00000	1.00000	1.00000
450	1.00000	1.00000	1.00000
500	1.00000	1.00000	1.00000
1000	1.00000	1.00000	1.00000

Los resultados muestran una rápida convergencia de la $P[TMIN = TA]$ a 1, lo que explica el resultado observado de que la distribución de $TMIN$ converge rápidamente a la distribución de TA .

Usando la tabla anterior podemos encontrar el verdadero nivel de significancia de la prueba para cada nivel α en la distribución de $TMIN$. Por ejemplo, para $n=30$, $K=5$ y $\alpha=0.05$, el verdadero nivel de significancia de la prueba es $\alpha^* = 1 - (0.95)(0.95433) = 0.093$. La siguiente tabla presenta los verdaderos niveles de significancia para los diferentes valores considerados de n , K y α .

Tabla 4.10 Verdaderos niveles de significancia α^* de la prueba $TMIN$

N	K=5			K=10			K=15		
	0.05	0.025	0.01	0.05	0.025	0.01	0.05	0.025	0.01
30	0.0934	0.0695	0.0552	0.0923	0.0684	0.0540	0.0876	0.0636	0.0492
40	0.0597	0.0349	0.0201	0.0608	0.0361	0.0213	0.0612	0.0364	0.0216
50	0.0517	0.0268	0.0118	0.0525	0.0276	0.0127	0.0526	0.0277	0.0127
60	0.0503	0.0253	0.0103	0.0504	0.0254	0.0104	0.0507	0.0257	0.0107
100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100
150	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100
200	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100
250	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100
300	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100
350	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100
400	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100
450	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100
500	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100
1000	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100	0.0500	0.0250	0.0100

Como era de esperar, para tamaños muestrales mayores que 60, los niveles de significancia nominales son los verdaderos. Para tamaños muestrales pequeños, entre 30 y 40, para tener un nivel de significancia verdadero α^* alrededor de 0.05 deberíamos usar un α aproximado de 0.01.

Ejemplo 4.5. Se simularon $n=60$ datos del modelo $Z_t = \beta_0 + Z_{t-1} + a_t$, con $\beta_0 = 1.5$ y $a_t \sim N(0,1)$. Para aplicar el procedimiento de la prueba, se obtuvo $TA = 4.267$, $TD = 382.746$, y entonces $TMIN = \text{Mínimo}\{TA, TD\} = 4.267$. Para $K=5$ y un nivel $\alpha = 0.025$, de la tabla 4.8 el valor crítico es $TMIN(0.025, 5, 60) = 24.4$. Como $TMIN = TA$ y $TMIN < 24.4$, no se rechaza H_0 , es decir, se concluye que los datos son generados por un proceso de paseo aleatorio con deriva, a un nivel de significancia $\alpha^* = 0.0253$.

4.3.2 Potencia de la prueba $TMIN$

La siguiente tabla presenta la potencia de la prueba para los diferentes valores considerados de n , K y α . Se usaron 20000 simulaciones.

Tabla 4.11 Potencia de la prueba $TMIN$

n	K=5			K=10			K=15		
	0.050	0.025	0.010	0.050	0.025	0.010	0.050	0.025	0.010
30	0.946	0.944	0.943	0.930	0.928	0.926	0.925	0.922	0.920
40	0.972	0.970	0.968	0.961	0.959	0.958	0.954	0.952	0.950
50	0.987	0.985	0.984	0.979	0.977	0.976	0.974	0.972	0.971
60	0.994	0.993	0.992	0.988	0.986	0.985	0.984	0.982	0.981
100	1.000	1.000	0.999	0.999	0.999	0.998	0.998	0.998	0.997
150	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
300	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
350	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
450	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

De la tabla anterior observamos que:

- Para cada valor de K y cada valor de α , la potencia de la prueba tiende a 1 a medida que crece.
- La prueba tiene una potencia muy alta aun en muestras pequeñas.
- Para $n < 40$, la potencia que aparece en la tabla no es exacta, puesto que el verdadero nivel de significancia usado es α^* y no α . Sin embargo, si empleamos $\alpha = 0.01$ para obtener un nivel $\alpha^* \approx 0.05$, la verdadera potencia está alrededor de 0.943 contra la nominal de 0.946.

4.3.3 Distribución Asintótica de la estadística $TMIN$

Proposición 6. Bajo H_0 , Cuando $n \rightarrow \infty$

$$TMIN \xrightarrow{d} 2\chi^2(K).$$

Demostración:

Cuando $n \rightarrow \infty$, $TA \xrightarrow{d} 2\chi^2(K)$

Además $\{TA - TMIN\} \xrightarrow{p} 0$

Por tanto $TMIN \xrightarrow{d} 2\chi^2(K)$.

Proposición 7. La prueba basada en $TMIN$ es equivalente a la prueba basada en el estadístico

$$QMIN = \text{Mínimo}\{Q^0, Q^1\}$$

donde

$$Q^0 = n \sum_{j=1}^K [\hat{\rho}_a^{(0)}(j)]^2$$

$$Q^1 = n \sum_{j=1}^K [\hat{\rho}_a^{(1)}(j)]^2$$

Donde $\hat{\rho}_a^{(0)}(j)$ y $\hat{\rho}_a^{(1)}(j)$ son los coeficientes de orden j de la función de autocorrelación muestral de los residuales del modelo bajo H_0 y H_1 , respectivamente.

El contraste puede ser llevado a cabo de la siguiente forma:

i) Para muestras pequeñas y moderadas,

No rechace H_0 a un nivel de significancia α^* si

$$QMIN = Q^0 \text{ y } QMIN \leq QMIN(\alpha, n, K)$$

Rechace H_0 si

$$QMIN = Q^1 \text{ o } QMIN > QMIN(\alpha, n, K)$$

Donde,

$$\alpha^* = P[QMIN = Q^1 \text{ o } QMIN > QMIN(\alpha, K, n) | H_0] = 1 - P[QMIN \leq QMIN(\alpha, K, n) | H_0] P[QMIN = Q^0 | H_0]$$

y donde $QMIN(\alpha, K, n) = 0.5TMIN(\alpha, K, n)$

ii) Si n es grande, $QMIN \xrightarrow{d} \chi^2(K)$. En la regla de decisión anterior, reemplace el percentil $QMIN(\alpha, K, n)$ por el percentil $\chi^2(\alpha, K)$.

Demostración:

Se basa en la nota 1) de la sección 4.2.3.1.

Observación:

Para una mejor aproximación a la distribución chi-cuadrado, se pueden emplear las modificaciones de Ljung-Box $TMIN^*$ y $QMIN^*$ a los estadísticos $TMIN$ y $QMIN$, dados por

$$TMIN^* = \text{Mínimo}\{TA^*, TD^*\}$$

donde

$$TA^* = n(n+2) \sum_{\substack{j=-K \\ j \neq 0}}^K (n-j)^{-1} [\hat{\rho}_{a\sqrt{Z}}^{(0)}(j)]^2 \quad \text{y} \quad TD^* = n(n+2) \sum_{\substack{j=-K \\ j \neq 0}}^K (n-j)^{-1} [\hat{\rho}_{a\sqrt{Z}}^{(1)}(j)]^2$$

y

$$QMIN^* = \text{Mínimo}\{Q^{*0}, Q^{*1}\}$$

donde

$$Q^{*0} = n(n+2) \sum_{j=1}^K (n-j)^{-1} [\hat{\rho}_a^{(0)}(j)]^2 \quad \text{y} \quad Q^{*1} = n(n+2) \sum_{j=1}^K (n-j)^{-1} [\hat{\rho}_a^{(1)}(j)]^2.$$

4.3.3.1 Tamaño de la prueba asintótica

La siguiente tabla presenta el tamaño de la prueba para la distribución asintótica de $TMIN$ para diferentes valores de α , K y n . Se usaron 20000 simulaciones.

Tabla 4.12 Tamaño de la prueba bajo la distribución asintótica de $TMIN$.

n	K=5			K=10			K=15		
	0.050	0.025	0.010	0.050	0.025	0.010	0.050	0.025	0.010
30	0.090	0.071	0.060	0.104	0.079	0.063	0.112	0.086	0.068
40	0.061	0.037	0.022	0.079	0.053	0.035	0.088	0.062	0.040
50	0.055	0.030	0.015	0.069	0.044	0.025	0.081	0.054	0.032
60	0.054	0.029	0.014	0.066	0.040	0.023	0.074	0.046	0.026
100	0.052	0.028	0.012	0.061	0.035	0.017	0.065	0.039	0.020
150	0.052	0.028	0.012	0.058	0.032	0.014	0.060	0.035	0.018
200	0.053	0.027	0.011	0.054	0.029	0.013	0.060	0.034	0.016
250	0.049	0.023	0.010	0.054	0.028	0.014	0.060	0.033	0.015
300	0.050	0.026	0.011	0.052	0.028	0.012	0.057	0.031	0.014
350	0.051	0.026	0.011	0.053	0.028	0.012	0.056	0.030	0.014
400	0.051	0.027	0.012	0.052	0.027	0.012	0.054	0.028	0.012
450	0.049	0.024	0.010	0.052	0.027	0.011	0.053	0.028	0.012
500	0.051	0.026	0.010	0.052	0.026	0.010	0.054	0.029	0.013
1000	0.048	0.025	0.010	0.050	0.025	0.010	0.052	0.027	0.011

De la tabla anterior se observa que:

- Para $K=5$, los niveles verdaderos están más próximos a los nominales.
- A medida que K crece, más lenta es la convergencia de los verdaderos niveles a los nominales.
- Para cada K y α , el tamaño de la prueba converge al nivel de significancia nominal cuando $n \rightarrow \infty$.

4.3.3.2. Potencia de la prueba asintótica.

La siguiente tabla presenta la potencia de la prueba asintótica para diferentes valores de α , K y n . Se Usaron 20000 simulaciones.

