



EDITORIAL DE LA
UNIVERSIDAD
TECNOLÓGICA
NACIONAL (UTN)



Aplicación de los Modelos Autorregresivos Integrados de Media Móvil (ARIMA) a las Series de Precipitaciones de Lluvia

Guillermo Daniel Scheidereiter
Omar Roberto Faure

*Tesis de graduación de Guillermo Daniel Scheidereiter
para la carrera Licenciatura en Matemática Aplicada
de la Universidad Nacional de La Matanza.*

Director de Tesis: Dr. Omar Roberto Faure.

©[Copyright]

edUTecNe, la Editorial de la U.T.N., recuerda que las obras publicadas en su sitio web son de libre acceso para fines académicos y como un medio de difundir la producción cultural y el conocimiento generados por autores universitarios o auspiciados por las universidades, pero que estos y edUTecNe se reservan el derecho de autoría a todos los fines que correspondan.

*A mi mamá, Geneveva.
(En memoria)*

*Aplicación de los Modelos Autorregresivos
Integrados de Media Móvil (ARIMA) a las Series de
Precipitaciones de Lluvia*

“No puede haber sino borradores. El concepto de texto definitivo no corresponde sino a la religión o al cansancio.”

(Jorge Luis Borges, Prefacio al cementerio marino de Valéry)

“Creía en infinitas series de tiempos, en la red creciente y vertiginosa de tiempos divergentes, convergentes, paralelos.”

(Jorge Luis Borges, El jardín de senderos que se bifurcan)

“Tengo la satisfacción de dedicar este espacio a las personas a quienes debo mi agradecimiento. En primer lugar, al Dr. Omar Faure por su profesionalidad, compromiso y generosidad para guiarme en este trabajo. Al Dr. Ignacio Díaz-Emparanza por su ayuda desinteresada con el uso de Gretl. A la Lic. Graciela Gay, por su amistad, inspiración y apoyo incondicional a lo largo de toda mi formación académica. Y muy especialmente a mi mamá, Genoveva, y a mis queridos hermanos, sin ellos no sería posible. A todos, muchísimas gracias.”

D. S.

Contenido

Capítulo 1	8
Introducción.....	8
Capítulo 2	12
Marco Metodológico	12
Capítulo 3	15
3.1. Procesos Estocásticos y Series Temporales	15
3.2. Función de Autocorrelación y Función de Autocorrelación Parcial.....	22
3.3. Autocorrelación simple y parcial muestrales	26
Capítulo 4	34
4.1. Modelos ARMA.....	34
4.2. Funciones de autocovarianza, autocorrelación y autocorrelación parcial de los procesos ARMA.....	49
Capítulo 5	61
5.1. No estacionariedad	61
5.2. Modelos ARIMA.....	69
5.3. Modelos ARIMA Estacionales	72
Capítulo 6	86
6.1. Disponibilidad de los datos y representación gráfica de las series de registros de lluvia acumulada mensualmente.....	86
6.2. Comportamiento de las series de registros de lluvia de acuerdo con las estaciones del año	87
6.3. Análisis de estacionariedad.....	88
6.3.1. Transformación de Box-Cox	89
6.3.2. Test HEGY.....	89
6.3.3. Test de Canova y Hansen (CH)	95
6.4. Identificación del tipo de modelo y órdenes asociados p , q , P y Q	100
Capítulo 7	103
Estimación y Validación.....	103
7.1. Estimación por Método de Momentos.....	103
7.2. Estimación por Método de Máxima Verosimilitud	106
7.3. Contrastes de Significación de los Parámetros Estimados.....	115
7.4. Análisis de Residuos	121

7.5. Criterios de Información.....	128
7.6. Modelos propuestos para las series de registros de lluvia mensual acumulada de Entre Ríos.....	138
Capítulo 8	144
Conclusiones	144
Capítulo 9	148
Apéndice I: Lucas Sur	148
Apéndice II: Octavo Distrito	151
Apéndice III: Puente de Hierro	158
Apéndice IV: San Víctor	163
Apéndice V: Santa Anita	168
Apéndice VI: Séptimo Distrito	175
Apéndice VII: Viale	180
Apéndice VIII: Villa Paranacito	184
Apéndice IX: Colón	190
Apéndice X: La Lila	196
Apéndice XI: Lucas González	203
Apéndice XII: Santa María del Tatutí	210
Apéndice XIII: Villa Elisa	214
Apéndice XIV: Antelo	217
Apéndice XV: Febre	221
Apéndice XVI: Feliciano	225
Apéndice XVII: San Salvador	230
Apéndice XVIII: La Paz	236
Apéndice XIX: San Gustavo	244
Apéndice XX: Paraná	249
Apéndice XXI: Series de precipitaciones de lluvia con evidencia de raíces unitarias en alguna frecuencia. Resultados de los contrastes HEGY y Canova-Hansen	254
1. Octavo Distrito.....	254
2. San Víctor.....	254
3. Santa Anita.....	255
4. Séptimo Distrito.....	255
5. Viale.....	256
6. Villa Paranacito.....	256
7. Feliciano.....	257

8. San Jaime	257
Apéndice XXII: Transformaciones de Doornik-Hansen	258
Bibliografía	259

Capítulo 1

Introducción

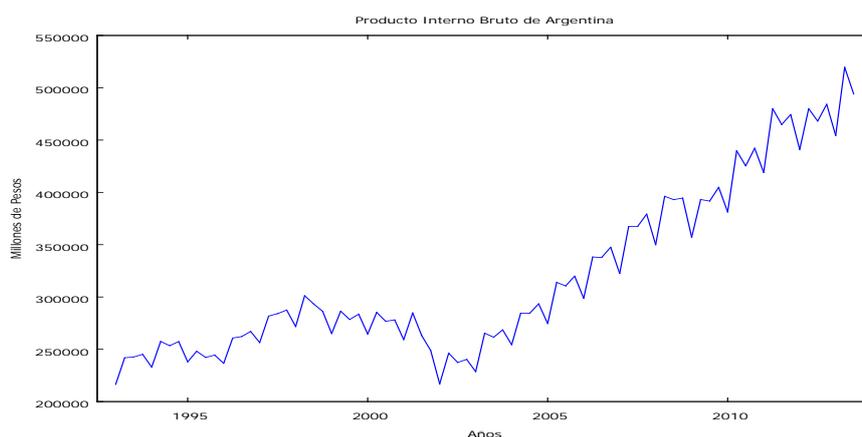
El estudio experimental de datos ordenados cronológicamente constituye uno de los problemas fundamentales del análisis estadístico y está presente en diversas áreas del conocimiento como economía, física, biología, epidemiología, entre otras. El enfoque sistemático mediante el cual se estudia los datos espaciados regularmente en el tiempo se denomina en estadística análisis de series de tiempo.

Una serie de tiempo es una sucesión de observaciones tomadas secuencialmente en el tiempo (Box, Jenkins, & Reinsel, 1994).

En economía, por ejemplo, los precios diarios de las acciones de una empresa, el valor mensual de las exportaciones de un determinado artículo, el producto bruto interno trimestral de un país, los beneficios netos mensuales de cierta entidad bancaria, las tasas anuales de empleo y desempleo, constituyen ejemplos clásicos de series temporales.

En el Gráfico 1.1 se observa la evolución temporal de la estimación trimestral del PIB en Argentina desde el primer trimestre de 1993 hasta el cuarto trimestre de 2013 (Fuente: INDEC). La representación gráfica de la serie permite apreciar las características primarias del fenómeno. Nótese que la serie no mantiene un nivel constante a lo largo del tiempo; entre 2002 y 2005 comienza a observarse un patrón de crecimiento sostenido. Este movimiento de la serie se conoce como *tendencia*. Por el contrario, cuando la serie oscila alrededor de un valor constante con una variabilidad estable se dice que es *estacionaria*.

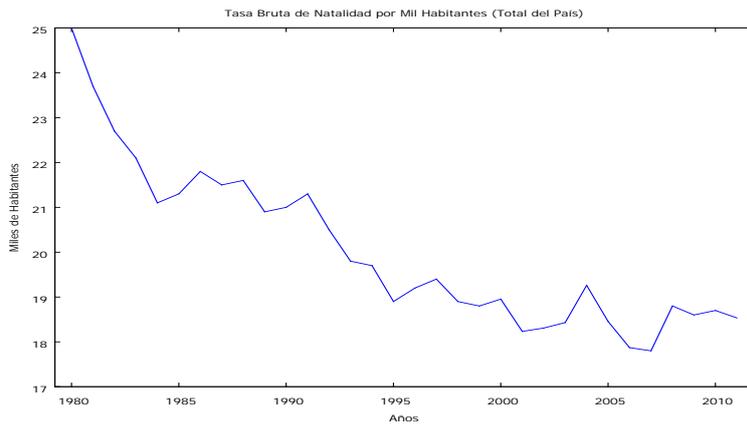
Gráfico 1.1



También en agronomía se emplea el análisis de series temporales, por ejemplo, en el pronóstico de la producción de cañas de azúcar (Ruiz-Ramírez, Hernández-Rodríguez, & Zulueta-Rodríguez, 2011) y su valor monetario por tonelada, o en el estudio de las variaciones de precios de venta del ganado en pie o los cambios anuales en la cotización de la soja, entre otros.

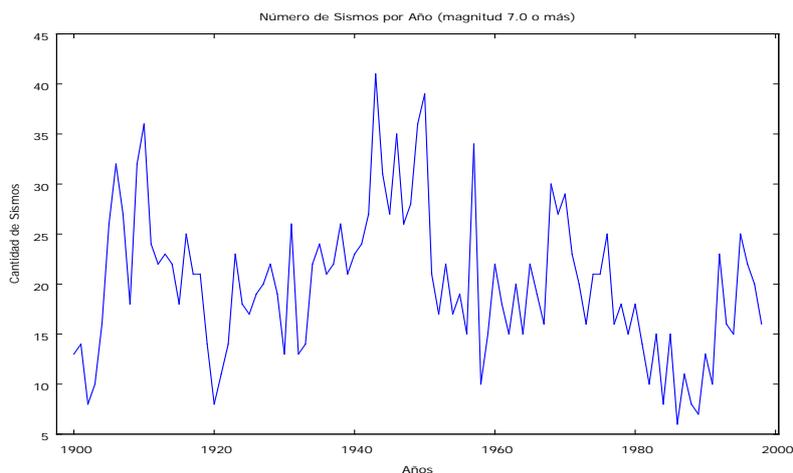
Por otro lado, un epidemiólogo podría estar interesado en el número de casos de influenza observada durante un determinado período de tiempo. En medicina, las mediciones de presión arterial trazadas en el tiempo podrían ser de utilidad para evaluar los fármacos utilizados en el tratamiento de la hipertensión (Shumway & Stoffer, 2011). En el siguiente gráfico se representa la evolución de la tasa bruta anual de natalidad en miles de habitantes en Argentina desde 1980 hasta 2011 (Fuente de datos: INDEC). Este es un clásico ejemplo de serie temporal cuyo estudio es de interés tanto en medicina como en demografía.

Gráfico 1.2



Las series de tiempo también son comunes en estudios ambientales. Recientes investigaciones están centradas en los cambios de temperatura producto del calentamiento global o en cómo los niveles de contaminación ambiental pueden afectar el número de muertes en personas mayores (Saez, y otros, 1999). También, son importantes las series de tiempo en Geofísica en la observación de los movimientos sísmicos (Gráfico 1.3), entre muchas otras aplicaciones que podrían citarse.

Gráfico 1.3



En este trabajo se presentan los aspectos esenciales de la teoría estadística de las series de tiempo, a los efectos de analizar los datos históricos correspondientes a la cantidad acumulada mensual de lluvia caída en distintos puntos de registro pluviométrico en la Provincia de Entre Ríos, Argentina.

Las lluvias, al igual que otros componentes del clima, se manifiestan de formas diferentes conforme a las estaciones del año. Por ejemplo, en primavera las precipitaciones son más frecuentes que en invierno. Un fenómeno similar se puede observar en relación con la temperatura. En general, todos los veranos presentan temperaturas superiores a cualquier otra estación del año. La característica según la cual una serie temporal repite un comportamiento similar con cierta periodicidad en el tiempo se conoce como *estacionalidad*. Esta conducta no es exclusiva de los fenómenos climáticos, también se encuentra presente en economía, agricultura, ecología, etc.

El estudio matemático de las series temporales permite describir sus características (como tendencia, estacionariedad, estacionalidad, entre otras), a través de modelos estadísticos. La teoría que se empleará para el estudio de las series de precipitaciones de lluvia es la metodología de modelización ARIMA (*Autoregressive Integrated Moving Average*). Esta fue propuesta en 1970 por George Box y Gwilym Jenkins en la obra *Time Series Analysis: Forecasting and Control* y constituye uno de los avances más notables en el estudio de las series temporales.

En base a esta metodología se construirán modelos matemáticos que describan el comportamiento de las series que conforman el campo muestral de estudio. El tratamiento de los temas necesarios para tal fin se ha dividido en capítulos. En el Capítulo 2 se exponen los lineamientos metodológicos a seguir. Se exhibe los objetivos, el problema y las hipótesis de investigación, como también la selección de los datos que conforman el campo de estudio, su organización y los pasos a seguir en la aplicación de la metodología ARIMA.

En los capítulos 3, 4 y 5 se presenta la teoría elemental de las series temporales en el marco de los procesos estocásticos. Puntualmente, en el capítulo 3 se da la definición de proceso estocástico y de serie temporal, la noción de modelo y las características primarias de un proceso, como la estacionariedad y las funciones de autocovarianza, autocorrelación simple y autocorrelación parcial. También se aborda una breve descripción de los momentos muestrales asociados a una serie de tiempo, y en consecuencia las funciones de autocovarianza muestral, autocorrelación simple muestral y autocorrelación parcial muestral.

En el capítulo 4 se inicia el estudio de los procesos autorregresivos y de media móvil. En principio, se comienza analizando los procesos autorregresivos llamados, abreviadamente, modelos *AR*, se introduce el operador de retardos y se define el polinomio autorregresivo. Luego, se presenta el modelo de media móvil o modelo *MA* y su polinomio asociado. Posteriormente, ambos modelos se integran para conformar los procesos autorregresivos y de media móvil llamados modelos *ARMA*, se analiza las condiciones de causalidad e invertibilidad de este tipo de proceso y se concluye el capítulo con una breve descripción de las funciones de autocorrelación simple y parcial que caracterizan a los mismos. En el transcurso del capítulo, al igual que en los capítulos siguientes, se muestran diversos ejemplos y simulaciones generadas con el software R. En la generación de gráficos a partir de datos concretos se trabaja con el software econométrico Gretl.

En el capítulo 5 se da inicio al estudio de los procesos no estacionarios. Se introducen distintas transformaciones para lograr la estacionariedad ya sea en media o en varianza, se estudia la tendencia de una serie y se define los procesos *ARIMA*, procesos autorregresivos integrados

de media móvil. Además, se explora las características de las funciones de autocorrelación simple y parcial para los procesos con tendencia. Luego, se analizan los modelos *ARIMA* estacionales puros y los multiplicativos. Estos modelos son de particular importancia en este trabajo ya que permiten el estudio de procesos que presentan cierta periodicidad en el tiempo, como es el caso de los registros de precipitaciones de lluvia que mantienen año a año similitudes que se repiten estacionalmente.

En el capítulo 6 se da inicio al trabajo de identificación de los modelos para las series de registros de lluvia. Se pone de manifiesto la estacionalidad de las series y se desarrollan los contrastes HEGY y Canova-Hansen para detectar la presencia de raíces unitarias. El método de momentos y el de máxima verosimilitud para la estimación de los parámetros de los modelos se desarrollan en el capítulo 7. Aquí también se discute los contrastes de significatividad de los parámetros y las técnicas de validación, basadas principalmente en los ajustes residuales. La ejemplificación de estos procedimientos se realiza directamente sobre las series objeto de estudio.

Por último, las conclusiones se exponen en el capítulo 8. Los gráficos, tablas y resultados de las estimaciones y validaciones de los modelos propuestos se consignan en el Capítulo 9, en veinte apéndices que llevan el título de cada estación de registro. Además, en este capítulo hay un apéndice dedicado a las series en las que se detectó evidencia de raíces unitarias en alguna frecuencia.

Capítulo 2

Marco Metodológico

El presente trabajo se trata de una investigación estadística en el marco de la teoría de los procesos estocásticos, donde las unidades de análisis son series de registros de precipitaciones históricas de lluvia de la provincia de Entre Ríos. La variable de estudio es la cantidad acumulada mensual de lluvia registrada en milímetros.

Estas series se estudiarán a partir de la metodología de modelización *ARIMA*. El objetivo principal es elaborar un modelo *ARIMA* para cada serie de registro histórico de precipitación pluviométrica, que describa adecuadamente su comportamiento, de manera que las implicaciones teóricas de los modelos se ajusten y sean compatibles con los datos muestrales.

Los modelos *ARIMA* están conformados por un conjunto de parámetros reunidos en una ecuación que permite expresar una observación de una serie de tiempo como una función lineal de observaciones anteriores, a lo que se agrega una serie de carácter puramente aleatorio. Estos modelos pueden contener, además, una componente tendencial y una componente estacional.

A partir de esta metodología se trata de dar respuesta a la siguiente pregunta: ¿Cuáles son los modelos *ARIMA* que explican los procesos estocásticos generadores de las series de registros históricos de lluvia en la Provincia de Entre Ríos?

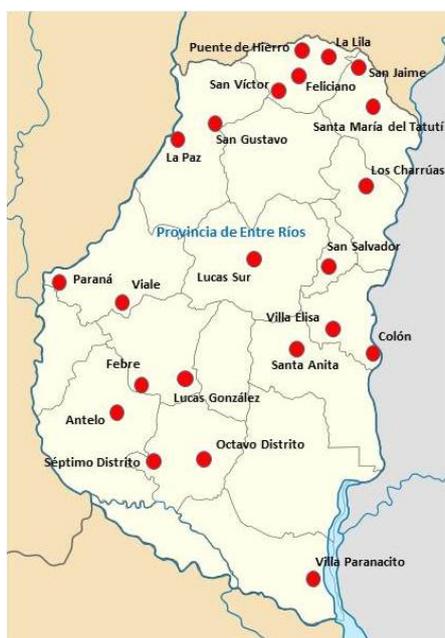
Dar solución a este problema, implica estudiar y desarrollar la teoría de series temporales en el marco de los procesos estocásticos, en particular los modelos univariantes, y analizar los aspectos esenciales de los modelos *ARIMA*. Además, en este contexto se espera aportar información de interés metodológico para el estudio de las series de precipitaciones históricas de lluvia, que pueda servir a futuras investigaciones.

La selección de los datos constituye el primer paso en cualquier investigación estadística. En este caso, se seleccionó un conjunto de estaciones de registro de precipitación mensual de lluvia cuyos datos se encuentran publicados en la página oficial de la Dirección de Hidráulica de Entre Ríos (<http://www.hidraulica.gob.ar/>). La selección de las series se realizó de la siguiente manera: en principio, se tomaron aquellas series con períodos completos de registros con continuidad mes a mes a través de los años. Si la serie de datos presentó al menos un año sin observaciones, se tomó desde el primer año de registro a partir del cual la continuidad de datos mes a mes se prolonga hasta años recientes. Además, no se incluyeron estaciones con registros inferiores a los 15 años. Resultaron 22 series que se agruparon en intervalos de tiempo medidos en años. En la siguiente tabla se muestran las series seleccionadas, reunidas en intervalos anuales; se indica a la derecha el departamento geográfico al que corresponde la estación, el período de inicio y fin de la serie y la cantidad de observaciones que la conforman. El mapa inferior muestra la distribución geográfica de las diferentes estaciones pluviométricas:

Tabla 2.1: Listado de las estaciones de registro pluviométrico seleccionadas y su agrupamiento de acuerdo a su longitud.

<i>Intervalo anual</i>	Estación de Registro	Departamento	Inicio - Fin	Longitud
<i>15 a 25 años</i>	Los Charrúas	Concordia	Desde agosto de 1993 hasta julio de 2012	n=228
	Lucas Sur	Villaguay	Desde agosto de 1993 hasta mayo de 2012	n=226
	Octavo Distrito	Guaaleguay	Desde agosto de 1993 hasta marzo de 2012	n=224
	Puente de Hierro	Feliciano	Desde Septiembre de 1993 hasta junio de 2012	n=226
	San Víctor	Feliciano	Desde Octubre de 1993 hasta julio de 2012	n=226
	Santa Anita	Uruguay	Desde enero de 1994 hasta julio de 2012	n=223
	Séptimo Distrito	Guaaleguay	Desde agosto de 1993 hasta julio de 2012	n=228
	Viale	Paraná	Desde Septiembre de 1995 hasta abril de 2012	n=200
	Villa Paranacito	Islas del Ibicuy	Desde enero de 1993 hasta febrero de 2012	n=230
<i>26 a 35 años</i>	Colón	Colón	Desde noviembre de 1981 hasta mayo de 2011	n=355
	La Lila	Feliciano	Desde mayo de 1980 hasta abril de 2012	n=384
	Lucas González	Nogoyá	Desde enero de 1982 hasta julio de 2012	n=367
	Santa María del Tatutí	Federación	Desde diciembre de 1978 hasta julio de 2012	n=404
	Villa Elisa	Colón	Desde junio de 1993 hasta mayo de 2012	n=348
<i>36 a 45 años</i>	Antelo	Victoria	Desde enero de 1970 hasta junio de 2012	n=510
	Febre	Nogoyá	Desde enero de 1970 hasta abril de 2012	n=508
	Feliciano	Feliciano	Desde enero de 1968 hasta marzo de 2012	n=531
	San Jaime	Federación	Desde enero de 1968 hasta julio de 2012	n=535
	San Salvador	San Salvador	Desde enero de 1976 hasta junio de 2012	n=438
<i>46 a 55 años</i>	La Paz	La Paz	Desde enero de 1958 hasta diciembre de 2011	n=648
	San Gustavo	La Paz	Desde enero de 1958 hasta diciembre de 2011	n=648
	Paraná	Paraná	Desde enero de 1953 hasta diciembre de 2008	n=662

Gráfico 2.1: Distribución geográfica de las estaciones de registro seleccionadas.



Mediante el análisis de los datos obtenidos se trabajarán las siguientes etapas metodológicas para la construcción del modelo:

- *Identificación del modelo:* Este paso consiste en utilizar los datos recogidos, y cualquier información de cómo se genera la serie temporal objeto de estudio, para sugerir un conjunto reducido de posibles modelos, que tengan muchas posibilidades de ajustarse a los datos (Pérez, 2001).
- *Estimación:* Una vez seleccionados los posibles modelos candidatos a explicar las series, se procede a estimar los valores numéricos de los parámetros del modelo.
- *Validación:* En esta fase se comprueba el ajuste del modelo a los datos muestrales.
- *Reformulación:* En caso de que el modelo no supere las pruebas de ajuste deberá realizarse una nueva formulación del mismo.

Un análisis exploratorio de los datos permite plantear las siguientes hipótesis de investigación:

Hipótesis I: Los modelos que explican las series de precipitaciones acumuladas mensuales de lluvia de la Provincia de Entre Ríos son modelos ARIMA estacionales, con periodicidad estacional de 12 meses.

Hipótesis II: Las series de precipitaciones de lluvia son estacionarias, es decir, no presentan cambios estructurales que señalen un comportamiento tendencial.

A lo largo del desarrollo del trabajo se intentará corroborar o refutar estas hipótesis.

Capítulo 3

La definición de proceso estocástico y la caracterización de sus propiedades primarias resultan fundamentales para dar marco al estudio de las series temporales. En este capítulo se introduce el concepto de proceso estocástico y las nociones de estacionariedad, tanto desde la estructura probabilística como a partir de los momentos de primer y segundo orden. Se definen además, tres funciones que serán claves en la identificación de modelos: la función de autocovarianza, la función de autocorrelación simple y la función de autocorrelación parcial. Estas funciones se analizan luego desde sus definiciones muestrales.

3.1. Procesos Estocásticos y Series Temporales

Las series de tiempo son una parte (o muestra) de un proceso estocástico. El objetivo principal del análisis de series temporales es desarrollar modelos matemáticos que provean una descripción plausible de los datos muestrales que conforman la serie. Para llegar a tal objetivo es de carácter primario y formal familiarizarse con los conceptos fundamentales.

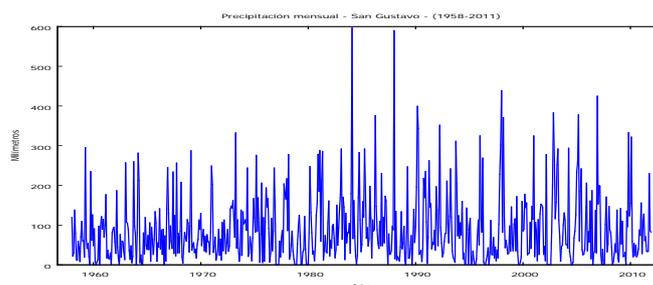
Definición 3.1.1: Una serie de tiempo es una sucesión de n observaciones x_t tomadas secuencialmente en el tiempo.

Una serie temporal es univariante cuando recoge la información de una sola variable en el tiempo. La serie solo se analiza a partir de la información que aporta su propio pasado. Por otro lado, en muchos estudios empíricos, los datos que conforman una serie temporal constan de observaciones de varias variables. En este caso la serie es multivariante o vectorial.

De acuerdo con la definición 3.1.1, una serie de tiempo es una secuencia de valores x_1, x_2, x_3, \dots , donde x_1 denota el primer valor tomado por la serie, x_2 designa el valor para el segundo período de tiempo, x_3 el valor para el tercer período temporal y así sucesivamente. A lo largo de este trabajo se representará de modo general una serie de tiempo univariante mediante la expresión $\{x_t\}$, donde t toma valores naturales entre 1 y n ($t = 1, 2, 3, \dots, n$). Las series multivariantes no serán objeto de estudio en esta investigación.

Ejemplo 3.1.1: A continuación se muestra la representación gráfica de la serie de observaciones de registros de lluvia caída mensualmente, registrada en la Localidad de San Gustavo, Departamento La Paz, Provincia de Entre Ríos, desde enero de 1958 hasta diciembre de 2011.

Gráfico 3.1



Las series temporales pueden ser discretas o continuas. Se dice que una serie temporal es *discreta* si las observaciones muestrales se recogen a intervalos iguales, en momentos determinados de tiempo. Por otro lado, si los datos se generan de forma continua y se observan de la misma manera, la serie temporal es *continua*. En este caso, la serie se denota $\{x_t\}$ con $t \in \mathbb{R}$.

En el análisis de series temporales se busca describir las características evolutivas de la serie, como por ejemplo, los cambios estructurales marcados por períodos de crecimiento o decrecimiento continuado (tendencia), o los patrones que se repiten de acuerdo a un ciclo estacional. Muchas investigaciones, sin embargo, tienden a un objetivo más ambicioso que es el de obtener previsiones futuras basadas en inferencias probabilísticas. La predicción es posible a partir de la confección de un modelo matemático que se formula en términos de los valores pasados de la serie.

Se mencionó anteriormente que el análisis de las series de tiempo se circunscribe al estudio de los procesos estocásticos. Un modelo que describe la estructura de probabilidad de una secuencia de observaciones es llamado un proceso estocástico (Box, Jenkins, & Reinsel, 1994). Para llegar a una definición rigurosa es preciso establecer de forma breve los conceptos elementales de probabilidad necesarios para tal fin.

En la investigación de diversos fenómenos naturales, en experimentos o en juegos de azar, es habitual contar con una representación de todos los resultados posibles. Los resultados individuales, denotados por ω , son llamados *eventos elementales*. El conjunto de todos los posibles eventos elementales se conoce como *evento seguro* o *espacio de sucesos elementales* y se denotará por Ω . Sean A un subconjunto de Ω y \mathcal{A} una colección de tales conjuntos. Si se observa un resultado ω y ese resultado está en A , se dice que A ha ocurrido. Es posible especificar la probabilidad de ocurrencia de A , que se denotará por $P(A)$.¹ El sistema de axiomas que caracteriza a la teoría de probabilidad está conformado por las siguientes proposiciones:

Axioma 1: $P(A) \geq 0$ para todo A en \mathcal{A} .

Axioma 2: $P(\Omega) = 1$.

Axioma 3: Si A_1, A_2, \dots es una secuencia numerable en \mathcal{A} y $A_i \cap A_j$ es el conjunto vacío para todo $i \neq j$, entonces $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

El conjunto de todos los posibles subconjuntos de Ω , se denota por $\mathcal{P}(\Omega)$. Desafortunadamente, por razones matemáticas técnicas, no siempre es posible la definición de $P(A)$ para todo A en $\mathcal{P}(\Omega)$; en particular, no siempre se verifica el axioma 3. Para eliminar esta dificultad, se restringe la clase de subconjuntos \mathcal{A} de Ω en la que se define P . Se requiere que la colección \mathcal{A} satisfaga las siguientes condiciones:

1. Si A está en \mathcal{A} , entonces el complemento A^c está también en \mathcal{A} .

¹ En el lenguaje del análisis real, una probabilidad P es una función de conjunto numerablemente aditiva y no negativa (es decir, una medida positiva), que cumple la condición adicional $P(\Omega) = 1$. (Petrov & Mordecki, 2008)

2. Si A_1, A_2, \dots es una secuencia numerable en \mathcal{A} , entonces $\bigcup_{i=1}^{\infty} A_i$ está en \mathcal{A} .

3. El conjunto vacío está en \mathcal{A} .

Una colección no vacía de subconjuntos de Ω que satisface las condiciones 1 a 3 se dice que es una *sigma-álgebra*. En este contexto, un *espacio de probabilidad*, representado por (Ω, \mathcal{A}, P) , es el suceso seguro Ω junto con una *sigma-álgebra* \mathcal{A} de subconjuntos de Ω y una función $P(A)$ definida en \mathcal{A} que satisface los axiomas 1 a 3². (Fuller, 1996)

En la práctica habitualmente no es posible enumerar todos los elementos de un experimento, sin embargo puede darse una función que asume valores reales tal que a cada resultado ω le asigna un número real $X(\omega)$. Formalmente, una *variable aleatoria* X es una función de valor real definida en Ω tal que el conjunto $\{\omega: X(\omega) \leq x\}$ es un subconjunto de \mathcal{A} para todo número real x . La función $F_X(x) = P(\{\omega: X(\omega) \leq x\})$ es llamada la *función de distribución* de la variable aleatoria X . (Fuller, 1996)³

En general, una colección de variables aleatorias, indexada por t y definidas en un espacio de probabilidad (Ω, \mathcal{A}, P) se conoce como proceso estocástico, donde t toma valores enteros ($t \in \mathbb{Z}$). Atendiendo las exigencias de la precisión y el rigor conceptual se da la siguiente definición:

Definición 3.1.2: Sea (Ω, \mathcal{A}, P) un espacio de probabilidad y sea T un conjunto de índices. Un *proceso estocástico* de valor real es una función de valor real $y(t, \omega)$ definida en $T \times \Omega$ tal que para cada t fijo, $y(t, \omega)$ es una variable aleatoria en (Ω, \mathcal{A}, P) . (Fuller, 1996)

De esta manera, un proceso estocástico es una familia de variables aleatorias indexadas en el tiempo $y(t, \omega)$ tal que, para un ω dado, $y(t, \omega)$, como una función de t , es llamada una función muestral o realización. (Wei, 2006)

En este marco conceptual, una serie de tiempo es una realización de un proceso estocástico (Box, Jenkins, & Reinsel, 1994). Es decir, se conforma a partir de un período muestral que constituye una parte de la historia del proceso estocástico del que deviene.

Ejemplo 3.1.2: Si se considera el proceso estocástico infinito *Cantidad anual de lluvia caída en San Gustavo* el cual se puede representar por $\{y_t\}$, la serie temporal, considerada para este ejemplo anualmente,

$$x_{1958} \quad x_{1959} \quad x_{1960} \quad \dots \quad x_{2011},$$

de tamaño $n = 53$, es una realización del proceso estocástico *Cantidad anual de lluvia caída en San Gustavo* que se observa desde 1958 hasta 2011.

En el estudio de una serie temporal es importante conocer el proceso estocástico que la genera y utilizar esta teoría para caracterizar su comportamiento y realizar predicciones futuras. Para esto resulta fundamental que el proceso estocástico presente una estructura

² En análisis real, un espacio de probabilidad es un espacio medible (Ω, \mathcal{A}) con una medida no negativa P , que verifica $P(\Omega) = 1$. (Petrov & Mordecki, 2008)

³ A los efectos de profundizar estos conceptos de probabilidad puede consultarse, entre otros, el libro *Teoría de la Probabilidad* de Valentín Petrov y Ernesto Mordecki (2008), Dirac, Montevideo.

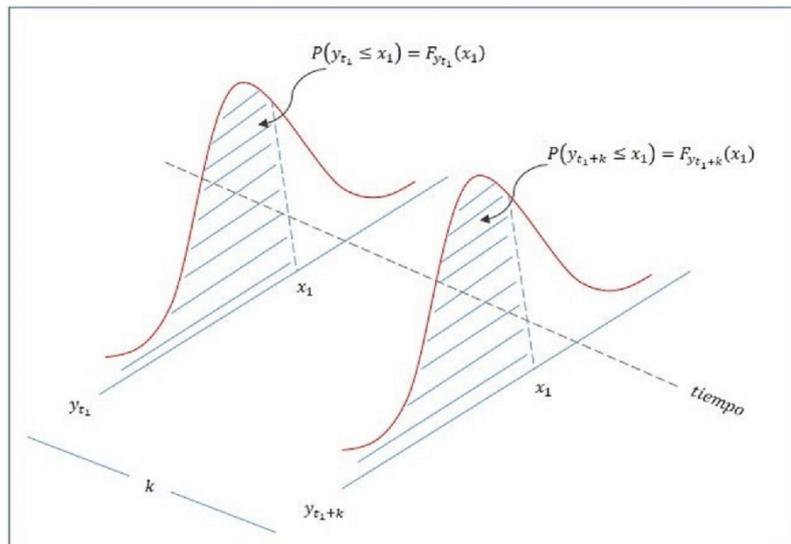
probabilística estable a lo largo del tiempo, ya que las predicciones se realizan en base a las regularidades manifiestas en la serie observada. Esta característica de ser estable en el tiempo, se conoce como *estacionariedad* del proceso.

Considérese un conjunto finito de variables aleatorias $\{y_{t_1}, y_{t_2}, \dots, y_{t_n}\}$ de un proceso estocástico $\{y(t, \omega): t = 0, \pm 1, \pm 2, \dots\}$. La *función de distribución n-dimensional* se define por

$$F_{y_{t_1}, \dots, y_{t_n}}(x_1, \dots, x_n) = P\{y(t, \omega): y_{t_1} \leq x_1, \dots, y_{t_n} \leq x_n\},$$

donde $x_i, i = 1, \dots, n$, son números reales cualesquiera. Un proceso estocástico se dice que es *estacionario de primer orden en distribución* si su función de distribución es invariante, es decir, $F_{y_{t_1}}(x_1) = F_{y_{t_1+k}}(x_1)$ para todo entero $t_1, k, t_1 + k$ (ver Gráfico 3.2); *estacionario de segundo orden en distribución* si $F_{y_{t_1}, y_{t_2}}(x_1, x_2) = F_{y_{t_1+k}, y_{t_2+k}}(x_1, x_2)$ para todo entero $t_1, t_2, k, t_1 + k$ y $t_2 + k$; y es *estacionario de orden n en distribución* si $F_{y_{t_1}, \dots, y_{t_n}}(x_1, \dots, x_n) = F_{y_{t_1+k}, \dots, y_{t_n+k}}(x_1, \dots, x_n)$ para toda *n-upla* (t_1, \dots, t_n) y k de enteros. Un proceso se dice que es *estrictamente estacionario* si $F_{y_{t_1}, \dots, y_{t_n}}(x_1, \dots, x_n) = F_{y_{t_1+k}, \dots, y_{t_n+k}}(x_1, \dots, x_n)$ es verdadera para todo $n = 1, 2, \dots$ (Wei, 2006)

Gráfico 3.2



En adelante, se hará referencia a un proceso estocástico simplemente indicando y_t o $\{y_t\}$, suprimiendo la variable ω y simplificando, de este modo, la escritura. Además, se considerará el estudio de procesos que toman solamente valores reales.

Los conceptos de estacionariedad, pueden especificarse también a partir de los distintos momentos asociados a un proceso estocástico, ya que no siempre es conocida la función de distribución del mismo. El primer momento viene dado por el conjunto de las esperanzas de todas las variables aleatorias que conforman el proceso:

Definición 3.1.3: Para un proceso de valor real $\{y_t: t = 0, \pm 1, \pm 2, \dots\}$ se define la función media del proceso como

$$\mu_y = E[y_t] < \infty.$$

La expresión $E[y_t]$ designa la esperanza matemática del proceso. El segundo momento está dado por el conjunto de las varianzas de todas las variables aleatorias y por las covarianzas entre todo par de variables aleatorias del proceso:

Definición 3.1.4: Para un proceso de valor real $\{y_t: t = 0, \pm 1, \pm 2, \dots\}$ se definen la varianza y la covarianza del mismo mediante las siguientes expresiones:

$$Var[y_t] = E[(y_t - \mu_y)^2] = \sigma_y^2 < \infty,$$

$$Cov[y_t, y_s] = E[(y_t - \mu_{y_t})(y_s - \mu_{y_s})] = \gamma_{t,s} \quad \forall t, s \quad (t \neq s).$$

Considérese una secuencia finita $\{y_{t_1}, y_{t_2}, \dots, y_{t_n}\}$ procedente de un proceso estocástico $\{y_t\}$. Pensada la secuencia como un vector aleatorio, donde cada y_{t_i} es, a su vez, un vector con los t_i elementos no necesariamente consecutivos del conjunto de índices, para $i = 1, \dots, n$, queda definida la esperanza matemática de la secuencia finita como $E[\{y_{t_1}, y_{t_2}, \dots, y_{t_n}\}] = \{E[y_{t_1}], E[y_{t_2}], \dots, E[y_{t_n}]\}$, cuando todas las esperanzas existen. Además, la matriz dada por

$$C_Y = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1n} \\ \vdots & \ddots & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nn} \end{pmatrix}$$

donde cada $\gamma_{i,j} = Cov[y_i, y_j]$ con $i, j = 1, \dots, n$ se llama *matriz de covarianza* de la secuencia o *matriz de segundos momentos*. Obsérvese que la matriz de covarianzas es simétrica, esto es $\gamma_{i,j} = \gamma_{j,i}$ para todo par i, j , y los elementos de su diagonal son las varianzas de las variables aleatorias y_1, \dots, y_n (Petrov & Mordecki, 2008).

La estacionariedad es un estado de cierto equilibrio estadístico que caracteriza la evolución temporal del proceso estocástico que ha generado una serie temporal. A continuación se explica la estacionariedad de un proceso a través de sus diferentes momentos:

Definición 3.1.5: Si $\{y_t: t \in T\}$ es un proceso estocástico *estrictamente estacionario* con $E[|y_t|] < \infty$, entonces el valor esperado de y_t es constante para todo t , ya que la función de distribución es la misma para todo t . Igualmente, si $E[y_t^2] < \infty$, entonces la varianza de y_t es constante para todo t . (Fuller, 1996)

Cuando la media del proceso es constante (no depende del tiempo) suele decirse que el proceso es *débilmente estacionario de primer orden*. Los segundos momentos de un proceso estocástico permiten definir la estacionariedad débil de segundo orden:

Definición 3.1.6: Un proceso estocástico $\{y_t\}$ con $E[y_t^2] < \infty$ para todo $t = 0, \pm 1, \pm 2, \dots$ es *débilmente estacionario de segundo orden* si satisface las siguientes dos condiciones:

1. El valor esperado de y_t es constante para todo t .
2. La matriz de covarianzas de $\{y_{t_1}, y_{t_2}, \dots, y_{t_n}\}$ es igual que la matriz de covarianza de $\{y_{t_1+k}, y_{t_2+k}, \dots, y_{t_n+k}\}$ para todo conjunto finito y no vacío de índices $\{t_1, t_2, \dots, t_n\}$ y todo k tal que $t_1, t_2, \dots, t_n, t_1+k, t_2+k, \dots, t_n+k$ están contenidos en el conjunto de índices. (Fuller, 1996)

Cuando un proceso estocástico es tal que su distribución de probabilidad conjunta es normal, entonces se dice que el proceso es *Gaussiano*.

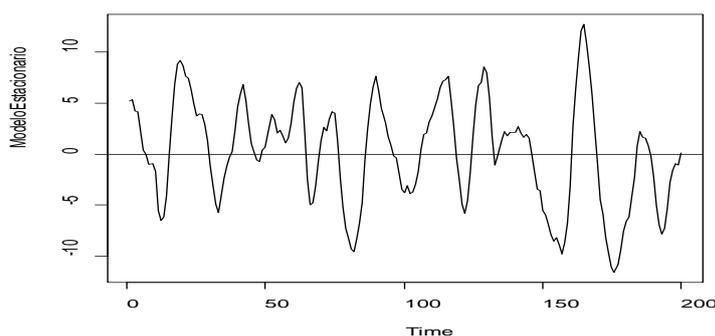
Definición 3.1.7: Un proceso estocástico $\{y_t\}$ es *Normal o Gaussiano* cuando para cualesquiera $n \geq 1$ momentos $t_1 < t_2 < \dots < t_n$ de su historia, la distribución de probabilidad conjunta de $\{y_{t_1}, y_{t_2}, \dots, y_{t_n}\}$ es una distribución normal n -variante. (Mauricio, 2007)

En general, la estacionariedad estricta resulta cuando se dan simultáneamente la estacionariedad en autocovarianza y, además, la distribución de probabilidad conjunta del proceso es normal.

La necesidad de contar con un proceso estocástico estacionario está relacionada con los supuestos de la regresión clásica. Los resultados de una regresión son confiables si los primeros momentos de un proceso son constantes y por lo tanto, los estimadores por mínimos cuadrados ordinarios serán insesgados y eficientes.⁴

Ejemplo 3.1.3: En la siguiente figura se muestra el gráfico de un proceso estacionario simulado con el *software R* para $n = 200$.