Tabla 4.13 Potencia de la prueba asintótica

n	K=5			K=10			K=15		
	0.050	0.025	0.010	0.050	0.025	0.010	0.050	0.025	0.010
30	0.944	0.941	0.940	0.922	0.918	0.916	0.915	0.911	0.909
40	0.971	0.970	0.968	0.960	0.957	0.955	0.946	0.943	0.940
50	0.985	0.983	0.982	0.977	0.974	0.972	0.969	0.966	0.964
60	0.995	0.993	0.992	0.987	0.984	0.983	0.984	0.982	0.980
100	1.000	1.000	0.100	0.999	0.999	0.998	0.999	0.998	0.998
150	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
300	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
350	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
450	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

De la tabla anterior observamos que:

- La prueba asintótica tiene una potencia mayor que 0.9 aun en muestras pequeñas.

- Alcanza rápidamente la máxima potencia de 1, independiente de K.
- Para cada K y α , la potencia de la prueba tiende a 1 cuando $n \rightarrow \infty$.

4.4 Generalización de la Prueba.

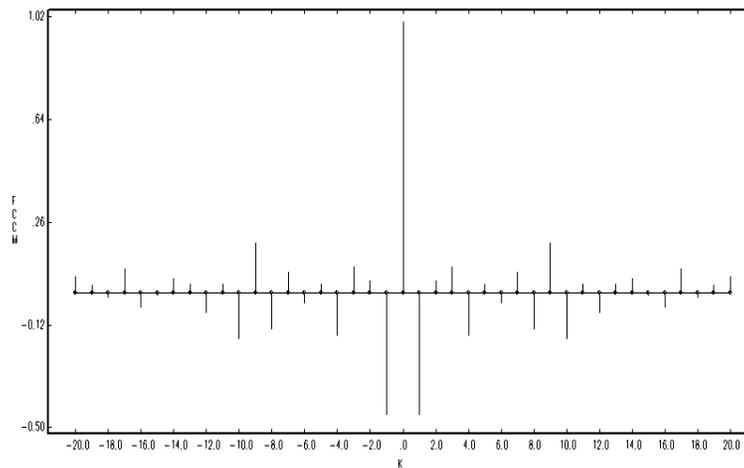
La prueba desarrollada en la sección anterior supone que el término de error de los modelos tanto bajo H_0 como H_1 no está autocorrelacionado. Sin embargo, en muchos casos suele existir autocorrelación que de no tenerla en cuenta, puede deteriorar el comportamiento de la prueba. Para observar los efectos sobre el tamaño de la prueba cuando la autocorrelación es ignorada, la siguiente tabla presenta los tamaños de la prueba suponiendo que bajo H_0 existe una simple estructura MA(1) con parámetro $\theta = 0.5$ en el ruido del modelo. Para K=5 se realizaron 20000 simulaciones para diferentes tamaños muestrales y una distribución N(0,1) para el término de error del modelo.

Tabla 4.14 Efectos de Errores correlacionados sobre el tamaño de la prueba

N	K=5		
	0.050	0.025	0.010
30	0.651	0.631	0.617
40	0.602	0.544	0.495
50	0.627	0.529	0.434
60	0.704	0.582	0.443
100	0.944	0.882	0.763
150	0.998	0.992	0.971
200	1.000	1.000	0.998
250	1.000	1.000	1.000
300	1.000	1.000	1.000
350	1.000	1.000	1.000
400	1.000	1.000	1.000
450	1.000	1.000	1.000
500	1.000	1.000	1.000
1000	1.000	1.000	1.000

Los resultados muestran el daño sobre el tamaño de la prueba debido a la existencia de autocorrelación en el término de error. Debido a la presencia de la autocorrelación, bajo H_0 , el estadístico de la prueba tiende a tomar valores grandes y a medida que n crece, debido a la consistencia de la FCCM, la probabilidad de rechazar H_0 , siendo cierta, crece hacia 1. El siguiente gráfico presenta el correlograma cruzado entre los residuales del modelo y ∇Z_t del proceso simulado cuando $n=100$. El lado derecho de la FCCM señala la presencia de MA(1), y el lado izquierdo no nula muestra la inadecuada especificación del modelo ajustado.

Gráfica 4.6 FCCM bajo H_0 entre \hat{a}_t y ∇Z_t , cuando ∇Z_t posee autocorrelación



En conclusión, si la existencia de autocorrelación no es tomada en cuenta la prueba rechazará H_0 , siendo cierta, con probabilidad más alta a medida que n crece.

Como la autocorrelación ocurre frecuentemente en las aplicaciones, a continuación se extiende la prueba a situaciones donde el término de ruido del modelo es generado por un proceso ARMA(p,q) estacionario e invertible.

Proposición 8: Generalización. Para contrastar $H_0 : Z_t = \beta_0 + Z_{t-1} + \frac{\theta(B)}{\phi(B)} a_t$ contra

$H_1 : Z_t = \beta_0 + \beta_1 t + \frac{\theta(B)}{\phi(B)} a_t$, donde a_t es ruido blanco $N(0, \sigma_a^2)$ y donde

$\phi(B)$ y $\theta(B)$ son los polinomios autorregresivo de orden p y de medias móviles de orden q , con raíces fuera del círculo unidad y sin raíces comunes entre ellos, se puede emplear el siguiente procedimiento:

i) Defina el estadístico $QMIN = \text{Mínimo}\{Q^0, Q^1\}$, donde

$$Q^0 = n \sum_{j=1}^K [\hat{\rho}_a^{(0)}(j)]^2$$

$$Q^1 = n \sum_{j=1}^K [\hat{\rho}_a^{(1)}(j)]^2$$

y donde $\hat{\rho}_a^{(0)}(j)$ y $\hat{\rho}_a^{(1)}(j)$ son los coeficientes de autocorrelación de orden j de la FACM de los residuales del modelo bajo H_0 y H_1 , respectivamente.

ii) Bajo H_0 y n grande, no rechace H_0 a un nivel de significancia aproximado α si

$$QMIN = Q^0 \text{ y } QMIN \leq \chi^2(\alpha, K - p - q)$$

Rechace H_0 si

$$QMIN = Q^1 \text{ o } QMIN > \chi^2(\alpha, K - p - q)$$

Donde $K - p - q > 0$

Demostración:

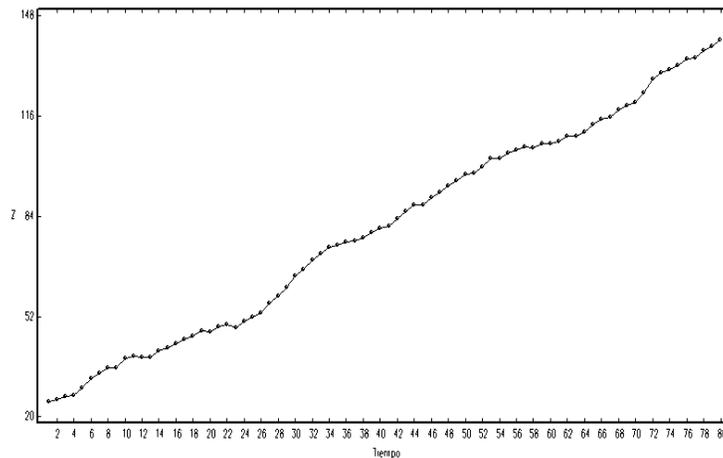
Se basa en la proposición 7.

Observaciones:

- Para una mejor aproximación a la distribución chi-cuadrado, se puede emplear la modificación de Ljung-Box al estadístico $QMIN$.
- En la mayoría de las aplicaciones el modelo ARMA(p,q) asociado al ruido del modelo es desconocido. En estos casos se puede emplear la aproximación de Said-Dickey (1984), para aproximarlos por medio de un modelo AR(p*) donde $p^* \leq n^{1/3}$.

Ejemplo 4.6. Se simularon n=80 datos del proceso $Z_t = \beta_0 + Z_{t-1} + \frac{(1-\theta_1 B)}{(1-\phi_1 B)} a_t$, donde $\theta_1 = 0.6$, $\phi_1 = 0.8$, $\beta_0 = 1.5$ y $a_t \sim N(0,1)$. La siguiente gráfica presenta la realización simulada.

Gráfica 4.7 Una realización del proceso ARIMA(1,1,1) del ejemplo 4.6



Para aplicar el procedimiento de la prueba, se calcularon $Q^0 = 3.018$ y $Q^1 = 4.722$ de los cuales se obtuvo $QMIN = 3.018$. Para $K=5$ y $\alpha = 0.05$, el valor crítico $\chi^2(\alpha, K - p - q)$ es $\chi^2(0.05, 3) = 7.815$. Como $QMIN = Q^0$ y

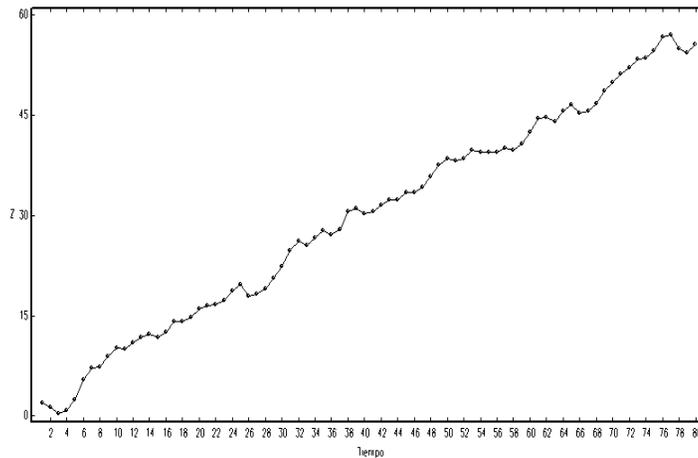
$$QMIN = 3.018 < 7.815 = \chi^2(0.05, 3),$$

no se rechaza H_0 , a un nivel de significancia aproximado $\alpha = 0.05$.

Ejemplo 4.7. Se simularon $n=80$ datos del proceso $Z_t = \beta_0 + \beta_0 t + \frac{(1-\theta_1 B)}{(1-\phi_1 B)} a_t$, donde

$\theta_1 = -0.6$, $\phi_1 = 0.8$, $\beta_0 = 2.0$ y $a_t \sim N(0,1)$. La siguiente gráfica presenta la realización simulada.

Gráfica 4.8 Una realización del proceso ARIMA(1,1,1) del ejemplo 4.7



Para aplicar el procedimiento de la prueba, se calcularon $Q^0 = 10.387$ y $Q^1 = 4.118$ de los cuales se obtuvo $QMIN = 4.118$. Para $K=5$ y $\alpha = 0.05$, el valor crítico $\chi^2(\alpha, K - p - q)$ es $\chi^2(0.05, 3) = 7.815$. Como $QMIN = Q^1$, se rechaza H_0 , es decir, hay evidencia de que el proceso generador de datos es un proceso con tendencia lineal determinística.

4.4.1 Tamaño de la prueba asintótica

A continuación se presentan los resultados del tamaño de la prueba para procesos diferentes procesos MA y AR, bajo diferentes tamaños muestrales y niveles de significancia nominales de $\alpha = 0.01, 0.025, 0.05$.

4.4.1.1 Tamaño de la prueba para procesos MA(1)

Para estudiar el comportamiento del tamaño de la prueba se tendrá en cuenta el signo del parámetro θ .

4.4.1.1.1 Para procesos MA(1) con θ positivo.

La siguiente tabla contiene los resultados de la simulación del tamaño de la prueba para modelos de medias móviles de orden 1 con parámetro positivo $\theta = 0.2, 0.5, 0.8, 0.95$ y valores de $\alpha = 0.01, 0.025, 0.05$. Para cada combinación de θ y α se realizaron 20000 simulaciones.

Tabla 4.15 Tamaño de la prueba para procesos MA(1) con $\theta > 0$.

n	$\theta=0.2$			$\theta=0.5$			$\theta=0.8$			$\theta=0.95$		
	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010
30	0.248	0.239	0.234	0.432	0.426	0.423	0.670	0.664	0.662	0.700	0.692	0.688
40	0.144	0.132	0.126	0.297	0.287	0.281	0.631	0.625	0.622	0.719	0.713	0.708
50	0.089	0.073	0.064	0.194	0.182	0.177	0.572	0.564	0.560	0.724	0.716	0.711
60	0.064	0.044	0.034	0.131	0.115	0.107	0.507	0.498	0.493	0.738	0.731	0.727
70	0.053	0.032	0.020	0.096	0.078	0.068	0.439	0.429	0.422	0.741	0.733	0.729
80	0.047	0.025	0.014	0.078	0.056	0.044	0.369	0.357	0.350	0.741	0.734	0.730
100	0.046	0.024	0.010	0.053	0.031	0.020	0.264	0.247	0.238	0.736	0.730	0.726
150	0.048	0.023	0.009	0.051	0.025	0.010	0.114	0.093	0.080	0.713	0.706	0.701
200	0.048	0.024	0.009	0.050	0.025	0.010	0.071	0.045	0.031	0.659	0.651	0.646
250	0.050	0.025	0.011	0.050	0.025	0.011	0.059	0.032	0.017	0.577	0.567	0.560
300	0.049	0.024	0.010	0.052	0.027	0.011	0.055	0.029	0.012	0.503	0.489	0.480
350	0.047	0.023	0.009	0.051	0.027	0.010	0.054	0.028	0.011	0.431	0.414	0.405
400	0.051	0.025	0.010	0.051	0.026	0.011	0.052	0.026	0.011	0.360	0.341	0.330
450	0.050	0.024	0.010	0.049	0.025	0.011	0.052	0.026	0.011	0.303	0.283	0.271
500	0.046	0.024	0.010	0.049	0.025	0.010	0.052	0.026	0.010	0.255	0.231	0.218
1000	0.048	0.026	0.010	0.049	0.024	0.010	0.053	0.027	0.012	0.086	0.053	0.031
5000	0.051	0.026	0.011	0.048	0.025	0.011	0.052	0.027	0.010	0.073	0.039	0.016

De la tabla anterior se observa que:

- Para cada α y $\theta > 0$, a medida que $n \rightarrow \infty$, el tamaño de la prueba converge al nivel de significancia nominal. La convergencia se hace más lenta a medida que $\theta \rightarrow 1$.
- Para tamaños muestrales pequeños y moderados, a mayor valor del parámetro θ , mayor es la distorsión del verdadero tamaño de la prueba con respecto al nivel de significancia nominal. Este resultado obedece a que en el modelo bajo H_0 dado por

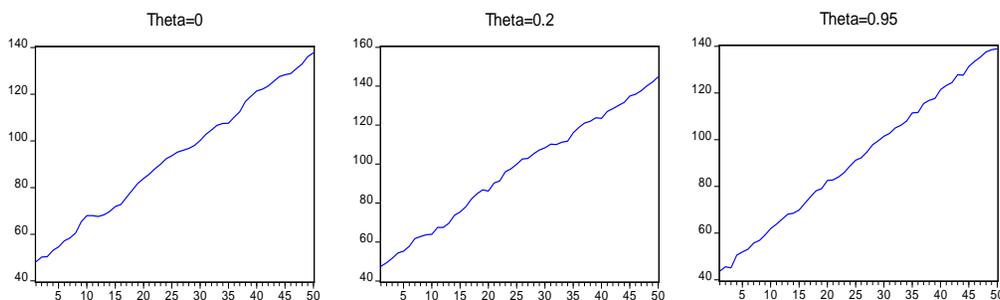
$$\nabla Z_t = \beta_0 + (1 - \theta B)a_t,$$

a medida que $\theta \rightarrow 1$, va apareciendo mayor información sobre una tendencia determinística lineal en el modelo, pues en el límite $\theta = 1$ el modelo es

$$\nabla Z_t = \beta_0 + (1 - B)a_t$$

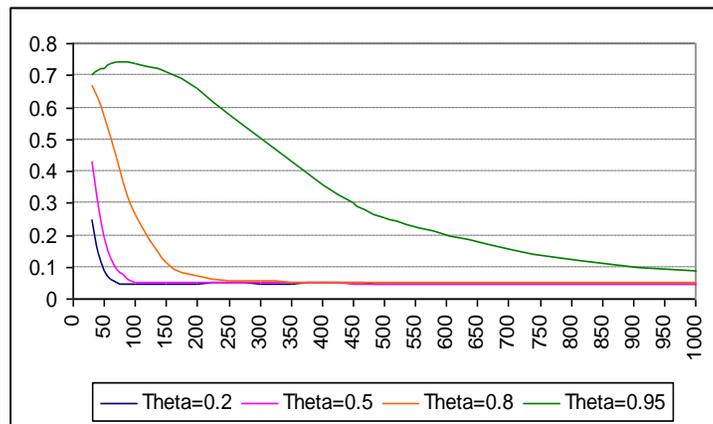
El cual corresponde al modelo de tendencia determinística lineal. Las siguientes gráficas ilustran el comportamiento para el modelo bajo H_0 para $\theta=0$, $\theta=0.2$ y $\theta=0.95$. A medida que $\theta \rightarrow 1$ es más clara la evidencia de una tendencia lineal determinística. En la simulación de las tres series se usó la misma semilla.

Gráfica 4.9 Modelo bajo H_0 para diferentes valores de $\theta > 0$ y $n=50$



- Para valores bajos de θ ($\theta=0.2$) los niveles de significancia nominales se mantienen para muestras de más de 70 observaciones.
- Para valores moderados de θ ($\theta=0.5$) los niveles se mantienen a partir de 100 observaciones.
- Para valores moderadamente grandes de θ ($\theta=0.8$) la distorsión es fuerte y una aproximación útil se da a partir de 200 observaciones.
- Para valores grandes de θ ($\theta=0.95$) la distorsión es mayor y, en general, la prueba tiene una alta probabilidad de rechazar H_0 siendo cierta. Esto indica que cuando existe una componente MA con una raíz próxima a la unidad, la prueba tiende a rechazar a H_0 siendo cierta con probabilidades altas.
- La siguiente gráfica ilustra el comportamiento del tamaño de la prueba para un nivel de significancia nominal de $\alpha = 0.05$ y diferentes valores de n y $\theta > 0$.

Gráfica 4.10 Tamaño de la prueba para $\alpha = 0.05$ y diferentes valores de n y $\theta > 0$



4.4.1.1.2 Para procesos MA(1) con θ negativo

La siguiente tabla presenta los resultados de la simulación del tamaño de la prueba para modelos de medias móviles de orden 1 con parámetro negativo $\theta = -0.2, -0.5, -0.8, -0.95$ y valores de $\alpha = 0.01, 0.025, 0.05$. Para cada combinación de θ y α se realizaron 20000 simulaciones.

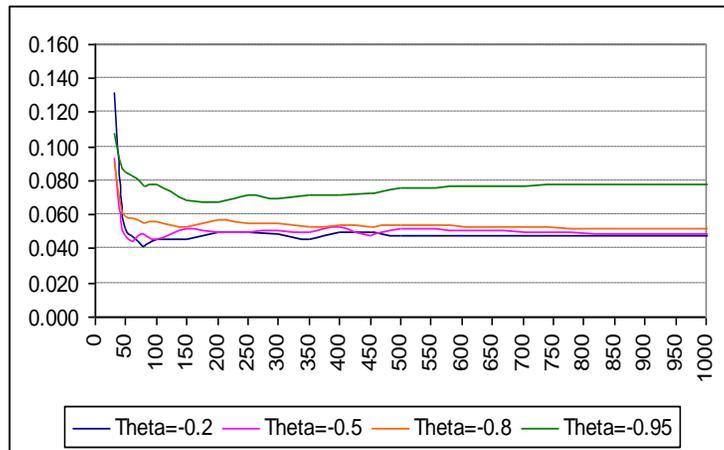
Tabla 4.16 Tamaño de la prueba para procesos MA(1) con $\theta < 0$.