Gráfico 3.3



Muchas series no se consideran provenientes de un proceso estacionario, ya sea porque en su estructura se manifiesta una marcada tendencia y/o porque existen períodos con variabilidad muy heterogénea, que hacen que sus momentos cambien con el tiempo.

Definición 3.1.8: Un proceso estocástico $\{y_t\}$ es *no estacionario* cuando las propiedades estadísticas de al menos una secuencia finita $y_{t_1}, y_{t_2}, \dots, y_{t_n}$, ($n \geq 1$) de componentes de $\{y_t\}$, son diferentes de las de la secuencia $y_{t_1+k}, y_{t_2+k}, \dots, y_{t_n+k}$ para al menos un número entero $k > 0$. (Mauricio, 2007)

Las propiedades estadísticas de un proceso no estacionario son más complicadas que las de un proceso estacionario, no obstante muchas series no estacionarias pueden transformarse en

⁴ Un tratamiento riguroso de la teoría de estimación y del método de mínimos cuadrados puede encontrarse en el libro *Probabilidad y Estadística* de Morris Degroot, (1988), Addison-Wesley Iberoamericana, Wilmington, EEUU.

estacionarias, y a partir de esa transformación se puede elaborar un modelo compatible con la hipótesis de estacionariedad.⁵

Anteriormente se mencionó que uno de los objetivos principales del estudio de las series de tiempo, consiste en dar una descripción matemática del fenómeno y, en algunos casos, sobrepasar la incertidumbre intrínseca del mismo y lograr establecer predicciones de su comportamiento futuro. Esto se logra a partir de la confección de un modelo. Aunque ya se habló de modelos, se caracterizarán en el contexto de este trabajo.

Definición 3.1.9: Un *modelo univariante* para un proceso estocástico univariante o escalar $\{y_t\}$ es cualquier conjunto de hipótesis bien definidas sobre ciertas propiedades teóricas de las distribuciones de probabilidad (conjuntas, marginales o condicionales) de los componentes del proceso $\{y_t\}$ del que se supone procede una serie temporal observada $\{x_t\}$. (Mauricio, 2007)

En términos técnicos, un modelo para una serie temporal $\{x_t\}$ se formula mediante una expresión matemática que contempla las propiedades teóricas que determinan los momentos de primer y segundo orden (medias, varianzas y covarianzas) de las distribuciones conjuntas de los componentes de $\{x_t\}$. Cuando dichas distribuciones conjuntas son normales, sus propiedades de primer y segundo orden caracterizan completamente la estructura probabilística de $\{x_t\}$ (Mauricio, 2007). El uso de estos modelos representa grandes ventajas cuando los parámetros que describen las series temporales son constantes en el tiempo. Por ejemplo, si la serie temporal se puede describir usando parámetros estacionales mensuales, entonces los parámetros estacionales para cada uno de los doce meses son iguales de un año al otro (Bowerman, O'Connell, & Koehler, 2007). Esta clase de series se estudiarán en el capítulo 5.

La función matemática (modelo) que caracteriza el proceso puede describirse, según el enfoque clásico o determinista, mediante cuatro componentes:

- (a) La *tendencia*. Movimiento o dirección general de la variable en períodos prolongados de tiempo.
- (b) La *estacionalidad*. Fluctuaciones periódicas de la variable, más o menos regulares, en períodos relativamente cortos de tiempo, con una oscilación repetitiva para lapsos de tiempo contiguos.
- (c) El *ciclo*. Movimientos de la variable similares a la estacionalidad, pero relativos a períodos de tiempo mucho más prolongados. Solo es posible detectar esta componente, de existir, en series suficientemente largas.
- (d) *Aleatoriedad*. Es el movimiento irregular de la variable, determinado por el azar, impredecible de forma determinística. Cuando en una serie se detectan y extraen, las componentes de tendencia, estacionalidad y ciclo, en la serie sólo debe quedar la componente aleatoria. Por ello, a esta componente se la suele llamar también “residuo aleatorio” o “ruido”. (Aguirre Jaime, 1994).

⁵ Estas transformaciones se discutirán en el capítulo 5.

3.2. Función de Autocorrelación y Función de Autocorrelación Parcial

Existen dos funciones que son muy importantes para estudiar las propiedades y características de un proceso en términos de la interrelación de sus valores: la *Función de Autocorrelación* y la *Función de Autocorrelación Parcial* del proceso. Tanto es así, que en los capítulos 4 y 5 se verá cómo estas funciones permiten decidir el tipo de modelo al cual se ajustan los datos observados.

Para un proceso estacionario $\{y_t\}$ la media $E[y_t] = \mu_y$ y la varianza $Var[y_t] = \sigma_y^2$ son constantes y las covarianzas, $Cov[y_t, y_s]$, son funciones solamente de la diferencia $|t - s|$. En este caso puede darse la siguiente definición:

Definición 3.2.1: Se llama *autocovarianza de orden $k > 0$* de un proceso estacionario $\{y_t\}$ a la expresión:

$$\gamma_k = Cov[y_t, y_{t+k}] = E[(y_t - \mu_y)(y_{t+k} - \mu_y)] \text{ con } k = 1, 2, \dots \quad (3.1)$$

Puede notarse que la *autocovarianza de orden k* no es otra cosa que la covarianza de cualesquiera dos componentes del proceso que están separadas entre sí por un intervalo de tiempo $k > 0$ conocido como *retardo*, y además, depende solamente de dicho retardo k y no de los momentos específicos referidos por los valores de $\{y_t\}$.

Aunque la autocovarianza se ha definido para un retardo $k > 0$, también puede definirse para $k = 0$. En este caso, la autocovarianza del proceso es la varianza del proceso estacionario $\{y_t\}$:

$$\gamma_0 = Cov[y_t, y_t] = Var[y_t]$$

Otra consecuencia inmediata de la definición 3.2.1 es que $\gamma_k = Cov[y_t, y_{t+k}] = Cov[y_{t-k}, y_t]$, ya que y_{t-k} y y_t están separados por el mismo retardo $k \geq 0$ que y_t y y_{t+k} .

Puesto que k puede variar tomando valores enteros positivos, γ_k puede considerarse una función del retardo k , denominada *función de autocovarianza* del proceso estacionario $\{y_t\}$.

Definición 3.2.2: Se llama *autocorrelación simple de orden $k > 0$* de un proceso estacionario $\{y_t\}$ a la correlación entre y_t e y_{t+k} dada por la expresión:

$$\rho_k = \frac{Cov[y_t, y_{t+k}]}{\sqrt{Var[y_t]}\sqrt{Var[y_{t+k}]}} = \frac{\gamma_k}{\gamma_0}. \quad (3.2)$$

De (3.2) se deduce que $\rho_0 = 1$. Además, claramente puede observarse que ρ_k es el coeficiente de correlación lineal entre dos valores cualesquiera y_t e y_{t+k} del proceso $\{y_t\}$, separados por el retardo $k > 0$ y por lo tanto está sujeto a las propiedades de éste, en particular, $-1 \leq \rho_k \leq 1$. En efecto, nótese que por ser la varianza positiva, se tiene:

$$Var[\lambda_1 y_t + \lambda_2 y_{t+k}] \geq 0$$

con λ_1 y λ_2 constantes. Por las propiedades de la varianza⁶:

$$\lambda_1^2 Var[y_t] + \lambda_2^2 Var[y_{t+k}] + 2\lambda_1\lambda_2 Cov[y_t, y_{t+k}] = (\lambda_1^2 + \lambda_2^2)\sigma^2 + 2\lambda_1\lambda_2\gamma_k$$

Cuando $\lambda_1 = \lambda_2 = 1$, se tiene:

$$\gamma_k \geq -\sigma^2$$

De modo que $\rho_k \geq -1$. Cuando $\lambda_1 = 1$ y $\lambda_2 = -1$, resulta

$$\sigma^2 \geq \gamma_k$$

Así que $\rho_k \leq 1$. (Chatfield, 1995)

Además, por la naturaleza de la definición de ρ_k en términos de la autocovarianza $\rho_k = \rho_{-k}$. En efecto, para $\{y_t\}$ estacionario $\gamma_k = Cov[y_t; y_{t+k}] = Cov[y_{t-k}; y_t] = \gamma_{-k}$. (Chatfield, 1995)

Considerada como una función del retardo k , la secuencia $\{\rho_k: k = 1, 2, \dots\}$ se denomina la *función de autocorrelación simple (ACF, del inglés AutoCorrelation Function)* del proceso estacionario $\{y_t\}$. Dado que cada $\{\rho_k\}$ es un coeficiente de correlación, suele decirse que la ACF de $\{y_t\}$ representa la duración y la intensidad de la memoria del proceso $\{y_t\}$. (Mauricio, 2007)

Pueden calcularse las funciones de autocovarianza y autocorrelación a partir de la ecuación matemática del modelo asociado a un proceso estocástico. Esto se realizará en los capítulos 4 y 5.

Relacionada con la autocorrelación simple, aparece la *autocorrelación parcial de orden k* que mide el exceso de correlación debida a y_{t-k} y se obtiene como la correlación que existe entre y_t e y_{t-k} después de eliminar el efecto de todas las variables aleatorias que están entre ellas.

Definición 3.2.3: Se llama *autocorrelación parcial de orden $k > 0$* de un proceso estacionario $\{y_t\}$ a la expresión:

$$\phi_{kk} = corr(y_t, y_{t-k} / y_{t-1}, y_{t-2}, \dots, y_{t-(k-1)}) \quad (3.3)$$

La *autocorrelación parcial de orden k* es una medida del grado de asociación lineal entre dos componentes cualesquiera de $\{y_t\}$ separados entre sí por un retardo $k \geq 1$ dado (como y_t e y_{t-k}) que no es debida a la posible correlación entre cada uno de ellos y todos los componentes de $\{y_t\}$ que se encuentran entre ambos $\{y_{t-1}, y_{t-2}, \dots, y_{t-k+1}\}$. (Mauricio, 2007)

Para obtener los valores de ϕ_{kk} se considera la regresión

$$y_t = \phi_{k1}y_{t-1} + \phi_{k2}y_{t-2} + \phi_{k3}y_{t-3} \dots + \phi_{kk}y_{t-k} + e_t$$

⁶ Puede encontrarse una demostración de que la varianza de una suma de dos variables aleatorias con covarianza finita es igual a la suma de las varianzas de cada variable aleatoria más dos veces la covarianza de ambas variables en el libro *Probabilidad y Estadística* de Morris Degroot, (1988), Addison Wesley Iberoamericana, Wilmington, U.S.A. pág. 205. Además, en pág. 186, se demuestra que $Var[aX + b] = a^2Var[X]$ para a y b constantes cualesquiera.

donde $y_{t-1}, y_{t-2}, \dots, y_{t-k}$ son independientes entre sí y también lo son de e_t , que es un término de error de media cero. Multiplicando la ecuación anterior por y_{t-i} :

$$y_t y_{t-i} = \phi_{k1} y_{t-1} y_{t-i} + \phi_{k2} y_{t-2} y_{t-i} + \phi_{k3} y_{t-3} y_{t-i} + \dots + \phi_{kk} y_{t-k} y_{t-i} + e_t y_{t-i}$$

Además, e_t e y_{t-i} están incorrelacionadas para $i \geq 1$ y se supone que la media de los y_i es cero por simplicidad. Por lo tanto,

$$E[y_t y_{t-i}] = E[\phi_{k1} y_{t-1} y_{t-i} + \phi_{k2} y_{t-2} y_{t-i} + \phi_{k3} y_{t-3} y_{t-i} + \dots + \phi_{kk} y_{t-k} y_{t-i} + e_t y_{t-i}]$$

La linealidad de la esperanza matemática permite escribir:

$$E[y_t y_{t-i}] = \phi_{k1} E[y_{t-1} y_{t-i}] + \phi_{k2} E[y_{t-2} y_{t-i}] + \phi_{k3} E[y_{t-3} y_{t-i}] \dots + \phi_{kk} E[y_{t-k} y_{t-i}] + E[e_t y_{t-i}]$$

Por la definición de autocovarianza y por los supuestos previos:

$$\gamma_i = \phi_{k1} \gamma_{i-1} + \phi_{k2} \gamma_{i-2} + \phi_{k3} \gamma_{i-3} + \dots + \phi_{kk} \gamma_{i-k}$$

Dividiendo por γ_0 se tiene:

$$\frac{\gamma_i}{\gamma_0} = \phi_{k1} \frac{\gamma_{i-1}}{\gamma_0} + \phi_{k2} \frac{\gamma_{i-2}}{\gamma_0} + \phi_{k3} \frac{\gamma_{i-3}}{\gamma_0} + \dots + \phi_{kk} \frac{\gamma_{i-k}}{\gamma_0}$$

y por definición de autocorrelación simple de orden k , resulta:

$$\rho_i = \phi_{k1} \rho_{i-1} + \phi_{k2} \rho_{i-2} + \phi_{k3} \rho_{i-3} + \dots + \phi_{kk} \rho_{i-k} \text{ con } i = 1, 2, 3, \dots$$

De este modo, si $i = 1$:

$$\rho_1 = \phi_{k1} \rho_0 + \phi_{k2} \rho_{-1} + \phi_{k3} \rho_{-2} \dots + \phi_{kk} \rho_{1-k} = \phi_{k1} \rho_0 + \phi_{k2} \rho_1 + \phi_{k3} \rho_2 + \dots + \phi_{kk} \rho_{k-1}$$

Si $i = 2$:

$$\rho_2 = \phi_{k1} \rho_1 + \phi_{k2} \rho_0 + \phi_{k3} \rho_{-1} + \dots + \phi_{kk} \rho_{2-k} = \phi_{k1} \rho_1 + \phi_{k2} \rho_0 + \phi_{k3} \rho_1 + \dots + \phi_{kk} \rho_{k-2}$$

Si $i = 3$:

$$\rho_3 = \phi_{k1} \rho_2 + \phi_{k2} \rho_1 + \phi_{k3} \rho_0 + \dots + \phi_{kk} \rho_{3-k} = \phi_{k1} \rho_2 + \phi_{k2} \rho_1 + \phi_{k3} \rho_0 + \dots + \phi_{kk} \rho_{k-3}$$

Continuando este proceso, si $i = k$:

$$\rho_k = \phi_{k1} \rho_{k-1} + \phi_{k2} \rho_{k-2} + \dots + \phi_{kk} \rho_{k-k} = \phi_{k1} \rho_{k-1} + \phi_{k2} \rho_{k-2} + \dots + \phi_{kk} \rho_0$$

Dado que k da el orden de desarrollo anterior, si $k = 1$ se obtiene el valor de ϕ_{11} :

$$\phi_{11} = \rho_1$$

Si $k = 2$ se obtiene el valor de ϕ_{22} resolviendo el sistema:

$$\begin{cases} \rho_1 = \phi_{21} + \phi_{22} \rho_1 \\ \rho_2 = \phi_{21} \rho_1 + \phi_{22} \end{cases}$$

Utilizando la regla de Cramer:

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

Para $k = 3$ se tiene el siguiente sistema:

$$\begin{cases} \rho_1 = \phi_{31}\rho_0 + \phi_{32}\rho_1 + \phi_{33}\rho_2 \\ \rho_2 = \phi_{31}\rho_1 + \phi_{32}\rho_0 + \phi_{33}\rho_1 \\ \rho_3 = \phi_{31}\rho_2 + \phi_{32}\rho_1 + \phi_{33}\rho_0 \end{cases}$$

Si se resuelve el sistema por la Regla de Cramer, resulta:

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

Y así sucesivamente, puede obtenerse el resto de los coeficientes.

En general, se puede escribir ϕ_{kk} como un cociente de determinantes,

$$\phi_{kk} = \frac{|A_k|}{|B_k|} \quad (3.4)$$

donde

$$A_k = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{k-4} & \rho_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{pmatrix} y$$

$$B_k = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{k-4} & \rho_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{pmatrix}$$

Las matrices A_k y B_k son matrices de autocorrelaciones, de manera que los coeficientes de correlación parcial son funciones de los coeficientes de autocorrelación simple de un proceso estrictamente estacionario. Considerada como una función del retardo k , la secuencia $\{\phi_{kk}: k = 1, 2, \dots\}$ se denomina la función de autocorrelación parcial (*PACF*, del inglés *Partial Autocorrelation Function*) del proceso $\{y_t\}$. (Mauricio, 2007)

A continuación se da la definición de un proceso que resultará de importancia relevante para el estudio de los temas a desarrollar en los próximos capítulos.

Definición 3.2.4: Un proceso $\{A_t\}$ se llama proceso de *ruido blanco univariante* si es una secuencia de variables aleatorias no correlacionadas con una distribución fija, con media

constante $E[A_t] = \mu_A$, usualmente 0, varianza $Var[A_t] = \sigma_A^2$ y $\gamma_k = Cov[A_t, A_{t+k}] = 0$ para todo $k \neq 0$. (Wei, 2006)

Habitualmente suele representarse un ruido blanco como $\{A_t\} \sim iid(0, \sigma_A^2)$, lo que significa que $\{A_t\}$ es una secuencia de variables aleatorias idéntica e independientemente distribuidas con media 0 y varianza constante. Cuando cada componente A_t de la secuencia $\{A_t\}$ tiene una distribución normal, entonces la sucesión se denomina proceso de *ruido blanco Normal o Gaussiano*, y se representa como $\{A_t\} \sim iidN(0, \sigma^2)$.

Un ruido blanco es estacionario si la varianza σ_A^2 es finita y su función de autocovarianzas es

$$\gamma_k = \begin{cases} \sigma_A^2 & \text{si } k = 0 \\ 0 & \text{si } k \neq 0 \end{cases}$$

Mientras que su función de autocorrelación es

$$\rho_k = \begin{cases} 1 & \text{si } k = 0 \\ 0 & \text{si } k \neq 0 \end{cases}$$

y su función de autocorrelación parcial es

$$\phi_{kk} = \begin{cases} 1 & \text{si } k = 0 \\ 0 & \text{si } k \neq 0 \end{cases}$$

El proceso de ruido blanco es muy útil en el análisis de series temporales porque es la base para la construcción de los modelos ARIMA que se estudiarán en los siguientes capítulos.

3.3. Autocorrelación simple y parcial muestrales

Cuando se dispone de una serie temporal de n observaciones, resulta imposible determinar media, varianza, covarianza, etc. del proceso estocástico del que proviene la serie. Sin embargo, utilizando las técnicas tradicionales de la estadística inferencial es posible estimar los distintos momentos y las k primeras autocorrelaciones simples y parciales de un proceso $\{y_t\}$ a partir de una serie temporal de n observaciones, cuando k es un número considerablemente menor que n . Generalmente, en la práctica es aconsejable que k no supere la cuarta parte de n ya que cuanto mayor sea k se dispone de menos información para estimar las funciones de autocorrelación. No obstante, desde el punto de vista de la estadística clásica hay un problema relacionado con la naturaleza de las variables presentes en un proceso estocástico ya que habitualmente no se cumple el supuesto de independencia y distribución idéntica (*iid*). Además, cuando se tienen datos muestrales, el supuesto de estacionariedad se vuelve imprescindible; de algún modo será necesario calcular una media sobre una realización particular para estimar los momentos poblacionales y las funciones de autocorrelación. Si el proceso estocástico $\{y_t\}$ del cual proviene la serie no es estacionario, resultará necesario estimar n medias y varianzas distintas, dependientes de la variable tiempo y un gran número de autocorrelaciones, lo que no será posible contando solamente con n observaciones muestrales. El problema no se soluciona aumentando el tamaño de la muestra, ya que habrá que estimar $n + m$ medias, varianzas, etc. más, con lo cual la situación permanecería igual, suponiendo que se hayan tomado m observaciones más.

Se supondrá, entonces, que dado un proceso estocástico estacionario $\{y_t\} = \{y_1, y_2, y_3 \dots\}$, la serie temporal asociada al mismo será $\{x_t\} = \{x_1, x_2, \dots, x_n\}$ donde cada x_i es una realización particular o valor observado de cada variable aleatoria y_i que conforma el proceso.

Definición 3.3.1: La *media muestral* de una muestra de tamaño n asociada a una serie temporal $\{x_t\}$ generada por un proceso estocástico estacionario $\{y_t\}$ está dada por la expresión:

$$\hat{\mu}_y = \frac{1}{n} \sum_{t=1}^n y_t \quad (3.5)$$

y suele emplearse como estimador de la media μ_y del proceso estacionario del que procede la muestra. Este estimador proporciona la estimación numérica

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t \quad (3.6)$$

cuando se sustituye en (3.5) cada variable aleatoria y_t del proceso por su valor observado x_t de la serie, con $t = 1, \dots, n$.

Es preciso, en este punto, plantearse si $\hat{\mu}_y$ es un buen estimador de la media del proceso. En principio es claro que

$$E[\hat{\mu}_y] = \frac{1}{n} \sum_{t=1}^n E[y_t] = \frac{1}{n} \cdot n \cdot \mu_y = \mu$$

Esto implica que $\hat{\mu}_y$ es un estimador insesgado de μ . Además⁷,

$$\begin{aligned} Var[\hat{\mu}_y] &= Var\left[\frac{1}{n} \sum_{t=1}^n y_t\right] = \frac{1}{n^2} \sum_{t=1}^n \sum_{s=1}^n Cov[y_t, y_s] = \frac{1}{n^2} \sum_{t=1}^n \sum_{s=1}^n \gamma_0 \rho_{t-s} \\ &= \frac{\gamma_0}{n^2} \sum_{k=-(n-1)}^{n-1} (n - |k|) \rho_k = \frac{\gamma_0}{n^2} \sum_{k=-(n-1)}^{n-1} \left(1 - \frac{|k|}{n}\right) \rho_k \end{aligned}$$

donde $k = t - s$. De esta manera

$$\lim_{n \rightarrow \infty} \left[\sum_{k=-(n-1)}^{n-1} \left(1 - \frac{|k|}{n}\right) \rho_k \right]$$

es finito, y por lo tanto, $Var[\hat{\mu}_y] \rightarrow 0$ cuando $n \rightarrow \infty$. Este resultado implica que $\hat{\mu}_y$ es un estimador consistente de μ_y . Esto es

⁷ Los resultados utilizados en este desarrollo pueden encontrarse en el libro *Probabilidad y Estadística* de Morris Degroot, (1988), Addison.Wesley Iberoamericana, Wilmington, U.S.A. en las páginas 206 y 207.

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n y_t = \mu_y$$

Se dice que el proceso es *ergódico* para la media si el resultado anterior se mantiene⁸. Una condición suficiente para que el proceso sea *ergódico* para la media es que $\rho_k \rightarrow 0$ cuando $k \rightarrow \infty$, porque esto implica que para todo $\varepsilon > 0$ puede encontrarse un N tal que $|\rho_k| < \frac{\varepsilon}{4}$ para todo $k > N$. Por lo tanto, para $n > N + 1$ resulta:

$$\left| \frac{1}{n} \sum_{k=-(n-1)}^{n-1} \rho_k \right| \leq \frac{2}{n} \sum_{k=0}^{n-1} |\rho_k| = \frac{2}{n} \sum_{k=0}^N |\rho_k| + \frac{2}{n} \sum_{k=N+1}^{n-1} |\rho_k| \leq \frac{2}{n} \sum_{k=0}^N |\rho_k| + \frac{\varepsilon}{2} \leq \varepsilon$$

Así,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=-(n-1)}^{n-1} \rho_k = 0,$$

y, por lo tanto,

$$\lim_{n \rightarrow \infty} \text{Var}[\hat{\mu}_y] = \lim_{n \rightarrow \infty} \frac{\gamma_0}{n^2} \sum_{k=-(n-1)}^{n-1} \left(1 - \frac{|k|}{n}\right) \rho_k = 0.$$

Siempre que se obtiene un estadístico calculado sobre una muestra, se requiere la estimación del parámetro poblacional, para lo cual es necesario utilizar el contraste de hipótesis. Para contrastar hipótesis sobre el parámetro μ_y que es la media de un proceso estacionario, como por ejemplo $\mu_y = h$ donde h es un número dado, contra las hipótesis alternativas que pueden plantearse como $\mu_y > h$, $\mu_y < h$ o bien $\mu_y \neq h$, dependiendo de la naturaleza del problema, se utiliza habitualmente un estadístico t basado en que

$$E[\hat{\mu}_y] = \mu_y \text{ y además, } \text{Var}[\hat{\mu}_y] \approx \frac{\sigma_y^2}{n} \left(1 + 2 \sum_{k=1}^{\infty} \rho_k\right) \quad (3.7)$$

Es así que el estadístico $t_y = \frac{\hat{\mu}_y - h}{\sqrt{\frac{\hat{\sigma}_y^2}{n}}}$ cuyo valor calculado puntual es $t = \frac{\bar{x} - h}{\sqrt{\frac{s_y^2}{n}}}$ sigue una distribución *t de student* con $n - 1$ grados de libertad. (Mauricio, 2007)

⁸ Un análisis más profundo del concepto de ergodicidad puede encontrarse en el libro *Elementos de econometría de los fenómenos dinámicos*, de Alberto Landro y Mirta González consignado en la bibliografía. Citando a los autores del libro mencionado, “se dice que un proceso es ergódico cuando sus momentos muestrales finitos, obtenidos a partir de una serie cronológica finita convergen, en cierta forma, a los correspondientes momentos muestrales infinitos, los cuales coinciden con los momentos poblacionales del proceso”.

Definición 3.3.2: La función de autocovarianza muestral⁹ de una muestra de tamaño n asociada a una serie temporal $\{x_t\}$ generada por un proceso estocástico estacionario $\{y_t\}$ está dada por

$$\hat{\gamma}_k = \frac{1}{n} \sum_{t=1}^{n-k} (y_t - \hat{\mu}_y)(y_{t+k} - \hat{\mu}_y) = \frac{1}{n} \sum_{t=k+1}^n (y_{t-k} - \hat{\mu}_y)(y_t - \hat{\mu}_y) \text{ con } k = 0, 1, 2, \dots \quad (3.8)$$

Para $k = 0$ se obtiene $\hat{\gamma}_0 = \hat{\sigma}_y^2$. Este estimador proporciona una estimación numérica puntual cuando se sustituye cada variable y_t del proceso por su valor observado x_t de la serie, con $t = 1, \dots, n$, la cual está dada por

$$c_k = \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x}) = \frac{1}{n} \sum_{t=k+1}^n (x_{t-k} - \bar{x})(x_t - \bar{x}) \text{ con } k = 0, 1, 2, \dots \quad (3.9)$$

Con respecto a la definición de la autocovarianza muestral surge una discusión basada en si se obtiene una mejor estimación al considerar $\hat{\gamma}_k = \frac{1}{n-k} \sum_{t=1}^{n-k} (y_t - \hat{\mu}_y)(y_{t+k} - \hat{\mu}_y)$ en lugar de $\hat{\gamma}_k$. En principio, obsérvese que

$$\begin{aligned} \sum_{t=1}^{n-k} (y_t - \hat{\mu}_y)(y_{t+k} - \hat{\mu}_y) &= \sum_{t=1}^{n-k} [(y_t - \mu_y) - (\hat{\mu}_y - \mu_y)] [(y_{t+k} - \mu_y) - (\hat{\mu}_y - \mu_y)] \\ &= \sum_{t=1}^{n-k} [(y_t - \mu_y)(y_{t+k} - \mu_y) - (y_t - \mu_y)(\hat{\mu}_y - \mu_y) \\ &\quad - (\hat{\mu}_y - \mu_y)(y_{t+k} - \mu_y) + (\hat{\mu}_y - \mu_y)^2] \\ &= \sum_{t=1}^{n-k} (y_t - \mu_y)(y_{t+k} - \mu_y) - (\hat{\mu}_y - \mu_y) \sum_{t=1}^{n-k} (y_t - \mu_y) \\ &\quad - (\hat{\mu}_y - \mu_y) \sum_{t=1}^{n-k} (y_{t+k} - \mu_y) + (n-k)(\hat{\mu}_y - \mu_y)^2 \end{aligned}$$

Si se aproximan los términos $\sum_{t=1}^{n-k} (y_t - \mu_y)$ y $\sum_{t=1}^{n-k} (y_{t+k} - \mu_y)$ por $(n-k)(\hat{\mu}_y - \mu_y)$ resulta:

$$\sum_{t=1}^{n-k} (y_t - \hat{\mu}_y)(y_{t+k} - \hat{\mu}_y) \approx \sum_{t=1}^{n-k} (y_t - \mu_y)(y_{t+k} - \mu_y) - (n-k)(\hat{\mu}_y - \mu_y)^2$$

De este modo,

⁹ Como caso particular, cuando $k = 0$ se obtiene un estimador para la varianza.

$$\begin{aligned}
E[\hat{\gamma}_k] &\approx \frac{1}{n} E \left[\sum_{t=1}^{n-k} (y_t - \mu_y)(y_{t+k} - \mu_y) - (n-k)(\hat{\mu}_y - \mu_y)^2 \right] \\
&= \frac{1}{n} E \left[\sum_{t=1}^{n-k} (y_t - \mu_y)(y_{t+k} - \mu_y) \right] - \left(\frac{n-k}{n} \right) E \left[(\hat{\mu}_y - \mu_y)^2 \right] \\
&\approx \frac{1}{n} (n-k) \gamma_k - \left(\frac{n-k}{n} \right) \text{Var}[\hat{\mu}_y]
\end{aligned}$$

Por lo tanto, $E[\hat{\gamma}_k] \approx \left(1 - \frac{k}{n}\right) \gamma_k - \left(\frac{n-k}{n}\right) \text{Var}[\hat{\mu}_y] = \gamma_k - \frac{k}{n} \gamma_k - \left(\frac{n-k}{n}\right) \text{Var}[\hat{\mu}_y]$

Y, por otro lado,

$$\begin{aligned}
E[\hat{\hat{\gamma}}_k] &= E \left[\frac{1}{n-k} \sum_{t=1}^{n-k} (y_t - \hat{\mu}_y)(y_{t+k} - \hat{\mu}_y) \right] \\
&\approx \frac{1}{n-k} E \left[\sum_{t=1}^{n-k} (y_t - \mu_y)(y_{t+k} - \mu_y) - (n-k)(\hat{\mu}_y - \mu_y)^2 \right] \\
&= \frac{1}{n-k} E \left[\sum_{t=1}^{n-k} (y_t - \mu_y)(y_{t+k} - \mu_y) \right] - \frac{n-k}{n-k} E \left[(\hat{\mu}_y - \mu_y)^2 \right] \\
&\approx \frac{1}{n-k} (n-k) \gamma_k - \text{Var}[\hat{\mu}_y] = \gamma_k - \text{Var}[\hat{\mu}_y]
\end{aligned}$$

Es evidente que ambos son estimadores sesgados de γ_k . Si se ignora el término $\text{Var}[\hat{\mu}_y]$, entonces $\hat{\hat{\gamma}}_k$ se convierte en un estimador insesgado de γ_k , pero $\hat{\gamma}_k$ sigue siendo insesgado. Sin embargo, cuando el proceso es ergódico para la media y $\lim_{n \rightarrow \infty} \text{Var}[\hat{\mu}_y] = 0$, ambos estimadores $\hat{\hat{\gamma}}_k$ y $\hat{\gamma}_k$ son asintóticamente insesgados. (Wei, 2006)

Además, es posible ver que la varianza de $\hat{\hat{\gamma}}_k$ es más grande que la varianza de $\hat{\gamma}_k$. En efecto, Bartlett (1946) mostró que cuando el proceso y_t es Gaussiano se tiene las siguientes aproximaciones (Wei, 2006):

$$\text{Cov}[\hat{\gamma}_k, \hat{\gamma}_{k+j}] \approx \frac{1}{n} \sum_{i=-\infty}^{\infty} (\gamma_i \gamma_{i+j} + \gamma_{i+k+j} \gamma_{i-k}) \quad \text{y} \quad \text{Var}[\hat{\gamma}_k] \approx \frac{1}{n} \sum_{i=-\infty}^{\infty} (\gamma_i^2 + \gamma_{i+k} \gamma_{i-k}).$$

Análogamente,

$$\text{Cov}[\hat{\hat{\gamma}}_k, \hat{\hat{\gamma}}_{k+j}] \approx \frac{1}{n-k} \sum_{i=-\infty}^{\infty} (\gamma_i \gamma_{i+j} + \gamma_{i+k+j} \gamma_{i-k}) \quad \text{y} \quad \text{Var}[\hat{\hat{\gamma}}_k] \approx \frac{1}{n-k} \sum_{i=-\infty}^{\infty} (\gamma_i^2 + \gamma_{i+k} \gamma_{i-k})$$

Es claro que la varianza de $\hat{\hat{\gamma}}_k$ es más grande que la varianza de $\hat{\gamma}_k$ pues, para un k grande, la varianza de $\hat{\hat{\gamma}}_k$ puede convertirse en un estimador inestable.

En resumen, para un k dado, la función de autocovarianza muestral es un estimador asintóticamente insesgado de γ_k , y por lo tanto es una condición suficiente para que $\hat{\gamma}_k$ sea consistente en media cuadrada y el proceso sea ergódico para la autocovarianza, que la autocovarianza sea absolutamente convergente, es decir $\sum_{-\infty}^{\infty} |\gamma_i| < \infty$ y en consecuencia $\lim_{n \rightarrow \infty} \text{Var}[\hat{\gamma}_k] = 0$. (Wei, 2006)

Definición 3.3.3: La *correlación simple muestral de orden k* de una muestra de tamaño n asociada a una serie temporal $\{x_t\}$ generada por un proceso estocástico estacionario $\{y_t\}$ está dada por

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \text{ con } k = 0, 1, 2, \dots \quad (3.10)$$

La estimación numérica puntual asociada a $\hat{\rho}_k$ cuando se sustituye cada variable y_t del proceso por su valor observado x_t de la serie, con $t = 1, \dots, n$, está dada por

$$r_k = \frac{c_k}{c_0} \text{ con } k = 0, 1, 2, \dots \quad (3.11)$$

De esta forma, la secuencia de valores r_k que se obtienen para $k = 1, 2, \dots$, conforman la *función de autocorrelación simple muestral de orden k* de la serie, *ACF muestral*.

La función de autocorrelación muestral tiene una distribución de muestreo que permite evaluar si los datos provienen de una serie completamente aleatoria o si las correlaciones son estadísticamente significativas en algunos rezagos.

Para un proceso estacionario Gaussiano, Bartlett (1946) demostró que para $k > 0$ y $k + j > 0$,

$$Cov[\hat{\rho}_k, \hat{\rho}_{k+j}] \approx \frac{1}{n} \sum_{i=-\infty}^{\infty} (\rho_i \rho_{i+j} + \rho_{i+j+k} \rho_{i-k} - 2\rho_k \rho_i \rho_{i-k-j} - 2\rho_{k+j} \rho_i \rho_{i-k} + 2\rho_k \rho_{k+j} \rho_i^2).$$

Para n suficientemente grande, $\hat{\rho}_k$ tiene una distribución aproximadamente normal con media ρ_k y varianza

$$Var[\hat{\rho}_k] \approx \frac{1}{n} \sum_{i=-\infty}^{\infty} (\rho_i^2 + \rho_{i+k} \rho_{i-k} - 4\rho_k \rho_i \rho_{i-k} + 2\rho_k^2 \rho_i^2).$$

Para el proceso en el que $\rho_k = 0$ para $k > m$, la aproximación de Bartlett se convierte en

$$Var[\hat{\rho}_k] \approx \frac{1}{n} (1 + 2\rho_1^2 + 2\rho_2^2 + \dots + 2\rho_m^2). \quad (3.12)$$

En la práctica, los ρ_i con $i = 1, \dots, m$, son desconocidos y se sustituyen por sus estimaciones muestrales $\hat{\rho}_i$, y el error estándar de $\hat{\rho}_i$ viene dado por (Wei, 2006) \square :

$$S_{\hat{\rho}_k} = \sqrt{\frac{1}{n} (1 + 2\hat{\rho}_1^2 + \dots + 2\hat{\rho}_m^2)}. \quad (3.13)$$

Si y_t es un ruido blanco, entonces para valores grandes de k , la función de autocorrelación muestral, $\hat{\rho}_k$, para $k = 1, \dots, m$, con m fijado arbitrariamente, tiene una distribución aproximadamente normal con media 0 y desviación estándar¹⁰ $S_{\hat{\rho}_k} = \frac{1}{\sqrt{n}}$.

¹⁰ Una demostración alternativa de este resultado puede encontrarse en el libro de Robert Shumway y David Stoffer, *Time Series Analysis and its Applications*, Apéndice A, Teorema A.7, Springer, New York, 2011.

Teniendo en cuenta el resultado anterior se obtiene un método para evaluar, aproximadamente, si los valores de $\hat{\rho}_k$ son significativos determinando si los valores observados se encuentran fuera del intervalo $\pm \frac{1.96}{\sqrt{n}}$ (esto es, más menos 2 veces el desvío estándar). Para una secuencia de ruidos blancos (cualquier secuencia puramente aleatoria), aproximadamente el 95% de las funciones de autocorrelación muestrales deben estar dentro de estos límites. Para determinar si las primeras k autocorrelaciones simples del proceso son conjuntamente significativas, se emplea habitualmente un estadístico llamado el *estadístico de Ljung-Box* dado por la expresión

$$Q = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k} \quad (3.14)$$

Bajo la hipótesis nula $\rho_1 = \rho_2 = \dots = \rho_h = 0$ el estadístico Q tiene aproximadamente una distribución χ^2 con $h - (p + q)$ grados de libertad.

La importancia de lo expuesto anteriormente se debe a que muchos procedimientos de modelización estadística dependen de la reducción de una serie de tiempo a una secuencia de *ruido blanco* usando diversos tipos de transformaciones.

Definición 4.3.4: La *correlación parcial muestral de orden k* ($k > 0$) de una muestra de tamaño n asociada a una serie temporal $\{x_t\}$ generada por un proceso estocástico estacionario $\{y_t\}$ está dada por

$$\hat{\phi}_{kk} = \frac{|\hat{A}_k|}{|\hat{B}_k|} \quad (3.15)$$

donde las matrices \hat{A}_k y \hat{B}_k son:

$$\hat{A}_k = \begin{pmatrix} 1 & \hat{\rho}_1 & \hat{\rho}_2 & \dots & \hat{\rho}_{k-2} & \hat{\rho}_1 \\ \hat{\rho}_1 & 1 & \hat{\rho}_1 & \dots & \hat{\rho}_{k-3} & \hat{\rho}_2 \\ \hat{\rho}_2 & \hat{\rho}_1 & 1 & \dots & \hat{\rho}_{k-4} & \hat{\rho}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{\rho}_{k-1} & \hat{\rho}_{k-2} & \hat{\rho}_{k-3} & \dots & \hat{\rho}_1 & \hat{\rho}_k \end{pmatrix} \text{ y}$$

$$\hat{B}_k = \begin{pmatrix} 1 & \hat{\rho}_1 & \hat{\rho}_2 & \dots & \hat{\rho}_{k-2} & \hat{\rho}_{k-1} \\ \hat{\rho}_1 & 1 & \hat{\rho}_1 & \dots & \hat{\rho}_{k-3} & \hat{\rho}_{k-2} \\ \hat{\rho}_2 & \hat{\rho}_1 & 1 & \dots & \hat{\rho}_{k-4} & \hat{\rho}_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{\rho}_{k-1} & \hat{\rho}_{k-2} & \hat{\rho}_{k-3} & \dots & \hat{\rho}_1 & 1 \end{pmatrix}$$

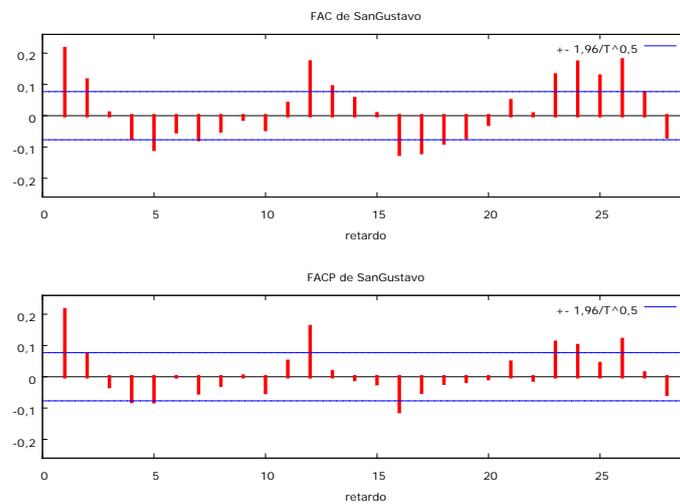
Considerada como una función del retardo k , la secuencia $\{\hat{r}_{kk}: k = 1, 2, \dots\}$ que se obtiene al aplicar los estimadores $\hat{\phi}_{11}, \hat{\phi}_{22}, \dots$, a una serie $\{x_t\}$ se llama la *función de autocorrelación parcial muestral*, *PACF* muestral de la serie.

En análogas condiciones que para la ACF muestral, resulta que cada $\hat{\phi}_{kk}$ tiene una distribución aproximadamente normal con media 0 y desviación estándar $\sigma_{\hat{\phi}_{kk}} = \frac{1}{\sqrt{n}}$. De este modo,

cualquier autocorrelación parcial $\hat{\phi}_{kk}$ ($k \geq 1$) puede considerarse individualmente significativa al 5% cuando $r_{kk} > \frac{1,96}{\sqrt{n}}$.

Una forma muy óptima de visualizar el comportamiento de la *ACF* y de la *PACF* muestrales es el *correlograma*, que es el gráfico de los coeficientes de correlación para cada retardo. Una forma comúnmente empleada para calcular el número de retardos consiste en tomar como cantidad de retardos la cuarta parte del total de observaciones. Es decir, si la serie de datos tiene n observaciones, una medida razonable para la cantidad de retardos a considerar es $\frac{n}{4}$ (Wei, 2006)¹¹. A continuación se muestran los correlogramas de la *ACF* y de la *PACF* asociadas a la serie de registros de lluvia de la Localidad de San Gustavo, Dpto. La Paz.

Gráfico 3.4



La función de autocorrelación y su representación gráfica (correlograma), son utilizadas como un instrumento fundamental para realizar un primer estudio sobre la estructura dinámica de las series a estudiar.

En los capítulos siguientes se profundizará en el estudio de las funciones de autocorrelación y de autocorrelación parcial y en sus correspondientes correlogramas, ya que sus estructuras permitirán identificar el tipo de modelo del proceso estocástico que genera la serie y el orden del mismo.

¹¹ Página 109.

Capítulo 4

La metodología de modelización ARIMA tiene como base los procesos autorregresivo y de media móvil. En este capítulo se introducen estos procesos y se definen los modelos ARMA (*Autoregressive Moving Average*). Además, se desarrollan los conceptos de causalidad e invertibilidad de un proceso y las condiciones que deben darse para el cumplimiento de estas propiedades. Por último, se analiza las características de las funciones de autocorrelación simple y parcial de estos modelos.

4.1. Modelos ARMA

A los efectos de establecer un modelo que explique la estructura del proceso estocástico del cual proviene una serie temporal, es necesario explicar el valor que toma una variable aleatoria que presenta dependencia temporal en un momento t . Para ello es preciso examinar el comportamiento pasado de la variable y explorar los patrones de regularidad a lo largo de su evolución. Es aquí donde resulta fundamental la obtención y estudio de la ACF y de la PACF introducidas en el capítulo anterior. En este sentido, los modelos ARMA, se construyen a partir de la información que aportan estas funciones, y se utilizan para explicar la estructura teórica temporal de la serie y predecir su comportamiento futuro a corto plazo.

Para los modelos que describen procesos estocásticos estacionarios, la representación matemática consiste en dos partes: una que permite recoger la regularidad del proceso (llamada a veces *parte sistemática*), y que constituye la parte predecible con el conjunto de datos constituido por la serie; y otra puramente aleatoria, denominada *innovación*, conformada por una secuencia de valores que no tienen ninguna relación o dependencia entre sí.

Una forma de hallar una ecuación que represente la influencia que los acontecimientos pasados tienen sobre el presente (y sobre el futuro), es escribir el valor actual del proceso como dependiente de forma lineal de valores pasados del propio proceso, más una perturbación aleatoria que hace que el modelo no sea determinista. Esta expresión caracteriza a los procesos autorregresivos.

Definición 4.1.1: Un *modelo autorregresivo de orden p* , abreviado como $AR(p)$, es de la forma

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + A_t \quad (4.1)$$

donde $\{y_t\}$ es un proceso estacionario, $\varphi_1, \varphi_2, \dots, \varphi_p$ son constantes ($\varphi_p \neq 0$) y $\{A_t\} \sim iid(0, \sigma^2)$. (Shumway & Stoffer, 2011)

En (4.1) la media de y_t es cero. Si la media, μ , de y_t no es cero, se reemplaza y_t por $y_t - \mu$ en (4.1)

$$y_t - \mu = \varphi_1 (y_{t-1} - \mu) + \varphi_2 (y_{t-2} - \mu) + \dots + \varphi_p (y_{t-p} - \mu) + A_t$$

Esto es

$$y_t = \mu + \varphi_1 y_{t-1} - \varphi_1 \mu + \varphi_2 y_{t-2} - \varphi_2 \mu + \dots + \varphi_p y_{t-p} - \varphi_p \mu + A_t$$

$$y_t = \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p) + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + A_t$$

o bien

$$y_t = \alpha + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + A_t \quad (4.2)$$

donde $\alpha = \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p)$. (Shumway & Stoffer, 2011)

Una forma alternativa de expresar la ecuación (4.1) es en términos del operador de retardos:

Definición 4.1.2: El operador de retardo B (del inglés *Backshift*, a veces también denotado por L de *Lag operator*), se define como:

$$B^d y_t = y_{t-d} \text{ con } d \in \mathbb{Z}_{\geq 1}$$

donde y_t es una variable referida a un momento t .

De (4.1):

$$y_t - \varphi_1 y_{t-1} - \varphi_2 y_{t-2} - \dots - \varphi_p y_{t-p} = A_t$$

Utilizando el operador de retardos:

$$y_t - \varphi_1 B y_t - \varphi_2 B^2 y_t - \dots - \varphi_p B^p y_t = A_t$$

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p) y_t = A_t$$

O más concisamente:

$$\varphi(B) y_t = A_t,$$

donde $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$.

Ejemplo 4.1.1: Un modelo autorregresivo de primer orden, está dado por la ecuación:

$$y_t = \varphi_1 y_{t-1} + A_t$$

O bien, expresado en términos del operador B :

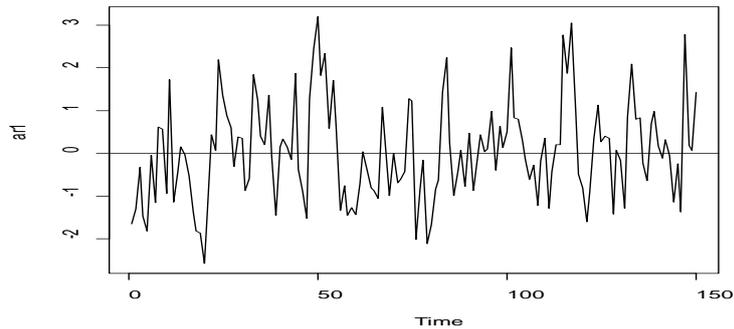
$$y_t - \varphi_1 B y_t = A_t$$

Escrito convenientemente queda:

$$(1 - \varphi_1 B) y_t = A_t$$

Por lo tanto, $\varphi(B) = 1 - \varphi_1 B$. El gráfico siguiente muestra un proceso $AR(1)$, simulado con R para $n = 150$, donde $\varphi_1 = 0,45$, por lo que su ecuación es $y_t = 0,45 y_{t-1} + A_t$:

Gráfico 4.1



Ejemplo 4.1.2: Un modelo autorregresivo de segundo orden o $AR(2)$ se expresa mediante la ecuación:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + A_t$$

Utilizando el operador de retardos, resulta:

$$y_t = \varphi_1 B y_t + \varphi_2 B^2 y_t + A_t$$

De la expresión anterior se obtiene:

$$y_t - \varphi_1 B y_t - \varphi_2 B^2 y_t = A_t$$

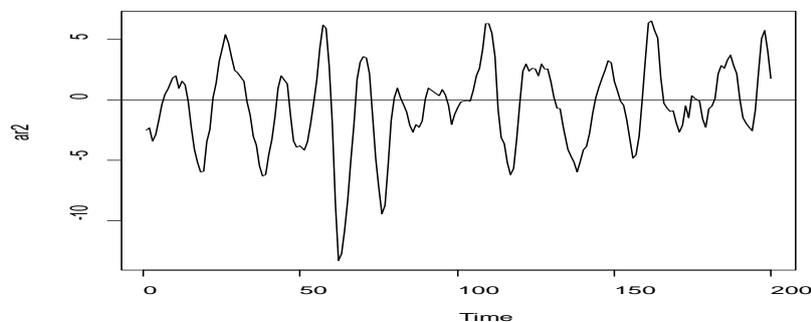
Factorizando y_t :

$$(1 - \varphi_1 B - \varphi_2 B^2) y_t = A_t$$

Luego, $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2$.

En la siguiente figura se representa un modelo autorregresivo de segundo orden, simulado para $n = 200$, con $\varphi_1 = 1,6$ y $\varphi_2 = -0,8$, es decir que la ecuación que caracteriza al modelo es $y_t = 1,6y_{t-1} - 0,8y_{t-2} + A_t$, y por lo tanto el polinomio autorregresivo en términos del operador de retardos es $\varphi(B) = 1 - 1,6B + 0,8B^2$.

Gráfico 4.2



La utilidad de los modelos autorregresivos está sujeta a las condiciones de estacionariedad discutidas en el capítulo anterior. Para mostrar que un proceso $AR(1)$ con $|\varphi_1| < 1$ es estacionario, puede emplearse una técnica de iteración “hacia atrás” como sigue:

$$y_t = \varphi_1 y_{t-1} + A_t$$

y como $y_{t-1} = \varphi_1 y_{t-2} + A_{t-1}$, resulta: $y_t = \varphi_1(\varphi_1 y_{t-2} + A_{t-1}) + A_t$. De donde: $y_t = \varphi_1^2 y_{t-2} + \varphi_1 A_{t-1} + A_t$. A su vez, $y_{t-2} = \varphi_1 y_{t-3} + A_{t-2}$. Por lo tanto, $y_t = \varphi_1^2(\varphi_1 y_{t-3} + A_{t-2}) + \varphi_1 A_{t-1} + A_t$. Desarrollando el producto:

$$y_t = \varphi_1^3 y_{t-3} + \varphi_1^2 A_{t-2} + \varphi_1 A_{t-1} + A_t$$

Siguiendo este proceso de iteración, se tiene:

$$y_t = \varphi_1^k y_{t-k} + \sum_{j=0}^{k-1} \varphi_1^j A_{t-j}$$

Este método sugiere que, por continuación de la iteración hacia atrás, y siempre que $|\varphi_1| < 1$ e y_t estacionario, podemos representar un modelo $AR(1)$ como un proceso lineal dado por:

$$y_t = \sum_{j=0}^{\infty} \varphi_1^j A_{t-j} \quad (4.3)$$

El proceso definido por (4.3) es estacionario con media cero. En efecto,

$$E[y_t] = E\left[\sum_{j=0}^{\infty} \varphi_1^j A_{t-j}\right] = \sum_{j=0}^{\infty} E[\varphi_1^j A_{t-j}] = \sum_{j=0}^{\infty} \varphi_1^j E[A_{t-j}] = \sum_{j=0}^{\infty} \varphi_1^j \cdot 0 = 0$$

Por otro lado, para probar la estacionariedad en autocovarianza, obsérvese que:

$$\gamma_k = Cov[y_{t+k}; y_t] = E\left[\left(\sum_{h=0}^{\infty} \varphi_1^h A_{t-h}\right)\left(\sum_{j=0}^{\infty} \varphi_1^j A_{t+k-j}\right)\right]$$

O bien:

$$\gamma_k = E\left[(A_t + \varphi_1 A_{t-1} + \varphi_1^2 A_{t-2} + \varphi_1^3 A_{t-3} + \dots)(A_{t+k} + \varphi_1 A_{t+k-1} + \varphi_1^2 A_{t+k-2} + \varphi_1^3 A_{t+k-3} + \dots + \varphi_1^k A_t + \varphi_1^{k+1} A_{t-1} + \dots)\right]$$

Desarrollando el producto y utilizando las propiedades de la esperanza:

$$\begin{aligned} \gamma_k = & E[A_t A_{t+k}] + \varphi_1 E[A_t A_{t+k-1}] + \varphi_1^2 E[A_t A_{t+k-2}] + \dots + \varphi_1^k E[A_t A_t] + \varphi_1^{k+1} E[A_t A_{t-1}] \\ & + \dots + \varphi_1 E[A_{t-1} A_{t+k}] + \varphi_1^2 E[A_{t-1} A_{t+k-1}] + \varphi_1^3 E[A_{t-1} A_{t+k-2}] + \dots \\ & + \varphi_1^{k+1} E[A_{t-1} A_t] + \varphi_1^{k+2} E[A_{t-1} A_{t-1}] + \dots + \varphi_1^2 E[A_{t-2} A_{t+k}] \\ & + \varphi_1^3 E[A_{t-2} A_{t+k-1}] + \varphi_1^4 E[A_{t-2} A_{t+k-2}] + \dots + \varphi_1^{k+4} E[A_{t-2} A_{t-2}] + \dots \end{aligned}$$

Y teniendo en cuenta que $\{A_t\} \sim iid(0, \sigma_A^2)$ resulta:

$$\gamma_k = \varphi_1^k \sigma_A^2 + \varphi_1^{k+1} \varphi_1 \sigma_A^2 + \varphi_1^{k+2} \varphi_1^2 \sigma_A^2 + \varphi_1^{k+3} \varphi_1^3 \sigma_A^2 + \dots$$

Y, por lo tanto:

$$\gamma_k = \sigma_A^2 \sum_{j=0}^{\infty} \varphi_1^{k+j} \varphi_1^j = \sigma_A^2 \sum_{j=0}^{\infty} \varphi_1^j \varphi_1^k \varphi_1^j = \sigma_A^2 \varphi_1^k \sum_{j=0}^{\infty} \varphi_1^{2j} = \frac{\sigma_A^2 \varphi_1^k}{1 - \varphi_1^2} \text{ con } k \geq 0$$

Nótese que γ_k sólo depende del retardo k . De la función de autocovarianza, se obtiene la función de autocorrelación *ACF*:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\frac{\sigma_A^2 \varphi_1^k}{1 - \varphi_1^2}}{\frac{\sigma_A^2}{1 - \varphi_1^2}} = \varphi_1^k \text{ con } k \geq 0$$

Además, para ρ_k vale la siguiente recursión:

$$\rho_{k-1} = \frac{\gamma_{k-1}}{\gamma_0} = \frac{\frac{\sigma_A^2 \varphi_1^{k-1}}{1 - \varphi_1^2}}{\frac{\sigma_A^2}{1 - \varphi_1^2}} = \varphi_1^{k-1} = \varphi_1^k \varphi_1^{-1} = \rho_k \varphi_1^{-1}$$

Y, en consecuencia: $\rho_k = \varphi_1 \rho_{k-1}$ (Shumway & Stoffer, 2011)

En la sección 4.2 se volverá al estudio de estas funciones.

Cabe preguntarse ahora si existe un modelo autorregresivo de orden 1 que sea estacionario con $|\varphi_1| > 1$. En efecto, tales modelos existen y son llamados *explosivos* ya que los valores de las series de tiempo que se ajustan a ellos tienen valores que rápidamente toman grandes magnitudes. Esto es así porque $|\varphi_1|^j$ crece sin límite cuando $j \rightarrow \infty$ y, por lo tanto, $\sum_{j=0}^{k-1} \varphi_1^j A_{t-j}$ diverge cuando $k \rightarrow \infty$.

Puede demostrarse la estacionariedad de un modelo de estas características modificando el razonamiento de “iteraciones hacia atrás” empleado para obtener (4.3).

Se considera $y_{t+1} = \varphi_1 y_t + A_t$ y en tal caso, $y_{t+1} - A_{t+1} = \varphi_1 y_t$. Entonces:

$$y_t = \varphi_1^{-1} y_{t+1} - \varphi_1^{-1} A_{t+1}$$

Como $y_{t+1} = \varphi_1^{-1} y_{t+2} - \varphi_1^{-1} A_{t+2}$ resulta:

$$y_t = \varphi_1^{-1} (\varphi_1^{-1} y_{t+2} - \varphi_1^{-1} A_{t+2}) - \varphi_1^{-1} A_{t+1}$$

y, por lo tanto, $y_t = \varphi_1^{-2} y_{t+2} - \varphi_1^{-2} A_{t+2} - \varphi_1^{-1} A_{t+1}$

Además, $y_{t+2} = \varphi_1^{-1} y_{t+3} - \varphi_1^{-1} A_{t+3}$ y así:

$$y_t = \varphi_1^{-2} (\varphi_1^{-1} y_{t+3} - \varphi_1^{-1} A_{t+3}) - \varphi_1^{-2} A_{t+2} - \varphi_1^{-1} A_{t+1}$$

Distribuyendo: $y_t = \varphi_1^{-3} y_{t+3} - \varphi_1^{-3} A_{t+3} - \varphi_1^{-2} A_{t+2} - \varphi_1^{-1} A_{t+1}$

Continuando con la iteración hacia adelante h pasos, se tiene:

$$y_t = \varphi_1^{-h} y_{t+h} - \sum_{j=1}^{h-1} \varphi_1^{-j} A_{t+j}$$

Porque $|\varphi_1|^{-1} < 1$, pues $|\varphi_1| > 1$, este resultado sugiere el futuro estacionario dependiente del modelo AR(1):

$$y_t = - \sum_{j=1}^{\infty} \varphi_1^{-j} A_{t+j} \quad (4.4)$$

Considerando que $A_t \sim iid(0, \sigma_A^2)$, el proceso definido por (4.4) es estacionario. En efecto,

$$E[y_t] = E \left[- \sum_{j=1}^{\infty} \varphi_1^{-j} A_{t+j} \right] = - \sum_{j=1}^{\infty} E[\varphi_1^{-j} A_{t+j}] = - \sum_{j=1}^{\infty} \varphi_1^{-j} E[A_{t+j}] = - \sum_{j=1}^{\infty} \varphi_1^{-j} \cdot 0 = 0$$

Y la función de autocovarianza es:

$$\gamma_k = Cov[y_{t+k}; y_t] = E \left[\left(- \sum_{h=1}^{\infty} \varphi_1^{-h} A_{t+h} \right) \left(- \sum_{j=1}^{\infty} \varphi_1^{-j} A_{t+k+j} \right) \right]$$

Desarrollando las sumatorias:

$$\begin{aligned} \gamma_k = E \left[\left(-\varphi_1^{-1} A_{t+1} - \varphi_1^{-2} A_{t+2} - \varphi_1^{-3} A_{t+3} - \varphi_1^{-4} A_{t+4} - \dots - \varphi_1^{-k} A_{t+k} - \varphi_1^{-(k+1)} A_{t+k+1} \right. \right. \\ \left. \left. - \varphi_1^{-(k+2)} A_{t+k+2} - \varphi_1^{-3} A_{t+k+3} - \varphi_1^{-4} A_{t+k+4} - \dots \right) \left(-\varphi_1^{-1} A_{t+k+1} \right. \right. \\ \left. \left. - \varphi_1^{-2} A_{t+k+2} - \varphi_1^{-3} A_{t+k+3} - \varphi_1^{-4} A_{t+k+4} - \dots \right) \right] \end{aligned}$$

De aquí resulta:

$$\begin{aligned} \gamma_k = E \left[\varphi_1^{-2} A_{t+1} A_{t+k+1} + \varphi_1^{-3} A_{t+1} A_{t+k+2} + \varphi_1^{-4} A_{t+1} A_{t+k+3} + \dots + \varphi_1^{-3} A_{t+2} A_{t+k+1} \right. \\ \left. + \varphi_1^{-4} A_{t+2} A_{t+k+2} + \dots + \varphi_1^{-k-1} A_{t+k} A_{t+k+1} \varphi_1^{-k-2} A_{t+k} A_{t+k+2} + \dots \right. \\ \left. + \varphi_1^{-k-2} A_{t+k+1} A_{t+k+1} + \dots + \varphi_1^{-k-4} A_{t+k+2} A_{t+k+2} + \dots \right. \\ \left. + \varphi_1^{-k-6} A_{t+k+3} A_{t+k+3} + \dots + \varphi_1^{-k-8} A_{t+k+4} A_{t+k+4} + \dots \right] \end{aligned}$$

Entonces:

$$\begin{aligned} \gamma_k = \varphi_1^{-k-2} E[A_{t+k+1} A_{t+k+1}] + \varphi_1^{-k-4} E[A_{t+k+2} A_{t+k+2}] + \varphi_1^{-k-6} E[A_{t+k+3} A_{t+k+3}] \\ + \varphi_1^{-k-8} E[A_{t+k+4} A_{t+k+4}] + \dots \end{aligned}$$

Por lo tanto:

$$\gamma_k = \varphi_1^{-k-2} \sigma_A^2 + \varphi_1^{-k-4} \sigma_A^2 + \varphi_1^{-k-6} \sigma_A^2 + \varphi_1^{-k-8} \sigma_A^2 + \dots$$

y factorizando, se tiene:

$$\begin{aligned} \gamma_k = \sigma_A^2 (\varphi_1^{-k-2} + \varphi_1^{-k-4} + \varphi_1^{-k-6} + \varphi_1^{-k-8} + \dots) = \sigma_A^2 \varphi_1^{-k} (\varphi_1^{-2} + \varphi_1^{-4} + \varphi_1^{-6} + \varphi_1^{-8} + \dots) \\ = \sigma_A^2 \varphi_1^{-k} \varphi_1^{-2} (1 + \varphi_1^{-2} + \varphi_1^{-4} + \varphi_1^{-6} + \dots) = \frac{\sigma_A^2 \varphi_1^{-k} \varphi_1^{-2}}{1 - \varphi_1^{-2}} \end{aligned}$$

Desafortunadamente, este modelo no es útil, ya que requiere que el futuro sea capaz de predecir el futuro. Cuando un proceso *no* es dependiente del futuro, tal como el $AR(1)$ cuando $|\varphi_1| < 1$, diremos que el proceso es *causal*. En el caso explosivo, cuando $|\varphi_1| > 1$, el proceso es estacionario, pero no causal ya que depende del futuro. (Shumway & Stoffer, 2011)

La exclusión de los modelos explosivos no representa un problema sustancial, puesto que estos modelos tienen contrapartes causales. Por ejemplo, si

$$y_t = \varphi_1 y_{t-1} + A_t \text{ con } |\varphi_1| > 1$$

y $A_t \sim iid(0, \sigma_A^2)$, entonces como se demostró, $\{y_t\}$ es un proceso estacionario no causal. Si se considera el proceso:

$$z_t = \varphi_1^{-1} z_{t-1} + v_t$$

donde $v_t \sim iid(0, \sigma_A^2 \cdot \varphi_1^{-2})$, resulta estocásticamente igual al proceso y_t (todas las distribuciones finitas de los procesos son las mismas). Por ejemplo, si $y_t = 2y_{t-1} + A_t$ con $\sigma_A^2 = 1$, entonces $z_t = \frac{1}{2}z_{t-1} + v_t$ con $\sigma_v^2 = \frac{1}{4}$ es un proceso causal equivalente. Esta idea es generalizable a órdenes superiores. (Shumway & Stoffer, 2011)

La técnica de iteración hacia atrás empleada para probar la estacionariedad de los modelos AR funciona correctamente cuando $p = 1$, pero no para órdenes mayores.

Una técnica general es la de coeficientes correspondientes. En principio, se escribe el modelo $AR(1)$ en términos del operador de retardo:

$$\varphi(B)y_t = A_t \quad (4.5)$$

donde $\varphi(B) = 1 - \varphi_1 B$, y $|\varphi_1| < 1$. O bien en términos de la ecuación (4.3) utilizando el operador de la siguiente manera:

$$y_t = \sum_{j=0}^{\infty} \psi_j A_{t-j} = \psi(B)A_t \quad (4.6)$$

donde $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ y $\psi_j = \varphi_1^j$. Supóngase que se desconoce que $\psi_j = \varphi_1^j$. Se puede substituir $\psi(B)A_t$ de (4.6) por y_t en (4.5) para obtener

$$\varphi(B)\psi(B)A_t = A_t \quad (4.7)$$

Los coeficientes de B en el lado izquierdo deben ser iguales a los del lado derecho de (4.7), lo que significa

$$(1 - \varphi_1 B)(1 + \psi_1 B + \psi_2 B^2 + \dots + \psi_j B^j + \dots) = 1$$

Reorganizando los coeficientes, resulta:

$$1 + (\psi_1 - \varphi_1)B + (\psi_2 - \psi_1 \varphi_1)B^2 + \dots + (\psi_j - \psi_{j-1} \varphi_1)B^j + \dots = 1$$

El coeficiente de B^j de la izquierda debe ser cero, ya que es cero a la derecha. El coeficiente de B en el lado izquierdo es $(\psi_1 - \varphi_1)$. Igualando este a cero, $\psi_1 - \varphi_1 = 0$, se tiene $\psi_1 = \varphi_1$.

Continuando este procedimiento, el coeficiente de B^2 es $\psi_2 - \psi_1\varphi_1$, entonces $\psi_2 = \psi_1\varphi_1$ y en consecuencia $\psi_2 = \varphi_1^2$. En general,

$$\psi_j = \psi_{j-1}\varphi_1$$

con $\psi_0 = 1$, lo cual conduce a la solución $\psi_j = \varphi_1^j$.

Otra forma de estudiar lo desarrollado es considerar el modelo $AR(1)$ en la forma (4.5). Si se asume que el operador inverso $\varphi^{-1}(B)$ existe y se multiplica (4.5) miembro a miembro, resulta:

$$\varphi^{-1}(B)\varphi(B)y_t = \varphi^{-1}(B)A_t$$

De donde:

$$y_t = \varphi^{-1}(B)A_t$$

Sabemos que

$$\varphi^{-1}(B) = 1 + \varphi_1B + \varphi_1^2B^2 + \dots + \varphi_1^jB^j + \dots$$

Es decir, $\varphi^{-1}(B)$ es $\psi(B)$ en (4.6).

Es posible ver que trabajar con operadores es comparable a trabajar con polinomios. En efecto, si se considera el polinomio $\varphi(z) = 1 - \varphi_1z$, donde z es un número complejo y $|\varphi_1| < 1$, se tiene,

$$\varphi^{-1}(z) = \frac{1}{1 - \varphi_1z} = 1 + \varphi_1z + \varphi_1^2z^2 + \dots + \varphi_1^jz^j + \dots, |z| \leq 1$$

y los coeficientes de B^j en $\varphi^{-1}(B)$ son iguales a los coeficientes de z^j en $\varphi^{-1}(z)$. En otras palabras, se puede interpretar al operador de retardo, B , como un número complejo, z .

En el caso de los modelos autorregresivos se asume que y_t está dado como una combinación lineal (expresada en el lado derecho de la ecuación que define al proceso autorregresivo) de los valores y_{t-k} de la serie. Como alternativa, en el modelo de media móvil, que abreviadamente se escribe $MA(q)$, se asume que los ruidos blancos A_t se combinan linealmente para conformar una representación del modelo.

Definición 4.1.3: El modelo de *media móvil de orden q* , o modelo $MA(q)$, está definido como:

$$y_t = A_t + \theta_1A_{t-1} + \theta_2A_{t-2} + \dots + \theta_qA_{t-q} \quad (4.8)$$

donde hay q rezagos en la media móvil y $\theta_1, \theta_2, \dots, \theta_q$ ($\theta_q \neq 0$) son los parámetros con $\{A_t\} \sim iid(0, \sigma_A^2)$.

Utilizando el operador de media móvil, es posible escribir el proceso $MA(q)$ de forma equivalente:

$$y_t = \theta(B)A_t$$

donde $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$.

Obsérvese de (4.8) que y_t está correlacionada con y_{t-1} pero no con y_{t-2}, y_{t-3}, \dots . Esto contrasta con el caso de los modelos *AR* en el cual la correlación entre y_t e y_{t-k} nunca es cero. Nótese que por ser suma de $q + 1$ procesos estacionarios un modelo *MA*(q) siempre es estacionario.

Es posible escribir el modelo *MA*(1) como $y_t = \theta(B)A_t$, donde $\theta(B) = 1 + \theta_1 B$. Si $|\theta_1| < 1$ (lo que provoca que la innovación pasada influya menos que la actual), entonces se puede expresar el modelo como $\pi(B)y_t = A_t$, donde $\pi(B) = \theta^{-1}(B)$. Si se considera $\theta(z) = 1 + \theta_1 z$, para $|z| \leq 1$, entonces $\pi(z) = \theta^{-1}(z) = \frac{1}{1 + \theta_1 z} = \sum_{j=0}^{\infty} (-\theta_1)^j z^j$, y se determina que $\pi(B) = \sum_{j=0}^{\infty} (-\theta_1)^j B^j$ (Shumway & Stoffer, 2011). Esta forma se generaliza luego para señalar la condición de *invertibilidad* de un proceso.

Definición 4.1.4: Un proceso estocástico estacionario $\{y_t\}$ sigue un modelo *autorregresivo de media móvil* de orden (p, q) , o *ARMA*(p, q) (del inglés *Autoregressive Moving Average*), si y sólo si

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + A_t + \theta_1 A_{t-1} + \theta_2 A_{t-2} + \dots + \theta_q A_{t-q} \quad (4.9)$$

para todo $t = 0, \pm 1, \pm 2, \dots$, donde $\{A_t\} \sim iid(0, \sigma_A^2)$, con $\varphi_p \neq 0$, $\theta_q \neq 0$ y $\sigma_A^2 > 0$. Los parámetros p y q son llamados los órdenes autorregresivo y de media móvil, respectivamente.

Si y_t tiene una media μ diferente de cero, se toma $\alpha = \mu(1 - \varphi_1 - \dots - \varphi_p)$ y se escribe el modelo como:

$$y_t = \alpha + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + A_t + \theta_1 A_{t-1} + \theta_2 A_{t-2} + \dots + \theta_q A_{t-q} \quad (4.10)$$

Una forma alternativa de expresar la ecuación (4.9) es en términos de los operadores *AR* y *MA*:

$$\varphi(B)y_t = \theta(B)A_t \quad (4.11)$$

A continuación se define los polinomios autorregresivo y de media móvil en términos de la variable compleja z :

Definición 4.1.5: Los polinomios *AR* y *MA* están definidos como

$\varphi(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p$, $\varphi_p \neq 0$, y $\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$, $\theta_q \neq 0$, respectivamente, donde z es un número complejo.

Ya fue comentada en párrafos anteriores la noción de causalidad. Se introduce, ahora, formalmente para procesos *ARMA*.

Definición 4.1.6: Un modelo $ARMA(p, q)$ se dice que es *causal* si el proceso estocástico $\{y_t\}$ puede escribirse como

$$y_t = \sum_{j=0}^{\infty} \psi_j A_{t-j} = \psi(B)A_t, \quad (4.12)$$

donde $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$, y $\sum_{j=0}^{\infty} |\psi_j| < \infty$, con $\psi_0 = 1$. (Shumway & Stoffer, 2011)

Wold (1938) demostró que un proceso que es puramente no determinístico (un proceso que no contiene ningún componente determinista que se pueda pronosticar o predecir exactamente de su propio pasado), siempre puede expresarse en la forma (4.12), la cual se conoce como *Representación de Wold*. De este modo un proceso que puede ser expresado mediante (4.12) se conoce como *proceso no determinístico*.

En el caso del proceso autorregresivo de orden 1, $y_t = \varphi_1 y_{t-1} + A_t$, solamente es causal cuando $|\varphi_1| < 1$. De modo equivalente, es posible decir que el proceso es causal solo cuando la raíz de $\varphi(z) = 1 - \varphi_1 z$ es en valor absoluto mayor que 1. La raíz de $\varphi(z)$ es $z_0 = \frac{1}{\varphi_1}$ y $|z_0| > 1$ pues $|\varphi_1| < 1$. En general tenemos el siguiente teorema:

Teorema 4.1.1: Un modelo $ARMA(p, q)$ se dice que es *causal* si y solo si $\varphi(z) \neq 0$ para $|z| \leq 1$. Los coeficientes del proceso lineal dado en (4.12) pueden determinarse mediante la resolución de

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\varphi(z)} \text{ con } |z| \leq 1. \quad (4.13)$$

Demostración: Supóngase primero que las raíces z_1, \dots, z_p de $\varphi(z)$, están fuera del círculo unitario. Se escribe las raíces en el siguiente orden, $1 < |z_1| \leq |z_2| \leq \dots \leq |z_p|$, teniendo en cuenta que z_1, \dots, z_p no son necesariamente únicas, y se escribe $|z_1| = 1 + \varepsilon$, para $\varepsilon > 0$. Así, $\varphi(z) \neq 0$ siempre que $|z| < |z_1| = 1 + \varepsilon$ y, por lo tanto, $\varphi^{-1}(z)$ existe y tiene la expansión en series de potencia,

$$\frac{1}{\varphi(z)} = \sum_{j=0}^{\infty} a_j z^j \text{ con } |z| < 1 + \varepsilon.$$

Ahora se elige un valor δ tal que $0 < \delta < \varepsilon$, y se establece $z = 1 + \delta$, que está dentro del radio de convergencia. A continuación, se deduce que

$$\varphi^{-1}(1 + \delta) = \sum_{j=0}^{\infty} a_j (1 + \delta)^j < \infty \quad (4.14)$$

De este modo podemos acotar cada uno de los términos en la suma (4.14) por una constante, es decir, $|a_j (1 + \delta)^j| < c$, para $c > 0$. A su vez, $|a_j| < c \cdot (1 + \delta)^{-j}$, de donde se deduce que

$$\sum_{j=0}^{\infty} |a_j| < \infty \quad (4.15)$$

Por lo tanto, $\varphi^{-1}(B)$ existe y se puede aplicar a ambos lados del modelo $ARMA$, $\varphi(B)y_t = \theta(B)A_t$, para obtener

$$y_t = \varphi^{-1}(B) \cdot \varphi(B)y_t = \varphi^{-1}(B) \cdot \theta(B) \cdot A_t.$$

Así, considerando $\psi(B) = \varphi^{-1}(B)\theta(B)$, tenemos

$$y_t = \psi(B) \cdot A_t = \sum_{j=0}^{\infty} \psi_j A_{t-j} \quad (4.16)$$

donde los pesos ψ , que son absolutamente sumables, pueden ser evaluados por $\psi(z) = \varphi^{-1}(z) \cdot \theta(z)$, para $|z| \leq 1$. (Ver (4.12)).

Ahora, supóngase que y_t es un proceso causal; es decir, éste tiene la representación

$$y_t = \sum_{j=0}^{\infty} \psi_j A_{t-j} \quad \text{con} \quad \sum_{j=0}^{\infty} |\psi_j| < \infty$$

En este caso se escribe

$$y_t = \psi(B)A_t$$

y multiplicando por $\varphi(B)$ resulta

$$\varphi(B)y_t = \varphi(B)\psi(B)A_t \quad (4.17)$$

En suma, para (4.17), el modelo es $ARMA$ y puede escribirse como

$$\varphi(B)y_t = \theta(B)A_t \quad (4.18)$$

De (4.17) y (4.18), se tiene

$$\varphi(B)\psi(B)A_t = \theta(B)A_t \quad (4.19)$$

Ahora, sea

$$a(z) = \varphi(z)\psi(z) = \sum_{j=0}^{\infty} a_j z^j \quad \text{con} \quad |z| \leq 1$$

y, por lo tanto, se escribe (4.19) como

$$\sum_{j=0}^{\infty} a_j A_{t-j} = \sum_{j=0}^q \theta_j A_{t-j} \quad (4.20)$$

A continuación, se multiplica ambos lados de (4.20) por A_{t-h} , para $h = 0, 1, 2, \dots$. Al hacer esto se obtiene

$$\begin{aligned} a_h &= \theta_h \quad \text{con} \quad h = 0, 1, \dots, q \\ a_h &= 0 \quad \text{con} \quad h > q \end{aligned} \quad (4.21)$$

De (4.21) se concluye que

$$\varphi(z)\psi(z) = a(z) = \theta(z) \text{ con } |z| \leq 1 \quad (4.22)$$

Si existe un número complejo en el círculo unitario, z_0 , para el cual $\varphi(z_0) = 0$, entonces por (4.22), $\theta(z_0) = 0$. Pero si tal z_0 existe, entonces $\varphi(z)$ y $\theta(z)$ tienen un factor común que no está permitido. Así, se puede escribir $\psi(z) = \frac{\theta(z)}{\varphi(z)}$. En suma, por hipótesis, tenemos que $|\psi(z)| < \infty$ para $|z| \leq 1$, y por lo tanto

$$|\psi(z)| = \left| \frac{\theta(z)}{\varphi(z)} \right| < \infty, \text{ para } |z| \leq 1 \quad (4.23)$$

Finalmente, (4.23) implica $\varphi(z) \neq 0$ para $|z| \leq 1$. Esto es, las raíces de $\varphi(z)$ están fuera del círculo unitario. (Shumway & Stoffer, 2011)

Del mismo modo que la representación (4.12) establece una forma de media móvil tal que es posible escribir en términos de la misma cualquier proceso causal, la siguiente expresión autorregresiva define una forma fundamental para los estudios de predicción en base a los modelos $ARMA(p, q)$:

Definición 4.1.7: Un modelo $ARMA(p, q)$ se dice que es *invertible*, si las series pueden escribirse como

$$\pi(B)y_t = \sum_{j=0}^{\infty} \pi_j y_{t-j} = A_t \quad (4.24)$$

donde $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$ y $\sum_{j=0}^{\infty} |\pi_j| < \infty$ con $\pi_0 = 1$.

Box y Jenkins (1976) argumentaron que en la predicción, un proceso no invertible carece de sentido. Desde ya que no todo proceso estacionario es invertible (ver ejemplo 4.1.4). Se presenta a continuación un teorema que establece la condición para que un modelo $ARMA$ sea invertible. La demostración se parece a la del teorema 4.1.1:

Teorema 4.1.2: Un modelo $ARMA(p, q)$ es *invertible* sí y sólo si $\theta(z) \neq 0$ para $|z| \leq 1$. Los coeficientes π_j de $\pi(B)$ dados en (5.20) pueden determinarse mediante la resolución de

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\varphi(z)}{\theta(z)} \text{ con } |z| \leq 1. \quad (4.25)$$

Ejemplo 4.1.3: Sea el siguiente modelo $ARMA(1,1)$ dado por la ecuación

$$y_t - 0,5y_{t-1} = A_t + 0,4A_{t-1}$$

donde A_t es un ruido blanco con media 0 y varianza σ_A^2 . En términos del operador de retardos, el modelo queda escrito en la forma:

$$(1 - 0,5B)y_t = (1 + 0,4B)A_t$$

Por lo tanto, los polinomios autorregresivo y de media móvil en variable compleja son:

$$\varphi(z) = 1 - 0,5z \quad \text{y} \quad \theta(z) = 1 + 0,4z$$

Es inmediato que $\varphi(z) = 0$ cuando $z = 2$ (fuera del círculo unitario), y $\theta(z) = 0$ si $z = -2,5$ (fuera del círculo unitario), por lo tanto se cumplen las condiciones de causalidad e invertibilidad. Los coeficientes ψ_j de la representación $MA(\infty)$ pueden determinarse directamente mediante el teorema 5.1.1, teniendo en cuenta que $\varphi(z)\psi(z) = \theta(z)$:

$$(1 - 0,5z)(\psi_0 + \psi_1z + \psi_2z^2 + \dots) = 1 + 0,4z$$

Distribuyendo se tiene:

$$\psi_0 + \psi_1z + \psi_2z^2 + \dots - 0,5\psi_0z - 0,5\psi_1z^2 - 0,5\psi_2z^3 - \dots = 1 + 0,4z$$

Agrupando y factorizando las potencias de z queda:

$$\psi_0 + (\psi_1 - 0,5\psi_0)z + (\psi_2 - 0,5\psi_1)z^2 + (\psi_3 - 0,5\psi_2)z^3 + \dots = 1 + 0,4z$$

De esta igualdad, se obtiene: $\psi_0 = 1$ y $\psi_1 - 0,5\psi_0 = 0,4$, por lo tanto $\psi_1 = 0,4 + 0,5 = 0,9$. Además, $\psi_2 - 0,5\psi_1 = \psi_2 - 0,5(0,4 + 0,5) = 0$, de donde $\psi_2 = 0,5(0,4 + 0,5) = 0,9 \cdot 0,5$. También, $\psi_3 - 0,5\psi_2 = \psi_3 - 0,5 \cdot 0,5(0,4 + 0,5) = 0$, y así $\psi_3 = 0,9 \cdot (0,5)^2$. Continuando con este razonamiento, $\psi_j = 0,9 \cdot (0,5)^{j-1}$. Por lo tanto:

$$\psi(z) = 1 + 0,9 \sum_{j=1}^{\infty} (0,5)^{j-1} z^j$$

Por (4.12) el modelo $MA(\infty)$ para y_t puede escribirse como

$$y_t = A_t + 0,9 \sum_{j=0}^{\infty} (0,5)^{j-1} A_{t-j}$$

De modo similar, puede obtenerse los coeficientes π_j de la representación $AR(\infty)$ mediante el Teorema 5.1.2, teniendo en cuenta que $\theta(z)\pi(z) = \varphi(z)$:

$$(1 + 0,4z)(\psi_0 + \psi_1z + \psi_2z^2 + \dots) = 1 - 0,5z$$

De aquí resulta:

$$\psi_0 + \psi_1z + \psi_2z^2 + \dots + 0,4\psi_0z + 0,4\psi_1z^2 + 0,4\psi_2z^3 + \dots = 1 - 0,5z$$

Agrupando y factorizando:

$$\psi_0 + (\psi_1 + 0,4\psi_0)z + (\psi_2 + 0,4\psi_1)z^2 + (\psi_3 + 0,4\psi_2)z^3 + \dots = 1 - 0,5z$$

Por igualación de coeficientes, se tiene $\psi_0 = 1$ y de $\psi_1 + 0,4\psi_0 = \psi_1 + 0,4 = -0,5$ resulta $\psi_1 = -0,5 - 0,4 = -0,9$. Además, $\psi_2 + 0,4\psi_1 = \psi_2 + 0,4 \cdot (-0,9) = 0$, de donde $\psi_2 = 0,4 \cdot 0,9$. De la misma manera, $\psi_3 + 0,4\psi_2 = \psi_3 + 0,4 \cdot 0,4 \cdot 0,9 = -0,9 \cdot (0,4)^2$. Continuando con este procedimiento:

$$\psi(z) = 1 + 0,9 \sum_{j=1}^{\infty} (-1)^j (0,4)^{j-1} z^j$$

De (5.20), el modelo $AR(\infty)$ para y_t queda:

$$y_t = 0,9 \sum_{j=1}^{\infty} (-1)^j (0,4)^{j-1} y_{t-j} + A_t$$

Ejemplo 4.1.4: Sea el siguiente modelo $ARMA(2,1)$ dado por la siguiente ecuación:

$$y_t - 0,75y_{t-1} + 0,5625y_{t-2} = A_t + 1,25A_{t-1}$$

En términos del operador de retardos:

$$(1 - 0,75B + 0,5625B^2)y_t = (1 + 1,25B)A_t$$

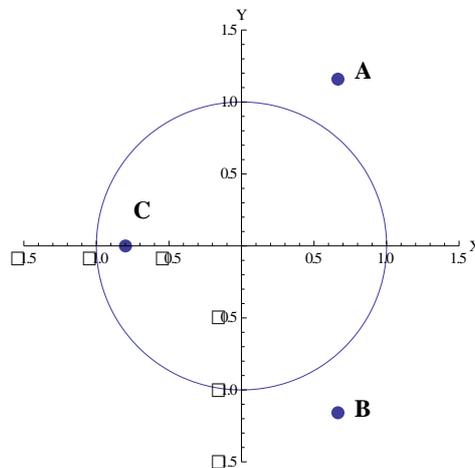
Por lo tanto, los polinomios autorregresivo y de media móvil en variable compleja son:

$$\varphi(z) = 1 - 0,75z + 0,5625z^2 \quad y \quad \theta(z) = 1 + 1,25z$$

Las soluciones para $\varphi(z) = 0$ son $z_1 = \frac{2}{3} + \frac{2}{3}\sqrt{3}i$ (punto A en la figura 4.3) y $z_2 = \frac{2}{3} - \frac{2}{3}\sqrt{3}i$ (punto B en el gráfico 4.3). Dado que las soluciones de la ecuación están fuera del círculo unitario, el modelo es causal.