n	$\theta = -0.2$			$\theta = -0.5$			$\theta = -0.8$			$\theta = -0.95$		
	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010
30	0.132	0.121	0.114	0.094	0.078	0.071	0.091	0.065	0.049	0.108	0.073	0.050
40	0.072	0.056	0.048	0.057	0.037	0.026	0.065	0.038	0.021	0.091	0.055	0.031
50	0.052	0.032	0.021	0.048	0.027	0.013	0.059	0.032	0.015	0.085	0.049	0.024
60	0.047	0.025	0.013	0.045	0.023	0.010	0.058	0.030	0.013	0.083	0.044	0.020
70	0.045	0.022	0.008	0.047	0.024	0.010	0.057	0.030	0.014	0.081	0.044	0.021
80	0.042	0.021	0.009	0.049	0.024	0.010	0.055	0.030	0.012	0.077	0.042	0.018
100	0.046	0.022	0.009	0.046	0.022	0.010	0.057	0.029	0.012	0.078	0.042	0.019
150	0.046	0.024	0.010	0.052	0.026	0.011	0.053	0.028	0.011	0.068	0.036	0.015
200	0.050	0.023	0.009	0.050	0.026	0.010	0.057	0.029	0.011	0.067	0.035	0.015
250	0.050	0.025	0.011	0.050	0.025	0.009	0.055	0.027	0.012	0.072	0.037	0.016
300	0.049	0.024	0.010	0.051	0.025	0.010	0.055	0.027	0.011	0.070	0.036	0.014
350	0.046	0.024	0.009	0.050	0.026	0.009	0.053	0.028	0.010	0.072	0.037	0.016
400	0.050	0.025	0.011	0.053	0.026	0.011	0.054	0.028	0.011	0.072	0.038	0.017
450	0.050	0.025	0.009	0.048	0.025	0.010	0.053	0.027	0.011	0.073	0.039	0.017
500	0.048	0.024	0.010	0.052	0.025	0.010	0.054	0.028	0.011	0.076	0.040	0.018
1000	0.047	0.024	0.010	0.049	0.024	0.009	0.052	0.025	0.010	0.078	0.042	0.019
5000	0.050	0.026	0.011	0.050	0.025	0.011	0.053	0.027	0.010	0.070	0.037	0.015

De la tabla anterior se observa que:

- Para cada α y $\theta > 0$, a medida que $n \rightarrow \infty$, el tamaño de la prueba converge a nivel de significancia nominal. La convergencia se hace más lenta a medida que $\theta \rightarrow -1$.
- Para $\theta = -0.2$ y $\theta = -0.5$, los niveles de significancia verdaderos se conservan para valores de $n > 60$ aproximadamente.
- Para $\theta = -0.8$ se obtiene una buena aproximación a partir de $n > 80$.
- Para $\theta = -0.95$, la convergencia es muy lenta pero, se pueden obtener niveles de significancia reales útiles.
- La siguiente gráfica ilustra el comportamiento del tamaño de la prueba para un nivel de significancia nominal de $\alpha = 0.05$ y diferentes valores de n y $\theta < 0$.

Gráfica 4.11 Tamaño de la prueba para $\alpha = 0.05$ y diferentes valores de n y $\theta < 0$



4.4.1.2 Tamaño de la prueba para el modelos AR(1).

Como en el caso del proceso MA(1), para estudiar el comportamiento del tamaño de la prueba se tendrá en cuenta el signo del parámetro ϕ .

4.4.1.2.1 Para procesos AR(1) con ϕ positivo.

La siguiente tabla contiene los resultados del tamaño de la prueba para modelos autorregresivos de orden 1 con parámetro $\phi = 0.2, 0.5, 0.8, 0.95$ y valores de $\alpha = 0.01, 0.025, 0.05$. Para cada combinación de ϕ y α se realizaron 20000 simulaciones.

Tabla 4.17 Tamaño de la prueba para procesos AR(1) con $\phi > 0$.

n	$\phi = 0.2$			$\phi = 0.5$			$\phi = 0.8$			$\phi = 0.95$		
	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010
30	0.293	0.280	0.274	0.126	0.106	0.095	0.074	0.050	0.036	0.076	0.050	0.035
40	0.273	0.261	0.255	0.088	0.067	0.055	0.058	0.036	0.019	0.062	0.037	0.019
50	0.239	0.227	0.220	0.062	0.040	0.029	0.049	0.026	0.013	0.057	0.031	0.015
60	0.215	0.200	0.192	0.058	0.034	0.019	0.051	0.026	0.013	0.059	0.031	0.015
70	0.188	0.172	0.163	0.056	0.031	0.015	0.055	0.030	0.011	0.065	0.035	0.014
80	0.161	0.146	0.137	0.050	0.027	0.012	0.050	0.027	0.012	0.060	0.032	0.013
100	0.126	0.107	0.097	0.049	0.025	0.010	0.052	0.027	0.011	0.064	0.034	0.013
150	0.078	0.057	0.045	0.048	0.024	0.010	0.050	0.026	0.010	0.063	0.031	0.012
200	0.066	0.040	0.025	0.052	0.027	0.010	0.054	0.027	0.011	0.068	0.035	0.015
250	0.057	0.032	0.017	0.052	0.027	0.011	0.055	0.028	0.012	0.069	0.036	0.015
300	0.053	0.026	0.013	0.051	0.025	0.010	0.052	0.026	0.010	0.066	0.034	0.014
350	0.051	0.026	0.011	0.052	0.026	0.010	0.053	0.026	0.010	0.068	0.033	0.014
400	0.048	0.024	0.010	0.047	0.024	0.009	0.050	0.024	0.010	0.064	0.032	0.013
450	0.051	0.024	0.009	0.051	0.025	0.010	0.054	0.027	0.011	0.068	0.035	0.014
500	0.052	0.026	0.010	0.050	0.026	0.011	0.050	0.027	0.011	0.067	0.034	0.015
1000	0.047	0.023	0.010	0.048	0.024	0.010	0.050	0.025	0.010	0.066	0.032	0.013
5000	0.048	0.024	0.010	0.051	0.026	0.011	0.052	0.026	0.010	0.069	0.035	0.014

De la tabla anterior se observa que:

- A medida que $n \rightarrow \infty$, el tamaño de la prueba converge a nivel de significancia nominal, aunque la convergencia es muy lenta para $\phi = 0.95$.
- Para muestras pequeñas, a mayor valor del parámetro ϕ , menor es la distorsión del nivel de significancia nominal con respecto al verdadero tamaño de la prueba, excepto para valores de ϕ próximos a 1 ($\phi = 0.95$).
- A medida que $\phi \rightarrow 1$ es más clara la evidencia de una tendencia aleatoria. Este resultado obedece a que en el modelo bajo H_0 dado por

$$\nabla Z_t = \beta_0 + \frac{1}{(1 - \phi B)} a_t,$$

el cual puede ser escrito como

$$(1 - \phi B)\nabla Z_t = (1 - \phi)\beta_0 + a_t$$

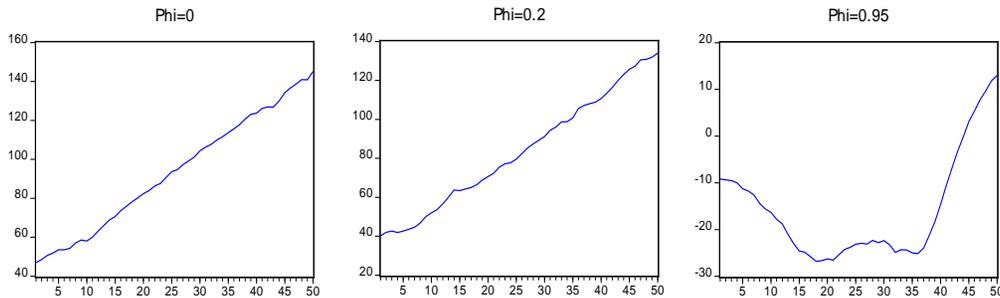
a medida que $\phi \rightarrow 0$, la tendencia determinística presente en el modelo bajo H_0 se fortalece (para $\phi > 0$), conduciendo a rechazar a H_0 más frecuentemente de lo esperado. Por otro lado, en el límite, cuando $\phi = 1$, el modelo es

$$\nabla^2 Z_t = a_t$$

en el cual no queda vestigio de la deriva y la tendencia del modelo es puramente aleatoria.

- Las siguientes gráficas ilustran el comportamiento para el modelo bajo H_0 para $\phi = 0$, $\phi = 0.2$ y $\phi = 0.95$.

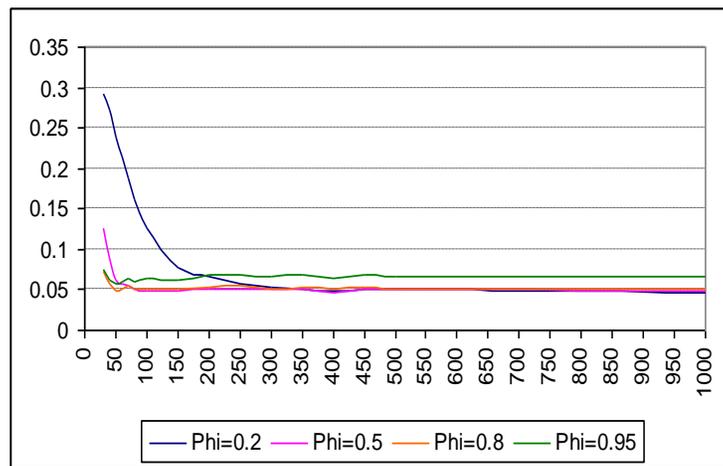
Gráfica 4.12 Modelo bajo H_0 para diferentes valores de ϕ y $n=50$



- Para valores moderadamente altos de ϕ ($\phi = 0.8$), los niveles de significancia nominales se mantienen para muestras de más de 50 observaciones.
- Para valores moderados de ϕ ($\phi = 0.5$) los niveles se mantienen a partir de 70 observaciones.
- Para valores pequeños ϕ ($\phi = 0.2$) se presenta la mayor distorsión y una aproximación útil se obtiene a partir de 150 observaciones.

- Para valores grandes de ϕ ($\phi=0.95$), el nivel nominal está cerca al tamaño de la prueba pero la convergencia es muy lenta y en general, las diferencias son mayores que para el caso de $\phi=0.8$.
- La siguiente gráfica ilustra el comportamiento del tamaño de la prueba para un nivel de significancia nominal de $\alpha = 0.05$ y diferentes valores de n y $\phi > 0$.

Gráfica 4.13 Tamaño de la prueba para $\alpha = 0.05$ y diferentes valores de $\phi > 0$



4.4.1.2.2 Para procesos AR(1) con ϕ negativo.

La siguiente tabla contiene los resultados del tamaño de la prueba para modelos autorregresivos de orden 1 con parámetro $\phi = -0.2, -0.5, -0.8, -0.95$ y valores de $\alpha = 0.01, 0.025, 0.05$. Para cada combinación de ϕ y α se realizaron 20000 simulaciones.