Por otro lado, la solución de la ecuación $\theta(z) = 0$ es $z = -0,8$ (punto C del gráfico 4.3). Como la solución se encuentra dentro del círculo unitario ($|z| = |-0,8| < 1$), el modelo no es invertible.

Gráfico 4.3



Ejemplo 4.1.5: Considérese el siguiente modelo dado por la ecuación:

$$y_t = 0,7y_{t-1} - 0,1y_{t-2} + A_t - 0,4A_{t-1} + 0,04A_{t-2}$$

Utilizando el operador de retardo, resulta

$$y_t = 0,7By_t - 0,1B^2y_t + A_t - 0,4BA_t + 0,04B^2A_t$$

Así

$$y_t(1 - 0,7B + 0,1B^2) = A_t(1 - 0,4B + 0,04B^2)$$

Aunque el proceso tiene apariencia de un $ARMA(2,2)$, al considerar los polinomios complejos asociados, se observa que $\varphi(z) = 1 - 0,7z + 0,1z^2 = 0,1(z - 5)(z - 2)$ y $\theta(z) = 1 - 0,4z + 0,04z^2 = 0,04(z - 5)^2$ tienen el factor $(z - 5)$ en común. Por lo tanto, en términos del operador de retardos:

$$y_t 0,1(B - 2) = A_t 0,04(B - 5)$$

O bien $0,1By_t - 0,2y_t = 0,04A_tB - 0,2A_t$. Así, se tiene:

$$y_t = 0,5y_{t-1} - 0,2A_{t-1} + A_t$$

Es posible ver que el modelo es causal porque $\varphi(z) = 0,1z - 0,2 = 0$ cuando $z = 2$, que está fuera del círculo unitario. Además, es invertible ya que la raíz de $\theta(z) = 0,04z - 0,2$ es $z = 5$, que también está fuera del círculo unitario.

Procediendo del mismo modo que en el ejemplo 5.1.2, se obtiene

$$\psi(z) = 1 + 0,3 \sum_{j=1}^{\infty} (0,5)^{j-1} z^j$$

La representación $MA(\infty)$ para el modelo es

$$y_t = A_t + 0,3 \sum_{j=1}^{\infty} (0,5)^{j-1} A_{t-j}$$

Además, de modo similar se obtiene

$$\pi(z) = 1 - 0,3 \sum_{j=1}^{\infty} (0,2)^{j-1} z^j$$

La forma $AR(\infty)$ queda

$$y_t = 0,3 \sum_{j=1}^{\infty} (0,2)^{j-1} y_{t-j} + A_t.$$

Cabe mencionar que una condición necesaria para que se verifique que todas las raíces del polinomio característico de un $AR(p)$ sean mayores que la unidad es que $\varphi_1 + \varphi_2 + \dots + \varphi_p < 1$. Como los φ_j pueden ser positivos o negativos, una condición suficiente para esta propiedad es que $|\varphi_1| + |\varphi_2| + \dots + |\varphi_p| < 1$. Si se verifica que $\varphi_1 + \varphi_2 + \dots + \varphi_p = 1$, se puede asegurar que la ecuación en diferencias que rige el comportamiento de la estructura temporal del proceso, $y_t - \varphi_1 y_{t-1} - \varphi_2 y_{t-2} - \dots - \varphi_p y_{t-p} = 0$, posee al menos una raíz unitaria. (Landro & González, 2009)

4.2. Funciones de autocovarianza, autocorrelación y autocorrelación parcial de los procesos ARMA

La información que aportan las funciones de autocorrelación y autocorrelación parcial resulta de fundamental importancia para decidir el tipo de proceso generador del conjunto de datos que conforman la serie de tiempo.

Considérese en principio un modelo $AR(p)$ dado por la ecuación:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + A_t$$

Multiplicando cada término por y_{t-k} , con $k > 0$, resulta:

$$y_t y_{t-k} = \varphi_1 y_{t-1} y_{t-k} + \varphi_2 y_{t-2} y_{t-k} + \cdots + \varphi_p y_{t-p} y_{t-k} + A_t y_{t-k}$$

Teniendo en cuenta las propiedades de la esperanza:

$$E[y_t y_{t-k}] = \varphi_1 E[y_{t-1} y_{t-k}] + \varphi_2 E[y_{t-2} y_{t-k}] + \cdots + \varphi_p E[y_{t-p} y_{t-k}] + E[A_t y_{t-k}]$$

Si se toma $E[y_t] = 0$, se tiene $E[y_{t-j} y_{t-k}] = \gamma_{k-j}$ para $j = 0, \dots, p$. Por lo tanto:

$$\gamma_k = \varphi_1 \gamma_{k-1} + \varphi_2 \gamma_{k-2} + \cdots + \varphi_p \gamma_{k-p} + E[A_t y_{t-k}]$$

Por (4.12) puede escribirse $y_{t-k} = \sum_{j=0}^{\infty} \psi_j A_{t-k-j}$ y así:

$$E[A_t y_{t-k}] = E \left[A_t \sum_{j=0}^{\infty} \psi_j A_{t-k-j} \right] = 0$$

Luego, la función de autocovarianza del proceso puede expresarse como

$$\gamma_k = \varphi_1 \gamma_{k-1} + \varphi_2 \gamma_{k-2} + \cdots + \varphi_p \gamma_{k-p} \quad (4.26)$$

Si se divide cada término de (4.26) por γ_0 se obtiene:

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + \cdots + \varphi_p \rho_{k-p}$$

O bien, como sistema de ecuaciones en diferencias, resulta:

$$\rho_k - \varphi_1 \rho_{k-1} - \varphi_2 \rho_{k-2} - \cdots - \varphi_p \rho_{k-p} = 0 \text{ para } k > p \quad (4.27)$$

El estudio de las características de (4.27), debe hacerse a la luz de las soluciones asociadas a las ecuaciones homogéneas en diferencias.

Si se tiene una ecuación homogénea en diferencias de primer orden: $\rho_k - \varphi_1 \rho_{k-1} = 0$ (4.28), con $k = 1, 2, \dots$, y $\varphi_1 \neq 0$, entonces $\rho_1 = \varphi_1 \rho_0$ y, por lo tanto para la siguiente expresión, resulta $\rho_2 = \varphi_1 \rho_1 = \varphi_1^2 \rho_0$. Siguiendo este esquema, $\rho_k = \varphi_1^k \rho_0$. Con condición inicial $\rho_0 = c$: $\rho_k = \varphi_1^k c$ (4.29).

Si se escribe (4.28) en términos del operador de retardos queda $(1 - \varphi_1 B)\rho_k = 0$. Por lo tanto, el polinomio asociado es $\varphi_1(z) = 1 - \varphi_1 z$. Es inmediato que $z_0 = \frac{1}{\varphi_1}$ satisface $\varphi(z) = 0$. Como $\varphi_1 = z_0^{-1}$, en (4.29): $\rho_k = (z_0^{-1})^k c = z_0^{-k} c$, que es la solución de (4.28) con condición inicial $\rho_0 = c$. Es inmediato que esta solución solo depende de la condición inicial y de la raíz inversa del polinomio asociado.

Considérese ahora la ecuación de segundo orden: $\rho_k - \varphi_1 \rho_{k-1} - \varphi_2 \rho_{k-2} = 0$ (4.30), con $\varphi_2 \neq 0$ y $k = 2, 3, \dots$. El polinomio asociado es $\varphi(z) = 1 - \varphi_1 z - \varphi_2 z^2$, el cual tiene dos raíces z_1 y z_2 .

Si $z_1 \neq z_2$, la solución general¹² de (4.30) es $\rho_k = c_1 z_1^{-k} + c_2 z_2^{-k}$ (4.31), donde c_1 y c_2 dependen de las condiciones iniciales dadas ρ_0 y ρ_1 . Puede verificarse esta solución fácilmente por sustitución en (4.30):

$$\begin{aligned} & (c_1 z_1^{-k} + c_2 z_2^{-k}) - \varphi_1 (c_1 z_1^{-(k-1)} + c_2 z_2^{-(k-1)}) - \varphi_2 (c_1 z_1^{-(k-2)} + c_2 z_2^{-(k-2)}) \\ &= c_1 z_1^{-k} + c_2 z_2^{-k} - \varphi_1 c_1 z_1^{-k} z_1 - \varphi_1 c_2 z_2^{-k} z_2 - \varphi_2 c_1 z_1^{-k} z_1^2 - \varphi_2 c_2 z_2^{-k} z_2^2 \\ &= c_1 z_1^{-k} (1 - \varphi_1 z_1 - \varphi_2 z_1^2) + c_2 z_2^{-k} (1 - \varphi_1 z_2 - \varphi_2 z_2^2) \\ &= c_1 z_1^{-k} \varphi(z_1) + c_2 z_2^{-k} \varphi(z_2) = 0. \end{aligned}$$

De este modo, las condiciones iniciales permitirán calcular c_1 y c_2 , pues $\rho_0 = c_1 + c_2$ y $\rho_1 = c_1 z_1^{-1} + c_2 z_2^{-1}$. Nótese que (4.31) es la combinación lineal de exponenciales decrecientes. Por lo tanto, ρ_k tiende a cero exponencialmente cuando $k \rightarrow \infty$.

En el caso en que $z_1 = z_2$, supóngase $z_1 = z_2 = z_0$, la solución general a (4.30) es $\rho_k = c_1 z_0^{-k} + c_2 k z_0^{-k} = z_0^{-k} (c_1 + c_2 k)$ (4.32), que también puede verificarse por sustitución en (4.30). Dadas las condiciones iniciales ρ_0 y ρ_1 resulta $\rho_0 = c_1$ y $\rho_1 = (c_1 + c_2) z_0^{-1}$. También, en este caso, ρ_k tiende a cero exponencialmente cuando $k \rightarrow \infty$.

Puede ocurrir, también, que las raíces sean un par de complejos conjugados, $z_1 = \bar{z}_2$. En este caso, puesto que ρ_k es real, $c_2 = \bar{c}_1$ y la solución a (4.30) es $\rho_k = c_1 z_1^{-k} + \bar{c}_1 \bar{z}_1^{-k}$ (4.33).

Si se considera la forma exponencial de un complejo, entonces $z_1 = |z_1| e^{i\alpha}$ y $c_1 = |c_1| e^{i\beta}$ donde α y β son, respectivamente, los argumentos de z_1 y c_1 . En este caso, se tendrá $\bar{z}_1 = |z_1| e^{-i\alpha}$ y $\bar{c}_1 = |c_1| e^{-i\beta}$, teniendo en cuenta que los argumentos principales¹³ de los conjugados son opuestos y que el módulo de un complejo y su conjugado son iguales. Sustituyendo en (4.33), resulta:

$$\rho_k = |c_1| e^{i\beta} |z_1|^{-k} e^{-ik\alpha} + |c_1| e^{-i\beta} |z_1|^{-k} e^{ik\alpha} = |c_1| |z_1|^{-k} [e^{i(k\alpha - \beta)} + e^{-i(k\alpha - \beta)}]$$

Teniendo en cuenta que $2 \cos \varepsilon = e^{i\varepsilon} + e^{-i\varepsilon}$, queda:

¹² Es importante tener en cuenta un lema asociado a las soluciones de las ecuaciones homogéneas en diferencias: Lema: Si z_1 y z_2 son soluciones de una ecuación homogénea, $c_1 z_1 + c_2 z_2$ es también solución de la ecuación, donde c_1 y c_2 son constantes arbitrarias. (Wei, 2006)

¹³ En la representación polar de un complejo, $z = |z| e^{i\theta}$, el número θ se llama argumento de z , denotado como $\arg(z)$, y es el ángulo, medido en radianes, que forma z con el eje real positivo cuando el complejo se interpreta como un radio-vector. Toma cualquier valor de entre infinitos posibles, que difieren dos a dos en múltiplos de 2π . El valor principal del argumento de z , se define como el único valor del argumento de z tal que $-\pi < \arg(z) \leq \pi$. (Churchill & Brown, 1992)

$$\rho_k = 2|c_1||z_1|^{-k} \cos(k\alpha - \beta)$$

Esta expresión permite ver que si $k \rightarrow \infty$, entonces ρ_k tiende a cero, decreciendo en forma sinusoidal amortiguada.

En general, si z_1, \dots, z_r denotan las raíces de $\varphi(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p$, que es el polinomio asociado a (4.27), cada una con multiplicidad m_1, \dots, m_r respectivamente, donde $m_1 + \dots + m_r = p$, entonces la solución para (4.27) es

$$\rho_k = z_1^{-k} P_1(k) + z_2^{-k} P_2(k) + \dots + z_r^{-k} P_r(k) \text{ con } k > p \quad (4.34)$$

donde $P_j(k)$ es un polinomio en k de grado $m_j - 1$. (Shumway & Stoffer, 2011)

Nótese que (4.34) está definida por una suma de funciones exponenciales decrecientes. Si el modelo estudiado es causal, todas las raíces estarán fuera del círculo unitario, $|z_i| > 1$, con $i = 1, \dots, r$. En caso de que todas las raíces sean reales, ρ_k tiende hacia cero exponencialmente cuando $k \rightarrow \infty$. Si alguna de las raíces son complejas, entonces estarán en pares conjugados y ρ_k decrecerá hacia cero pero en forma sinusoidal amortiguada cuando $k \rightarrow \infty$. En el caso en el que todas las raíces son complejas, se evidenciará un comportamiento de naturaleza cíclica.

En el caso de la función de autocorrelación parcial de un proceso autorregresivo, por (4.27), puede verse fácilmente que la última columna de la matriz en el numerador de ϕ_{kk} en (3.4) puede escribirse como una combinación lineal de las columnas previas de la misma matriz. Así, la *PACF* se hará nula después del retardo p .

Para un *AR*(1), por ejemplo, la *PACF* queda

$$\phi_{kk} = \begin{cases} \rho_1 = \varphi_1 & \text{si } k = 1 \\ 0 & \text{si } k \geq 2 \end{cases}$$

Si el proceso es un *AR*(2) se sabe que $\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2}$ para $k \geq 1$, es inmediato que

$$\phi_{11} = \rho_1 = \frac{\varphi_1}{1 - \varphi_2}$$

De (3.4):

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \varphi_2$$

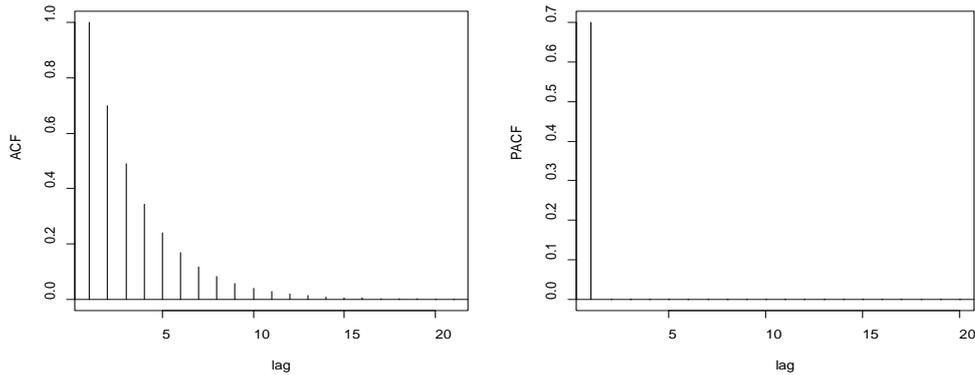
Por último,

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & \rho_1 & \varphi_1 + \varphi_2 \rho_1 \\ \rho_1 & 1 & \varphi_1 \rho_1 + \varphi_2 \\ \rho_2 & \rho_1 & \varphi_1 \rho_2 + \varphi_2 \rho_1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = 0.$$

Este último resultado se cumple para φ_{kk} cuando $k \geq 3$, puesto que la última columna del determinante del numerador se puede escribir como una combinación lineal de las columnas anteriores. (Wei, 2006)

Ejemplo 4.2.1: En a figura inferior se observan la *ACF* y la *PACF* del modelo *AR(1)* $y_t = 0,7y_{t-1} + A_t$. Nótese el decrecimiento exponencial de la función de autocorrelación y el comportamiento truncado de la función de autocorrelación parcial.

Gráfico 4.4



Ejemplo 4.2.2: Tómesese ahora en consideración el siguiente modelo *AR(2)* dado por la ecuación:

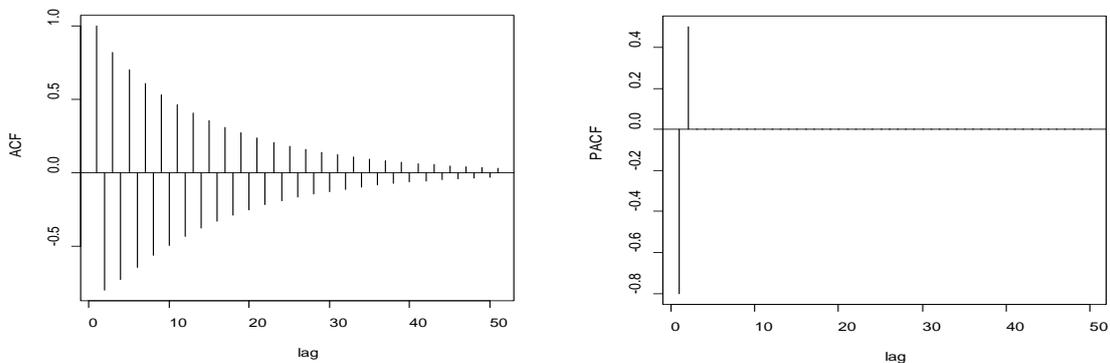
$$y_t = -0,4y_{t-1} + 0,5y_{t-2} + A_t$$

En términos del operador de retardos, resulta:

$$(1 + 0,4B - 0,5B^2)y_t = A_t$$

Por lo tanto, el polinomio autorregresivo en variable compleja es $\varphi(z) = -0,5z^2 + 0,4z + 1$. Los valores que satisfacen la ecuación $\varphi(z) = 0$ son los reales $z_1 = -\frac{1}{5}(2 - 3\sqrt{6})$ y $z_2 = -\frac{1}{5}(2 + 3\sqrt{6})$. Se presenta en la siguiente figura el gráfico de los correlogramas de la *ACF* y de la *PACF*.

Gráfico 4.5

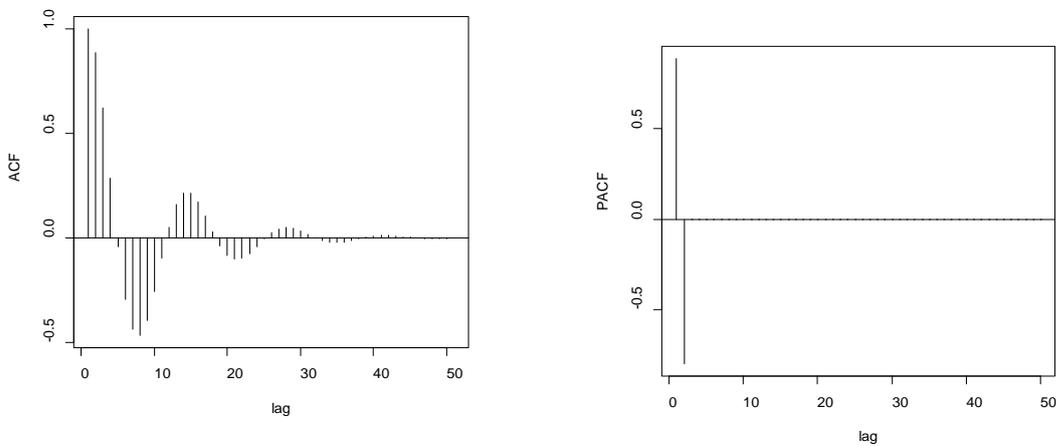


Ejemplo 4.2.3: Sea $y_t = 1,6y_{t-1} - 0,8y_{t-2} + A_t$ que es un $AR(2)$. En términos del operador de retardos puede escribirse el modelo como:

$$(1 - 1,6B + 0,8B^2)y_t = A_t$$

De aquí se deduce que $\varphi(z) = 0,8z^2 - 1,6z + 1$. Las raíces del polinomio autorregresivo $\varphi(z)$ son los complejos conjugados $z_1 = 1 + \frac{\sqrt{17}}{8}i$ y $z_2 = 1 - \frac{\sqrt{17}}{8}i$. La siguiente figura muestra los correlogramas de las funciones de autocorrelación y autocorrelación parcial. Nótese que la ACF presenta un decrecimiento con forma sinusoidal amortiguada, mientras que la $PACF$ se trunca después del retardo $k = 2$.

Gráfico 4.6



Considérese ahora un proceso $MA(q)$ dado por la ecuación

$$y_t = A_t + \theta_1 A_{t-1} + \theta_2 A_{t-2} + \dots + \theta_q A_{t-q} = \sum_{j=0}^q \theta_j A_{t-j}$$

Es inmediato que $E[y_t] = \sum_{j=0}^q \theta_j E[A_{t-j}] = 0$, puesto que y_t es una combinación lineal de ruidos blancos con media cero y varianza σ_A^2 . Por lo tanto,

$$\gamma_k = Cov(y_{t+k}, y_t) = Cov\left(\sum_{j=0}^q \theta_j A_{t+k-j}; \sum_{h=0}^q \theta_h A_{t-h}\right)$$

Por la definición de covarianza, se tiene:

$$\gamma_k = E\left[\left(\sum_{j=0}^q \theta_j A_{t+k-j}\right)\left(\sum_{h=0}^q \theta_h A_{t-h}\right)\right]$$

Desarrollando el producto $\left(\sum_{j=0}^q \theta_j A_{t+k-j}\right)\left(\sum_{h=0}^q \theta_h A_{t-h}\right)$ y tomando las esperanzas de cada término, es inmediato que

$$\gamma_k = \begin{cases} \sigma_A^2 \sum_{j=0}^{q-k} \theta_j \theta_{j+k} & \text{si } 0 \leq k \leq q \\ 0 & \text{si } k > q \end{cases}$$

Téngase en cuenta que sólo se considera $k \geq 0$ porque $\gamma_k = \gamma_{-k}$. (Shumway & Stoffer, 2011)

Dividiendo γ_k por γ_0 se tiene

$$\rho_k = \begin{cases} \frac{\sum_{j=0}^{q-k} \theta_j \theta_{j+k}}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_j^2} & \text{si } 0 \leq k \leq q \\ 0 & \text{si } k > q \end{cases}$$

Es posible ver que la función de autocorrelación de un $MA(q)$ es una función truncada en el retardo q . Esta importante propiedad permite identificar cuándo una serie de tiempo es generada por un proceso de media móvil.

Sea un proceso de media móvil de orden 1 dado por la ecuación $y_t = A_t + \theta_1 A_{t-1}$, con $|\theta_1| < 1$. Se sabe que la *función de autocovarianza* del proceso viene dada por la expresión:

$$\gamma_k = E[(y_t - E[y_t])(y_{t-k} - E[y_{t-k}])], \text{ con } k = 0, 1, 2, \dots$$

Como $E[y_t] = E[A_t + \theta_1 A_{t-1}] = E[A_t] + \theta_1 E[A_{t-1}] = 0$ resulta:

$$\gamma_k = E[(y_t - 0)(y_{t-k} - 0)] = E[y_t \cdot y_{t-k}] = E[(A_t + \theta_1 A_{t-1})(A_{t-k} + \theta_1 A_{t-k-1})].$$

Entonces $\gamma_k = E[A_t A_{t-k} + \theta_1 A_t A_{t-k-1} + \theta_1 A_{t-1} A_{t-k} + \theta_1^2 A_{t-1} A_{t-k-1}]$, y por la linealidad de la esperanza:

$$\gamma_k = E[A_t A_{t-k}] + \theta_1 E[A_t A_{t-k-1}] + \theta_1 E[A_{t-1} A_{t-k}] + \theta_1^2 E[A_{t-1} A_{t-k-1}].$$

Ahora bien, si $k = 0$ resulta:

$$\gamma_0 = E[A_t A_t] + \theta_1 E[A_t A_{t-1}] + \theta_1 E[A_{t-1} A_t] + \theta_1^2 E[A_{t-1} A_{t-1}]$$

$$\gamma_0 = E[A_t^2] + 2\theta_1 E[A_t A_{t-1}] + \theta_1^2 E[A_{t-1}^2].$$

Como la secuencia $\{A_t\}$ constituye un conjunto de variables aleatorias independientes e idénticamente distribuidas y, por lo tanto, la esperanza de cualquier productoria del conjunto es igual a la productoria de las esperanzas de cualesquiera dos variables aleatorias de la secuencia¹⁴, resulta:

$E[A_t A_{t-1}] = E[A_t] E[A_{t-1}] = 0$, y en consecuencia:

$$\gamma_0 = E[A_t^2] + \theta_1^2 E[A_{t-1}^2] = \sigma_A^2 + \theta_1^2 \sigma_A^2 = \sigma_A^2 (1 + \theta_1^2)$$

Por otro lado, si $k = 1$:

$$\gamma_1 = E[A_t A_{t-1}] + \theta_1 E[A_t A_{t-2}] + \theta_1 E[A_{t-1} A_{t-1}] + \theta_1^2 E[A_{t-1} A_{t-2}]$$

¹⁴ Una demostración de este resultado puede encontrarse en Degroot, Morris H., (1988), Probabilidad y Estadística, Addison.Wesley Iberoamericana, Wilmington, U.S.A. pág. 182.

Además, $E[A_t A_{t-1}] = E[A_t A_{t-2}] = E[A_{t-1} A_{t-2}] = 0$ y $E[A_{t-1} A_{t-1}] = E[A_{t-1}^2] = \sigma_A^2$ resulta:

$$\gamma_1 = \theta_1 \sigma_A^2$$

Si tomamos $k = 2$ tendremos:

$$\gamma_2 = E[A_t A_{t-2}] + \theta_1 E[A_t A_{t-3}] + \theta_1 E[A_{t-1} A_{t-2}] + \theta_1^2 E[A_{t-1} A_{t-3}] = 0$$

pues todas las esperanzas son nulas. Si se considera valores superiores a 2 para k , esto es, $k = 3, 4, 5, \dots$, puede comprobarse que $\gamma_k = 0$.

Luego, la función de autocovarianza del proceso es:

$$\gamma_k = \begin{cases} \sigma_A^2(1 + \theta_1^2) & \text{si } k = 0 \\ \theta_1 \sigma_A^2 & \text{si } k = 1 \\ 0 & \text{si } k \geq 2 \end{cases}$$

Para determinar la función de autocorrelación simple del proceso, partimos de que

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

Si $k = 0$, entonces

$$\rho_0 = \frac{\gamma_0}{\gamma_0} = 1$$

Si $k = 1$, resulta

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\theta_1 \sigma_A^2}{\sigma_A^2(1 + \theta_1^2)} = \frac{\theta_1}{(1 + \theta_1^2)}$$

Para $k = 2, 3, \dots$ se tiene

$$\rho_k = \frac{0}{\gamma_0} = 0 \text{ pues } \gamma_2 = \gamma_3 = \dots = 0$$

Luego, la función de autocorrelación simple, ACF, es

$$\rho_k = \begin{cases} 1 & \text{si } k = 0 \\ \frac{\theta_1}{(1 + \theta_1^2)} & \text{si } k = 1 \\ 0 & \text{si } k \geq 2 \end{cases}$$

Obsérvese que se trunca para el retardo $k \geq 2$.

Se determinará ahora, los primeros tres coeficientes de la función de autocorrelación parcial, PACF. En principio:

$$\phi_{11} = \rho_1 = \frac{\theta_1}{1 + \theta_1^2} = \frac{\theta_1(1 - \theta_1^2)}{(1 + \theta_1^2)(1 - \theta_1^2)} = \frac{\theta_1(1 - \theta_1^2)}{1 - \theta_1^4}$$

El segundo coeficiente viene dado por:

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{-\rho_1^2}{1 - \rho_1^2} = -\frac{\frac{\theta_1^2}{(1 + \theta_1^2)^2}}{1 - \frac{\theta_1^2}{(1 + \theta_1^2)^2}} = -\frac{\frac{\theta_1^2}{(1 + \theta_1^2)^2}}{\frac{(1 + \theta_1^2)^2 - \theta_1^2}{(1 + \theta_1^2)^2}}$$

Por lo tanto,

$$\phi_{22} = -\frac{\theta_1^2}{1 + \theta_1^2 + \theta_1^4} = -\frac{\theta_1^2(1 - \theta_1^2)}{(1 + \theta_1^2 + \theta_1^4)(1 - \theta_1^2)} = -\frac{\theta_1^2(1 - \theta_1^2)}{1 - \theta_1^6}$$

Por último:

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & 0 \\ 0 & \rho_1 & 0 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & 0 \\ 0 & \rho_1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & 0 \\ \rho_1 & 1 & \rho_1 \\ 0 & \rho_1 & 1 \end{vmatrix}} = \frac{\rho_1(\rho_1^2 - 0)}{1(1 - \rho_1^2) - \rho_1(\rho_1 - 0)} = \frac{\rho_1^3}{1 - 2\rho_1^2}$$

De donde, reemplazando y trabajando algebraicamente, se obtiene:

$$\phi_{33} = \frac{\frac{\theta_1^3}{(1 + \theta_1^2)^3}}{1 - 2\frac{\theta_1^2}{(1 + \theta_1^2)^2}} = \frac{\theta_1^3}{(1 + \theta_1^4)(1 + \theta_1^2)} = \frac{\theta_1^3(1 - \theta_1^2)}{(1 + \theta_1^4)(1 + \theta_1^2)(1 - \theta_1^2)} = \frac{\theta_1^3(1 - \theta_1^2)}{(1 - \theta_1^8)}$$

En general, se tiene:

$$\phi_{kk} = \frac{-(-\theta)^h(1 - \theta_1^2)}{1 - \theta_1^{2(k+1)}} \text{ para } k \geq 1$$

Es inmediato observar que la *PACF* de un *MA*(1) decrece exponencialmente en una de dos formas dependiendo del signo de θ_1 (consecuentemente de ρ_1). Si $\theta_1 > 0$, la *PACF* crece desde valores negativos en forma exponencial hacia cero. Si $\theta_1 < 0$, la *PACF* comienza con un valor positivo y decrece hacia cero alternadamente (en forma exponencial) tomando valores positivos y negativos.

Ejemplo 4.2.4: Sean los siguientes modelos de media móvil de primer orden: $y_t = A_t + 0,85A_{t-1}$ de parámetro positivo, e $y_t = A_t - 0,85A_{t-1}$ de parámetro negativo. En la figura siguiente puede verse las características descritas para la *ACF* y la *PACF* de acuerdo con el signo del parámetro (en la parte superior las correspondientes al parámetro positivo y en la inferior las de parámetro negativo).

Gráfico 4.7

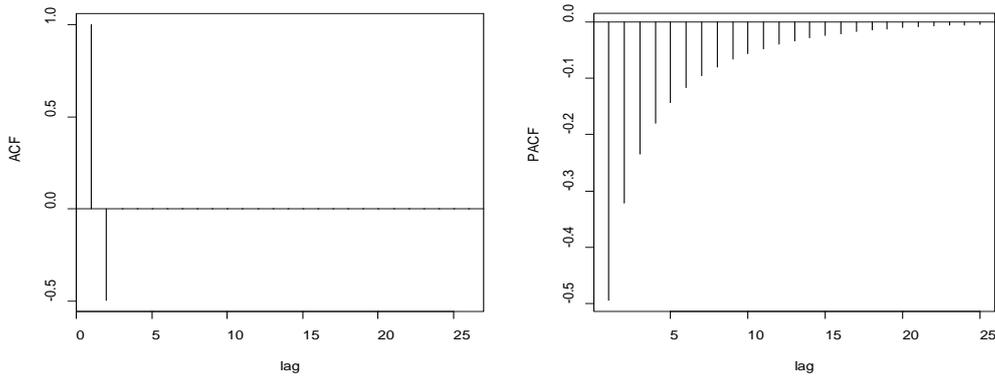
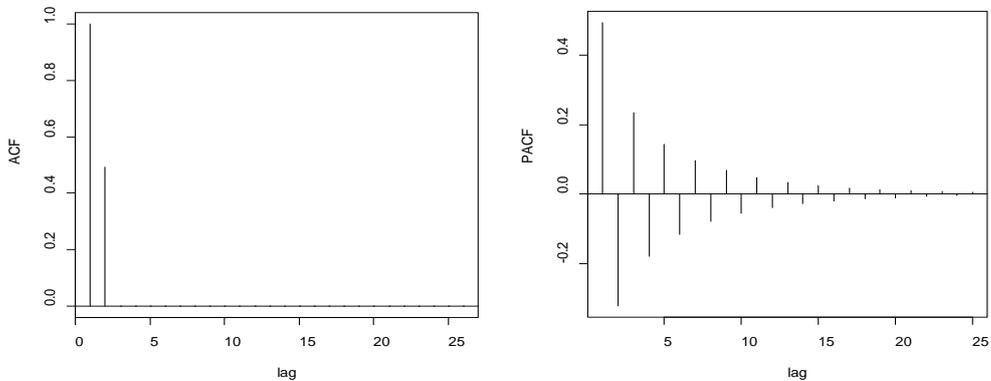


Gráfico 4.8

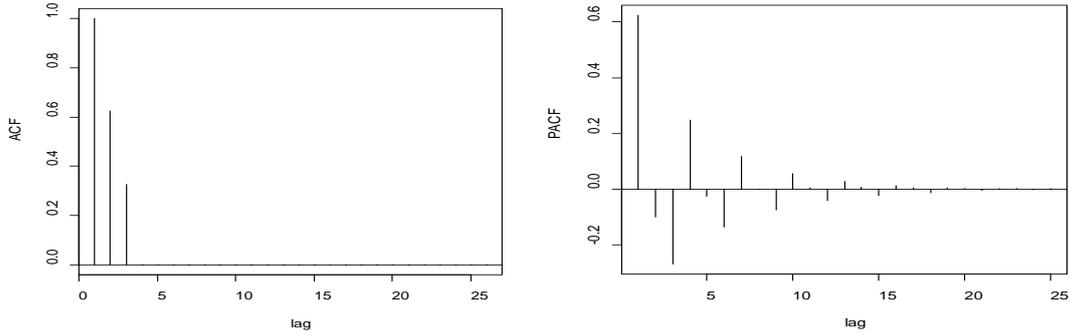


Procediendo del mismo modo puede mostrarse que si el proceso es un $MA(2)$, $y_t = A_t + \theta_1 A_{t-1} + \theta_2 A_{t-2}$, éste contiene un $MA(1)$ como caso particular. En este caso, la $PACF$ presenta un comportamiento exponencial decreciente o una forma sinusoidal amortiguada hacia cero dependiendo de la magnitud y signo de θ_1 y θ_2 , o equivalentemente de las raíces del polinomio en variable compleja de segundo grado asociado $\theta(z) = 0$. En particular, cuando las raíces de la ecuación $\theta(z) = 0$ sean complejas, la $PACF$ decrecerá hacia cero en forma sinusoidal amortiguada.

En general, La función de autocorrelación parcial de un proceso $MA(q)$ se comporta como una mezcla de decrecimiento exponencial y/u onda sinusoidal amortiguada, dependiendo de la naturaleza de las raíces de $\theta(z) = 0$. La función de autocorrelación parcial mostrará un patrón de onda sinusoidal amortiguada si algunas de las raíces son complejas.

Ejemplo 4.2.5: Sea $y_t = A_t + 0,75A_{t-1} + 0,65A_{t-2}$ un modelo $MA(2)$ con parámetros positivos. Escrito en términos del operador de retardos queda $y_t = (1 + 0,75B + 0,65B^2)A_t$ y, por lo tanto, el polinomio de media móvil en variable compleja es $\theta(z) = 1 + 0,75z + 0,65z^2$. Las raíces de $\theta(z) = 0$ son los complejos conjugados $z_1 = -\frac{15}{26} + \frac{5}{26}\sqrt{\frac{163}{5}}i$ y $z_2 = -\frac{15}{26} - \frac{5}{26}\sqrt{\frac{163}{5}}i$. La figura inferior muestra las características de las funciones de autocorrelación y de la autocorrelación parcial del proceso.

Gráfico 4.9



Si se toma un modelo $ARMA(p, q)$ causal dado por la ecuación $\varphi(B)y_t = \theta(B)A_t$, entonces puede escribirse, por (5.10), como

$$y_t = \sum_{j=0}^{\infty} \psi_j A_{t-j}$$

Es inmediato que $E[y_t] = 0$. Además, por definición de proceso $ARMA(p, q)$ es posible expresar y_{t+k} en la forma:

$$y_{t+k} = \sum_{j=1}^p \varphi_j y_{t+k-j} + \sum_{j=0}^q \theta_j A_{t+k-j}$$

De esta última expresión, la función de autocovarianza del proceso viene dada por:

$$\begin{aligned} \gamma_k &= Cov[y_{t+k}, y_t] = Cov \left[\sum_{j=1}^p \varphi_j y_{t+k-j} + \sum_{j=0}^q \theta_j A_{t+k-j}, y_t \right] \\ &= E \left[\left(\sum_{j=1}^p \varphi_j y_{t+k-j} \right) y_t + \left(\sum_{j=0}^q \theta_j A_{t+k-j} \right) y_t \right] \\ &= E \left[\sum_{j=1}^p \varphi_j y_t y_{t+k-j} \right] + E \left[\sum_{j=0}^q \theta_j y_t A_{t+k-j} \right] \\ &= \sum_{j=1}^p \varphi_j E[y_t y_{t+k-j}] + \sum_{j=0}^q \theta_j E[y_t A_{t+k-j}] \end{aligned}$$

Es claro que $E[y_t y_{t+k-j}] = Cov[y_{t+k-j}, y_t] = \gamma_{k-j}$. Además, utilizando la forma (5.10) para y_t resulta

$$\begin{aligned} E[y_t A_{t+k-j}] &= Cov[A_{t+k-j}, y_t] = Cov \left[A_{t+k-j}, \sum_{h=0}^{\infty} \psi_h A_{t-h} \right] = E \left[A_{t+k-j} \sum_{h=0}^{\infty} \psi_h A_{t-h} \right] \\ &= \sum_{h=0}^{\infty} \psi_h E[A_{t+k-j} A_{t-h}] \end{aligned}$$

Si se toma $h = j - k$ se tiene $E[y_t A_{t+k-j}] = \psi_{j-k} \sigma_A^2$ para $k \geq 0$. Por lo tanto,

$$\gamma_k = \sum_{j=1}^p \varphi_j \gamma_{k-j} + \sigma_A^2 \sum_{j=0}^q \theta_j \psi_{j-k} \quad (5.39)$$

Tomando como condiciones iniciales

$$\gamma_k - \sum_{j=1}^p \varphi_j \gamma_{k-j} = \sigma_A^2 \sum_{j=0}^q \theta_j \psi_{j-k} \text{ con } 0 \leq k < \max(p, q + 1)$$

puede escribirse (5.39) como un sistema de ecuaciones en diferencias:

$$\gamma_k - \varphi_1 \gamma_{k-1} - \varphi_2 \gamma_{k-2} - \dots - \varphi_p \gamma_{k-p} = 0 \text{ con } k \geq \max(p, q + 1) \quad (5.40)$$

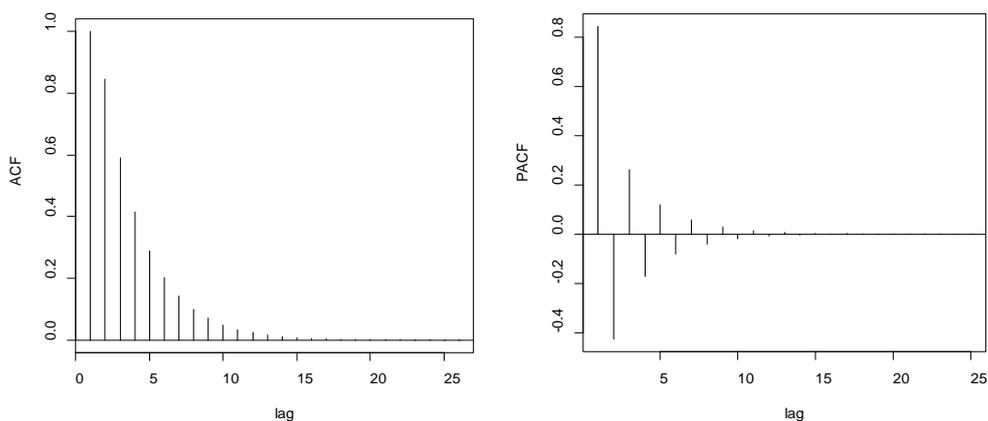
Dividiendo (5.40) por γ_0 se obtiene ρ_k (Shumway & Stoffer, 2011).

Es claro que la función de autocorrelación de un $ARMA(p, q)$ reúne las características de las funciones de autocorrelación de un $AR(p)$ y de un $MA(q)$ ya que ambos modelos están contenidos en un $ARMA$. Igualmente, para el caso de la función de autocorrelación parcial. La misma presenta una mezcla de decrecimiento exponencial o de forma sinusoidal amortiguada dependiendo de las raíces de los polinomios asociados $\varphi(z)$ y $\theta(z)$.

Ejemplo 4.2.2: A continuación se presenta algunos procesos $ARMA(1,1)$ acompañados de los correlogramas de sus correspondientes ACF y $PACF$.