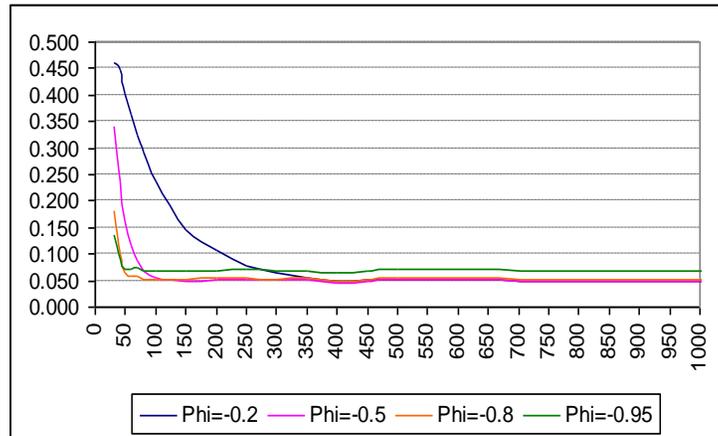
Tabla 4.18 Tamaño de la prueba para procesos AR(1) con $\phi < 0$.

n	$\phi = -0.2$			$\phi = -0.5$			$\phi = -0.8$			$\phi = -0.95$		
	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010
30	0.461	0.455	0.452	0.340	0.333	0.331	0.181	0.168	0.161	0.138	0.114	0.101
40	0.447	0.441	0.437	0.239	0.228	0.224	0.100	0.080	0.070	0.091	0.055	0.031
50	0.404	0.396	0.393	0.162	0.149	0.142	0.064	0.042	0.029	0.071	0.039	0.020
60	0.365	0.357	0.353	0.115	0.097	0.089	0.058	0.032	0.018	0.071	0.037	0.017
70	0.326	0.316	0.311	0.089	0.069	0.056	0.060	0.032	0.013	0.074	0.040	0.017
80	0.293	0.284	0.278	0.068	0.048	0.036	0.052	0.026	0.011	0.067	0.035	0.014
100	0.237	0.223	0.216	0.054	0.032	0.018	0.053	0.027	0.013	0.067	0.036	0.016
150	0.146	0.129	0.120	0.049	0.024	0.010	0.053	0.026	0.011	0.068	0.034	0.015
200	0.106	0.086	0.072	0.051	0.027	0.011	0.054	0.028	0.011	0.070	0.036	0.014
250	0.079	0.057	0.044	0.052	0.026	0.011	0.055	0.028	0.012	0.070	0.037	0.016
300	0.065	0.041	0.027	0.051	0.026	0.010	0.051	0.027	0.011	0.067	0.036	0.014
350	0.056	0.034	0.020	0.051	0.025	0.010	0.054	0.027	0.011	0.069	0.036	0.015
400	0.050	0.027	0.014	0.047	0.024	0.009	0.049	0.025	0.010	0.065	0.034	0.013
450	0.052	0.026	0.012	0.049	0.024	0.010	0.052	0.025	0.011	0.069	0.034	0.014
500	0.051	0.026	0.011	0.052	0.025	0.010	0.054	0.028	0.011	0.073	0.037	0.015
1000	0.049	0.023	0.010	0.050	0.024	0.010	0.051	0.025	0.010	0.067	0.032	0.013
5000	0.048	0.025	0.010	0.050	0.025	0.011	0.051	0.025	0.010	0.069	0.035	0.014

De la tabla anterior se observa que:

- Para cada α y $\phi < 0$, a medida que $n \rightarrow \infty$, el tamaño de la prueba converge al nivel de significancia nominal. La convergencia se hace más lenta para $\phi = -0.95$.
- A medida que la magnitud de ϕ tiende a cero, mayor es la distorsión del verdadero tamaño de la prueba con respecto a su nivel nominal.
- Para $\phi = -0.2$, se obtienen tamaños útiles a partir de $n > 200$.
- Para $\phi = -0.5$, se obtienen tamaños útiles a partir de $n > 70$.
- Para $\phi = -0.8$ y $\phi = -0.95$, se obtienen tamaños útiles a partir de $n > 40$, aunque para $\phi = -0.95$ la convergencia es más lenta.
- La siguiente gráfica ilustra el comportamiento del tamaño de la prueba para un nivel de significancia nominal de $\alpha = 0.05$ y diferentes valores de n y $\phi < 0$.

Gráfica 4.14 Tamaño de la prueba para $\alpha = 0.05$ y diferentes valores de $\phi < 0$



4.4.2 Potencia de la prueba asintótica

A continuación se presentan los resultados de la potencia de la prueba para diferentes procesos MA y AR, para $K=5$ y bajo diferentes tamaños muestrales y niveles de significancia nominales de $\alpha = 0.01, 0.025, 0.05$.

4.4.2.1 Potencia para modelos MA(1)

Para estudiar el comportamiento de la potencia de la prueba se tendrá en cuenta el signo del parámetro θ .

4.4.2.1.1 Para modelos MA(1) con θ positivo

La siguiente tabla contiene los resultados de la potencia de la prueba para modelos de medias móviles de orden 1 con parámetro $\theta = 0.2, 0.5, 0.8, 0.95$ y valores de $\alpha = 0.01, 0.025, 0.05$. Para cada combinación de θ y α se realizaron 20000 simulaciones.

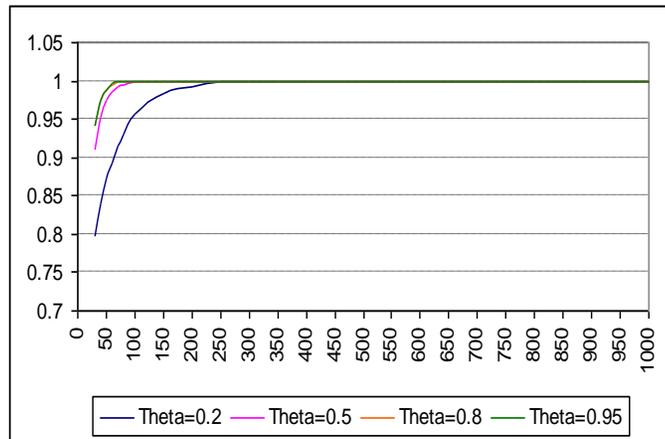
Tabla 4.19 Potencia de la prueba para procesos MA(1) con $\theta > 0$.

n	$\theta=0.2$			$\theta=0.5$			$\theta=0.8$			$\theta=0.95$		
	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010
30	0.798	0.791	0.788	0.910	0.905	0.903	0.942	0.937	0.934	0.942	0.936	0.933
40	0.839	0.834	0.831	0.952	0.949	0.946	0.974	0.972	0.969	0.974	0.971	0.967
50	0.873	0.868	0.865	0.974	0.971	0.969	0.989	0.986	0.984	0.989	0.985	0.982
60	0.892	0.888	0.884	0.986	0.983	0.982	0.995	0.993	0.992	0.996	0.993	0.991
70	0.914	0.908	0.906	0.992	0.990	0.989	0.998	0.997	0.996	0.998	0.997	0.996
80	0.930	0.925	0.922	0.995	0.993	0.992	1.000	0.998	0.997	1.000	0.999	0.997
100	0.956	0.953	0.951	0.999	0.998	0.997	1.000	1.000	1.000	1.000	1.000	1.000
150	0.984	0.982	0.980	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	0.993	0.992	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
250	0.997	0.997	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
300	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
350	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
450	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

De la tabla anterior se observa que:

- La prueba es consistente, ya que para cada valor α y de θ , la potencia de la prueba converge a 1 a medida que $n \rightarrow \infty$.
- Para $\theta=0.2$, la prueba alcanza una potencia >0.9 para $n > 60$
- Para $\theta=0.5$, la prueba alcanza una potencia >0.9 para $n \geq 30$
- Para $\theta=0.8$, la prueba alcanza una potencia >0.9 para $n \geq 30$
- Para $\theta=0.95$, la prueba alcanza una potencia >0.9 para $n \geq 30$
- La siguiente gráfica ilustra el comportamiento de la potencia de la prueba para un nivel de significancia nominal de $\alpha = 0.05$ y diferentes valores de n y $\theta > 0$.

Gráfica 4.15 Potencia de la prueba para $\alpha = 0.05$ y diferentes valores de $\theta > 0$.



4.4.2.1.2 Para modelos MA(1) con θ negativo

La siguiente tabla contiene los resultados de la potencia de la prueba para modelos de medias móviles de orden 1 con parámetro $\theta = -0.2, -0.5, -0.8, -0.95$ y valores de $\alpha = 0.01, 0.025, 0.05$. Para cada combinación de θ y α se realizaron 20000 simulaciones.