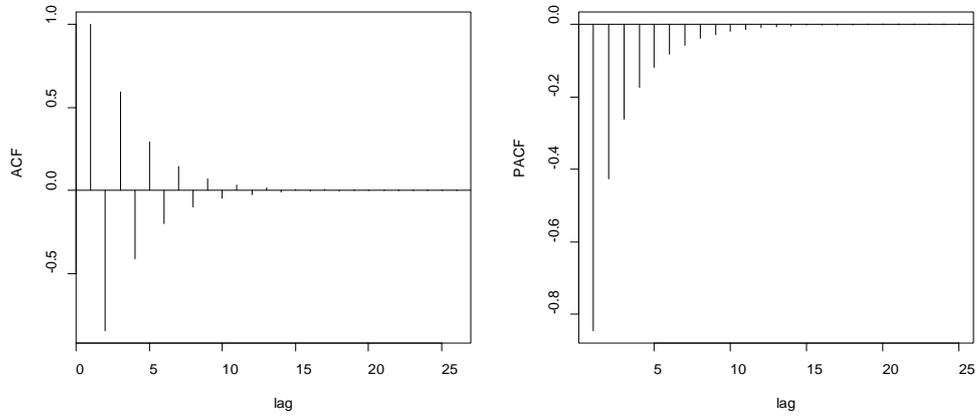
$$a) y_t = 0,7y_{t-1} + A_t + 0,7A_{t-1}$$

Gráfico 4.10



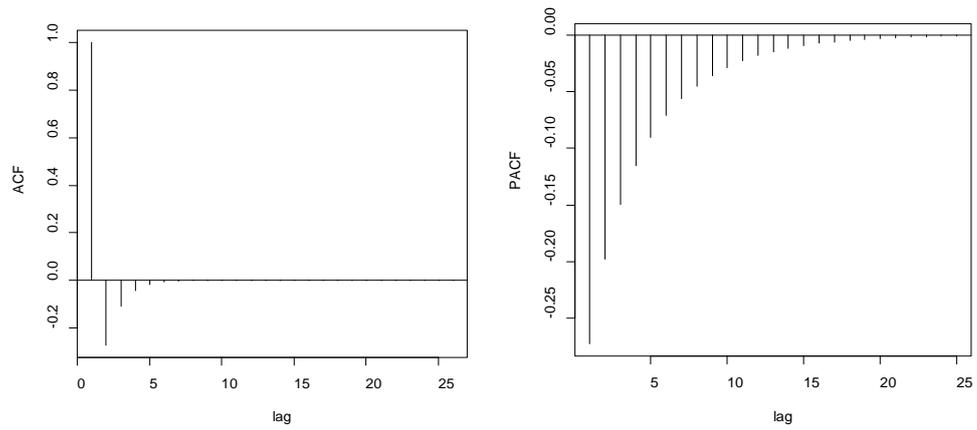
$$b) y_t = -0,7y_{t-1} + A_t - 0,7A_{t-1}$$

Gráfico 4.11



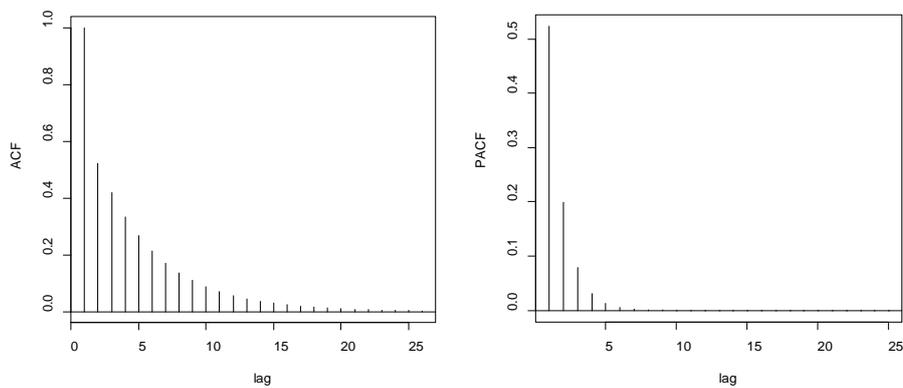
$$c) y_t = 0,4y_{t-1} + A_t - 0,8A_{t-1}$$

Gráfico 4.12



$$d) y_t = 0,8y_{t-1} + A_t - 0,4A_{t-1}$$

Gráfico 4.13



Capítulo 5

Los modelos expuestos en el capítulo anterior permiten estudiar los procesos estacionarios. Es frecuente, sin embargo, encontrar series que provienen de procesos no estacionarios. En este capítulo se estudian las características de los procesos con tendencia, que no mantienen una media y/o una varianza constante e independiente del tiempo. Se discuten los procesos integrados y los modelos estacionales con periodicidad regular en el tiempo, que son de principal importancia en este trabajo. Además, se examina las particularidades de la ACF y de la PACF de estos modelos.

5.1. No estacionariedad

Muchas series temporales no provienen de procesos estocásticos estacionarios y, por lo tanto, son series no estacionarias. No obstante, es posible convertir una serie no estacionaria en una estacionaria aplicando una transformación adecuada que establezca el comportamiento de la serie, para que presente un movimiento afín hacia algún valor constante, con una dispersión estacionaria.

Habitualmente, se emplean dos clases de transformaciones: una para estabilizar la dispersión conocida como *transformación de Box-Cox* y otra para inducir estacionariedad en media. En el caso de los procesos que carecen de estacionariedad en varianza, la solución en el ámbito de la dinámica lineal consiste en aplicar transformaciones que conviertan al proceso original en un proceso condicionadamente homocedástico. (Landro & González, 2009)

Definición 5.1.1: Sea $\{y_t\}$ un proceso estocástico no estacionario tal que la media y la varianza dependen de t (*no son constantes*), esto es

$$E[y_t] = \mu_t \quad y \quad Var[y_t] = \sigma_t^2$$

Si $\sigma_t^2 = \sigma^2 \cdot h(\mu_t)^2$, donde $\sigma^2 > 0$ es una constante y h es una función real diferente de cero para cualquier valor de μ_t , entonces *una transformación estabilizadora de la varianza* del proceso $\{y_t\}$ es cualquier función real g tal que

$$g' = \frac{1}{h}, \quad \forall \mu_t. \quad (5.1)$$

La expresión g' es la primer derivada de g . Esto muestra que el objetivo es encontrar una función g que transforme el proceso tal que $g(y_t)$ tenga varianza constante.

Utilizando la aproximación de Taylor de primer orden alrededor de μ_t resulta:

$$g(y_t) \approx g(\mu_t) + g'(\mu_t)(y_t - \mu_t)$$

Por lo tanto, la varianza de la transformación $g(y_t)$ se puede aproximar como¹⁵:

$$\begin{aligned} Var[g(y_t)] &\approx Var[g(\mu_t) + g'(\mu_t)(y_t - \mu_t)] = Var[g(\mu_t)] + Var[g'(\mu_t)(y_t - \mu_t)] \\ &= Var[g'(\mu_t)y_t] - Var[g'(\mu_t)\mu_t] = [g'(\mu_t)]^2 Var[y_t] = \sigma_t^2 [g'(\mu_t)]^2 \end{aligned}$$

Por lo tanto,

$$Var[g(y_t)] \approx \sigma^2 \cdot [h(\mu_t)]^2 [g'(\mu_t)]^2$$

Por (5.1), $Var[g(y_t)] \approx \sigma^2$ que es un valor constante.

Ejemplo 5.1.1: Consideremos un proceso no estacionario $\{y_t\}$ tal que $E[y_t] = \mu_t$ y $Var[y_t] = \sigma_t^2$ no son constantes. Supongamos que $\sigma_t^2 = \sigma^2 \mu_t^2$, es decir que σ_t^2 es proporcional a μ_t . Por la definición 5.1.1, $h(\mu_t) = \mu_t$ y en consecuencia por (5.1), se tiene $g'(\mu_t) = \mu_t^{-1}$. Por lo tanto, $g(\mu_t) = \ln \mu_t$ lo que implica que la transformación logarítmica proporciona una varianza constante para $\{y_t\}$, es decir el proceso $\{\ln y_t\}$ tiene una varianza aproximadamente constante.

El siguiente ejemplo muestra otra transformación posible para un proceso con varianza no constante.

Ejemplo 5.1.2: Sea $\{y_t\}$ un proceso con media μ_t y varianza σ_t^2 no constantes. Si $\sigma_t^2 = \sigma^2 \mu_t$, es decir que σ_t^2 es proporcional a μ_t , entonces $h(\mu_t) = \sqrt{\mu_t}$. Por lo tanto, $g'(\mu_t) = \frac{1}{2\sqrt{\mu_t}}$ esto implica que $g(\mu_t) = 2\sqrt{\mu_t}$. Esta transformación, es tal que la varianza del proceso $\{2\sqrt{y_t}\}$ es aproximadamente constante.

Las dos transformaciones presentadas en los ejemplos anteriores, están contempladas en una expresión general conocida como *transformación de Box-Cox* de parámetros λ y m :

$$g(y_t) = \frac{(y_t + m)^\lambda - 1}{\lambda} \quad (5.2)$$

Donde $|\lambda| \leq 2$ y el -1 del numerador se utiliza debido a que $\lim_{\lambda \rightarrow 0} \frac{(y_t + m)^\lambda - 1}{\lambda} = \ln(y_t + m)$. En particular, si $\lambda = 0$ y $m = 0$ proporcionan la transformación $g(y_t) = \ln y_t$ en tanto que si $\lambda = \frac{1}{2}$ y $m = 0$ se obtiene la transformación $g(y_t) = 2\sqrt{y_t}$. (Mauricio, 2007)

Por otro lado, la forma típica de no-estacionariedad en media está caracterizada por la presencia de tendencia estocástica en el comportamiento del proceso, la cual se origina en la falla de las condiciones de estacionariedad referidas al comportamiento de las raíces de la ecuación que define su componente autorregresiva. (Landro & González, 2009)

Definición 5.1.2: El operador *diferencia regular de orden d* ($d \in \mathbb{Z}_{\geq 1}$) se denota con el símbolo ∇^d y se define como $\nabla^d = (1 - B)^d$, donde B es el operador de retardo, de modo que si $\{y_t\}$ es un proceso estocástico, la primera diferencia está definida por

¹⁵ Entre otros, puede encontrarse una demostración de las propiedades básicas de la varianza en el libro "Probabilidad y Estadística. La ciencia de la incertidumbre" de M. Evans y J. Rosenthal, Cap.3, pág.170, Reverté, Barcelona, 2004.

$$\nabla y_t = y_t - y_{t-1}, \quad \text{con } t = 1, 2, \dots \quad (5.3)$$

y la diferencia de orden n está dada por

$$\nabla^d y_t = \nabla^{d-1} y_t - \nabla^{d-1} y_{t-1} = \sum_{k=0}^d (-1)^k \binom{d}{k} y_{t-k}, \quad t = n, n+1, \dots \quad (5.4)$$

donde $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ son los coeficientes binomiales. (Fuller, 1996)

La definición 5.1.2 permite dar la de proceso integrado:

Definición 5.1.3: Un proceso estocástico $\{y_t\}$ es *integrado de orden d* ($d \in \mathbb{Z}_{\geq 1}$), o $I(d)$, si y sólo si el proceso $\{\nabla^d y_t\}$ sigue un modelo $ARMA(p, q)$ causal e invertible.

En particular se presentan las siguientes situaciones:

1) Un proceso estocástico $\{y_t\}$ es integrado de orden 0, si y sólo si $\{y_t\}$ sigue un modelo $ARMA(p, q)$ estacionario e invertible (para lo cual es necesario que $\{y_t\}$ sea un proceso estacionario).

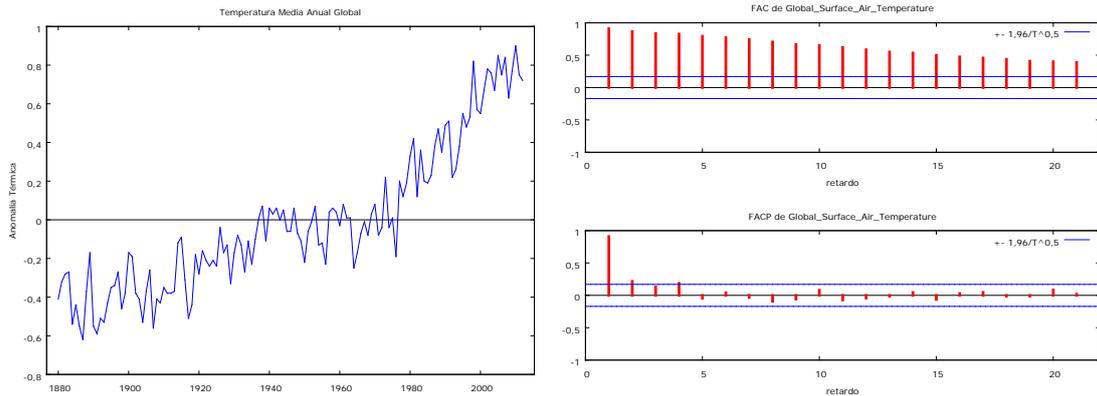
2) Un proceso estocástico $\{y_t\}$ es integrado de orden 1, si y sólo si $\{\nabla y_t\}$ es integrado de orden 0.

3) Un proceso estocástico $\{y_t\}$ es integrado de orden 2, si y sólo si $\{\nabla^2 y_t\}$ es integrado de orden 0.

Ejemplo 5.1.3: Como ya se mencionó, una característica presente en algunas series no estacionarias en media es la *tendencia estocástica*. La misma se trata del movimiento a largo plazo de la serie cuando los ciclos y el término irregular se han eliminado. En la figura 5.1 se muestra el gráfico de la serie de observaciones de la temperatura media anual global desde 1880 hasta 2012, con período base 1951 – 1980, derivada de la red de estaciones meteorológicas [Esta es una actualización de la Figura 6 (b) en Hansen et al. (2001)]. En particular los datos son desviaciones, medidas en grados centígrados, respecto a la media 1951 – 1980. Es posible observar una fuerte tendencia en crecimiento a partir del siglo xx que sustenta la hipótesis del calentamiento global. El correlograma de la temperatura media global reúne las características típicas de un correlograma de una serie con marcada tendencia. (Fuente: *National Aeronautics and Space Administration (NASA) - Goddard Institute for Space Studies* http://data.giss.nasa.gov/gistemp/graphs_v3/).

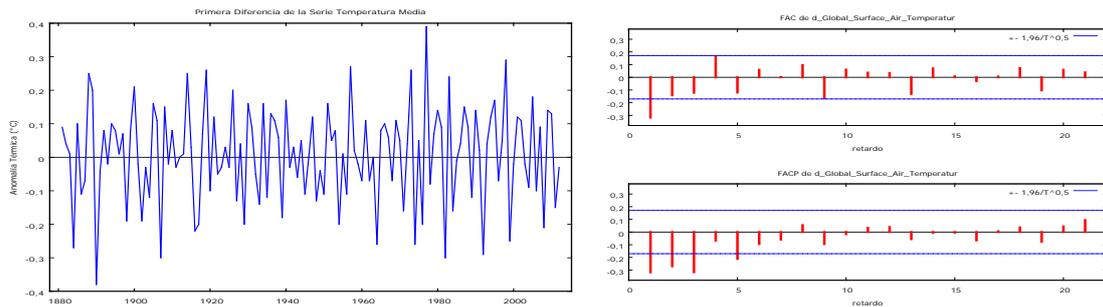
Esta serie es un claro ejemplo de proceso integrado. Una diferencia de orden 1 puede convertirla en estacionaria como se muestra en el siguiente ejemplo.

Gráfico 5.1



Ejemplo 5.1.4: En el gráfico 5.2 se representa la primera diferencia o diferencia regular de orden 1 de la serie “Temperatura Media Anual Global”. En principio, el gráfico ahora oscila alrededor de un valor constante y el correlograma de la ACF cambió considerablemente. El decrecimiento lento se modificó y en su lugar la ACF tiene un solo retardo significativamente distinto de cero. Puede mostrarse que el proceso diferenciado es un $ARMA(0,1)$ cuya ecuación es $y_t = 0,00895785 + (1 - 0,711913B)A_t$.

Gráfico 5.2



Ejemplo 5.1.5: Un modelo que resulta apropiado para el análisis de la tendencia en una serie temporal, es el modelo llamado “paseo aleatorio con deriva”, dado por la expresión

$$y_t = \delta + y_{t-1} + A_t \quad (5.5)$$

para $t = 1, 2, 3, \dots$ con condición inicial $y_0 = 0$, y donde A_t es un ruido blanco. La constante δ se denomina *deriva* y cuando $\delta = 0$, la expresión (5.5) es llamada “paseo aleatorio” (*random walk*). Es importante señalar que cuando $\delta = 0$, el valor del proceso en t es igual al valor en el instante $t - 1$ más un movimiento completamente aleatorio dado por A_t .

Considerando que $y_{t-1} = \delta + y_{t-2} + A_{t-1}$, sustituyendo en (5.5) resulta:

$$y_t = \delta + \delta + y_{t-2} + A_{t-1} + A_t$$

Además $y_{t-2} = \delta + y_{t-3} + A_{t-2}$. De donde:

$$y_t = \delta + \delta + \delta + y_{t-3} + A_{t-2} + A_{t-1} + A_t$$

Continuando con este procedimiento, sumando hasta un tiempo t , se tiene:

$$y_t = \sum_{j=1}^t \delta + y_0 + \sum_{j=1}^t A_j \quad (5.6)$$

Con condición inicial $y_0 = 0$, puede escribirse (5.5) como una suma acumulativa de ruidos blancos, de la forma:

$$y_t = \delta t + \sum_{j=1}^t A_j$$

para $t = 1, 2, \dots$. Puesto que A_t es un ruido blanco, $E[A_t] = 0$, resulta entonces:

$$E[y_t] = E\left[\delta t + \sum_{j=1}^t A_j\right] = \delta t + \sum_{j=1}^t E[A_j] = \delta t$$

En consecuencia, la media del proceso es una recta con pendiente δ y variable independiente t , lo que muestra que el proceso no es estacionario en media.

Calculando la autocovarianza del proceso, resulta:

$$\begin{aligned} \gamma_k &= cov[y_t, y_{t+k}] = E[(y_t - \delta t)(y_{t+k} - \delta(t+k))] \\ &= E\left[\left(\delta t + \sum_{j=1}^t A_j - \delta t\right)\left(\delta(t+k) + \sum_{j=1}^{t+k} A_j - \delta(t+k)\right)\right] \\ &= E\left[\left(\sum_{j=1}^t A_j\right)\left(\delta k + \sum_{j=1}^{t+k} A_j\right)\right] \end{aligned}$$

Distribuyendo y teniendo en cuenta que $E[A_j A_{j+1}] = 0$ y $E[A_j A_j] = E[A_j^2] = \sigma_A^2$ se tiene:

$$\gamma_k = \sum_{j=1}^t \sigma_A^2 = t\sigma_A^2$$

Y también $Var[y_t] = t\sigma_A^2$. Es decir que la media, la varianza y la covarianza son funciones de t . Este resultado puede mirarse desde el polinomio autorregresivo del modelo, ya que el mismo, en términos del operador de retardo, es $(1 - B)y_t = \delta + A_t$. Por lo tanto, el polinomio autorregresivo de variable compleja del proceso $\varphi(z) = 1 - z$ tiene raíz $z = 1$, por lo que no se trata de un proceso causal estacionario.

Sin embargo, aplicando el operador diferencia de orden 1 a (5.5), resulta:

$$\nabla y_t = y_t - y_{t-1} = \delta + A_t \quad (5.7)$$

El proceso $\nabla y_t = \delta + A_t$ sigue un modelo $ARMA(0,0)$ estacionario e invertible (nótese que $E[\nabla y_t] = \delta$ y $Var[\nabla y_t] = Var[\delta + A_t] = Var[A_t] = \sigma_A^2$). El modelo diferenciado ∇y_t resulta

un proceso $I(0)$ lo cual implica que un paseo aleatorio con deriva es un proceso integrado de orden 1, $I(1)$.

La ACF de un paseo aleatorio tiene un decrecimiento lento que es característico de cualquier proceso integrado de orden $d \geq 1$. En general, un proceso integrado de orden $d \geq 1$ (el paseo aleatorio es un caso particular con $d = 1$), recibe el nombre de *proceso homogéneamente no estacionario*, ya que el proceso presenta una *tendencia* en su evolución temporal, esto es un nivel hacia el cual tiende el proceso desde un origen temporal suficientemente anterior al instante t y una *componente irregular* que es independiente de dicha tendencia.

Tomando en (5.6) un origen temporal $k < t$ resulta:

$$y_t = (t - k)\delta + y_k + \sum_{j=0}^{t-k} A_{t-j}$$

Esta igualdad puede escribirse como:

$$y_t = (y_k - \delta k) + A_t + \sum_{j=1}^{t-(k+1)} A_{t-j} + \delta t$$

Llamando $T_t[y_t]$ a la tendencia del proceso e $I_t[y_t]$ a la componente irregular se tiene:

$$T_t[y_t] = (y_k - \delta k) + \sum_{j=1}^{t-(k+1)} A_{t-j} + \delta t \quad \text{y} \quad I_t[y_t] = A_t$$

Está claro que la tendencia de un paseo aleatorio con deriva tiene una componente aleatoria dada por la expresión $(y_k - \delta k) + \sum_{j=1}^{t-(k+1)} A_{t-j}$ y una componente determinista asociada al término δt (cuando $\delta = 0$ la componente determinista no está presente). Por otro lado, de (5.7) se obtiene:

$$T_t[\nabla y_t] = \delta \quad \text{y} \quad I_t[\nabla y_t] = A_t$$

lo que muestra que el proceso ∇y_t carece de tendencia.

Ejemplo 5.1.6: Considérese el proceso estocástico cuyo modelo responde a la ecuación $y_t = y_{t-1} + A_t - \theta_1 A_{t-1}$ (5.8), para cualquier $t \in \mathbb{Z}$ donde el parámetro θ_1 es tal que $|\theta_1| < 1$ y A_t es un ruido blanco ($E[A_t] = 0$ y $Var[A_t] = \sigma_A^2$).

De (5.8), que caracteriza el modelo, se tiene $y_{t-1} = y_{t-2} + A_{t-1} - \theta_1 A_{t-2}$. Sustituyendo en (5.8) resulta: $y_t = y_{t-2} + A_{t-1} - \theta_1 A_{t-2} + A_t - \theta_1 A_{t-1}$. A su vez, $y_{t-2} = y_{t-3} + A_{t-2} - \theta_1 A_{t-3}$. Reemplazando en la ecuación anterior:

$$\begin{aligned} y_t &= y_{t-3} + A_{t-2} - \theta_1 A_{t-3} + A_{t-1} - \theta_1 A_{t-2} + A_t - \theta_1 A_{t-1} \\ &= y_{t-3} + A_t - \theta_1 A_{t-3} + (1 - \theta_1)A_{t-2} + (1 - \theta_1)A_{t-1}. \end{aligned}$$

Continuando con este razonamiento hasta un retardo $k < t$, se obtiene:

$$y_t = y_k - \theta_1 A_k + A_t + \sum_{i=1}^{t-k-1} (1 - \theta_1) A_{t-i} \quad (5.9)$$

Si $k = 0$, resulta:

$$y_t = y_0 - \theta_1 A_0 + A_t + \sum_{i=1}^{t-1} (1 - \theta_1) A_{t-i}$$

Si además, $y_0 = A_0 = 0$:

$$y_t = A_t + \sum_{i=1}^{t-1} (1 - \theta_1) A_{t-i}$$

Es claro que $E[y_t] = E[A_t] + \sum_{i=1}^{t-1} (1 - \theta_1) E[A_{t-i}] = 0$. Sin embargo, para la varianza se tiene:

$$Var[y_t] = E \left[\left(A_t + \sum_{i=1}^{t-1} (1 - \theta_1) A_{t-i} \right) \left(A_t + \sum_{i=1}^{t-1} (1 - \theta_1) A_{t-i} \right) \right]$$

Desarrollando las sumas, queda:

$$Var[y_t] = E[(A_t + (1 - \theta_1)A_{t-1} + (1 - \theta_1)A_{t-2} + \dots + (1 - \theta_1)A_2 + (1 - \theta_1)A_1) \cdot (A_t + (1 - \theta_1)A_{t-1} + (1 - \theta_1)A_{t-2} + \dots + (1 - \theta_1)A_2 + (1 - \theta_1)A_1)]$$

Al desarrollar el producto, para $i \neq j$ se tiene:

$$E[(1 - \theta_1)^2 A_{t-i} A_{t-j}] = (1 - \theta_1)^2 E[A_{t-i} A_{t-j}] = (1 - \theta_1)^2 \cdot 0 = 0. \text{ Por lo tanto:}$$

$$Var[y_t] = E[A_t^2] + (1 - \theta_1)^2 E[A_{t-1}^2] + \dots + (1 - \theta_1)^2 E[A_2^2] + (1 - \theta_1)^2 E[A_1^2]$$

Como además, $E[A_{t-i}^2] = \sigma_A^2$ para $i = 1, \dots, t - 1$, resulta:

$$\begin{aligned} Var[y_t] &= \sigma_A^2 + (1 - \theta_1)^2 \sigma_A^2 + \dots + (1 - \theta_1)^2 \sigma_A^2 = \sigma_A^2 + (1 - \theta_1)^2 \sum_{i=1}^{t-1} \sigma_A^2 \\ &= \sigma_A^2 + (1 - \theta_1)^2 (t - 1) \sigma_A^2 \end{aligned}$$

Luego, la varianza del proceso depende de t por lo que el proceso no es estacionario en varianza.

De la ecuación (5.9), resulta:

$$T_t[y_t] = y_k - \theta_1 A_k + A_{t-1} + \sum_{i=2}^{t-k-1} (1 - \theta_1) A_{t-i} \quad y \quad I_t[y_t] = A_t - \theta_1 A_{t-1}$$

Es claro que la tendencia es esencialmente estocástica (sin componente determinista). La diferencia regular de orden 1 del proceso produce:

$$\nabla y_t = y_t - y_{t-1} = A_t - \theta_1 A_{t-1} \text{ donde } t \in \mathbb{Z}$$

Se observa entonces que ∇y_t es un proceso $MA(1)$. El polinomio de media móvil en variable compleja es $\theta(z) = 1 - \theta_1 z$, cuya raíz, $z = \frac{1}{\theta_1}$, está fuera del círculo unitario ya que $|\theta_1| < 1$. Por lo tanto ∇y_t es invertible. Se puede concluir que y_t es un proceso integrado de orden 1, $I(1)$, pues ∇y_t es integrado de orden 0. Además, $T_t[\nabla y_t] = 0$ y $I_t[\nabla y_t] = A_t - \theta_1 A_{t-1}$, de modo que $\{\nabla y_t\}$ carece de tendencia.

En conclusión, se observa que como metodología inmediata para detectar si un proceso $\{y_t\}$ es o no estacionario, puede estudiarse la forma en la cual decrece su función de autocorrelación a partir de las características gráficas del correlograma. Si del análisis del correlograma se concluye que el proceso no es estacionario, se toma en consideración el correlograma de su primera diferencia $\{\nabla y_t\}$, si el proceso diferenciado tampoco es estacionario, se considera la segunda diferencia $\{\nabla^2 y_t\}$, y se continua este procedimiento hasta obtener un proceso cuya función de autocorrelación tenga un correlograma con decrecimiento exponencial o con forma sinusoidal amortiguada.¹⁶

No obstante las ventajas de la diferenciación, a veces se corre el riesgo de caer en el problema de la *sobre-diferenciación*.

Ejemplo 5.1.7: Sea un proceso $\{y_t\}$ que admite una representación $MA(1)$ invertible de la forma $y_t = A_t - \theta_1 A_{t-1}$, donde A_t es un ruido blanco con media 0 y varianza σ_A^2 . En términos del operador de retardo se puede escribir el proceso como $y_t = (1 - \theta_1 B)A_t$. Aplicando el operador diferencia regular, la primera diferencia resulta:

$$\nabla y_t = y_t - y_{t-1} = A_t - \theta_1 A_{t-1} - A_{t-1} + \theta_1 A_{t-2}$$

Aplicando el operador de retardo:

$$\nabla y_t = y_t - B y_t = A_t - \theta_1 B A_t - B A_t + \theta_1 B^2 A_t$$

Factorizando:

$$\nabla y_t = (1 - B)y_t = (1 - \theta_1 B - B + \theta_1 B^2)A_t = [1 - (1 + \theta_1)B + \theta_1 B^2]A_t$$

Llamando $\theta_1^* = 1 + \theta_1$ y $\theta_2^* = -\theta_1$ se tiene:

$$(1 - B)y_t = (1 - \theta_1^* B - \theta_2^* B^2)A_t$$

Luego, $\varphi(B) = 1 - B$ y $\theta^*(B) = 1 - \theta_1^* B - \theta_2^* B^2$ y puede escribirse $\varphi(B)y_t = \theta^*(B)A_t$. Sin embargo, obsérvese que la expresión obtenida es más complicada que la original puesto que aparecen dos coeficientes θ_1^* y θ_2^* , en lugar de uno y, además, la suma entre ellos da por resultado 1; es decir, $\theta_1^* + \theta_2^* = 1 + \theta_1 - \theta_1 = 1$, lo que tiene como consecuencia (ver capítulo anterior) que la ecuación $\theta^*(z) = 1 - \theta_1^* z - \theta_2^* z^2 = 0$ tiene una raíz unitaria y, por lo tanto, el proceso diferenciado no es invertible y su parte autorregresiva estacionaria no existe. (Landro & González, 2009)

¹⁶ Debe tenerse en cuenta que no siempre las funciones de autocorrelaciones exhiben las características comentadas precedentemente. A veces esta función decrece en forma más rápida que la lineal, pero no exponencialmente. Un caso particular, es el del decrecimiento hiperbólico de la función de autocorrelaciones que da origen a representaciones con orden de diferenciación fraccionario, conocidas como ARFIMA. (Landro & González, 2009)

El proceso y_t tiene media cero pues $E[y_t] = E[A_t - \theta_1 A_{t-1}] = E[A_t] - \theta_1 E[A_{t-1}] = 0 - \theta_1 \cdot 0 = 0$. Además, calculando la varianza del proceso se tiene: $Var[y_t] = E[(A_t - \theta_1 A_{t-1})^2] = E[A_t^2 - 2A_t \theta_1 A_{t-1} + \theta_1^2 A_{t-1}^2] = E[A_t^2] - 2\theta_1 E[A_t A_{t-1}] + \theta_1^2 E[A_{t-1}^2] = \sigma_A^2 + \theta_1^2 \sigma_A^2 = (1 + \theta_1^2) \sigma_A^2$.

Por otro lado, la media del proceso ∇y_t también es cero ya que $E[A_t - (1 + \theta_1)A_{t-1} + \theta_1 A_{t-2}] = E[A_t] - (1 + \theta_1)E[A_{t-1}] + \theta_1 E[A_{t-2}] = 0$, la cual coincide con la media del proceso original. Además, la varianza es:

$$\begin{aligned} Var[\nabla y_t] &= E[(A_t - (1 + \theta_1)A_{t-1} + \theta_1 A_{t-2})^2] \\ &= E\{[A_t - (1 + \theta_1)A_{t-1}]^2 + 2[A_t - (1 + \theta_1)A_{t-1}]\theta_1 A_{t-2} + \theta_1^2 A_{t-2}^2\} \\ &= E[A_t^2] - 2(1 + \theta_1)E[A_t A_{t-1}] + (1 + \theta_1)^2 E[A_{t-1}^2] + 2\theta_1 E[A_t A_{t-2}] \\ &\quad - 2(1 + \theta_1)\theta_1 E[A_{t-1} A_{t-2}] + \theta_1^2 E[A_{t-2}^2] = \sigma_A^2 + (1 + \theta_1)^2 \sigma_A^2 + \theta_1^2 \sigma_A^2 \\ &= 2(1 + \theta_1 + \theta_1^2) \sigma_A^2 \end{aligned}$$

Comparando las varianzas de ambos procesos, puede notarse que

$$\begin{aligned} Var[\nabla y_t] - Var[y_t] &= 2(1 + \theta_1 + \theta_1^2) \sigma_A^2 - (1 + \theta_1^2) \sigma_A^2 = (1 + 2\theta_1 + \theta_1^2) \sigma_A^2 \\ &= (1 + \theta_1)^2 \sigma_A^2 > 0. \end{aligned}$$

De donde es posible concluir que la varianza del proceso diferenciado es mayor que la del proceso original.

Un resultado útil, para determinar el orden de diferenciación adecuado, es que las varianzas de los sucesivos procesos diferenciados decrecerán hasta que el proceso alcance el estado estacionario. En caso de sobre-diferenciación las varianzas muestrales tenderán a crecer. (Landro & González, 2009)

5.2. Modelos ARIMA

Un proceso estocástico homogéneamente no estacionario puede ser reducido a un proceso estacionario tomando un grado adecuado de diferenciación. Se da así inicio al estudio de los modelos *ARIMA* (*Autoregressive Integrated Moving Average*).

Definición 5.2.1: Un proceso y_t se dice que es *ARIMA*(p, d, q) si

$$\nabla^d y_t = (1 - B)^d y_t \quad (5.10)$$

es *ARMA*(p, q). En general se escribe el modelo como

$$\varphi(B)(1 - B)^d y_t = \theta(B)A_t. \quad (\text{Shumway \& Stoffer, 2011}) \square$$

Si $E[\nabla^d y_t] = \mu$, resulta

$$\varphi(B)(1 - B)^d y_t = \delta + \theta(B)A_t, \quad (5.11)$$

El parámetro δ juega un papel muy diferente para $d = 0$ y $d > 0$. Cuando $d = 0$ el proceso original es estacionario, y entonces $\delta = \mu(1 - \varphi_1 - \dots - \varphi_p)$. Cuando $d \geq 1$, sin embargo, δ es llamado el término de tendencia determinista. (Wei, 2006)

La expresión (p, d, q) representa el orden del modelo. Cuando $p = 0$, el modelo $ARIMA(p, d, q)$ es también llamado el modelo integrado de media móvil de orden (d, q) y es denotado como el modelo $IMA(d, q)$.

Ejemplo 5.2.1: El modelo de paseo aleatorio discutido en el capítulo anterior es un modelo $ARIMA(0,1,0)$. En efecto, en (5.11) con $p = 0$, $d = 1$ y $q = 0$, resulta:

$$(1 - B)y_t = A_t.$$

O bien

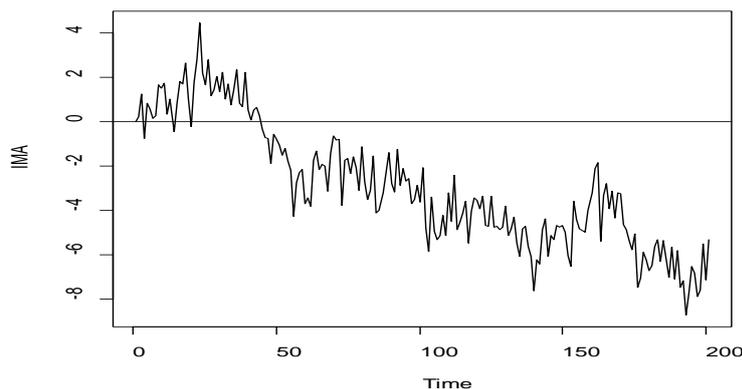
$$y_t = y_{t-1} + A_t.$$

Nótese que el modelo de paseo aleatorio es un proceso limitante del proceso $AR(1)$, $(1 - \varphi_1 B)y_t = A_t$ con $\varphi_1 \rightarrow 1$. Porque la función de autocorrelación del proceso $AR(1)$ es $\rho_k = \varphi_1^k$, cuando $\varphi_1 \rightarrow 1$, el fenómeno del modelo de paseo aleatorio puede caracterizarse por picos grandes, distintos de cero, de la función de autocorrelación de la muestra del proceso original y por una función de autocorrelación prácticamente nula para la serie diferenciada $\{(1 - B)y_t\}$. (Wei, 2006).

Ejemplo 5.2.2: El modelo $y_t = y_{t-1} + A_t - \theta_1 A_{t-1}$ con $|\theta_1| < 1$, analizado en el ejemplo 5.1.6, es un modelo $IMA(1,1)$ o $ARIMA(0,1,1)$. En efecto, aplicando el operador de retardo, resulta: $y_t = B y_t + A_t - \theta_1 B A_t$. O bien, $y_t - B y_t = A_t - \theta_1 B A_t$. De esta expresión, se puede obtener: $(1 - B)y_t = (1 - \theta_1 B)A_t$. De aquí es posible ver que $\varphi(B) = 1$, esto es $p = 0$, además, $\theta(B) = 1 - \theta_1 B$, es decir, $q = 1$, y $d = 1$. La serie del ejemplo 5.1.3 es un caso particular de este tipo de modelo.

En la figura siguiente se muestra una simulación para un modelo $IMA(1,1)$ con $n = 200$:

Gráfico 5.3



De lo estudiado en la sección 5.1 puede concluirse que un proceso que es estacionario en media, no necesariamente lo es en varianza y autocovarianza. Pero un proceso que es no estacionario en media, es también no estacionario en varianza y autocovarianza. Si se retoma el ejemplo 5.1.6, donde se trabaja con un modelo $IMA(1,1)$, es posible ver que:

$$Var[y_t] = \sigma_A^2 + (1 - \theta_1)^2 (t - 1) \sigma_A^2$$

la cual ya fue determinada en dicho ejemplo. Además, para un determinado retardo $k \neq 0$:

$$Var[y_{t-k}] = \sigma_A^2 + (1 - \theta_1)^2(t - k - 1)\sigma_A^2$$

que es la varianza de y_{t-k} . Por otro lado, del cálculo de la covarianza entre y_t e y_{t-k} resulta:

$$Cov[y_{t-k}, y_t] = (1 - \theta_1)\sigma_A^2 + (1 - \theta_1)^2(t - k - 1)\sigma_A^2$$

Las funciones anteriores permiten calcular la función de correlación entre dos valores del proceso separados entre sí por el retardo k :

$$\begin{aligned} \rho_k = Corr[y_{t-k}, y_t] &= \frac{Cov[y_{t-k}, y_t]}{\sqrt{Var[y_{t-k}]Var[y_t]}} \\ &= \frac{(1 - \theta_1)\sigma_A^2 + (1 - \theta_1)^2(t - k - 1)\sigma_A^2}{\sqrt{[(1 - \theta_1)\sigma_A^2 + (1 - \theta_1)^2(t - k - 1)\sigma_A^2][\sigma_A^2 + (1 - \theta_1)^2(t - k - 1)\sigma_A^2]}} \end{aligned}$$

En general, puede observarse que la varianza de un proceso *ARIMA* es una función dependiente del tiempo, no acotada para $t \rightarrow \infty$ y diferente para dos valores que difieren en un retardo $k \neq 0$; es decir $Var[y_t] \neq Var[y_{t-k}]$ para $k \neq 0$, es decir que no son procesos estacionarios en varianza. También la autocovarianza y la autocorrelación son funciones dependientes de la variable tiempo (no permanecen invariantes ante traslaciones del tiempo). Nótese que la función de autocorrelación ρ_k de un proceso *ARIMA* decrece lentamente en la medida en que k se incrementa.

Ejemplo 5.2.3: A continuación se simula para $n = 200$, un modelo *ARIMA*(1,1,2) con $\varphi_1 = 0,8$, $\theta_1 = 0,5$ y $\theta_2 = -0,7$. Además se muestra los correlogramas de la *ACF* y de la *PACF* del modelo:

Gráfico 5.4

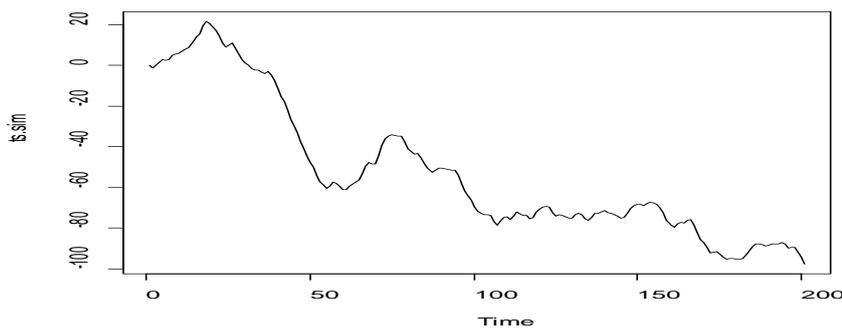
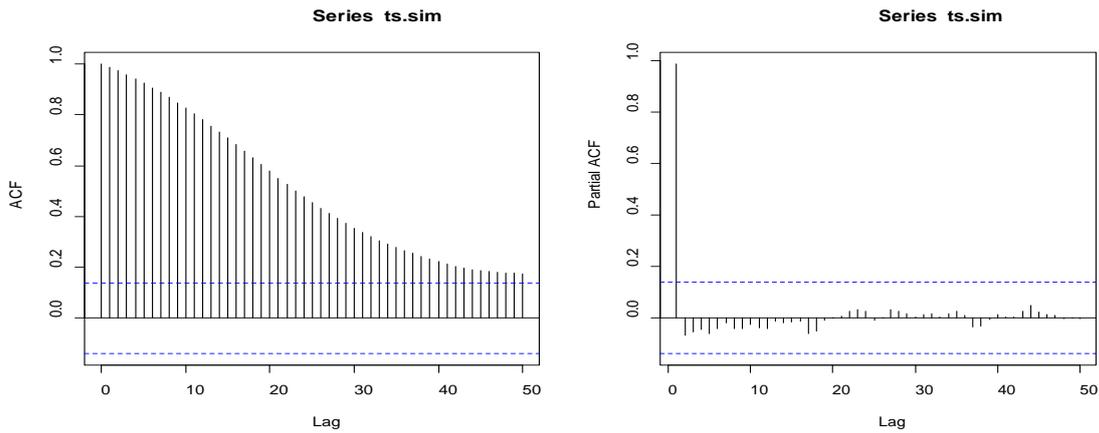


Gráfico 5.5



5.3. Modelos ARIMA Estacionales

Muchas series temporales provienen de procesos con características estacionales que repiten el mismo comportamiento con cierta periodicidad en el tiempo. El período de tiempo más pequeño en el cual se repite el fenómeno se conoce como período estacional de la serie.

La importancia de estudiar las características estacionales de un proceso radica en que las observaciones registradas en un determinado período de tiempo (pueden ser series mensuales o trimestrales) y las registradas en el mismo período de tiempo correspondiente al año anterior presentan patrones similares de comportamiento, por lo que están temporalmente correlacionadas. Sin embargo, con el propósito de comenzar el estudio de estos procesos, se supone que no existe correlación entre observaciones consecutivas y se elabora un *modelo estacional puro*, que aunque no son modelos útiles en la práctica (ya que el supuesto de no correlación es poco adaptable a la realidad), serán de utilidad para determinar en qué retardos de la función de autocorrelación se mostrará la estructura estacional del proceso.

Se inicia el estudio de estos procesos con la definición de los polinomios autorregresivo y de media móvil característicos de los modelos estacionales.

Definición 5.3.1: Los operadores $\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}$ y $\Theta_Q(B^s) = 1 + \theta_1 B^s + \theta_2 B^{2s} + \dots + \theta_Q B^{Qs}$ son el *operador autorregresivo estacional* y el *operador de media móvil estacional* de órdenes P y Q , respectivamente, con *período estacional* s . (Shumway & Stoffer, 2011)

Utilizando estos operadores puede escribirse el modelo estacional puro $ARMA(P, Q)_s$ de la siguiente manera:

$$\Phi_P(B^s)y_t = \Theta_Q(B^s)A_t \quad (5.12)$$

Como ocurre en el caso de los procesos $ARMA(p, q)$ estudiados en el capítulo 4, el modelo $ARMA(P, Q)_s$ es causal solo cuando las raíces de $\Phi_P(z^s)$ están fuera del círculo unitario y es invertible cuando las raíces de $\Theta_Q(z^s)$ están fuera del círculo unitario.

Ejemplo 5.3.1: Supóngase que se cuenta con datos mensuales correspondientes a n años, los cuales se muestran dispuestos en la siguiente tabla:

Tabla 5.1

<i>Año/Mes</i>	<i>1</i>	<i>2</i>	<i>...</i>	<i>12</i>
<i>1</i>	y_1	y_2	\dots	y_{12}
<i>2</i>	y_{13}	y_{14}	\dots	y_{24}
<i>3</i>	y_{25}	y_{26}	\dots	y_{36}
\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot
<i>n</i>	$y_{1+12(n-1)}$	$y_{2+12(n-1)}$	\dots	$y_{12+12(n-1)}$

Cada columna es una realización de una serie temporal. Supóngase que cada una de estas doce series es generada por un modelo $ARMA(P, Q)$, o más específicamente, la serie correspondiente al j -ésimo mes, y_{j+12t} , con $t = 0, 1, \dots, n - 1$, satisface la ecuación en diferencias:

$$y_{j+12t} = \Phi_1 y_{j+12(t-1)} + \dots + \Phi_P y_{j+12(t-P)} + A_{j+12t} + \Theta_1 A_{j+12(t-1)} + \dots + \Theta_Q A_{j+12(t-Q)} \quad (5.13)$$

donde $\{A_{j+12t}, t \in \mathbb{Z}\}$ es un ruido blanco Gaussiano con media 0 y varianza σ_A^2 .

Luego, se supone que el mismo modelo $ARMA(P, Q)$ se aplica a cada mes, es decir que (5.13) se mantiene para cada $j = 1, \dots, 12$. Nótese, sin embargo, que $E(A_t A_{t+k})$ no es necesariamente cero excepto cuando k es un entero múltiplo de 12. (Brockwell & Davis, Introduction to Time Series and Forecasting, 2002)

Utilizando el operador de retardos resulta:

$$y_{j+12t} = \Phi_1 B^{12} y_{j+12t} + \dots + \Phi_P B^{12P} y_{j+12t} + A_{j+12t} + \Theta_1 B^{12} A_{j+12t} + \dots + \Theta_Q B^{12Q} A_{j+12t}$$

De donde:

$$y_{j+12t} - \Phi_1 B^{12} y_{j+12t} - \dots - \Phi_P B^{12P} y_{j+12t} = A_{j+12t} + \Theta_1 B^{12} A_{j+12t} + \dots + \Theta_Q B^{12Q} A_{j+12t}$$

De cada miembro de la ecuación se obtiene:

$$(1 - \Phi_1 B^{12} - \dots - \Phi_P B^{12P}) y_{j+12t} = (1 + \Theta_1 B^{12} + \dots + \Theta_Q B^{12Q}) A_{j+12t}$$

De la definición 5.3.1:

$$\Phi_P(B^{12}) y_{j+12t} = \Theta_Q(B^{12}) A_{j+12t}$$

quedando de este modo expresado el modelo (5.13) en términos de los operadores autorregresivo y de media móvil estacional. En general:

$$\Phi_P(B^{12}) y_t = \Theta_Q(B^{12}) A_t \quad (5.14)$$

que es la ecuación (5.12) con $s = 12$.

Si en (5.14) se toma $P = 0$ y $Q = 1$ resulta:

$$y_t = (1 + \Theta B^{12})A_t = A_t + \Theta B^{12}A_t$$

Y puesto que el operador de retardos produce $B^{12}A_t = A_{t-12}$ se obtiene $y_t = A_t + \Theta_1 A_{t-12}$ el cual es un modelo $MA(1)_{12}$.

En la siguiente figura se observa los correlogramas de la ACF y la $PACF$ de un $MA(1)_{12}$ con $\Theta = 0,5$ (figura superior), y luego (figura inferior) con $\Theta = -0,5$.

Gráfico 5.6

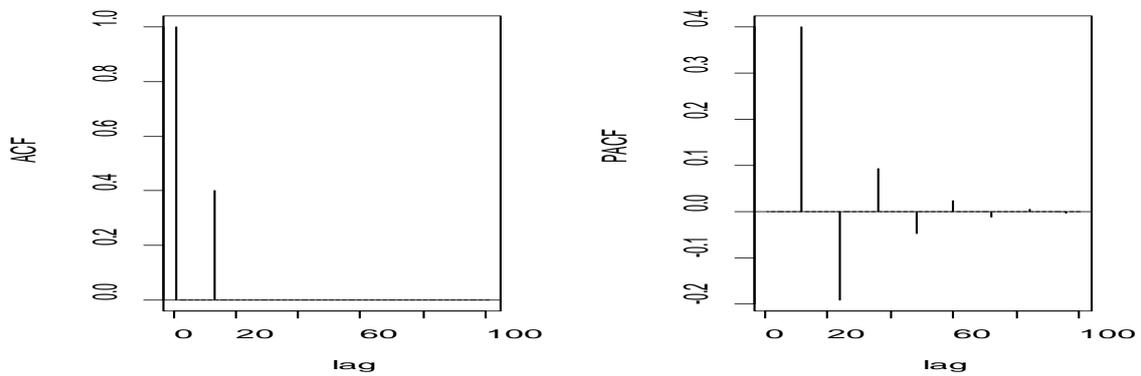
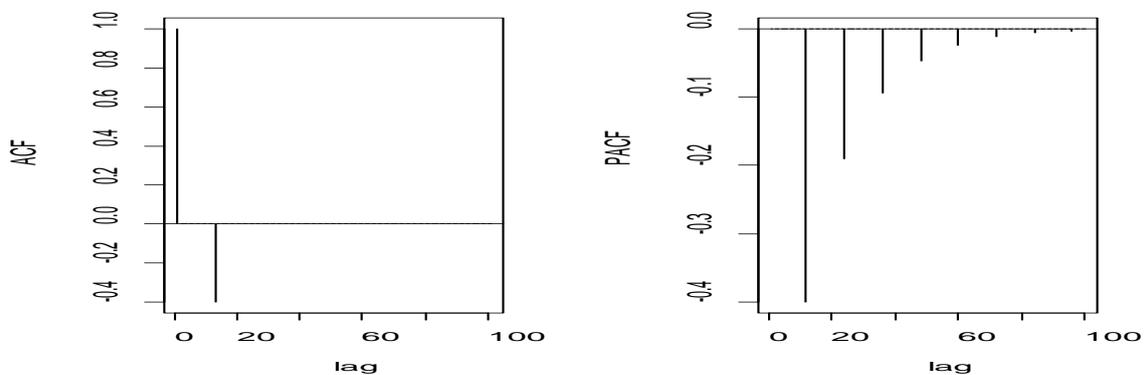


Gráfico 5.7



Si se elige en (5.14) $P = 1$ y $Q = 0$ se tiene:

$$(1 - \Phi B^{12})y_t = A_t$$

de donde

$$y_t = \Phi B^{12}y_t + A_t = \Phi y_{t-12} + A_t$$

En este caso, se trata de un modelo $AR(1)_{12}$.

En la siguiente figura se representa los correlogramas de la ACF y de la $PACF$ de un $AR(1)_{12}$ con $\Phi = 0,4$ (figura superior), y $\Phi = -0,6$ (figura inferior):

Gráfico 5.8

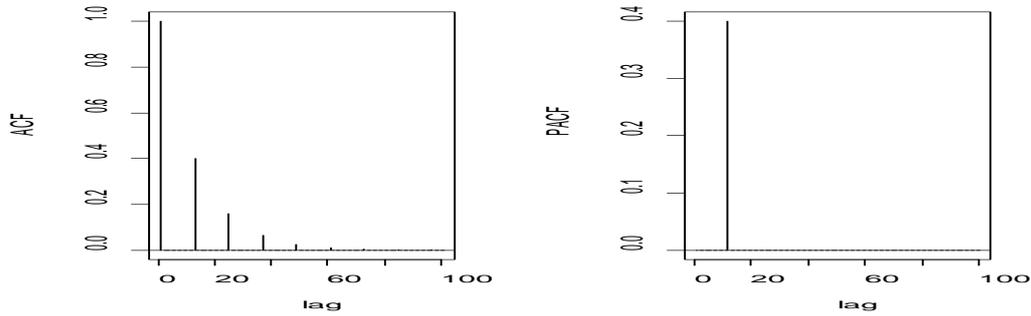
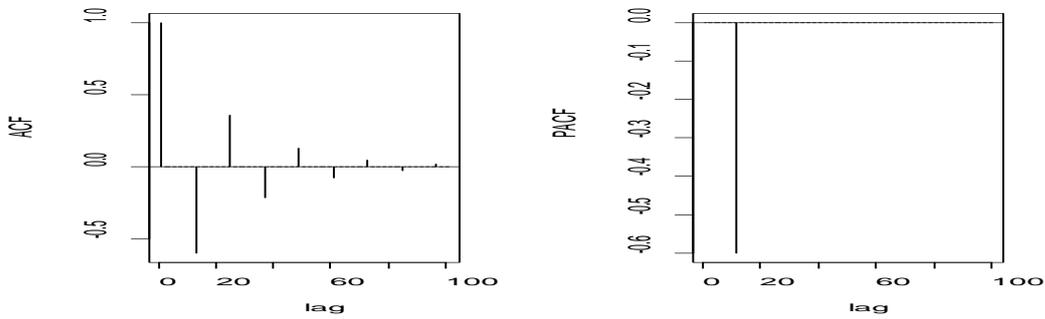


Gráfico 5.9



Ejemplo 5.3.2: Se considera en (5.14), $P = 1$ y $Q = 1$, resultando el siguiente modelo autorregresivo y de media móvil de primer orden y período estacional anual $s = 12$ meses ($ARMA(1,1)_{12}$) dado por la ecuación $(1 - \Phi B^{12})y_t = (1 + \Theta B^{12})A_t$. O bien desarrollando los operadores $y_t = \Phi y_{t-12} + A_t + \Theta A_{t-12}$.

Teniendo en cuenta que $y_{t-12} = \Phi y_{t-24} + A_{t-12} + \Theta A_{t-24}$ y sustituyendo en el modelo original:

$$\begin{aligned} y_t &= \Phi(\Phi y_{t-24} + A_{t-12} + \Theta A_{t-24}) + A_t + \Theta A_{t-12} \\ &= \Phi^2 y_{t-24} + \Phi A_{t-12} + \Phi \Theta A_{t-24} + A_t + \Theta A_{t-12} \end{aligned}$$

Además, $y_{t-24} = \Phi y_{t-36} + A_{t-24} + \Theta A_{t-36}$, así que reemplazando en la última expresión para y_t :

$$y_t = \Phi^3 y_{t-36} + \Phi^2 A_{t-24} + \Phi^2 \Theta A_{t-36} + \Phi A_{t-12} + \Phi A_{t-24} + A_t + \Theta A_{t-12}$$

Continuando con este procedimiento, se obtiene:

$$y_t = \Phi^k y_{t-12k} + \sum_{j=0}^{k-1} \Phi^j [A_{t-12j} + \Theta A_{t-12(j+1)}]$$

Luego, el procedimiento conduce a la siguiente forma para el proceso:

$$y_t = \sum_{j=0}^{\infty} \Phi^j [A_{t-12j} + \Theta A_{t-12(j+1)}]$$

La media del proceso es cero, pues:

$$\begin{aligned} E[y_t] &= E \left\{ \sum_{j=0}^{\infty} \Phi^j [A_{t-12j} + \Theta A_{t-12(j+1)}] \right\} = \sum_{j=0}^{\infty} \Phi^j E[A_{t-12j} + \Theta A_{t-12(j+1)}] \\ &= \sum_{j=0}^{\infty} \Phi^j \{E[A_{t-12j}] + \Theta E[A_{t-12(j+1)}]\} = \sum_{j=0}^{\infty} \Phi^j \cdot 0 = 0 \end{aligned}$$

Por otro lado, al calcular la función de autocovarianza del proceso, se tiene:

$$\gamma_k = E[(y_t - E[y_t])(y_{t+k} - E[y_{t+k}])]$$

Como la esperanza del proceso es cero, resulta:

$$\gamma_k = E \left\{ \left(\sum_{j=0}^{\infty} \Phi^j [A_{t-12j} + \Theta A_{t-12(j+1)}] \right) \left(\sum_{j=0}^{\infty} \Phi^j [A_{t+k-12j} + \Theta A_{t+k-12(j+1)}] \right) \right\}$$

Al desarrollar el producto se tiene:

$$\begin{aligned} \gamma_k &= E[A_t A_{t+k}] + \Theta E[A_t A_{t+k-12}] + \Phi E[A_t A_{t+k-12}] + \Phi \Theta E[A_t A_{t+k-12}] + \dots \\ &\quad + \Phi^h E[A_t A_{t+k-12h}] + \Phi^h \Theta E[A_t A_{t+k-12(k+1)}] + \dots + \Theta E[A_{t-12} A_{t+k}] \\ &\quad + \Theta^2 E[A_{t-12} A_{t+k-12}] + \Theta \Phi E[A_{t-12} A_{t+k-12}] + \Theta^2 \Phi E[A_{t-12} A_{t+k-12}] \\ &\quad + \Theta \Phi^2 E[A_{t-12} A_{t+k-24}] + \Phi^2 \Theta^2 E[A_{t-12} A_{t+k-36}] + \dots \\ &\quad + \Phi^h \Theta E[A_{t-12} A_{t+k-12h}] + \Phi^h \Theta^2 E[A_{t-12} A_{t+k-12(k+1)}] + \dots \\ &\quad + \Phi E[A_{t-12} A_{t+k}] + \Phi \Theta E[A_{t-12} A_{t+k-12}] + \Phi^2 E[A_{t-12} A_{t+k-12}] \\ &\quad + \Phi^2 \Theta E[A_{t-12} A_{t+k-24}] + \Phi^3 E[A_{t-12} A_{t+k-24}] + \Phi^3 \Theta E[A_{t-12} A_{t+k-36}] \\ &\quad + \Phi^4 E[A_{t-12} A_{t+k-36}] + \Phi^4 \Theta E[A_{t-12} A_{t+k-48}] + \dots \\ &\quad + \Phi^{h+1} E[A_{t-12} A_{t+k-12h}] + \Phi^{h+1} \Theta E[A_{t-12} A_{t+k-12(h+1)}] + \dots \end{aligned}$$

De esta expresión puede observarse lo siguiente:

Si $k = 0$:

$$\begin{aligned} \gamma_0 &= \sigma_A^2 + \Theta^2 \sigma_A^2 + \Phi \Theta \sigma_A^2 + \Phi \Theta \sigma_A^2 + \Phi^2 \sigma_A^2 + \Phi^2 \Theta^2 \sigma_A^2 + \Phi^3 \Theta \sigma_A^2 + \Phi^3 \Theta \sigma_A^2 + \Phi^4 \sigma_A^2 \\ &\quad + \Phi^4 \Theta^2 \sigma_A^2 + \Phi^5 \Theta \sigma_A^2 + \dots \end{aligned}$$

Factorizando, resulta:

$$\begin{aligned} \gamma_0 &= \sigma_A^2 (1 + \Theta^2 + \Phi \Theta + \Phi \Theta + \Phi^2 + \Phi^2 \Theta^2 + \Phi^3 \Theta + \Phi^3 \Theta + \Phi^4 + \Phi^4 \Theta^2 + \Phi^5 \Theta + \dots) \\ &= \sigma_A^2 [1 + \Theta^2 + 2\Phi \Theta + (1 + \Theta^2)\Phi^2 + 2\Phi \Theta^3 + (1 + \Theta^2)\Phi^4 + 2\Phi \Theta^5 + \dots] \end{aligned}$$

Agrupando términos, se obtiene:

$$\begin{aligned} \gamma_0 &= \sigma_A^2 \{ [1 + \Theta^2 + (1 + \Theta^2)\Phi^2 + (1 + \Theta^2)\Phi^4 + (1 + \Theta^2)\Phi^6 + \dots] \\ &\quad + [2\Phi \Theta + 2\Phi \Theta^3 + 2\Phi \Theta^5 + 2\Phi \Theta^7 + \dots] \} \end{aligned}$$

Y, por lo tanto:

$$\gamma_0 = \sigma_A^2 \left(\frac{1 + \theta^2}{1 - \Phi^2} + \frac{2\theta\Phi}{1 - \Phi^2} \right) = \frac{\sigma_A^2(1 + 2\theta\Phi + \theta^2)}{1 - \Phi^2}$$

Si $k = 12$:

$$\gamma_{12} = \theta\sigma_A^2 + \Phi\sigma_A^2 + \Phi\theta^2\sigma_A^2 + \Phi^2\theta\sigma_A^2 + \Phi^2\theta\sigma_A^2 + \Phi^3\sigma_A^2 + \Phi^3\theta^2\sigma_A^2 + \Phi^4\theta\sigma_A^2 + \Phi^4\theta\sigma_A^2 + \Phi^5\sigma_A^2 + \Phi^5\theta^2\sigma_A^2 + \dots$$

Factorizando y agrupando:

$$\gamma_{12} = \sigma_A^2 \{ (\theta + 2\theta\Phi^2 + 2\theta\Phi^4 + 2\theta\Phi^6 + \dots) + [(1 + \theta^2)\Phi + (1 + \theta^2)\Phi^3 + (1 + \theta^2)\Phi^5 + \dots] \}$$

Y, en consecuencia, se tiene:

$$\gamma_{12} = \left[\frac{2\theta}{1 - \Phi^2} + \frac{(1 + \theta^2)\Phi}{1 - \Phi^2} \right] = \frac{\sigma_A^2(2\theta + \Phi + \Phi\theta^2)}{1 - \Phi^2}$$

Si $k = 24$:

$$\gamma_{24} = \Phi\theta\sigma_A^2 + \Phi^2\sigma_A^2 + \Phi^2\theta^2\sigma_A^2 + \Phi^3\theta\sigma_A^2 + \Phi^3\theta\sigma_A^2 + \Phi^4\sigma_A^2 + \Phi^4\theta^2\sigma_A^2 + \Phi^5\theta\sigma_A^2 + \dots$$

Se obtiene de la expresión anterior:

$$\gamma_{24} = \Phi\sigma_A^2 \{ (\theta + 2\theta\Phi^2 + 2\theta\Phi^4 + \dots) + \Phi[1 + \theta^2 + (1 + \theta^2)\Phi^2 + \dots] \}$$

Y, luego:

$$\gamma_{24} = \Phi\sigma_A^2 \left(\frac{2\theta}{1 - \Phi^2} + \Phi \frac{1 + \theta^2}{1 - \Phi^2} \right) = \Phi\sigma_A^2 \frac{\Phi + 2\theta + \Phi\theta^2}{1 - \Phi^2} = \Phi\gamma_{12}$$

Generalizando y resumiendo, puede escribirse:

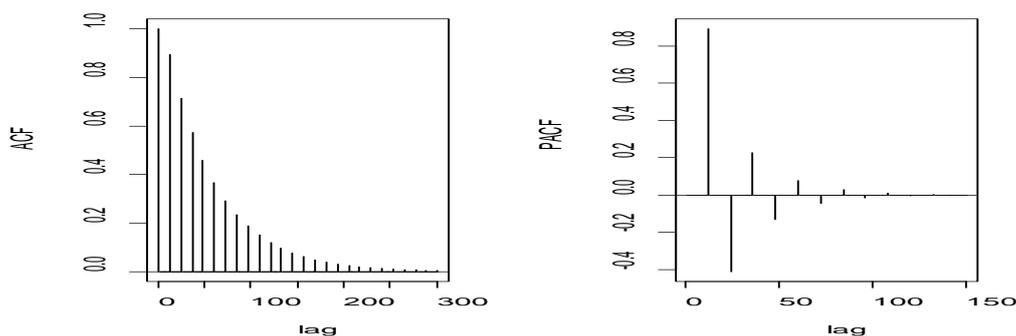
$$\gamma_k = \begin{cases} \frac{\sigma_A^2(1 + 2\theta\Phi + \theta^2)}{1 - \Phi^2} & \text{si } k = 0 \\ \frac{\sigma_A^2(2\theta + \Phi + \Phi\theta^2)}{1 - \Phi^2} & \text{si } k = \pm 12 \\ \Phi\gamma_{k-12} & \text{si } k = \pm 12.j \text{ con } j = 2,3,4, \dots \\ 0 & \text{si } k \neq \pm 12.j \text{ con } j = 1,2,3 \dots \end{cases}$$

Por otro lado, puede determinarse la función de autocorrelación, ACF , del proceso:

$$\rho_k = \begin{cases} 1 & \text{si } k = 0 \\ \frac{2\theta + \Phi + \Phi\theta^2}{1 + 2\theta\Phi + \theta^2} & \text{si } k = \pm 12 \\ \Phi\rho_{k-12} & \text{si } k = \pm 12.j \text{ con } j = 2,3,4, \dots \\ 0 & \text{si } k \neq \pm 12.j \text{ con } j = 1,2,3 \dots \end{cases}$$

Nótese que resulta dificultoso el cálculo de las funciones de autocovarianza y autocorrelación de un proceso estacional. En el gráfico siguiente se muestran los correlogramas de la ACF y de la $PACF$ del modelo del ejemplo 5.3.2 cuando $\Phi = 0,8$ y $\theta = 0,6$.

Gráfico 5.10



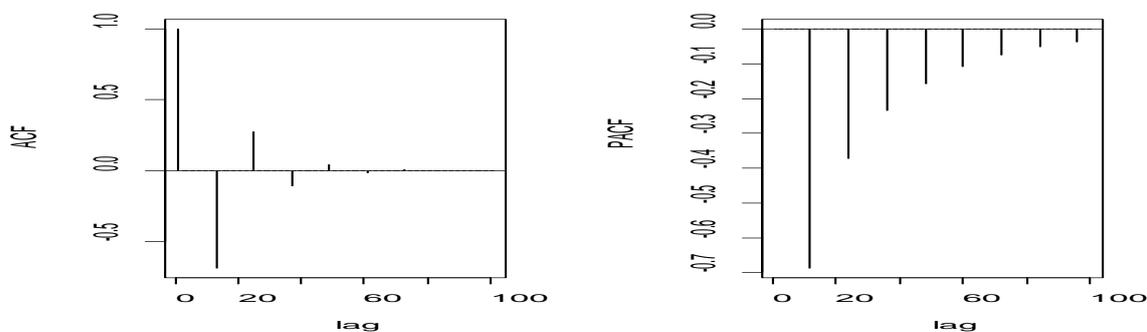
De modo general, si el modelo es un $MA(Q)_s$ la ACF tendrá Q coeficientes de autocorrelación diferentes de cero para los primeros Q retardos estacionales $k = s, 2s, 3s, \dots, Qs$, mientras que los demás coeficientes son cero. La $PACF$ presenta decrecimiento exponencial o con forma sinusoidal amortiguada.

Si el modelo es un $AR(P)_s$ la ACF decrece exponencialmente o en forma sinusoidal amortiguada con valores distintos de cero para los retardos estacionales $k = s, 2s, 3s, \dots, Ps$, mientras que la $PACF$ se trunca a partir del P -ésimo retardo estacional, es decir para $k = Ps$.

Si el modelo es un $ARMA(P, Q)_s$ la ACF es infinita con decrecimiento exponencial y diferente de cero para los retardos estacionales $k = s, 2s, 3s, \dots$. Los parámetros autorregresivos y de medias móviles influyen sobre los primeros Q -ésimos retardos estacionales; a partir del retardo $k = (Q + 1)s$ la función de autocorrelación viene determinada exclusivamente por la parte autorregresiva del modelo. Las raíces del polinomio autorregresivo determinarán el tipo de decrecimiento de la ACF , ya sea exponencial o en forma sinusoidal amortiguada. La $PACF$ existe solamente en los retardos estacionales; en los primeros P -ésimos retardos estacionales, los coeficientes de la $PACF$ dependen de la estructura autorregresiva y de media móvil, mientras que a partir del retardo $k = (P + 1)s$ quedan determinadas solamente por la estructura de media móvil.

En la siguiente figura se muestra los correlogramas de la ACF y de la $PACF$ del modelo del ejemplo 5.3.2 con $\Phi = -0,4$ y $\theta = -0,7$:

Gráfico 5.11



En general, los operadores estacional y no estacional pueden combinarse en una expresión multiplicativa conocida como modelo autorregresivo de media móvil estacional multiplicativo, que se denota $ARMA(p, q) \times (P, Q)_s$ y se caracteriza por la ecuación:

$$\Phi_P(B^s)\varphi(B)y_t = \Theta_q(B^s)\theta(B)A_t \quad (5.15)$$

Ejemplo 5.3.3: Considérese el siguiente modelo $ARMA(0,1) \times (1,0)_{12}$:

$$y_t = \Phi y_{t-12} + A_t + \theta A_{t-1}$$

donde $|\Phi| < 1$ y $|\theta| < 1$. Por un procedimiento análogo al empleado en el ejemplo anterior y teniendo en cuenta que y_t es estacionario, y_{t-12} , A_t y A_{t-1} no tienen correlación, se obtiene:

$$\gamma_0 = \frac{1 + \theta^2}{1 - \Phi^2} \sigma_A^2.$$

Además, puede mostrarse que si se multiplica el modelo por y_{t-k} para $k > 0$ y calculando las esperanzas, resulta $\gamma_1 = \Phi\gamma_{11} + \theta\sigma_A^2$ y $\gamma_k = \Phi\gamma_{k-12}$, para $k \geq 2$. Por lo tanto, la función de autocorrelación del modelo se resume en las siguientes ecuaciones (Shumway & Stoffer, 2011):

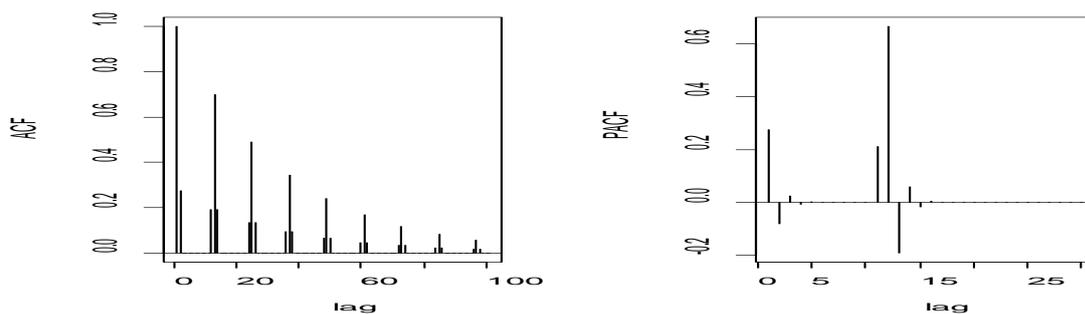
$$\rho_{12k} = \Phi^k \text{ con } k = 1, 2, \dots$$

$$\rho_{12k-1} = \rho_{12k+1} = \frac{\theta}{1 + \theta^2} \Phi^k \text{ con } k = 0, 1, 2, \dots$$

$$\rho_k = 0 \text{ en otro caso.}$$

Supóngase que se eligen los valores $\Phi = 0,7$ y $\theta = 0,3$ para el $ARMA(0,1) \times (1,0)_{12}$ de este ejemplo. En la siguiente figura se muestran la *ACF* y la *PACF* del modelo:

Gráfico 5.12



En los procesos $ARMA(p, q) \times (P, Q)_s$ los operadores $\varphi(B)$ y $\theta(B)$ (llamados *operadores regulares*), permiten modelar la correlación entre pares de componentes de y_t separados entre sí por $k = 1, 2, 3, \dots$ períodos básicos, como pueden ser meses o trimestres. Los operadores $\Phi(B)$ y $\Theta(B)$ (denominados *operadores anuales*), se emplean para describir la correlación entre pares de componentes de y_t separados entre sí por $k = s, 2s, 3s, \dots$ períodos básicos (Mauricio, 2007).

El correlograma de la ACF de un modelo $ARMA(p, q) \times (P, Q)_s$ refleja en los retardos $k = 1, 2, 3, \dots$ la estructura dada por la parte regular $ARMA(p, q)$, mientras que en los retardos estacionales $k = s, 2s, 3s, \dots$ se refleja la estructura de la parte estacional $ARMA(P, Q)_s$. Igual comportamiento presenta la $PACF$ de estos modelos, es decir, en los retardos $k = 1, 2, 3, \dots$ la forma dada por la parte regular y en los retardos estacionales la forma dada por la parte estacional.

Ejemplo 5.3.4: Considérese el siguiente modelo $ARMA(1,1) \times (1,1)_{12}$ que expresado en términos de la ecuación (5.15) queda

$$(1 - \Phi B^{12})(1 - \varphi B)y_t = (1 + \Theta B^{12})(1 + \theta B)A_t$$

Desarrollando los productos, resulta:

$$(1 - \varphi B - \Phi B^{12} + \varphi \Phi B^{13})y_t = (1 + \theta B + \Theta B^{12} + \theta \Theta B^{13})A_t$$

Por lo tanto, el modelo puede expresarse mediante la siguiente ecuación:

$$y_t = \varphi y_{t-1} + \Phi y_{t-12} - \varphi \Phi y_{t-13} + A_t + \theta A_{t-1} + \Theta A_{t-12} + \theta \Theta A_{t-13}$$

Nótese que la naturaleza multiplicativa del modelo hace que los coeficientes de y_{t-13} y de A_{t-13} sean el resultado del producto entre los coeficientes de y_{t-1} e y_{t-12} y de A_{t-1} y A_{t-12} , respectivamente, en lugar de parámetros libres.

Supóngase que en el modelo se elige $\Phi = 0,2$, $\Theta = 0,6$, $\varphi = 0,7$ y $\theta = 0,5$, entonces puede simularse para $n = 200$ el proceso junto con la ACF y la $PACF$ asociadas al mismo, lo cual se observa en la siguiente figura:

Gráfico 5.13

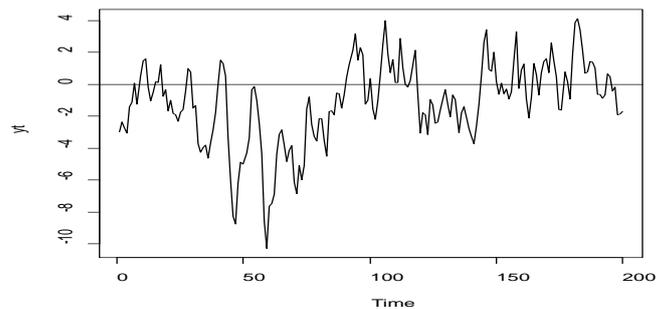
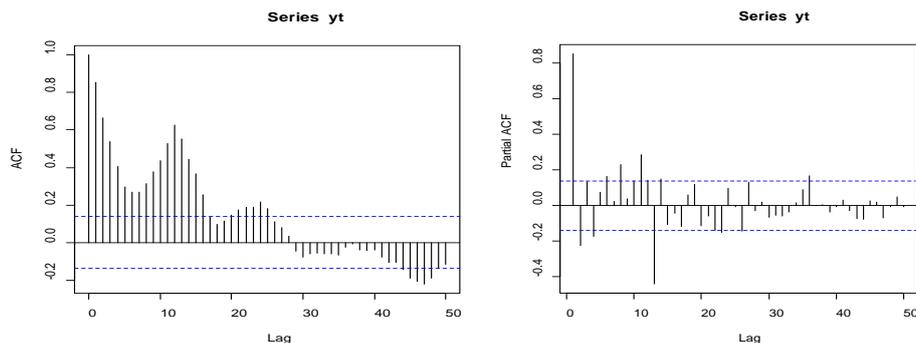


Gráfico 5.14



Con frecuencia es posible encontrar series que provienen de procesos estacionales no estacionarios. Por ejemplo, series mensuales donde la media del proceso varía mes a mes y, por lo tanto, la serie no es estacionaria en media, lo que se refleja en cambios sistemáticos en el nivel de la serie. En estos casos, se supone que la estacionalidad es sólo aproximadamente constante y presenta evolución estocástica. Estos procesos pueden modelarse mediante la expresión:

$$y_t = S_t + A_t,$$

donde S_t es un componente estacional que varía lentamente de un año a otro, de acuerdo con un paseo aleatorio y tal que

$$S_t = S_{t-s} + v_t.$$

En este modelo, A_t y v_t son procesos de ruido blanco no correlacionados.

A los efectos de eliminar la no estacionariedad en media generada por la componente estacional, se toman diferencias estacionales.

Definición 5.3.2: El operador *diferencia estacional de período s y orden D* está dado por la expresión

$$\nabla_s^D y_t = (1 - B^s)^D y_t \quad (5.16)$$

donde D toma valores enteros positivos.

De este modo, si y_t es un proceso estacional de período s , entonces $\nabla_s^D y_t$ no presenta estacionalidad. En general $D = 1$ es suficiente para lograr la estacionariedad en media. Utilizando el operador (7.5) con $D = 1$, se obtiene:

$$\nabla_s y_t = (1 - B^s) y_t = y_t - y_{t-s} = S_t + A_t - S_{t-s} + A_{t-s} = v_t + A_t + A_{t-s}.$$

Este modelo es un $MA(1)_s$ estacionario e invertible.

Las ideas expresadas en estos párrafos permiten dar la siguiente definición:

Definición 5.3.3: El *modelo autorregresivo integrado de media móvil estacional multiplicativo*, o modelo *SARIMA*, está dado por

$$\Phi_P(B^s) \varphi_p(B) \nabla_s^D \nabla^d y_t = \delta + \Theta_Q(B^s) \theta_q(B) A_t \quad (5.17)$$

donde A_t es un proceso de ruido blanco Gaussiano. El modelo general se denota como $ARIMA(p, d, q) \times (P, D, Q)_s$. Los polinomios autorregresivo y de media móvil habituales son $\varphi_p(B)$ y $\theta_q(B)$, de órdenes p y q respectivamente, los polinomios estacionales autorregresivo y de media móvil son $\Phi_P(B^s)$ y $\Theta_Q(B^s)$ de órdenes P y Q respectivamente, el operador diferencia regular de orden d es $\nabla^d = (1 - B)^d$ y $\nabla_s^D = (1 - B^s)^D$ es el operador diferencia estacional de orden D .

Ejemplo 5.3.5: De acuerdo con (5.17), un $ARIMA(1,1,1) \times (1,1,1)_{12}$ se expresa mediante la ecuación:

$$(1 - \Phi B^{12})(1 - \varphi B)(1 - B^{12})(1 - B)y_t = (1 + \Theta B^{12})(1 + \theta B)A_t$$

Donde $\Phi(B^{12}) = 1 - \Phi B^{12}$, $\varphi(B) = 1 - \varphi B$, $\nabla_{12}^{D=1} = 1 - B^{12}$, $\nabla^{d=1} = 1 - B$, $\Theta(B^{12}) = 1 + \Theta B^{12}$ y $\theta(B) = 1 + \theta B$. Desarrollando los productos resulta:

$$(1 - B - B^2 + B^{13} - \varphi B + \varphi B^2 + \varphi B^{13} - \varphi B^{14} - \Phi B^{12} + \Phi B^{13} + \Phi B^{24} - \Phi B^{25} + \varphi \Phi B^{13} - \varphi \Phi B^{14} - \varphi \Phi B^{25} + \varphi \Phi B^{26})y_t = (1 + \theta B + \Theta B^{12} + \theta \Theta B^{13})A_t$$

Agrupando, se obtiene:

$$[1 - (1 + \varphi)B + \varphi B^2 - (1 + \Phi)B^{12} + (1 + \varphi + \Phi + \varphi \Phi)B^{13} - (\varphi + \varphi \Phi)B^{14} + \Phi B^{24} - (\Phi + \varphi \Phi)B^{25} + \varphi \Phi B^{26}]y_t = (1 + \theta B + \Theta B^{12} + \theta \Theta B^{13})A_t$$

Por lo tanto:

$$y_t = (1 + \varphi)y_{t-1} - \varphi y_{t-2} + (1 + \Phi)y_{t-12} - (1 + \varphi + \Phi + \varphi \Phi)y_{t-13} + (\varphi + \varphi \Phi)y_{t-14} - \Phi y_{t-24} + (\Phi + \varphi \Phi)y_{t-25} - \varphi \Phi y_{t-26} + A_t + \theta A_{t-1} + \Theta y_{t-12} + \theta \Theta A_{t-13}$$

Si en la expresión (6.15) $p = 0, q = 0$ y $d = 0$, entonces se obtiene el modelo $ARIMA(P, D, Q)_s$ que es un modelo estacional puro que en términos generales se expresa como

$$\Phi_P(B^s)\nabla_s^D y_t = \delta + \Theta_Q(B^s)A_t.$$

Ejemplo 5.3.6: Considérese un período estacional $s = 4$ (lo que corresponde a una serie trimestral). Un modelo $ARIMA(1,1,2)_4$ estacional puro se puede escribir mediante la siguiente ecuación:

$$(1 - \Phi B^4)(1 - B^4)y_t = (1 + \Theta_1 B^4 + \Theta_2 B^8)A_t$$

En este caso $\Phi(B^4) = 1 - \Phi B^4$, $\nabla_4^{D=1} = 1 - B^4$ y $\Theta(B^4) = 1 + \Theta_1 B^4 + \Theta_2 B^8$.

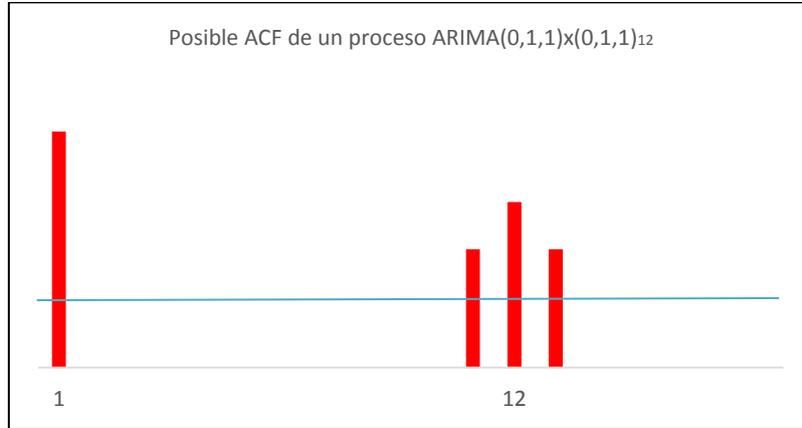
Las características de las funciones de autocorrelación de los modelos ARIMA estacionales multiplicativos se discuten en una muy amplia variedad de libros y documentos. Seguidamente se exponen algunas de estas particularidades acompañadas de un posible correlograma asociado a ciertos modelos que suelen encontrarse frecuentemente en la práctica¹⁷.

(a) El siguiente correlograma corresponde a la función de autocorrelación de un modelo $ARIMA(0,1,1)x(0,1,1)_{12}$. Nótese que en la estructura que refleja la parte regular del modelo $\rho_k \neq 0$ si $k = 1$ y $\rho_k = 0$ cuando $k = 2, \dots, 10$. Recién en el primer retardo estacional y alrededor del mismo la ACF es distinta de cero. Esto es $\rho_{k-1} = \rho_{k+1} \neq 0$ para $k = 12$ y tomando un valor inferior a $\rho_{12} \neq 0$. En el resto de los retardos estacionales la ACF vale cero. Los valores de los coeficientes en los retardos que son significativamente distintos de cero pueden ser positivos, negativos o de signos alternados dependiendo a su vez de los signos de los parámetros del modelo. Además, es importante señalar que en la práctica, un valor se considera cero cuando su módulo es inferior a $\frac{2}{\sqrt{n}}$ donde n es el tamaño de la serie (Pérez,

¹⁷ En el libro de Alberto Landro y Mirta González, *Elementos de Econometría de los Fenómenos Dinámicos*, citado en este documento, se exponen algunas de estas funciones en su expresión formal, pág. 355 en adelante.

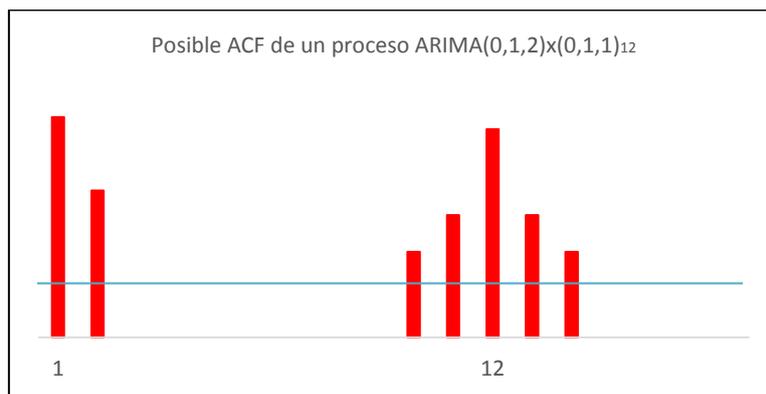
2001). En el caso de Gretl, por ejemplo, las “líneas” de confianza reflejadas en el correlograma se estiman como $\pm \frac{1,96}{\sqrt{n}}$ de cero.