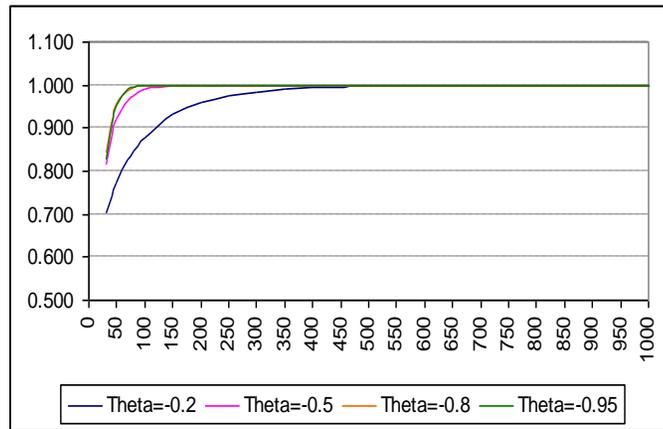
Tabla 4.20 Potencia de la prueba para procesos MA(1) con $\theta < 0$.

n	$\theta = -0.2$			$\theta = -0.5$			$\theta = -0.8$			$\theta = -0.95$		
	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010
30	0.704	0.698	0.694	0.814	0.808	0.806	0.843	0.835	0.830	0.827	0.815	0.808
40	0.740	0.733	0.729	0.882	0.877	0.873	0.917	0.911	0.907	0.909	0.900	0.895
50	0.774	0.768	0.764	0.923	0.919	0.917	0.954	0.949	0.946	0.951	0.944	0.939
60	0.804	0.798	0.794	0.946	0.942	0.939	0.975	0.970	0.968	0.975	0.970	0.965
70	0.828	0.823	0.820	0.964	0.961	0.959	0.988	0.984	0.982	0.989	0.984	0.982
80	0.847	0.842	0.839	0.976	0.974	0.972	0.994	0.991	0.989	0.995	0.991	0.988
100	0.878	0.875	0.872	0.989	0.988	0.987	0.999	0.998	0.996	1.000	0.999	0.997
150	0.931	0.928	0.927	0.998	0.997	0.997	1.000	1.000	1.000	1.000	1.000	1.000
200	0.960	0.959	0.958	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000
250	0.974	0.974	0.973	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
300	0.984	0.983	0.982	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
350	0.990	0.989	0.989	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	0.995	0.994	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
450	0.996	0.996	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	0.998	0.998	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

De la tabla anterior se observa que:

- La prueba es consistente, ya que para cada valor α y de θ , la potencia de la prueba converge a 1 a medida que $n \rightarrow \infty$.
- Para cada α , a medida que $\theta \rightarrow -1$ la potencia se incrementa.
- Para $\theta = -0.2$, la prueba alcanza una potencia > 0.9 para $n > 100$
- Para $\theta = -0.5$, la prueba alcanza una potencia > 0.9 para $n \geq 40$
- Para $\theta = -0.8$, la prueba alcanza una potencia > 0.9 para $n \geq 40$
- Para $\theta = -0.95$, la prueba alcanza una potencia > 0.9 para $n \geq 40$
- La siguiente gráfica ilustra el comportamiento de la potencia de la prueba para un nivel de significancia nominal de $\alpha = 0.05$ y diferentes valores de n y $\theta < 0$.

Gráfica 4.16 Potencia de la prueba para $\alpha = 0.05$ y diferentes valores de $\theta < 0$.



4.4.2.2 Potencia de la prueba para modelos AR(1)

Para estudiar el comportamiento de la potencia de la prueba se tendrá en cuenta el signo del parámetro ϕ .

4.4.2.2.1 Para modelos AR(1) con ϕ positivo

La siguiente tabla contiene los resultados de la potencia de la prueba para modelos autorregresivos de orden 1 con parámetro $\phi = 0.2, 0.5, 0.8, 0.95$ y valores de $\alpha = 0.01, 0.025, 0.05$. Para cada combinación de ϕ y α se realizaron 20000 simulaciones.

Tabla 4.21 Potencia de la prueba para procesos AR(1) con $\phi > 0$.

n	$\phi = 0.2$			$\phi = 0.5$			$\phi = 0.8$			$\phi = 0.95$		
	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010
30	0.833	0.827	0.825	0.728	0.719	0.715	0.544	0.535	0.530	0.444	0.436	0.432
40	0.908	0.905	0.903	0.826	0.819	0.815	0.619	0.611	0.607	0.462	0.454	0.451
50	0.944	0.941	0.939	0.884	0.876	0.872	0.679	0.672	0.668	0.471	0.465	0.462
60	0.963	0.961	0.959	0.918	0.913	0.910	0.714	0.707	0.704	0.468	0.461	0.458
70	0.978	0.976	0.975	0.943	0.939	0.935	0.747	0.741	0.738	0.464	0.456	0.452
80	0.987	0.985	0.984	0.962	0.958	0.956	0.772	0.767	0.764	0.458	0.451	0.447
100	0.994	0.993	0.992	0.978	0.977	0.975	0.814	0.809	0.807	0.468	0.460	0.457
150	0.999	0.999	0.999	0.994	0.993	0.993	0.879	0.875	0.872	0.475	0.468	0.464
200	1.000	1.000	1.000	0.998	0.997	0.997	0.916	0.914	0.913	0.492	0.485	0.480
250	1.000	1.000	1.000	0.999	0.999	0.999	0.942	0.941	0.940	0.507	0.501	0.497
300	1.000	1.000	1.000	1.000	1.000	1.000	0.959	0.957	0.956	0.524	0.516	0.511
350	1.000	1.000	1.000	1.000	1.000	1.000	0.971	0.970	0.969	0.541	0.533	0.529
400	1.000	1.000	1.000	1.000	1.000	1.000	0.979	0.977	0.977	0.562	0.555	0.550
450	1.000	1.000	1.000	1.000	1.000	1.000	0.985	0.984	0.984	0.577	0.569	0.565
500	1.000	1.000	1.000	1.000	1.000	1.000	0.988	0.987	0.987	0.594	0.586	0.582
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.703	0.696	0.693
5000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.956	0.953	0.952

De la tabla anterior se observa que:

- La prueba es consistente, ya que para cada valor de α y de ϕ , la potencia de la prueba converge a 1 a medida que $n \rightarrow \infty$, aunque la convergencia para $\phi = 0.95$ es muy lenta.
- Para cada α , a medida que $\phi \rightarrow 1$ la potencia disminuye.
- Para $\phi = 0.2$, la prueba alcanza potencia > 0.9 para $n > 40$
- Para $\phi = 0.5$, la prueba alcanza potencia > 0.9 para $n > 60$
- Para $\phi = 0.8$, la prueba alcanza potencia > 0.9 para $n > 200$
- Para $\phi = 0.95$, la prueba alcanza potencia > 0.7 para $n > 1000$.
- La disminución de la potencia a medida que $\phi \rightarrow 1$, obedece a que en el modelo bajo H_1 dado por

$$Z_t = \beta_0 + \beta_1 t + \frac{1}{(1-\phi B)} a_t,$$

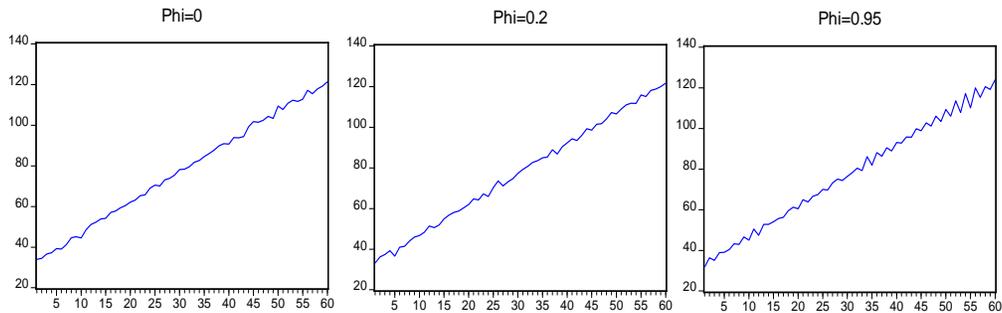
el cual puede escribirse como

$$(1-\phi B)Z_t = \gamma + \beta_1(1-\phi)t + a_t$$

donde $\gamma = (1-\phi)\beta_0 + \phi\beta_1$, cuando $\phi > 0$ y $\phi \rightarrow 1$, la componente de tendencia determinística lineal va desapareciendo y va apareciendo una tendencia aleatoria.

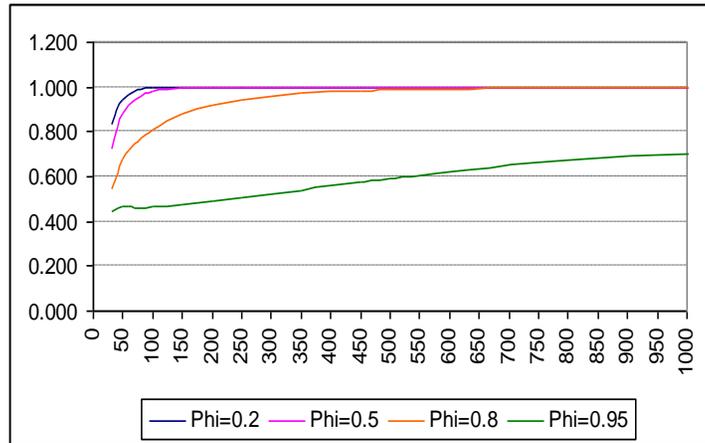
- Las siguientes gráficas ilustran el comportamiento para el modelo bajo H_1 para $\phi=0$, $\phi=0.2$ y $\phi=0.95$.

Gráfica 4.17 Modelo bajo H_1 para diferentes valores de ϕ y $n=60$



- La siguiente gráfica ilustra el comportamiento de la potencia de la prueba para un nivel de significancia nominal de $\alpha = 0.05$ y diferentes valores de n y $\phi > 0$.

Gráfica 4.18 Potencia de la prueba para $\alpha = 0.05$ y diferentes valores de $\phi > 0$.



4.4.2.2.2 Para modelos AR(1) con ϕ negativo

La siguiente tabla contiene los resultados de la potencia de la prueba para modelos autorregresivos de orden 1 con parámetro $\phi = -0.2, -0.5, -0.8, -0.95$ y valores de $\alpha = 0.01, 0.025, 0.05$. Para cada combinación de ϕ y α se realizaron 20000 simulaciones.