Gráfico 5.15



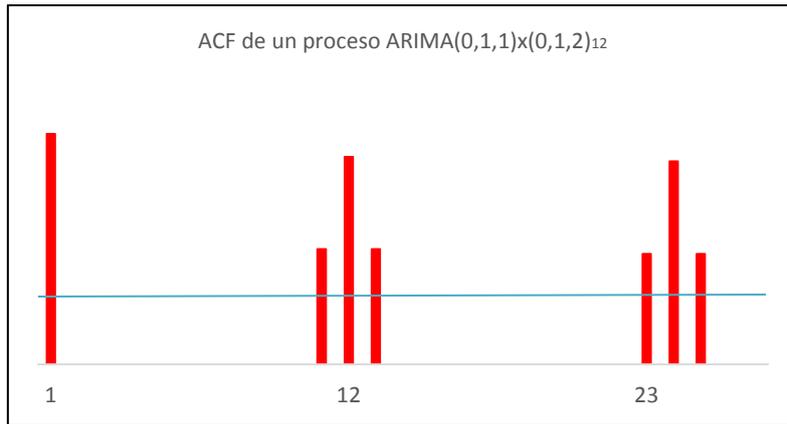
(b) Si se observa la ACF de un modelo $ARIMA(0,1,2)x(0,1,1)_{12}$ la diferencia con el caso anterior es que aparecen los dos primeros retardos asociados a la parte regular distintos de cero, $\rho_k \neq 0$ cuando $k = 1,2$, mientras que $\rho_k = 0$ cuando $k = 3, \dots, 9$. En la parte estacional se observa que $\rho_{12} \neq 0$ mientras que $\rho_{11} = \rho_{13} \neq 0$ y $\rho_{10} = \rho_{14} \neq 0$, mostrando estos coeficientes un patrón de comportamiento similar a la parte regular alrededor del retardo estacional. El resto de los retardos estacionales es cero. Si existen retardos diferentes de cero entre el período, en general no se tienen en cuenta.

Gráfico 5.16



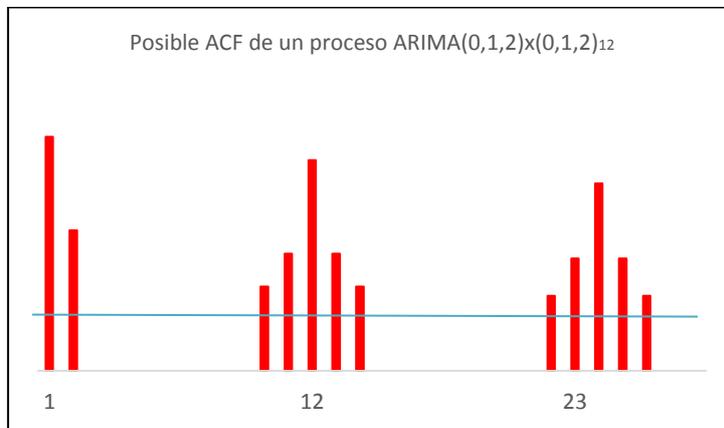
(c) En la ACF de un modelo $ARIMA(0,1,1)x(0,1,2)_{12}$ la parte regular tiene el primer retardo diferente de cero y el resto desde $k = 2, \dots, 10$ valen cero, lo que se asocia al orden $q = 1$. Hay dos “bloques estacionales” que identifican el orden estacional $Q = 2$, donde los retardos alrededor de $\rho_{12} \neq 0$ y de $\rho_{24} \neq 0$ son distintos de cero y tal que $\rho_{k-1} = \rho_{k+1}$ y $\rho_{2k-1} = \rho_{2k+1}$ para $k = 12$. El decrecimiento lento de ρ_1, ρ_{12} y ρ_{24} caracteriza los órdenes de diferenciación regular y estacional $d = D = 1$.

Gráfico 5.17



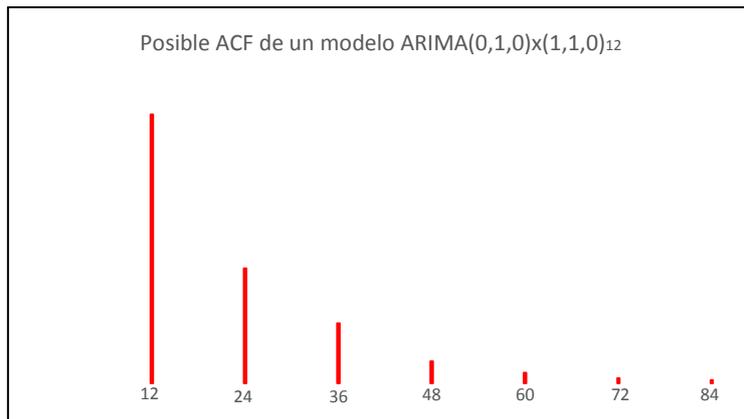
(d) Si el modelo en cuestión se trata de un $ARIMA(0,1,2)x(0,1,2)_{12}$ se observa que la parte del correlograma que refleja la estructura regular del proceso tiene los dos primeros retardos distintos de cero (lo que está determinado por el orden de media móvil regular $q = 2$), y el resto, para $k = 3, \dots, 9$, es cero. La estructura estacional tiene dos “bloques estacionales” asociados al orden estacional $Q = 2$, con los retardos ρ_{12} y de ρ_{24} son diferentes de cero. Además, se observa que $\rho_{k-1} = \rho_{k+1} \neq 0$ y $\rho_{k-2} = \rho_{k+2} \neq 0$ para $k = 12$. La función de autocorrelación vale cero cuando $k = 15, \dots, 21$ mientras que, nuevamente, $\rho_{2k-1} = \rho_{2k+1} \neq 0$ y $\rho_{2k-2} = \rho_{2k+2} \neq 0$ para $k = 12$. El decrecimiento lento caracteriza los órdenes de diferenciación regular y estacional $d = D = 1$.

Gráfico 5.18



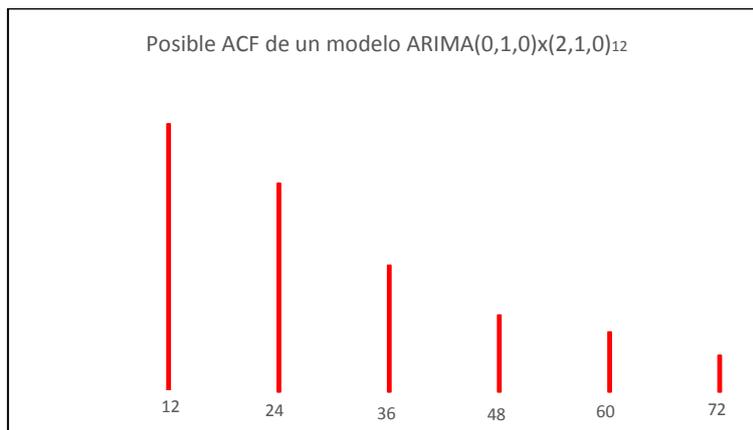
(e) El caso de la ACF de un modelo $ARIMA(0,1,0)x(1,1,0)_{12}$ difiere notablemente de los casos anteriores. En general, la función de autocorrelación toma valores distintos de cero en los retardos estacionales siguiendo un decrecimiento lento de tipo exponencial donde el primer retardo estacional es significativamente distinto de cero.

Gráfico 5.19



(f) El correlograma siguiente representa a la ACF de un modelo $ARIMA(0,1,0) \times (2,1,0)_{12}$. Como en el caso anterior se observa el decrecimiento exponencial lento en los retardos estacionales, donde los dos primeros son significativamente distintos de cero. En los retardos no estacionales, tanto en este caso como en el anterior, la ACF vale cero.

Gráfico 5.20



Capítulo 6

Lo estudiado en los capítulos anteriores, permite decir que un modelo *ARIMA* describe una observación como una función lineal de datos anteriores más una serie de errores que se deben al azar. Además, estos modelos pueden contener un componente cíclico o un componente estacional como también un componente tendencial. En este capítulo se analizará la metodología de modelización *ARIMA*, concretamente sobre las series de tiempo que conforman el campo de estudio, es decir, las series seleccionadas de registros de precipitaciones mensuales de lluvia de la Provincia de Entre Ríos.

En secciones previas se estableció que una serie de tiempo es una realización muestral particular de un proceso estocástico. Un mismo proceso puede dar origen a infinitas series temporales, las cuales, aunque diferentes unas de otras, comparten la misma estructura dinámica subyacente, que es la que le confiere el proceso.

Según (5.17) la ecuación que define un modelo *ARIMA* estacional de período s es

$$\Phi_P(B^s)\phi_p(B)\nabla_s^D\nabla^d y_t = \delta + \Theta_Q(B^s)\theta_q(B)A_t.$$

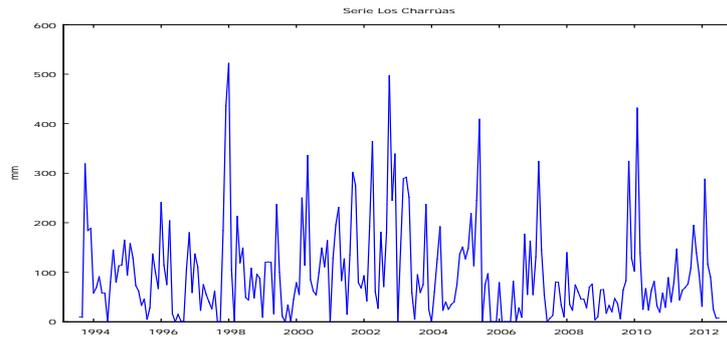
El objetivo de la modelización *ARIMA* consiste en identificar el modelo generador de la serie y estimar los parámetros del mismo a partir de las observaciones que conforman la serie temporal.

El procedimiento de modelización se caracteriza por obtener todos los elementos necesarios del modelo pero conservando los *mínimos necesarios* para describir el fenómeno estudiado, característica que se conoce como *parsimonia*. En este procedimiento se trabajará de acuerdo a las siguientes fases: disponibilidad de datos, representación gráfica de las series a estudiar, análisis de estacionariedad en media y en varianza, identificación del modelo, estimación de los coeficientes del modelo y contraste, análisis de los errores, selección del modelo.

6.1. Disponibilidad de los datos y representación gráfica de las series de registros de lluvia acumulada mensualmente

Las series temporales a estudiar corresponden a registros mensuales de precipitaciones de lluvia de 22 estaciones pluviométricas de la Provincia de Entre Ríos. Estas series de datos se tomaron de la página oficial de la Dirección de Hidráulica de la Provincia de Entre Ríos y fueron clasificadas en cuatro grupos según la longitud anual de las mismas. A continuación, a modo ilustrativo, se muestra el gráfico de la serie procedente del registro de Los Charrúas. Los gráficos de las demás series se pueden consultar en el Capítulo 9 en los apéndices correspondientes.

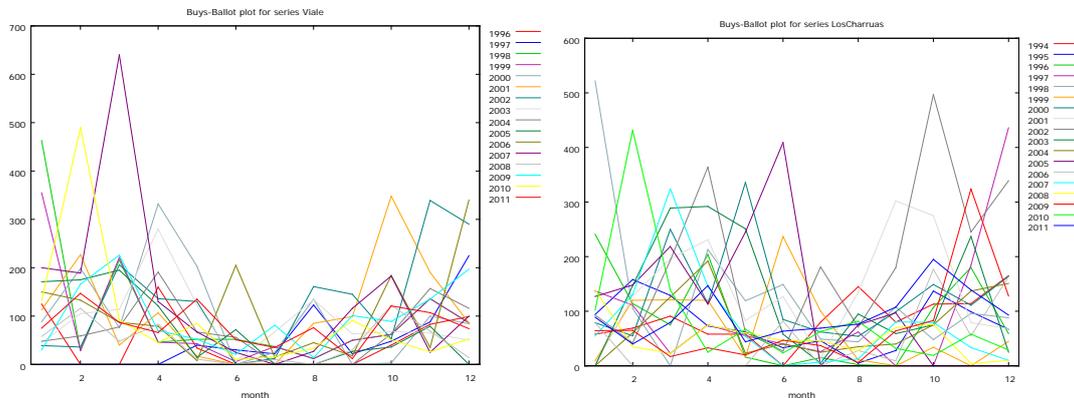
Gráfico 6.1



6.2. Comportamiento de las series de registros de lluvia de acuerdo con las estaciones del año

La representación gráfica de la serie no siempre permite observar con claridad las particularidades que subyacen al fenómeno, como por ejemplo las influencias que las estaciones del año tienen sobre la serie. Un tipo de gráfico empleado en estos casos es el de *Buys-Ballot* que muestra la evolución de la serie en cada mes a lo largo de los años. Estos gráficos son útiles a veces para tener una idea preliminar del comportamiento estacional de un fenómeno.

Gráfico 6.2



En el gráfico de Buys-Ballot para las seriea de Viale y de Los Charrúas, aunque las curvas asociadas a cada año son muchas y se enredan unas con otras, se puede observar que en los meses de primavera, verano y parte de otoño se alcanzan picos más altos, mientras que en invierno se llega, en líneas generales, a los registros acumulados más bajos. Esto muestra cómo las series de registros de lluvia están afectadas al cambio estacional.

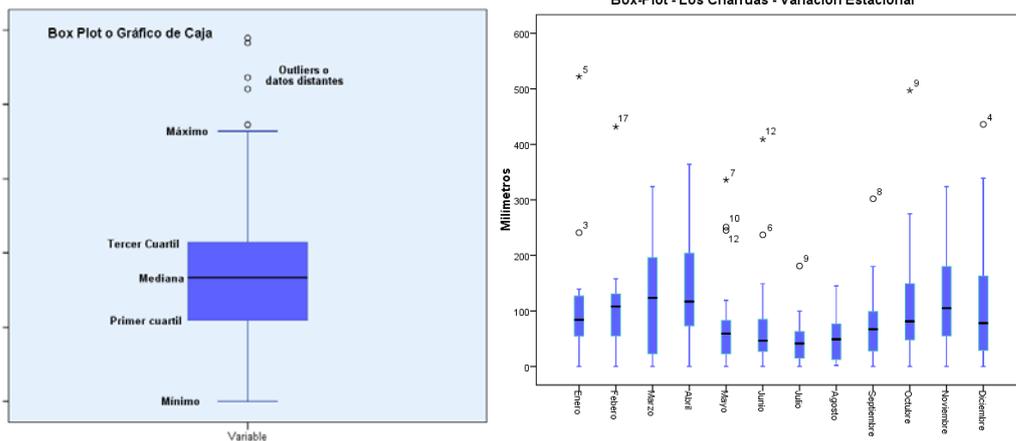
Otra forma de observar los mismos patrones consiste en representar cada mes de la serie a lo largo de los años a través de un gráfico de caja o box-plot. De la tabla del ejemplo 5.1, puede tomarse, por ejemplo, la secuencia de observaciones correspondientes al mes 1 a lo largo de los años:

$$y_1, y_{13}, y_{25}, \dots, y_{1+12(n-1)}$$

Este conjunto de datos presenta un valor máximo y un mínimo, y pueden calcularse la mediana y el primer y tercer cuartil¹⁸ que se denotan q_1 y q_3 , respectivamente. El rango intercuartílico, $RI = q_3 - q_1$, mide la dispersión del 50% central de los datos, por lo tanto dará una idea de la variación de los mismos con respecto a su centro. El gráfico de caja o Box-Plot, se construye con estas cinco medidas y, además permite detectar la presencia de outliers o datos distantes.¹⁹ El box-plot resulta de mucha utilidad a la hora de comparar la variabilidad de un grupo de variables.

En el Gráfico 6.3, a la izquierda se aprecia un diagrama de caja general que indica las cinco medidas asociadas y a la derecha el box-plot para los registros correspondientes a la estación de Los Charrúas. Nótese la característica estacional del fenómeno. Al igual que lo reflejado en el gráfico de Buys-Ballot en los meses correspondientes a primavera, verano y parte de otoño, la mediana muestra un nivel más alto que en los meses de invierno y las cajas están definidas por rangos intercuartilares mayores y hay incremento de los valores máximos. Todas las series estudiadas, muestran un patrón de comportamiento similar, reflejando las características estacionales del fenómeno a nivel general (ver Capítulo 9).

Gráfico 6.3



6.3. Análisis de estacionariedad

En esta sección se procede a aplicar la transformación de Box-Cox discutida en el Capítulo 5 y luego se estudia la presencia de raíces unitarias mediante dos contrastes comúnmente empleados en series con presencia de algún tipo de estacionalidad, a saber, el Test HEGY y el Test de Canova y Hansen.

¹⁸ El primer cuartil deja por debajo de sí el 25% de las observaciones cuando estas están ordenadas de menor a mayor; la mediana, que coincide con el segundo cuartil, separa el 50% inferior de los datos ordenados del 50% superior; el tercer cuartil separa el 75% inferior del 25% superior de los datos ordenados.

¹⁹ Estos valores se identifican por quedar fuera de las barreras internas superior e inferior. La barrera interna inferior se calcula restándole al primer cuartil 1,5 veces el rango intercuartílico, mientras que para la barrera interna superior se suma al tercer cuartil 1,5 veces el rango intercuartílico.

6.3.1. Transformación de Box-Cox

A los efectos de estudiar la estacionariedad en varianza, se procede a trabajar con la transformación de Box-Cox. En principio, recuérdese que la transformación de Box-Cox de parámetros λ y m está dada por la expresión:

$$g(y_t) = \frac{(y_t + m)^\lambda - 1}{\lambda}$$

En la siguiente tabla se muestra los valores de λ y de m obtenidos empleando el software gretl, mientras que en el Capítulo 9 se presentan los gráficos de las transformaciones de las series.

Cabe señalar que hay casos en los que m es muy cercano a cero y el parámetro λ muy próximo a 0,5.

Tabla 6.1

<i>Estación de Registro</i>	<i>Transformación de Box-Cox</i>	
	λ	Skewness
<i>Los Charrúas</i>	0,44	0,00513
<i>Lucas Sur</i>	0,44	0,00456
<i>Octavo Distrito</i>	0,4	0,0122
<i>Puente de Hierro</i>	0,45	0,0116
<i>San Víctor</i>	0,4	0,0138
<i>Santa Anita</i>	0,43	0,0151
<i>Séptimo Distrito</i>	0,42	0,0231
<i>Viale</i>	0,43	0,00913
<i>Villa Paranacito</i>	0,5	0,00202
<i>Colón</i>	0,46	0,0119
<i>La Lila</i>	0,42	0,0152
<i>Lucas González</i>	0,37	0,00982
<i>Santa María del Tatutí</i>	0,47	0,00663
<i>Villa Elisa</i>	0,44	0,0136
<i>Antelo</i>	0,41	0,00738
<i>Febre</i>	0,42	0,0142
<i>Feliciano</i>	0,35	0,00562
<i>San Jaime</i>	0,46	0,0158
<i>San Salvador</i>	0,37	0,00674
<i>La Paz</i>	0,42	0,0205
<i>San Gustavo</i>	0,42	0,0037
<i>Paraná</i>	0,38	0,0164

6.3.2. Test HEGY

El estudio de los tests de estacionariedad lo inician Dickey y Fuller (1979) a través de un contraste que permite detectar la presencia de raíces unitarias. Los autores del test

consideran, en principio, un proceso $AR(1)$ dado por la ecuación $y_t = \varphi_1 y_{t-1} + A_t$. Cuando $|\varphi_1| < 1$ el proceso es integrado de orden cero, $I(0)$, (en otras palabras, causal y estacionario); cuando $\varphi_1 = 1$, y_t es integrado de orden uno, $I(1)$, es decir un paseo aleatorio. Con el propósito de averiguar si el proceso en cuestión es estacionario, es suficiente realizar un test formal de hipótesis, donde se contrasten la hipótesis nula, $H_0: y_t \text{ es } I(1)$, frente a la alternativa $H_1: y_t \text{ es } I(0)$. O bien, $H_0: \varphi_1 = 1$ frente a $H_1: |\varphi_1| < 1$.²⁰ Utilizando mínimos cuadrados ordinarios, Dickey y Fuller proponen un estadístico conocido como estadístico de Dickey-Fuller que, bajo la hipótesis nula de no estacionariedad, no sigue ninguna distribución conocida, por lo que sus creadores, empleando el método de simulación de Monte Carlo, confeccionaron una tabla con los niveles críticos adecuados en base al tamaño de muestra n y al nivel de significación α . El contraste, que abreviadamente se denota DF , consiste en rechazar la hipótesis nula H_0 en favor de la alternativa H_1 cuando el valor calculado del estadístico de contraste es menor que el valor crítico de la distribución de Dickey-Fuller para un determinado nivel de significación α .

El contraste fue ampliado al caso en que el modelo incluya constante y tendencia. Esta ampliación del test se conoce como contraste de Dickey-Fuller aumentado, ADF .²¹ Posteriormente, Dickey, Fuller y Hasza (1984), desarrollan una extensión del test que permite contrastar la presencia de raíces unitarias estacionales²², aunque el test no tiene en cuenta la posibilidad de que dicha raíz esté presente solamente en alguna de las frecuencias estacionales. Este problema lo soluciona el test HEGY.

Esta prueba fue desarrollada por Hylleberg, Engle, Granger y Yoo en 1990 y tiene la ventaja de testear las raíces unitarias estacionales separadamente para cada frecuencia, por lo que resulta un contraste aplicado ampliamente en el tratamiento de series estacionales. El test HEGY fue propuesto en principio para series con estacionalidad trimestral. Luego, Franses (1990), y Beaulieu y Miron (1993) lo extienden para series mensuales.

Para comenzar se considera que la serie de tiempo está generada por un proceso autorregresivo de orden p de la forma $\varphi(B)y_t = A_t$, donde B es el operador de retardos y A_t es un ruido blanco. Si $\varphi(B) = 1 - B^s$ que es el operador diferencia estacional, entonces s es el número que señala la periodicidad de la serie. La ecuación $\varphi(z) = 0$ tiene s raíces en el círculo unitario:

$$z_k = e^{\frac{2\pi i}{s}k} = \cos\left(\frac{2k\pi}{s}\right) + i \sin\left(\frac{2k\pi}{s}\right), k = 0, 1, 2, \dots, s-1, \quad (6.1)$$

²⁰ Obsérvese que H_1 significa que $0 < \varphi_1 < 1$ puesto que $\varphi_1 > 1$ caracteriza a un proceso explosivo, como se analizó en el capítulo 4, mientras que $\varphi_1 < 0$ se da con muy poca frecuencia en la práctica. Nótese, además, que no se puede emplear el test t habitual ya que la hipótesis nula que generalmente se contrasta es la nulidad del parámetro $\varphi_1 = 0$, y en este caso, por el contrario lo que se quiere contrastar es $H_0: \varphi_1 = 1$. Así, bajo hipótesis nula cierta la varianza del proceso no sería estacionaria y la distribución del parámetro (y del estadístico t), no puede caracterizarse de modo efectivo. La distribución del parámetro es una función de procesos Brownianos. Ver (Wei, 2006).

²¹ Además del contraste de Dickey-Fuller existen otros de aplicación habitual. Entre los más conocidos puede mencionarse el contraste de Kwiatkowski, Phillips, Schmidt y Shin (1992), $KPSS$, o el test de Phillips y Perron, entre otros.

²² Un desarrollo formal de los tests de Dickey-Fuller y Dickey-Fuller-Hasza puede encontrarse en Wei, W. (2006), *Time Series Analysis. Univariate and multivariate methods*. Pearson Addison Wesley. USA.

donde i es la unidad imaginaria²³. Cada z_k está relacionada con una frecuencia específica $\frac{2k\pi}{s}$. Cuando $k = 0$, la raíz z_k en (6.1) se llama raíz unitaria no estacional (asociada a la componente regular). Las otras raíces z_k en (6.1) se llaman raíces estacionales. Con excepción de las raíces en la frecuencia 0 y π , las raíces en (6.1) son pares conjugados.

En particular, si $s = 4$ (serie trimestral), puede escribirse $\varphi(z)$ en forma factorizada de la siguiente manera:

$$\varphi(z) = 1 - z^4 = (1 - z)(1 + z)(1 - iz)(1 + iz)$$

De (6.1) la raíz asociada al factor $(1 - z)$ tiene frecuencia cero, pues $z_0 = \cos\left(\frac{2 \cdot 0 \cdot \pi}{4}\right) + i \sin\left(\frac{2 \cdot 0 \cdot \pi}{4}\right) = 1 + i \cdot 0 = 1$. Por lo tanto, es la que corresponde a la estructura regular.

Cada raíz está asociada a un determinado ciclo y el tiempo necesario para alcanzar ese ciclo puede estudiarse aplicando cada factor de $\varphi(z)$ sobre el proceso y_t . En efecto, el factor asociado a la raíz $z_2 = \cos\left(\frac{2 \cdot 2 \cdot \pi}{4}\right) + i \sin\left(\frac{2 \cdot 2 \cdot \pi}{4}\right) = \cos(\pi) + i \sin(\pi) = -1 + i \cdot 0 = -1$ es $(1 + z)$. Si se escribe utilizando el operador de retardo y se aplica sobre y_t produce:

$$(1 + B)y_t = y_t + By_t = y_t + y_{t-1} = 0$$

Entonces $y_t = -y_{t-1}$. Incrementando una unidad el índice temporal del proceso: $y_{t+1} = -y_t$ y por la ecuación anterior $y_{t+1} = -(-y_{t-1}) = y_{t-1}$. Si se vuelve a aumentar una unidad el índice temporal: $y_{t+2} = y_t$. Obsérvese que en este caso se alcanza el ciclo completo en dos períodos (cada período es un trimestre). Por este motivo, se asocia esta raíz de frecuencia π con el semestre (dos trimestres).

Por último, las raíces $z_1 = \cos\left(\frac{2 \cdot 1 \cdot \pi}{4}\right) + i \sin\left(\frac{2 \cdot 1 \cdot \pi}{4}\right) = \cos\left(\frac{\pi}{2}\right) + i \sin\left(\frac{\pi}{2}\right) = 0 + i \cdot 1 = i$ y $z_3 = \cos\left(\frac{2 \cdot 3 \cdot \pi}{4}\right) + i \sin\left(\frac{2 \cdot 3 \cdot \pi}{4}\right) = \cos\left(\frac{3}{2}\pi\right) + i \sin\left(\frac{3}{2}\pi\right) = 0 + i \cdot (-1) = -i$ están asociadas al factor $(1 - iz)(1 + iz) = 1 + z^2$. Si se escribe en términos del operador de retardo y se aplica sobre y_t resulta: $(1 + B^2)y_t = y_t + B^2y_t = y_t + y_{t-2} = 0$, por lo que $y_t = -y_{t-2}$. Analizando la variación unitaria del índice temporal, se obtiene:

Primer incremento unitario del índice temporal: $y_{t+1} = -y_{t-1}$

Segundo incremento unitario del índice temporal: $y_{t+2} = -y_t$. Luego, $y_{t+2} = -(-y_{t-2}) = y_{t-2}$.

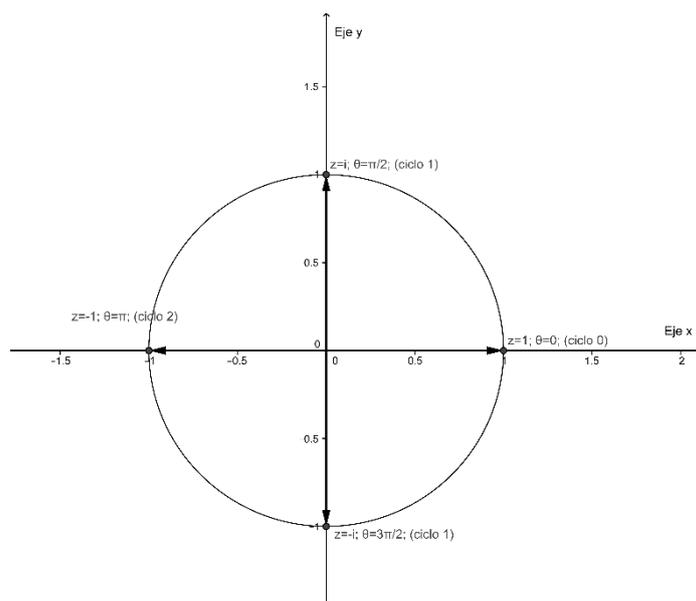
Tercer incremento unitario del índice temporal: $y_{t+3} = y_{t-1}$

Cuarto incremento del índice temporal: $y_{t+4} = y_t$.

Obsérvese que se alcanza un ciclo en cuatro incrementos del índice temporal. Cuatro trimestres equivalen a un año. Por lo que la frecuencia asociada a estas raíces es anual y como no es posible distinguir el efecto que cada raíz imaginaria tiene se toman conjuntamente.

²³ Nótese que la ecuación $1 - z^s = 0$ es equivalente a la ecuación $e^{i\theta s} = 1$, pues $z = e^{i\theta}$, cuyas soluciones, naturalmente, son las z_k de (6.1).

Gráfico 6.4



El test HEGY consiste en contrastar las siguientes hipótesis:

$$H_0: \varphi(1) = 0 \text{ versus } H_1: \varphi(1) > 0$$

$$H_0: \varphi(-1) = 0 \text{ versus } H_1: \varphi(-1) > 0$$

$$H_0: \varphi(\pm i) = 0 \text{ versus } H_1: \varphi(\pm i) > 0$$

Si en el contraste no se rechaza la hipótesis nula de que $\varphi(1) = 0$, entonces se concluye que existe una raíz unitaria en la componente regular, es decir una raíz unitaria no estacional. Si la hipótesis que no puede ser rechazada es $\varphi(-1) = 0$, significa que la serie tiene una raíz unitaria estacional en la frecuencia semestral. Por último, La existencia de raíz unitaria estacional en la frecuencia trimestral, viene condicionada a la prueba de significación conjunta de la hipótesis $\varphi(\pm i) = 0$. En caso de no rechazarse la hipótesis nula, se concluye la existencia de raíz unitaria estacional en la frecuencia trimestral. Los autores realizan el contraste empleando una formulación basada en variables auxiliares y los tests se llevan a cabo sobre la significatividad de los coeficientes asociados a esas variables. Se remite al lector al trabajo original, *Seasonal Integration and Cointegration* de Hylleberg Engle, Granger y Yoo para interiorizarse sobre los detalles.

Las dos primeras hipótesis se contrastan a través de un estadístico t mientras que la hipótesis conjunta se prueba a partir de un estadístico F . Los autores tabulan los valores críticos para las pruebas t y F que permiten realizar los contrastes.

Franses (1991) y Beaulieu y Miron (1993) extienden el test HEGY para series con período estacional $s = 12$. En este caso, $\varphi(z) = 1 - z^{12}$ y puede escribirse de forma factorizada:

$$1 - z^{12} = (1 - z)(1 + z)(1 + z^2)(1 + z + z^2)(1 - z + z^2)(1 + \sqrt{3}z + z^2)(1 - \sqrt{3}z + z^2)$$

De acuerdo con (6.1) las raíces unitarias estacionales son:

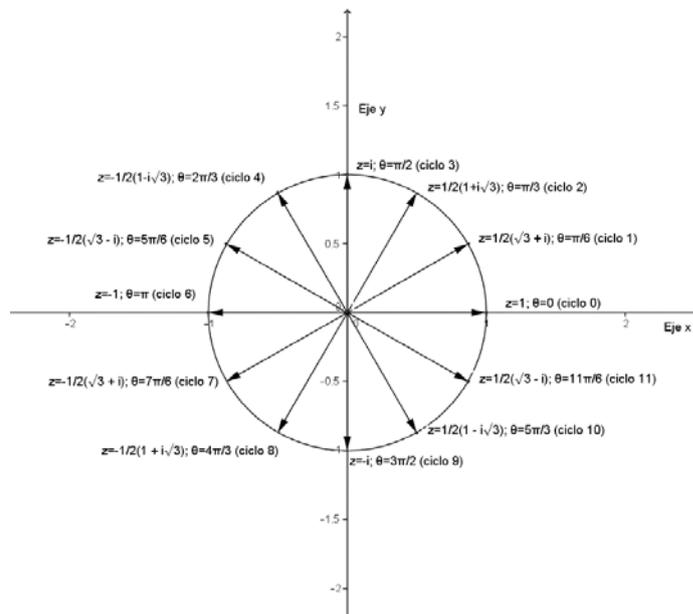
$$-1; \pm i; -\frac{1}{2}(1 \pm i\sqrt{3}); \frac{1}{2}(1 + i\sqrt{3}); -\frac{1}{2}(\sqrt{3} \pm i); \frac{1}{2}(\sqrt{3} \pm i)$$

Estas raíces están asociadas, respectivamente, a 6, 3, 9, 8, 4, 2, 10, 7, 5, 1 y 11 ciclos por año. Las frecuencias en radianes para cada una de ellas son:

$$\left\{ \pi; \pm \frac{\pi}{2}; \pm \frac{2}{3}\pi; \pm \frac{\pi}{3}; \pm \frac{5}{6}\pi; \pm \frac{\pi}{6} \right\}$$

En el siguiente gráfico se representan las raíces complejas asociadas a su frecuencia angular²⁴:

Gráfico 6.5



Las hipótesis a contrastar son:

$$H_0: \varphi(z_k) = 0 \text{ versus } H_1: \varphi(z_k) > 0 \text{ con } k = 0, 6.$$

$$H_0: \varphi(z_k) = 0 \text{ versus } H_1: \varphi(z_k) > 0 \text{ con } k = 1, 2, 3, 4, 5, 7, 8, 9, 10, 11.$$

Al igual que para el caso trimestral, si $\varphi(z_k) = 0$ entonces las raíces asociadas serán unitarias.

Franses y Hobijn (1997), utilizando el método de Monte Carlo, confeccionaron sendas tablas de valores críticos para realizar los tests de hipótesis. Este contraste también se plantea en términos de los coeficientes de variables auxiliares, extendiendo lo realizado por Hylleberg Engle, Granger y Yoo, y el lector interesado puede consultar el documento original de Beaulieu y Miron, *Seasonal Unit Roots in Aggregate U.S. Data*. Los test de significación cuando $k = 0$ y $k = 6$ se realizan mediante una prueba t , mientras que los contrastes de significación conjunta se realizan mediante una prueba F .

²⁴ En el gráfico se tomó $0 \leq \theta \leq 2\pi$. Nótese que la raíz no estacional $z = 1$ está asociada a $\theta = 0$ y le corresponde el ciclo 0.

A través de gretl se puede acceder a una herramienta de distribución libre que permite aplicar el contraste HEGY para testear la presencia de raíces unitarias estacionales, que ofrece los estadísticos t y F observados para cada frecuencia estacional y el *valor-p*²⁵ asociado.

Ejemplo 6.3.2.1: Con propósito ilustrativo, en el siguiente cuadro se presenta la salida de gretl para la serie de registros de lluvia correspondiente a San Jaime, donde se incluye una constante y el orden de retardos del polinomio autorregresivo se determina de modo automático:

Deterministic component: constant + (s-1) trigonometric terms

Dof (T-k) = 499

Statistic	p-value	Ang. Frequency	Period
t1= -5,45	0,00000 ***	zero	infinity
F1= 31,88	0,00000 ***	$+-\pi/6$	12
F2= 29,85	0,00000 ***	$+-\pi/3$	6
F3= 36,01	0,00000 ***	$+-\pi/2$	4
F4= 31,56	0,00000 ***	$+2*\pi/3$	3
F5= 42,95	0,00000 ***	$+5*\pi/6$	2,40
t2= -6,18	0,00000 ***	π	2
Fs= 39,83	0,00000	All the seasonal cycles	
Ft= 37,97	0,00000	Delta_s (all the seas. + zero freq.)	

En la salida de gretl puede verse en la primera columna los valores de los estadísticos t para π_1 y π_2 , que son $t_1 = -5,45$ y $t_2 = -6,18$ (primera fila y última respectivamente). En las filas intermedias de la primera columna, en orden creciente, los estadísticos F que permiten testear las pruebas conjuntas. Además, se muestran el valor-p asociado, la frecuencia y el período.

Ejemplo 6.3.2.2: A continuación se muestra la salida de gretl del test HEGY para el registro procedente de San Víctor:

Statistic	p-value	Ang. Frequency	Period
t1= -3,94	0,00091 ***	zero	infinity
F1= 4,99	0,00871 ***	$+-\pi/6$	12
F2= 4,62	0,01232 **	$+-\pi/3$	6
F3= 6,17	0,00283 ***	$+-\pi/2$	4
F4= 8,34	0,00036 ***	$+2*\pi/3$	3
F5= 16,22	0,00000 ***	$+5*\pi/6$	2,40
t2= -4,09	0,00000 ***	π	2
Fs= 9,45	0,00000	All the seasonal cycles	
Ft= 10,31	0,00000	Delta_s (all the seas. + zero freq.)	

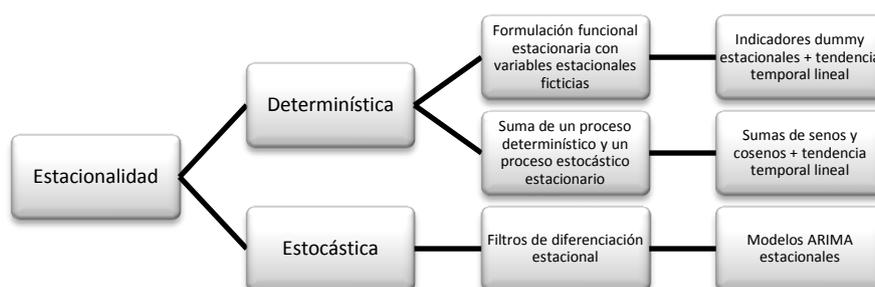
Puede observarse que en los casos de los estadísticos F1 y F2 el valor p señalado es aproximadamente 0,01. Es decir, que no existe evidencia suficiente para rechazar H_0 a un nivel de significancia del 1%. Aunque esta probabilidad es muy poco significativa, sugiere la posibilidad de que exista raíz unitaria en la frecuencia estacional indicada.

²⁵ El *valor-p* es la probabilidad de obtener un valor del estadístico de contraste tan extremo como el observado considerando cierta la hipótesis nula. Así, por ejemplo, si el *valor-p* es inferior a 0,05, se concluye que no existe evidencia suficiente para aceptar H_0 a un nivel de significancia del 5%.

6.3.3. Test de Canova y Hansen (CH)

En el test HEGY, rechazar la hipótesis nula implica que la serie no presenta raíces unitarias estacionales, sin embargo, es un contraste con baja potencia²⁶ cuando se aplica a muestras de reducido tamaño. Es decir, que el no rechazo de la hipótesis nula no puede ser interpretado como evidencia a favor de la presencia de raíces unitarias estacionales. Por lo que Canova y Hansen proponen una alternativa a esta prueba.

Para ello tienen en cuenta las diferentes formas de estacionalidad que puede presentar una serie temporal y las expresiones a partir de las cuales ésta puede modelarse. Una manera, la más sencilla, es emplear una formulación funcional determinística con variables estacionales ficticias, y otra, es modelar la estacionalidad mediante la suma de un proceso determinístico y un proceso estocástico estacionario. Por otro lado, cuando la estacionalidad de una serie cambia con el tiempo, entonces es estocástica y una forma de modelarla es emplear las técnicas discutidas en el capítulo 5 en relación a los modelos ARIMA estacionales. Desde ya que constituye una dificultad importante conocer a priori qué tipo de estacionalidad presentan los datos que se estudian. El siguiente diagrama arroja luz sobre estas diferencias:



Canova y Hansen (1995) proponen el siguiente contraste:

$$H_0: \text{La estacionalidad es estacionaria}$$

$$H_1: \text{La estacionalidad es no estacionaria}$$

El rechazo de la hipótesis nula implica que la serie tiene estacionalidad estocástica. En su trabajo original, *Are Seasonal Patterns Constant Over Time? A Test for Seasonal Stability*, Canova y Hansen comienzan con un modelo para la serie y_t con estacionalidad estacionaria:

$$y_t = g(x_t) + S_t, \quad t = 1, 2, \dots, n \quad (6.2)$$

²⁶ La potencia de una prueba se refiere a la probabilidad de rechazar la hipótesis nula cuando la alternativa es cierta. Una potencia baja implica que el contraste no es capaz de detectar lo establecido en la hipótesis alternativa cuando esta es cierta. Complementariamente, puede decirse que es alta la probabilidad de no rechazar la hipótesis nula cuando es falsa.

En esta ecuación $g(x_t)$ es una regresión auxiliar²⁷ y S_t es una componente estacional determinística de período s . Con el propósito de distinguir la no estacionariedad en la frecuencia cero y en las frecuencias estacionales, Canova y Hansen requieren que y_t no tenga raíz unitaria en la frecuencia cero. Por lo que, en la práctica, si existe la hipótesis de que la serie tratada tenga una raíz unitaria en la estructura regular, una opción aceptable es aplicar el test sobre la serie diferenciada regularmente.

La especificación propuesta para la estacionalidad S_t consiste en una representación trigonométrica dada por la siguiente formulación²⁸:

$$S_t = \sum_{j=1}^q f'_{jt} \gamma_j, \quad (6.3)$$

donde $q = \frac{s}{2}$ y para $j < q$, $f'_{jt} = \left(\cos\left(\frac{j}{q}\pi t\right), \sin\left(\frac{j}{q}\pi t\right) \right)$, mientras que para $j = q$, $f'_{qt} = \cos(\pi t) = (-1)^t$ pues $\sin(\pi t) = 0$ cualquiera sea el entero t .

Las series de registros de lluvia se tomaron mensualmente por lo que $s = 12$ y entonces $q = 6$. Por lo tanto, $f'_{jt} = (f'_{1t}, f'_{2t}, f'_{3t}, f'_{4t}, f'_{5t}, f'_{6t})_{1 \times 11}$ donde $f'_{1t} = \left(\cos\left(\frac{\pi}{6}t\right); \sin\left(\frac{\pi}{6}t\right) \right)$, $f'_{2t} = \left(\cos\left(\frac{\pi}{3}t\right); \sin\left(\frac{\pi}{3}t\right) \right)$, $f'_{3t} = \left(\cos\left(\frac{\pi}{2}t\right); \sin\left(\frac{\pi}{2}t\right) \right)$, $f'_{4t} = \left(\cos\left(\frac{2\pi}{3}t\right); \sin\left(\frac{2\pi}{3}t\right) \right)$, $f'_{5t} = \left(\cos\left(\frac{5\pi}{6}t\right); \sin\left(\frac{5\pi}{6}t\right) \right)$, $f'_{6t} = (\cos(\pi t); 0)$. En tal caso γ_j será un vector de 11 filas por 1 columna.