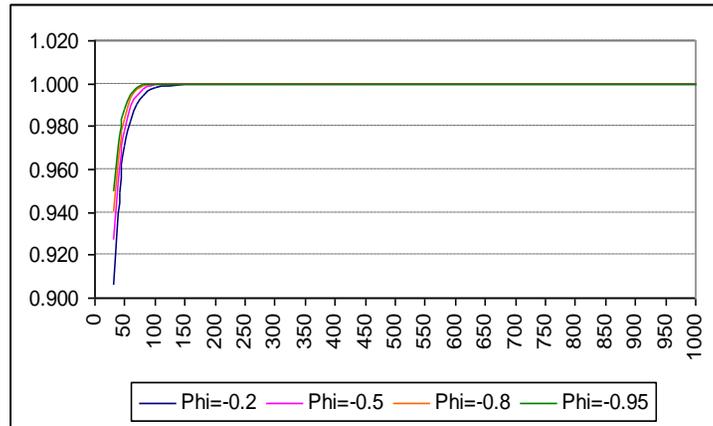
Tabla 4.22 Potencia de la prueba para procesos AR(1) con $\phi < 0$.

n	$\phi = -0.2$			$\phi = -0.5$			$\phi = -0.8$			$\phi = -0.95$		
	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010	0.05	0.025	0.010
30	0.906	0.903	0.902	0.928	0.925	0.924	0.941	0.938	0.936	0.950	0.947	0.945
40	0.949	0.947	0.946	0.962	0.960	0.959	0.970	0.967	0.966	0.976	0.974	0.972
50	0.971	0.970	0.969	0.978	0.976	0.976	0.983	0.981	0.981	0.988	0.986	0.984
60	0.983	0.982	0.981	0.990	0.988	0.987	0.994	0.992	0.991	0.995	0.994	0.993
70	0.990	0.989	0.989	0.995	0.993	0.992	0.997	0.995	0.995	0.998	0.997	0.996
80	0.995	0.993	0.993	0.997	0.996	0.996	0.999	0.998	0.997	1.000	0.999	0.998
100	0.999	0.998	0.997	1.000	0.999	0.999	1.000	1.000	0.999	1.000	1.000	1.000
150	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
300	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
350	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
450	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

De la tabla anterior se observa que:

- Para cada valor de α y de ϕ , la potencia de la prueba converge a 1 a medida que $n \rightarrow \infty$.
- Para cada α , a medida que $\phi \rightarrow -1$ la potencia aumenta.
- Para cada ϕ la prueba alcanza potencia > 0.9 para $n > 30$.
- La siguiente gráfica ilustra el comportamiento de la potencia de la prueba para un nivel de significancia nominal de $\alpha = 0.05$ y diferentes valores de n y $\phi > 0$.

Gráfica 4.19 Potencia de la prueba para $\alpha = 0.05$ y diferentes valores de $\phi > 0$.



CAPÍTULO 5

CONCLUSIONES

La principal contribución de esta tesis consiste en el empleo de la función de correlación cruzada muestral entre la serie de tiempo y los residuales de un modelo ARIMA para:

- i) Elaborar un diagnóstico gráfico que ayude a la validación de un modelo ARIMA estimado.
- ii) Elaborar un diagnóstico gráfico que permita la discriminación entre un paseo aleatorio con deriva y un proceso con tendencia determinística lineal.
- iii) Elaborar una prueba de la hipótesis nula de que el proceso generador de datos es un paseo aleatorio con deriva contra la hipótesis alternativa: a) de que el proceso no es un paseo aleatorio con deriva, b) de que el proceso tiene tendencia lineal determinística, bajo el supuesto de que el término de error del modelo es ruido blanco, y c) de que el proceso tiene tendencia lineal determinística, bajo el supuesto de que el término de error del modelo es ARMA(p,q) .

Los resultados obtenidos permiten concluir que:

- La FCCM entre los datos del proceso estacionarizado y los residuales de un modelo preliminar ajustado, contiene información relevante sobre: a) el verdadero modelo y b) la posible falla en la especificación del modelo preliminar.
- El empleo de este diagnóstico gráfico sobre la serie G del número de pasajeros de líneas aéreas de Box y Jenkins (1976), muestra que funciona adecuadamente comparándolo con el método tradicional de la función de autocorrelación muestral y la función de autocorrelación parcial muestral.

- Tomando como base el resultado anterior, se puede construir una herramienta adecuada para diferenciar si los datos que se están analizando provienen de un modelo de paseo aleatorio con deriva o de un modelo con tendencia lineal determinística, cuando el término de error del modelo es ruido blanco.
- La FCCM entre la serie estacionarizada y los residuales también puede ser empleada para la derivación de una prueba de la hipótesis $H_0 : Z_t = \beta_0 + Z_{t-1} + a_t$ contra la alternativa no informativa $H_1 : Z_t$ no es un paseo aleatorio, usando teoría tradicional. Para la prueba exacta, los resultados muestran que la prueba usando el estadístico TA tiene una alta potencia, aun en muestras pequeñas, contra las alternativas de un modelo de tendencia determinística lineal y contra un modelo de ruido blanco con cambio estructural, alcanzando potencias mayores que el contraste de Dickey y Fuller. Para la prueba asintótica, se obtuvo que la distribución límite del estadístico TA es $2\chi^2(K)$ y por tanto se puede utilizar la teoría estándar para realizar la prueba en muestras grandes. Se mostró además que la prueba puede ser llevada a cabo equivalentemente usando el estadístico Q de Box-Pierce, o su modificación Q^* de Ljung y Box. La prueba exacta puede ser llevada a cabo calculando el estadístico Q y comparándolo con la mitad del respectivo percentil de la distribución exacta de TA . Para muestras grandes, el valor de Q (o de Q^*) se compara con el percentil de la distribución $\chi^2(K)$. Para este caso, los resultados muestran que el tamaño de la prueba no presenta grandes distorsiones con respecto a su nivel nominal. La aplicación de la prueba, tanto exacta como asintótica, a datos simulados mostró un funcionamiento adecuado.

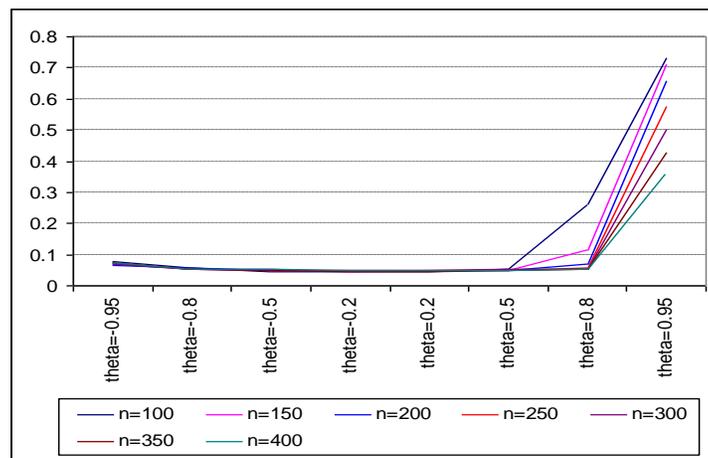
Cuando el término de error del modelo es ruido blanco normal, la FCCM entre la serie estacionarizada y los residuales también puede ser empleada para la derivación de una prueba para la hipótesis nula $H_0 : Z_t = \beta_0 + Z_{t-1} + a_t$ contra la alternativa informativa $H_1 : Z_t = \beta_0 + \beta_1 t + a_t$. En este caso, se trata de buscar evidencia sobre

el tipo de tendencia presente en la serie de tiempo. La especificación de dicha componente es de mucho interés en áreas como la economía, pues como se mencionó antes, la naturaleza temporal o permanente de los ‘shocks’ macroeconómicos sobre la serie depende del tipo de tendencia que posea. Los resultados obtenidos muestran que: i) En muestras pequeñas, el tamaño de la prueba está muy cerca del nivel de significancia nominal y que converge rápidamente a dicho nivel a medida que $n \rightarrow \infty$; ii) La prueba tiene una gran potencia (mayor que 0.9) aun en muestras pequeñas y que rápidamente alcanza la máxima potencia de 1; iii) La distribución asintótica del estadístico de la prueba es $2\chi^2(K)$; iv) El tamaño de la prueba asintótica no tiene grandes distorsiones en muestras pequeñas y converge rápidamente a su nivel nominal; v) La prueba tiene una gran potencia aun en muestras pequeñas (mayor que 0.9) y que rápidamente alcanza la máxima potencia de 1; vi) La prueba puede ser llevada a cabo equivalentemente en términos del estadístico Q de Box y Pierce, o del estadístico Q* de Ljung y Box.

- Finalmente, cuando el término de error del modelo está autocorrelacionado, se generalizó la prueba asintótica con el fin de adaptarla a esta situación más general. Se mostró que cuando el término de error del modelo no es ruido blanco, el tamaño de la prueba tiende a crecer a medida que n aumenta y la prueba tiende a rechazar H_0 siendo cierta con alta probabilidad. Para este caso, el estadístico de la prueba puede estar basado en el estadístico Q o Q* y se mostró que su distribución asintótica es $\chi^2(K - p - q)$, donde p y q son, respectivamente, los órdenes de los polinomios autorregresivo y de medias móviles del modelo ARMA asociado al término de error. Para explorar las propiedades de la prueba se simuló diferentes modelos AR(1) y MA(1) para K=5. Para estos casos los resultados muestran que: i) El comportamiento del tamaño de la prueba asintótica está relacionado con el tipo del modelo; para las situaciones simuladas se encontró que cuando el modelo es un MA(1), el tamaño de la prueba, en general, no sufre fuertes distorsiones cuando el parámetro θ es negativo;

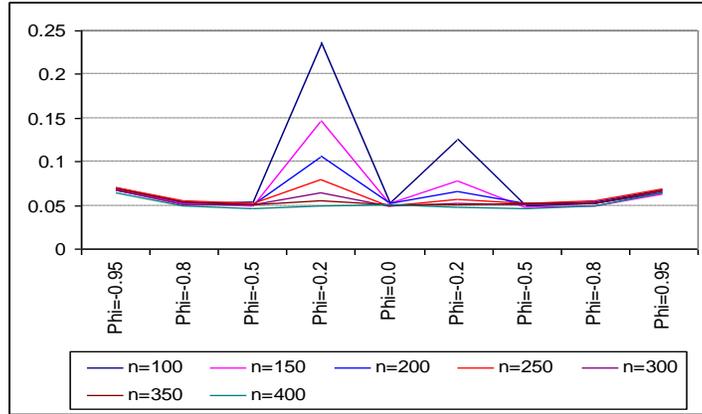
para valores positivos de θ , a medida que $\theta \rightarrow 1$, las distorsiones son más severas y la probabilidad de rechazar H_0 , siendo cierta, es cada vez mayor; en todos los casos, excepto en las situaciones de casi no invertibilidad, se presenta convergencia rápida a los niveles de significancia nominales. La siguiente gráfica resume el comportamiento del tamaño de la prueba para el MA(1) y para algunos valores de n .

Gráfica 5.1 Tamaño de la prueba para modelos MA(1)



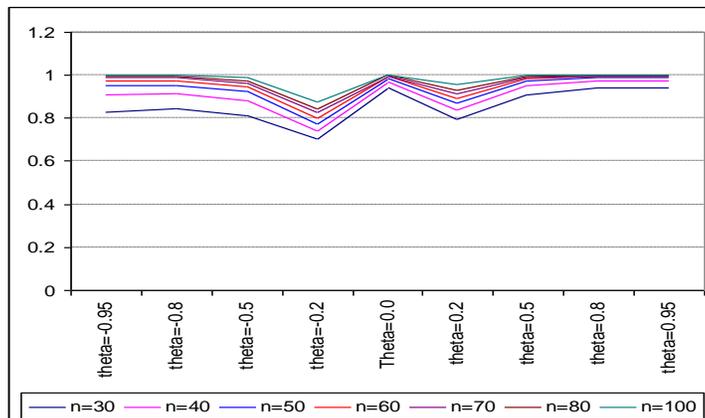
- Cuando el modelo es AR(1), para valores pequeños de n y magnitudes pequeñas de ϕ , hay distorsiones grandes en el tamaño de la prueba; en todos los casos, excepto en las situaciones de casi no estacionaridad, se presenta convergencia relativamente rápida a los niveles de significancia nominales. El siguiente gráfico resume el comportamiento del tamaño de la prueba para el AR(1) y para algunos valores de n .

Gráfica 5.2 Tamaño de la prueba para modelos AR(1)



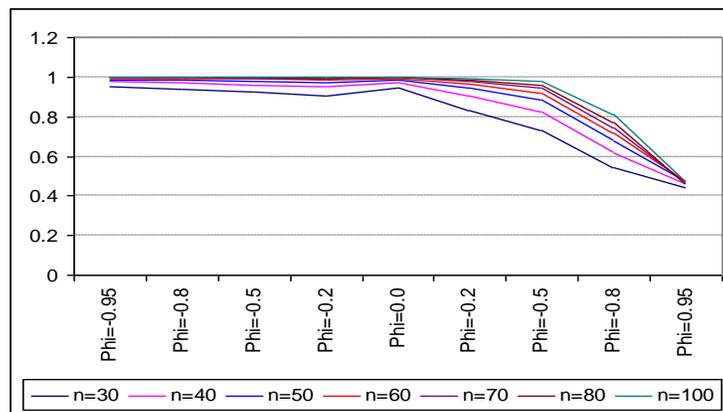
ii) El comportamiento de la potencia de la prueba también está relacionada con el tipo de modelo: para modelos MA(1), en todos los casos donde $-1 < \theta < 1$ la potencia tiende a 1 a medida que n crece; para valores positivos de θ , a medida que $\theta \rightarrow 1$, la potencia aumenta; para valores negativos de θ , a medida que $\theta \rightarrow -1$, la potencia aumenta. El siguiente gráfico resume el comportamiento de la potencia de la prueba para los modelos MA(1) estudiados y algunos valores de n .

Gráfica 5.3 Potencia de la prueba para modelos MA(1)



Para modelos AR(1), en todos los casos donde $-1 < \phi < 1$, la potencia tiende a 1 a medida que n crece, aunque la convergencia es muy lenta para valores altos de ϕ (casi no estacionaridad); para valores positivos de ϕ , a medida que $\phi \rightarrow 1$, la potencia disminuye; para valores negativos de ϕ , a medida que $\phi \rightarrow -1$, la potencia aumenta. El siguiente gráfico resume el comportamiento de la potencia de la prueba para los modelos AR(1) estudiados y algunos valores de n .

Gráfica 5.4 Potencia de la prueba para modelos AR(1)



BIBLIOGRAFÍA

- AKAIKE, H. (1974). A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, AC-19. 716-723.
- ARELLANO, C. and S.G. Pantula (1990). Trend Stationarity versus Difference Stationarity, in: Proceedings of the Business and Economic Statistics Section (American Statistical Association), 188-196.
- BARTLETT, M.S. (1955) Stochastic Processes. Cambridge University Press, Cambridge.
- BEGUIN, J.M.; Gouriéroux C. and Monfort A. (1980). Identification of a Mixed Autoregressive-Moving Average Process: The Corner Method, Time Series (Ed. O.D. Anderson), 423-436, North-Holland, Amsterdam.
- BOX, G.E.P., and Pierce, D.A. (1970). Distribution of the residual autocorrelations in autoregressive-integrated moving average time series models. *JASA*, 65, 1509-1526.
- BOX, G.E.P., and Jenkins, G.M. (1976). *Time Series Analysis Forecasting and Control*, Holden-Day, San Francisco.
- BROCKWELL, P.J. and Davis, R.A. (1987) Time Series: Theory and Methods, Springer-Verlag, New York.
- CASTAÑO, E. y Rosales L.F. (2005). Un diagnóstico adicional en la selección de modelos ARIMA usando la función de correlación cruzada, por publicar.
- CLEVELAND, W.S. (1993) Visualizing Data, Hobart Press.
- CLEVELAND, W.S. (1994) The Elements of Graphing Data, Hobart Press.
- CHOI, I. (1992a). Durbin-Hausman Tests for a Unit Roots, *Oxford Bulletin of Economics and Statistics*, 54, 289-304.
- CHOI, I., and B.C. Yu (1997). A general Framework for testing $I(m)$ contra $I(m+k)$, *Journal of Economic Theory and Econometrics*, 3, 103-138.
- COCHRANE, J.H. (1991). A Critique of the Application of the Unit Root Tests", *Journal of Economic Dynamics and Control*, 15, 275-284.

- DeJONG, D.N., J.C. Nankervis, N.E. Savin, and C.H. Whiteman (1992a). The Powers Problems of the Unit Root Tests in Time Series with Autoregressive Errors, *Journal of Econometrics*, 53, 323-343.
- DeJONG, D.N., J.C. Nankervis, N.E. Savin, and C.H. Whiteman (1992b). Integration Versus Trend Stationarity in Time Series, *Econometrica* , 60, 423-433.
- DICKEY, D.A., and W.A. Fuller (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association*, 76, pp. 427-31.
- FLORES de Frutos, R. and Jerez, M. (2002). Testing for Invertibility in Univariate Arima Process, Departamento de Economía Cuantitativa. Facultad de Ciencias Económicas y Empresariales. Universidad Complutense de Madrid.
- HÄRDLE, W. (1991) Smoothing Techniques with Implementation in S, Springer_Verlag.
- HALL, A. (1989). Testing for a Unit Root in the Presence of Moving Average Errors, *Biometrika*, 76, 49-56.
- HARVEY, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.
- HARVEY, A.C. (2000). Trend Analysis, Mimeo, Faculty of Economics and Politics, University of Cambridge.
- HUDAK, G.B. and Liu, L. (1992). Reference Manual for Forecasting and Time series Analysis. Scientific Computing Associates. Illinois.
- JONES, R.H.(1993). *Longitudinal Data with Serial Correlation: A State Space Approach*. London: Chapman and Hall.
- KITAGAWA, G., and W Gersch (1996). *Smoothness Priors Analysis of Time Series*. Berlin: Springer-Verlag.
- LJUNG, G.M. and Box, G.E.P. (1978) On a measure of lack of fit in time series models, *Biometrika*, 65, 297-303.
- LJUNG, G.M. (1986) Diagnostic testing of univariate time series models. *Biometrika*, 73, 725-730.

- NELSON, C.R., and C.I. Plosser (1982). Trends and Random Walks in Macroeconomic Time Series, *Journal of Monetary Economics*, 10, 139-162.
- PERRON, P. (1989). The Great Cash, the Oil Price Stock and the Unit Root Hypothesis, *Econometrica*, 57, 1361-1401.
- PHILLIPS, P.C.B., and P. Perron (1988). Testing for a Unit Root in Time Series Regression, *Biometrika*, 75, 335-346.
- ROSALES, L.F. (2004). La función de Correlación Cruzada como Elemento de Diagnóstico para los Modelos ARMA(p,q). Tesis de pregrado en Estadística, Universidad Nacional de Colombia, Sede Medellín.
- SAID, S. and Dickey, D. (1984) Testing Unit Roots in Autoregressive Moving Average Models with Unknown Order. *Biometrika*, 71, 599-607.
- SAIKKONEN, P. and R. Luukkonen (1993). Testing for a Moving Average Unit Root in Autoregressive Integrated Moving Average Models, *Journal of the American Statistical Association*, 88, 422, 596- 601.
- SARGAN, J.D., and A. Bhargava (1983). Maximum Likelihood Estimation of Regression Models with First Order Moving Average Errors When the Root Lies on the Unit Circle, *Econometrica*, 51, 799-820.
- SCA Corp. (2001) The SCA Statistical System, Version VI.3a, Dekalb, Illinois.
- SCHWARTZ, G. (1978). Estimating the Dimension of a Model, *Ann. Statist*, 6, 461-464.
- SCHWERT, G.W. (1989). Tests for Unit Roots: A Monte Carlo Investigation, *Journal of Business and Economics Statistics*, 7, 147-159.
- SILVERMAN, B.W. (1986). Density Estimation for Statistics and Data Analysis, Chapman & Hall.
- SHEPHARD, N. (1993). Distribution of the ML Estimator of an MA(1) Model and a Local level Model, *Econometric Theory*, 9, 377-401.
- SHEPHARD, N. and Harvey, A.C. (1990). On the Probability of Estimating a Deterministic Component in the local Level Model. *Journal of Time Series Analysis*, 11, 339-247.
- TANAKA, K. (1990). Testing for a Moving Average Unit Root, *Econometric Theory*, 6, 433-444.

- TSAY, R.S. (1993). Testing for Noninvertible Models With Applications, *Journal of Business and Economic Statistics* 11, 2, 225-233.
- TSAY, R.S. and Tiao, G.C. (1984). Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Non-stationary ARMA Models, *Journal of the American Statistical Association*, 79, 84-96.
- TSAY, R.S. and Tiao, G.C. (1985). Use of Canonical Analysis in Time Series Model Identification, *Biometrika*, 72, 299-315.
- WEI, W.W.S. (1990). Time Series Analysis, Univariate and Multivariate Methods. Addison-Wesley. Redwood, California.
- WEST, M. and J. Harrison, (1989). *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York.
- YOUNG, P. (1984). *Recursive Estimation and Time-Series Analysis*. Berlin:Springer-Verlag.