De acuerdo con (6.3):

$$S_t = \left(\cos\frac{\pi}{6}t; \sin\frac{\pi}{6}t \right) \gamma_1 + \left(\cos\frac{\pi}{3}t; \sin\frac{\pi}{3}t \right) \gamma_2 + \left(\cos\frac{\pi}{2}t; \sin\frac{\pi}{2}t \right) \gamma_3 + \left(\cos\frac{2\pi}{3}t; \sin\frac{2\pi}{3}t \right) \gamma_4 \\ + \left(\cos\frac{5\pi}{6}t; \sin\frac{5\pi}{6}t \right) \gamma_5 + \cos(\pi t) \gamma_6 \text{ con } t = 1, 2, 3, \dots, n$$

Los componentes de γ_j se ven como el peso o contribución que cada vector de f'_{jt} tiene sobre el término estacional S_t . Además, las propiedades periódicas de las funciones trigonométricas permiten ver que $\sum_{t=1}^n f_t = 0$ por lo que f_t es un proceso estacionario de media cero. Por lo tanto, que S_t sea estacionaria o estocástica depende del comportamiento de γ_j , no de las f'_{jt} .

En virtud de ello, Canova y Hansen proponen que si γ_j varía como un paseo aleatorio, S_t es estocástica. Escribiendo los componentes de (6.3) en forma vectorial:

$$\gamma = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_q \end{pmatrix}, \quad f_t = \begin{pmatrix} f_{1t} \\ \vdots \\ f_{qt} \end{pmatrix} \quad (6.4)$$

²⁷En la formulación original del trabajo $g(x_t) = \mu + x'_t \beta + e_t$ donde μ es una constante, x_t es un vector de variables explicativas de dimensión $k \times 1$, y e_t es un término de error no correlacionado con x_t y S_t , con media 0 y varianza σ_e^2 .

²⁸ Canova y Hansen proponen, además, una formulación basada en variables ficticias (indicadores dummy) estacionales y muestran que ambas formulaciones son equivalentes. En esta exposición sólo se toma la formulación trigonométrica.

donde γ y f_t tienen ambos $s - 1$ elementos, la igualdad (6.3) queda $S_t = f_t' \gamma$. Sustituyendo en (6.2) resulta:

$$y_t = g(x_t) + f_t' \gamma, \quad t = 1, 2, \dots, n \quad (6.5)$$

De acuerdo con la definición de paseo aleatorio (ver ejemplo 5.1.5), se propone que los coeficientes γ_j varíen de acuerdo con la siguiente formulación:

$$\gamma_t = \gamma_{t-1} + u_t, \quad (6.6)$$

donde $u_t \sim iid$. Cuando la matriz de covarianza de u_t sea idénticamente nula, el modelo (6.5) se reduce a un modelo con estacionalidad estacionaria, pues γ_t deja de variar de acuerdo con un paseo aleatorio (esto permite concluir que no existe evidencia suficiente para rechazar la hipótesis nula). En caso contrario, la estacionalidad es estocástica y se rechaza H_0 . Si se denota $C(u_t)$ a la matriz de covarianzas²⁹ de u_t , entonces el test de hipótesis puede plantearse a groso modo de la siguiente manera:

$$H_0: C(u_t) = 0 \quad vs. \quad H_1: C(u_t) \neq 0$$

Para realizar el contraste los autores proponen un estadístico L_f que permite testear simultáneamente en todas las frecuencias la existencia de raíces unitarias y un conjunto de estadísticos denotados como $L_{\left(\frac{j\pi}{q}\right)}$ para cada frecuencia individual. Canova y Hansen sugieren que los test individuales son complementarios del test conjunto. Si en éste se rechaza la hipótesis nula es porque podrían existir raíces unitarias en cualquiera de las frecuencias estacionales. Por su parte, los test individuales permiten detectar la frecuencia precisa donde surge la no estacionariedad.

“The $L_{\left(\frac{j\pi}{q}\right)}$ tests are useful complements to the joint test L_f . If the joint test rejects, it could be due to unit roots at any of the seasonal frequencies. The $L_{\left(\frac{j\pi}{q}\right)}$ tests are specifically designed to detect at which particular seasonal frequency non stationarity emerges.” (Canova & Hansen, 1995)

Las características específicas del contraste pueden consultarse en el trabajo original de Canova y Hansen citado en este documento³⁰.

²⁹ Recuérdese del Capítulo 3 que los elementos de la diagonal de la matriz de covarianzas son las varianzas de las variables aleatorias.

³⁰ El test CH es una generalización para series estacionales del test de Kwiatkowski, Phillips, Schmidt y Shin, conocido como KPSS. En éste se asume que la serie temporal puede modelarse como la suma de una tendencia determinística, un camino aleatorio y un error estacionario: $y_t = \xi t + r_t + \varepsilon_t$ donde $r_t = r_{t-1} + u_t$ con $u_t \sim iid(0, \sigma_u^2)$. De esta manera, la hipótesis de estacionariedad es sencillamente $\sigma_u^2 = 0$. (Kwiatkowski, Phillips, Schmidt, & Shin, 1992)

Ejemplo 6.3.3.1: A continuación se muestra la salida de gretl para la serie procedente de San Jaime:

Regressors in the auxiliary regression: Trigonometric terms

Degrees of freedom (T-k) = 522, lag order = 7

Statistic	p-value	Ang. Frequency	Period
L_1 = 0,5396	0,14140	$\pm\pi/6$	12
L_2 = 0,8152	0,03347 **	$\pm\pi/3$	6
L_3 = 0,1601	0,84189	$\pm\pi/2$	4
L_4 = 0,9612	0,01513 **	$\pm 2\pi/3$	3
L_5 = 0,1531	0,86135	$\pm 5\pi/6$	2,40
L_pi = 0,3139	0,12657	π	2

L_f = 2,3717	0,08439 *	Joint test	

Obsérvese que se rechaza la hipótesis nula en la frecuencia $\pm \frac{\pi}{3}$ y $\pm \frac{2\pi}{3}$. El valor-p asociado al estadístico de contraste $L_2 = 0,8152$ es $0,03347 < 0,05$ y el asociado a $L_4 = 0,9612$ es $0,01513 < 0,05$, por lo que la serie presentaría, según el test CH, estacionalidad estocástica no estacionaria. Este resultado, no concuerda con los resultados del test HEGY del Ejemplo 6.3.2.1.

Ejemplo 6.3.3.2: En el caso de la serie procedente de San Víctor, los resultados del test CH parecen arribar a una conclusión similar a la obtenida con el test HEGY:

Regressors in the auxiliary regression: Trigonometric terms

Degrees of freedom (T-k) = 213, lag order = 5

Statistic	p-value	Ang. Frequency	Period
L_1 = 0,1027	0,97659	$\pm\pi/6$	12
L_2 = 0,2664	0,56102	$\pm\pi/3$	6
L_3 = 0,7655	0,04140 **	$\pm\pi/2$	4
L_4 = 0,2784	0,53123	$\pm 2\pi/3$	3
L_5 = 0,3436	0,38915	$\pm 5\pi/6$	2,40
L_pi = 0,0969	0,63650	π	2

L_f = 2,0134	0,17774	Joint test	

La siguiente tabla muestra los resultados obtenidos con Gretl para cada serie con ambos contrastes. Para el test HEGY las pruebas se realizaron por separado con constante y con constante más términos trigonométricos, ya que se desconoce el verdadero proceso generador de los datos. No se tiene en cuenta la componente de tendencia ya que en los gráficos de las series no se evidencian marcados crecimientos y/o decrecimientos. Además, los resultados de la tabla están basados en el criterio de Schwarz para las estimaciones del orden autorregresivo, aunque también se realizaron las estimaciones mediante los criterios AIC y HQC (consultar la sección 7.5). El test HEGY y el CH en muchos casos no son coincidentes en sus resultados, por lo que podrá considerarse evidencia fuerte de la presencia de raíces unitarias cuando ambos contrastes permitan arribar a conclusiones similares.

Tabla 6.2

Estación de Registro	Test HEGY		Test de Canova-Hansen (CH)
	Constante	Constante + términos trigonométricos	
Los Charrúas	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
Lucas Sur	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
Octavo Distrito	La serie no presenta raíces unitarias estacionales ³¹	La serie no presenta raíces unitarias estacionales	Valor - p < 0,10 en las frecuencias $\pm \frac{\pi}{3}$ y $\pm \frac{\pi}{2}$
Puente de Hierro	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
San Víctor	Valor - p > 0,01 en la frecuencia $\pm \frac{\pi}{3}$	La serie no presenta raíces unitarias estacionales ³³	Valor - p < 0,05 en la frecuencia $\pm \frac{\pi}{2}$
Santa Anita	La serie no presenta raíces unitarias estacionales ³⁴	La serie no presenta raíces unitarias estacionales ³⁵	Valor - p < 0,10 en la frecuencia $\pm \frac{\pi}{2}$
Séptimo Distrito	Valor - p > 0,01 en la frecuencia $\pm \frac{\pi}{6}$	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
Viale	Valor - p > 0,01 en la frecuencia cero ³⁷ .	La serie no presenta raíces unitarias estacionales ³⁸	Valor - p < 0,05 en las frecuencias $\pm \frac{\pi}{3}$ y $\pm \frac{\pi}{2}$
Villa Paranacito	Valor - p > 0,01 en la frecuencia cero ³⁹ .	Valor - p > 0,01 en la frecuencia cero ⁴⁰ .	La serie no presenta raíces unitarias estacionales
Colón	La serie no presenta raíces unitarias estacionales ⁴¹	La serie no presenta raíces unitarias estacionales ⁴²	La serie no presenta raíces unitarias estacionales
La Lila	La serie no presenta raíces unitarias estacionales ⁴³	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
Lucas González	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
Santa María del Tatutí	La serie no presenta raíces unitarias estacionales ⁴⁴	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
Villa Elisa	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
Antelo	La serie no presenta raíces	La serie no presenta raíces	La serie no presenta raíces

³¹ Con constante y criterio AIC el valor-p>0,01 en la frecuencia $\pm \frac{\pi}{6}$.

³² Con constante y criterio AIC el valor-p>0,05 en las frecuencias $\pm \frac{\pi}{6}$ y $\pm \frac{\pi}{3}$ y el valor-p>0,01 en las frecuencias $\pm \frac{\pi}{2}$ y $\pm \pi$. Con constante y criterio HQC, el valor-p>0,01 en la frecuencia $\pm \frac{\pi}{3}$.

³³ Con constante y términos trigonométricos y criterio AIC, el valor-p>0,01 en las frecuencias $\pm \frac{\pi}{2}$ y $\pm \pi$. Con criterio HQC, valor-p>0,01 en la frecuencia $\pm \frac{\pi}{2}$.

³⁴ Con constante y criterio AIC el valor-p>0,05 en las frecuencias 0, $\pm \frac{\pi}{6}$ y $\pm \frac{\pi}{3}$.

³⁵ Con criterio AIC, el valor-p>0,05 en las frecuencias $\pm \frac{\pi}{2}$. El valor-p>0,01 en las frecuencias 0 y $\pm \frac{\pi}{3}$.

³⁶ Con criterio AIC se confirma el resultado. Con criterio HQC se obtiene igual resultado.

³⁷ Con criterio AIC, el valor-p>0,05. Con criterio HQC, el valor-p>0,01.

³⁸ Con criterio AIC, el valor-p>0,01 en las frecuencias $\pm \frac{\pi}{2}$ y π .

³⁹ Igual resultado se obtiene con los criterios AIC y HQC.

⁴⁰ Se obtiene igual resultado con criterio HQC.

⁴¹ Con criterio AIC el valor-p>0,05 en la frecuencia $\pm \frac{\pi}{6}$ y valor-p>0,01 en la frecuencia $\pm \frac{\pi}{3}$.

⁴² Con criterio AIC valor-p>0,01 en la frecuencia $\pm \frac{\pi}{3}$.

⁴³ Con criterio AIC, valor-p>0,01 en las frecuencias 0 y $\pm \frac{\pi}{6}$.

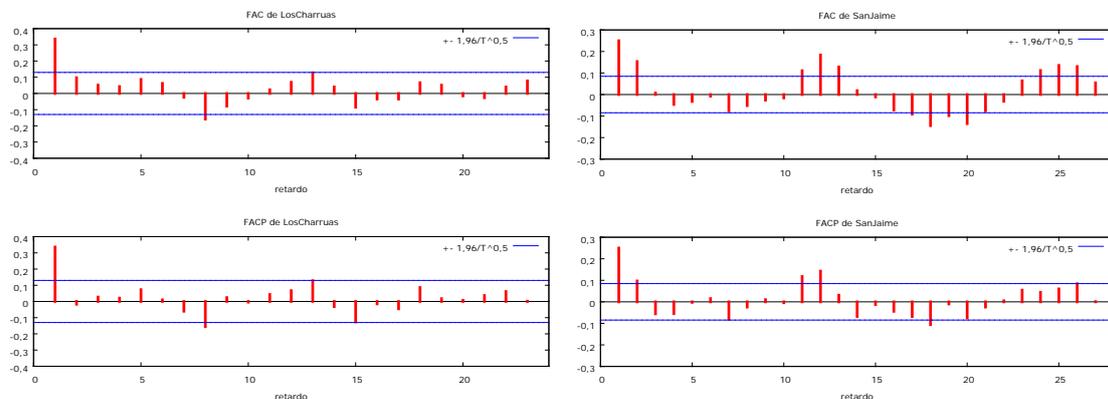
⁴⁴ Con criterio AIC, valor-p>0,05 en la frecuencia $\pm \frac{\pi}{6}$ y el valor-p>0,01 en la frecuencia $\pm \frac{\pi}{3}$.

	unitarias estacionales ⁴⁵	unitarias estacionales	unitarias estacionales
Febre	La serie no presenta raíces unitarias estacionales ⁴⁶	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
Feliciano	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales	Valor $-p < 0,05$ en las frecuencias $\pm \frac{\pi}{6}, \pm \frac{\pi}{3}$ y $\frac{2\pi}{3}$
San Jaime	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales	Valor $-p < 0,05$ en las frecuencias $\pm \frac{\pi}{3}$ y $\pm \frac{2\pi}{3}$
San Salvador	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
La Paz	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
San Gustavo	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales
Paraná	La serie no presenta raíces unitarias estacionales ⁴⁷	La serie no presenta raíces unitarias estacionales	La serie no presenta raíces unitarias estacionales

6.4. Identificación del tipo de modelo y órdenes asociados p, q, P y Q

La identificación de los órdenes del modelo ARIMA que explica el comportamiento de una serie de tiempo, se realiza a partir del estudio de las funciones de autocorrelación y de autocorrelación parcial. En general, es un trabajo de observación muy minucioso y en ningún modo concluyente, es decir, un correlograma puede sugerir varios modelos candidatos a recoger la estructura del proceso generador de los datos.

Gráfico 6.6



Si se observa, a modo de ejemplo ilustrativo, la *ACF* y la *PACF* de la serie procedente del registro de Los Charrúas (figura izquierda del Gráfico 6.6), se ve las características de un proceso autorregresivo de orden 1, es decir, un *AR(1)*. En efecto, nótese el decrecimiento exponencial en los primeros retardos de la *ACF*, mientras que la *PACF* muestra el comportamiento truncado después del primer retardo, propio de los procesos *AR(1)*. El resto de los retardos no parecen ser significativamente diferentes de cero.

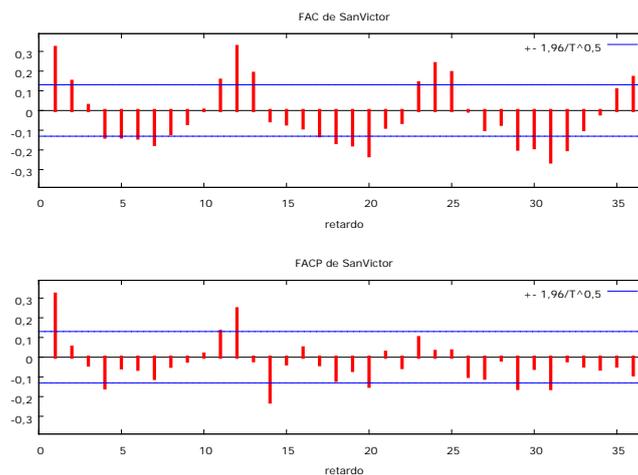
⁴⁵ Con criterio AIC, valor- $p > 0,01$ en la frecuencia $\pm \frac{\pi}{6}$.

⁴⁶ Con criterio AIC el valor- $p > 0,05$ en la frecuencia $\pm \frac{\pi}{6}$.

⁴⁷ Con criterio AIC, valor- $p > 0,05$ en la frecuencia $\pm \frac{\pi}{6}$ y superior a 0,01 en la frecuencia cero.

Siguiendo esta línea de observación si se investiga los órdenes del modelo para la serie de registros de lluvia procedente de San Jaime (figura derecha del Gráfico 6.6), se ve que en la *ACF* de la serie, es clara la existencia de dos retardos significativamente distintos de cero en la parte regular y luego se trunca. En la estructura estacional el retardo $\hat{\rho}_{12}$ es también significativamente distinto de cero y luego hay un coeficiente diferente de cero a ambos lados de cada retardo estacional, $\hat{\rho}_{11} \approx \hat{\rho}_{13}$. Estas características reflejan dos posibilidades: los órdenes de media móvil son $q = 2$ y $Q = 1$ o bien $q = 1$ y $Q = 1$ (recuérdese del capítulo 6 que cuando $q = 2$ hay dos retardos significativos a ambos lados del retardo estacional y no es este caso, sin embargo la parte regular tiene dos retardos distintos de cero). Los retardos estacionales de la *ACF* parecen decrecer lentamente, lo que puede significar un orden de diferenciación estacional igual a 1, sustentado por los resultados arrojados por el test de Canova y Hansen (no así por el contraste HEGY). La *PACF* tiene decrecimiento rápido en la parte donde se recoge la estructura regular, el primer retardo distinto de cero y el segundo apenas significativo, que pueden sugerir $p = 2$. El retardo $s = 12$ también es diferente de cero. En base a estas consideraciones se proponen los siguientes modelos: $ARIMA(0,0,2) \times (1,0,1)_{12}$, o $ARIMA(2,0,1) \times (1,0,1)_{12}$ o $ARIMA(2,0,1) \times (1,1,1)_{12}$.

Gráfico 6.7



Para el caso del registro procedente de San Víctor, si se compara el correlograma de la *ACF* (Gráfico 6.7) con las características de los correlogramas de los modelos *ARIMA* consignados al final del capítulo 6 y con los resultados de los test de raíces unitarias, pueden sugerirse los siguientes modelos:

Modelo 1: $ARIMA(1,0,0) \times (1,1,1)_{12}$

Modelo 2: $ARIMA(1,0,1) \times (1,1,1)_{12}$

La elección del modelo que mejor recoja la estructura de la serie deberá decidirse a través de procesos de validación y diagnóstico que se tratarán en el siguiente capítulo.

La tabla siguiente muestra los modelos que en base a las características de la *ACF*, de la *PACF* y de los resultados del test HEGY y del contraste de Canova y Hansen, se propusieron. En el Capítulo 9, en los anexos, puede consultarse los correlogramas de cada serie.

Tabla 6.3

Estación de Registro	Modelo 1	Modelo 2	Modelo 3	Modelo 4
<i>Los Charrúas</i>	$ARIMA(1,0,0)x(0,0,0)_{12}$	-	-	-
<i>Lucas Sur</i>	$ARIMA(0,0,2)x(1,0,0)_{12}$	-	-	-
<i>Octavo Distrito</i>	$ARIMA(0,0,2)x(0,1,1)_{12}$	$ARIMA(0,0,2)x(1,1,1)_{12}$	$ARIMA(0,0,2)x(1,0,1)_{12}$	$ARIMA(2,0,0)x(1,1,1)_{12}$
<i>Puente de Hierro</i>	$ARIMA(0,0,5)x(1,0,1)_{12}$	$ARIMA(0,0,2)x(1,0,1)_{12}$	-	-
<i>San Víctor</i>	$ARIMA(1,0,0)x(1,1,1)_{12}$	$ARIMA(1,0,1)x(1,1,1)_{12}$	-	-
<i>Santa Anita</i>	$ARIMA(0,1,1)x(0,1,1)_{12}$	$ARIMA(1,1,1)x(1,1,1)_{12}$	$ARIMA(1,0,0)x(0,1,1)_{12}$	$ARIMA(1,0,1)x(1,1,1)_{12}$
<i>Séptimo Distrito</i>	$ARIMA(1,0,2)x(1,0,1)_{12}$	$ARIMA(1,0,2)x(1,1,1)_{12}$	-	-
<i>Viale</i>	$ARIMA(1,0,0)x(0,1,1)_{12}$	$ARIMA(1,0,0)x(1,0,1)_{12}$	-	-
<i>Villa Paranacito</i>	$ARIMA(0,0,1)x(1,0,0)_{12}$	$ARIMA(0,1,1)x(1,0,0)_{12}$	$ARIMA(1,0,0)x(1,0,0)_{12}$	-
<i>Colón</i>	$ARIMA(1,0,0)x(0,1,1)_{12}$	$ARIMA(1,0,0)x(1,0,1)_{12}$	$ARIMA(1,0,0)x(1,1,1)_{12}$	-
<i>La Lila</i>	$ARIMA(0,0,2)x(0,1,1)_{12}$	$ARIMA(1,0,2)x(1,0,1)_{12}$	$ARIMA(1,0,2)x(1,1,1)_{12}$	-
<i>Lucas González</i>	$ARIMA(0,0,0)x(0,1,1)_{12}$	$ARIMA(0,0,1)x(0,0,2)_{12}$	$ARIMA(0,0,8)x(1,0,1)_{12}$	-
<i>Santa María del Tatutí</i>	$ARIMA(1,0,0)x(1,0,1)_{12}$	$ARIMA(1,0,0)x(0,1,1)_{12}$	-	-
<i>Villa Elisa</i>	$ARIMA(1,0,0)x(0,0,0)_{12}$	-	-	-
<i>Antelo</i>	$ARIMA(1,0,0)x(1,0,1)_{12}$	$ARIMA(1,0,1)x(1,0,1)_{12}$	-	-
<i>Febre</i>	$ARIMA(1,0,1)x(1,0,1)_{12}$	$ARIMA(1,0,1)x(0,1,1)_{12}$	-	-
<i>Feliciano</i>	$ARIMA(0,0,2)x(1,1,1)_{12}$	$ARIMA(1,0,2)x(1,1,1)_{12}$	-	-
<i>San Jaime</i>	$ARIMA(0,0,2)x(0,1,1)_{12}$	$ARIMA(0,0,1)x(0,1,1)_{12}$	-	-
<i>San Salvador</i>	$ARIMA(0,0,1)x(1,0,1)_{12}$	$ARIMA(1,0,0)x(2,0,0)_{12}$	$ARIMA(1,0,1)x(2,0,0)_{12}$	-
<i>La Paz</i>	$ARIMA(1,0,2)x(1,0,1)_{12}$	$ARIMA(0,0,2)x(1,0,1)_{12}$	$ARIMA(0,0,5)x(1,0,1)_{12}$	$ARIMA(0,0,5)x(2,0,0)_{12}$
<i>San Gustavo</i>	$ARIMA(1,0,1)x(2,0,0)_{12}$	$ARIMA(1,0,10)x(1,0,1)_{12}$	-	-
<i>Paraná</i>	$ARIMA(2,0,2)x(1,0,1)_{12}$	$ARIMA(0,0,6)x(1,0,1)_{12}$	-	-

Los modelos propuestos para cada serie son tentativos en el sentido de que son posibles candidatos a recoger la estructura del proceso que genera las series. Las características presentes en los correlogramas (fundamentalmente de la ACF), no concuerdan en todos los casos con los resultados obtenidos mediante los contrastes de raíces unitarias (HEGY y CH). Más aún, hay series para las que tampoco es unánime la conclusión de los test. Por lo tanto, los modelos se sugirieron teniendo en cuenta todas las alternativas. La decisión sobre cuál de los modelos propuestos es el que mejor se ajusta al proceso, está sujeta a los procedimientos de estimación y validación que se tratan en el siguiente capítulo.

Capítulo 7

Estimación y Validación

Una vez propuestos los modelos tentativos que explican el comportamiento de los procesos generadores de las series, se procede a estimar los parámetros del modelo. Concretamente y para simplificar el trabajo, dado un modelo $ARMA(p, q)$, se discutirá la estimación de los parámetros $\varphi_1, \varphi_2, \dots, \varphi_p$, también de la media $\mu = E[y_t]$, de $\theta_1, \theta_2, \dots, \theta_q$ y $\sigma_A^2 = E[A_t^2]$ en el modelo

$$y_t = \alpha + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + A_t + \theta_1 A_{t-1} + \theta_2 A_{t-2} + \dots + \theta_q A_{t-q}$$

donde $\alpha = \mu(1 - \varphi_1 - \dots - \varphi_p)$.

Como primer paso se tratará el método de los momentos y después el método de máxima verosimilitud. Seguidamente, es preciso validar el modelo obtenido. En principio, esta etapa se efectúa mediante contrastes de significación sobre los parámetros estimados; si existen parámetros que no pueden considerarse significativamente distintos de cero, entonces se eliminan del modelo. En segundo lugar, es necesario verificar el cumplimiento de las condiciones de causalidad e invertibilidad discutidas en el capítulo 4 y tener en cuenta la matriz de correlaciones entre los parámetros para estudiar la posible colinealidad entre los mismos.

7.1. Estimación por Método de Momentos

El método de momentos consiste en resolver un sistema de ecuaciones que resulta de sustituir los momentos muestrales como la media $\hat{\mu}_y$, la varianza muestral $\hat{\gamma}_0$ y los coeficientes $\hat{\rho}_i$ de la ACF por sus contrapartes teóricas. A los efectos de ilustrar el procedimiento, se considera un proceso $AR(p)$ dado por la siguiente ecuación:

$$y_t = \alpha + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + A_t \quad (7.1)$$

donde $\alpha = \mu(1 - \varphi_1 - \dots - \varphi_p)$. La media $\mu = E(y_t)$ es estimada por la media muestral $\hat{\mu}_y$ que es la media de la serie observada. Para estimar los parámetros $\varphi_1, \varphi_2, \dots, \varphi_p$, se tiene en cuenta, la igualdad $\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + \dots + \varphi_p \rho_{k-p}$, que para $k \geq 1$ permite obtener el siguiente sistema conocido como sistema de ecuaciones de Yule-Walker:

$$\begin{aligned} \rho_1 &= \varphi_1 + \varphi_2 \rho_1 + \varphi_3 \rho_2 + \dots + \varphi_p \rho_{p-1} \\ \rho_2 &= \varphi_1 \rho_1 + \varphi_2 + \varphi_3 \rho_1 + \dots + \varphi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \varphi_1 \rho_{p-1} + \varphi_2 \rho_{p-2} + \varphi_3 \rho_{p-3} + \dots + \varphi_p \end{aligned}$$

Realizando la sustitución de los ρ_i por los $\hat{\rho}_i$ y resolviendo el sistema se obtiene los estimadores $\hat{\varphi}_1, \hat{\varphi}_2, \dots, \hat{\varphi}_p$. Concretamente:

$$\begin{pmatrix} \hat{\varphi}_1 \\ \hat{\varphi}_2 \\ \vdots \\ \hat{\varphi}_p \end{pmatrix} = \begin{pmatrix} 1 & \hat{\rho}_1 & \hat{\rho}_2 & \cdots & \hat{\rho}_{p-2} & \hat{\rho}_{p-1} \\ \hat{\rho}_1 & 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_{p-3} & \hat{\rho}_{p-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \hat{\rho}_{p-1} & \hat{\rho}_{p-2} & \hat{\rho}_{p-3} & \cdots & \hat{\rho}_1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\rho}_1 \\ \hat{\rho}_2 \\ \vdots \\ \hat{\rho}_p \end{pmatrix}$$

Estos estimadores son habitualmente llamados estimadores de Yule-Walker.

Por otro lado, Desarrollando α y asociando términos, puede escribirse (7.1) como en la sección (4.1):

$$y_t - \mu = \varphi_1(y_{t-1} - \mu) + \varphi_2(y_{t-2} - \mu) + \cdots + \varphi_p(y_{t-p} - \mu) + A_t \quad (7.2)$$

Considerando $w_t = y_t - \mu$ resulta⁴⁸:

$$w_t = \varphi_1 w_{t-1} + \varphi_2 w_{t-2} + \cdots + \varphi_p w_{t-p} + A_t \quad (7.3)$$

Multiplicando ambos miembros por w_t queda:

$$w_t w_t = \varphi_1 w_t w_{t-1} + \varphi_2 w_t w_{t-2} + \cdots + \varphi_p w_t w_{t-p} + w_t A_t$$

Por lo tanto, teniendo en cuenta las propiedades de la esperanza matemática:

$$E[w_t w_t] = \varphi_1 E[w_t w_{t-1}] + \varphi_2 E[w_t w_{t-2}] + \cdots + \varphi_p E[w_t w_{t-p}] + E[w_t A_t]$$

De esta igualdad y por lo definido en el capítulo 3, se obtiene:

$$\gamma_0 = \varphi_1 \gamma_1 + \varphi_2 \gamma_2 + \cdots + \varphi_p \gamma_p + E[w_t A_t] \quad (7.4)$$

Además, $E[w_t A_t] = E[(\varphi_1 w_{t-1} + \varphi_2 w_{t-2} + \cdots + \varphi_p w_{t-p} + A_t) A_t]$.

Multiplicando adecuadamente y teniendo en cuenta otra vez las propiedades de la esperanza, resulta:

$$E[w_t A_t] = \varphi_1 E[w_{t-1} A_t] + \varphi_2 E[w_{t-2} A_t] + \cdots + \varphi_p E[w_{t-p} A_t] + E[A_t A_t]$$

Como se mostró en la sección 4.2, $E[w_{t-k} A_t] = 0$, mientras que $E[A_t A_t] = \sigma_A^2$. Sustituyendo en (7.4), se tiene:

$$\gamma_0 = \varphi_1 \gamma_1 + \varphi_2 \gamma_2 + \cdots + \varphi_p \gamma_p + \sigma_A^2 \quad (7.5)$$

De (7.5), se obtiene:

$$\sigma_A^2 = \gamma_0 - \varphi_1 \gamma_1 - \varphi_2 \gamma_2 - \cdots - \varphi_p \gamma_p$$

Factorizando:

$$\sigma_A^2 = \gamma_0 \left(1 - \varphi_1 \frac{\gamma_1}{\gamma_0} - \varphi_2 \frac{\gamma_2}{\gamma_0} - \cdots - \varphi_p \frac{\gamma_p}{\gamma_0} \right)$$

Recordando que $\rho_k = \frac{\gamma_k}{\gamma_0}$ resulta:

⁴⁸ Nótese que $E[w_t] = 0$.

$$\sigma_A^2 = \gamma_0(1 - \varphi_1\rho_1 - \varphi_2\rho_2 - \dots - \varphi_p\rho_p) \quad (7.6)$$

Teniendo en cuenta los estimadores de Yule-Walker, se obtiene el estimador para σ_A^2 :

$$\hat{\sigma}_A^2 = \hat{\gamma}_0(1 - \hat{\varphi}_1\hat{\rho}_1 - \hat{\varphi}_2\hat{\rho}_2 - \dots - \hat{\varphi}_p\hat{\rho}_p) \quad (7.7)$$

Ejemplo 7.1.1: En el capítulo anterior se propuso para el registro procedente de Los Charrúas un modelo $ARIMA(1,0,0)x(0,0,0)_{12}$ cuya ecuación (recordar que se aplicó la transformación g de Box-Cox), responde a la forma $g(y_t) - \mu = \varphi_1[g(y_{t-1}) - \mu] + A_t$. Empleando el método de los momentos, el estimador de Yule-Walker para φ_1 es $\hat{\varphi}_1 = \hat{\rho}_1 = 0,3360$. El estimador para la media es $\hat{\mu}_y = 12,5196$ y $\hat{\sigma}_A^2 = 59,931(1 - 0,3360 \cdot 0,3360) = 53,1650$, donde $\hat{\gamma}_0 = 59,931$.

El método de los momentos sugiere la siguiente ecuación:

$$g(y_t) - 12,5196 = 0,3360[g(y_{t-1}) - 12,5196] + A_t$$

Luego, resulta el modelo: $g(y_t) = 8,3130 + 0,3360 \cdot g(y_{t-1}) + A_t$ para el proceso generador de la serie procedente del registro de Los Charrúas.

Si se considera un modelo $MA(1)$, dado por $w_t = A_t + \theta_1 A_{t-1}$, donde $w_t = y_t - \mu$, de la sección 4.2 se sabe que

$$\rho_1 = \frac{\theta_1}{1 + \theta_1^2}$$

De aquí se obtiene la ecuación cuadrática $\rho_1\theta_1^2 - \theta_1 + \rho_1 = 0$, cuya solución general es

$$\theta_1 = \frac{1 \pm \sqrt{1 - 4\rho_1^2}}{2\rho_1}$$

Es necesario notar que si $\hat{\rho}_1 = \frac{1}{2}$, entonces $\hat{\theta}_1 = \frac{1 \pm \sqrt{1 - 4 \cdot \frac{1}{4}}}{2 \cdot \frac{1}{2}} = 1$ mientras que si $\hat{\rho}_1 = -\frac{1}{2}$, resulta

$\hat{\theta}_1 = \frac{1 \pm \sqrt{1 - 4 \cdot \frac{1}{4}}}{2 \cdot (-\frac{1}{2})} = -1$. Por lo tanto, en cualquier caso el proceso resultante no es invertible. Si

$|\hat{\rho}_1| > \frac{1}{2}$, la ecuación $\hat{\rho}_1\hat{\theta}_1^2 - \hat{\theta}_1 + \hat{\rho}_1 = 0$ carece de solución, por lo que $\hat{\theta}_1$ no existe. Sólo en el caso en que $|\hat{\rho}_1| < \frac{1}{2}$ la ecuación $\hat{\rho}_1\hat{\theta}_1^2 - \hat{\theta}_1 + \hat{\rho}_1 = 0$ tiene dos soluciones reales distintas y siempre es posible elegir aquel que satisface la condición de invertibilidad del proceso. Una vez que se tiene $\hat{\theta}_1$, para obtener el estimador de σ_A^2 , se utiliza:

$$\hat{\sigma}_A^2 = \frac{\hat{\gamma}_0}{1 + \hat{\theta}_1^2}$$

Tratándose de modelos autorregresivos, el método de los momentos no presenta mayores inconvenientes, no así para modelos de media móvil o procesos $ARMA$, donde, como puede verse en el apartado precedente, el grado de dificultad es mayor. En general, los estimadores obtenidos a partir del método de los momentos, no se recomiendan como estimaciones finales y no deben utilizarse si el proceso está cerca de ser no estacionario o no invertible. (Wei, 2006)

7.2. Estimación por Método de Máxima Verosimilitud

El método de máxima verosimilitud fue propuesto por R. A. Fisher en 1920 y actualmente es uno de los métodos más eficaces en la teoría de la estimación estadística, ya que produce estimadores suficientes, cuando existen, y además, éstos estimadores son asintóticamente insesgados de varianza mínima. (Freund & Walpole, 1990)

La característica fundamental de este método es que dado un conjunto de valores observados de una muestra aleatoria (es decir, un vector aleatorio, como en este caso pueden ser los registros de lluvia observados), se selecciona como estimador de un determinado parámetro desconocido de la población, el valor para el cual la probabilidad de lograr las observaciones obtenidas es máxima.

Supóngase que $\{x_t\} = \{x_1, x_2, \dots, x_n\}$ es el vector observado de una muestra, es decir la serie de tiempo, realización de un proceso estocástico $\{y(t, \omega): t = 0, \pm 1, \pm 2, \dots\}$. Si $\{y_{t_1}, y_{t_2}, \dots, y_{t_n}\}$ es un conjunto finito de variables aleatorias del proceso, entonces la distribución de probabilidad conjunta para el caso discreto será

$$f(x_1, x_2, \dots, x_n | \eta) = P(y_{t_1} = x_1, y_{t_2} = x_2, \dots, y_{t_n} = x_n)$$

que es sencillamente la probabilidad conjunta de que las variables aleatorias $y_{t_1}, y_{t_2}, \dots, y_{t_n}$ tomen el punto de muestra observado $\{x_1, x_2, \dots, x_n\}$. Puesto que los valores de la serie (muestra del proceso), son constantes fijas (vector observado), se considera a $f(x_1, x_2, \dots, x_n | \eta)$ como el valor de una función del parámetro η que se llama *función de verosimilitud*. Lo que se pretende en este método es encontrar la estimación de η para la cual la función de verosimilitud se hace máxima. Aquí η puede ser un parámetro real o un vector de parámetros. Es preciso notar que para el caso continuo f será la función de densidad de probabilidad conjunta.

Definición 7.2.1: Para cada posible vector observado $\{x_t\}$, sea $\vartheta(x_t)$ un valor observado del parámetro η cuya función de verosimilitud $f(x_1, x_2, \dots, x_n | \eta)$ es un máximo y sea $\hat{\eta} = \vartheta(y_t)$ el estimador de η . El estimador $\hat{\eta}$ se denomina *estimador máximo verosímil* de η .⁴⁹

Existen casos en los que, para ciertos vectores observados no puede alcanzarse el valor máximo de f para un punto η . En ese caso se dice que η carece de estimador máximo verosímil. También puede ocurrir que en algunos problemas, dado un vector observado, el valor máximo de f se alcanza en más de un punto η . En ese caso, el estimador máximo verosímil no tiene definición única y cualquiera de los valores obtenidos puede elegirse para $\hat{\eta}$. En algunos textos la función de verosimilitud aparece designada mediante la expresión $L(\eta)$ y el estimador máximo verosímil suele denotarse $E.M.V$ o bien $M.L.E$ en inglés. Es

⁴⁹ Aquí $\vartheta(y_t)$ puede tratarse de la media o la varianza del proceso, por ejemplo. Sendos desarrollos de la estimación por método de máxima verosimilitud puede hallarse en el libro *Probabilidad y Estadística* de Morris Degroot, (1988), Addison-Wesley Iberoamericana, Wilmington, EEUU., o también en el libro "Probabilidad y Estadística. La ciencia de la incertidumbre" de M. Evans y J. Rosenthal, (2004), Reverté, Barcelona., entre otros.

importante remarcar que el valor de $f(x_1, x_2, \dots, x_n | \eta)$ es la probabilidad de x_1, x_2, \dots, x_n dado que η es el valor verdadero, y no la probabilidad de η dado que se ha observado x_1, x_2, \dots, x_n .

Supóngase, en principio y para simplificar, que se tiene un proceso $AR(1)$:

$$y_t = \mu + \varphi_1(y_{t-1} - \mu) + A_t$$

donde $|\varphi_1| < 1$ y $A_t \sim iidN(0; \sigma_A^2)$. Dados los valores y_1, y_2, \dots, y_n se buscará la función de verosimilitud $f(y_1, y_2, \dots, y_n | \mu, \varphi_1, \sigma_A^2)$.

En el caso de un proceso $AR(1)$ como el que se está considerando, la función de verosimilitud puede escribirse como

$$f(y_1, y_2, \dots, y_n | \mu, \varphi_1, \sigma_A^2) = f(y_1) \cdot f(y_2 | y_1) \cdot \dots \cdot f(y_n | y_{n-1})$$

donde de cada f se suprimieron los parámetros por una cuestión de simplicidad en la notación⁵⁰.

Es claro que $A_t + \mu + \varphi_1(y_{t-1} - \mu) \sim N(\mu + \varphi_1(y_{t-1} - \mu); \sigma_A^2)$, que es, a su vez, la distribución de la variable y_t condicionada a y_{t-1} . Es decir, $y_t | y_{t-1} \sim N(\mu + \varphi_1(y_{t-1} - \mu); \sigma_A^2)$ y reescribiendo el proceso $AR(1)$ como $A_t = (y_t - \mu) - \varphi_1(y_{t-1} - \mu)$, se tiene:

$$f(y_t | y_{t-1}) = f_A[(y_t - \mu) - \varphi_1(y_{t-1} - \mu)]$$

donde f_A es la función de densidad de probabilidad de A_t , que es la densidad normal con media 0 y varianza σ_A^2 . Luego, es posible expresar la función de verosimilitud de la siguiente manera:

$$f(y_1, y_2, \dots, y_n | \mu, \varphi_1, \sigma_A^2) = f(y_1) \cdot \prod_{t=2}^n f_A[(y_t - \mu) - \varphi_1(y_{t-1} - \mu)] \quad (7.8)$$

Si a y_1 se la escribe mediante la forma causal

$$y_1 = \mu + \sum_{j=0}^{\infty} \varphi_1^j A_{1-j},$$

puede observarse que se trata de la combinación lineal de variables aleatorias independientes e idénticamente distribuidas normalmente. Por lo que y_1 tiene una distribución normal con media μ y varianza $\frac{\sigma_A^2}{1-\varphi_1^2}$. Por lo tanto, siguiendo la expresión que define a una densidad normal:

⁵⁰ Nótese que cada $f(y_i | y_{i-1})$ es una función de densidad de probabilidad condicional en el caso continuo o función de probabilidad condicional en el caso discreto. Por definición, $f(y_i | y_{i-1}) = \frac{f(y_{i-1}; y_i)}{f(y_{i-1})}$ que justifica la expresión $f(y_{i-1}; y_i) = f(y_{i-1}) \cdot f(y_i | y_{i-1})$. En términos estrictos no debe emplearse la misma f ya que $f(y_i | y_{i-1})$ es condicional, $f(y_{i-1}; y_i)$ es conjunta y $f(y_{i-1})$ es marginal. Un tratamiento riguroso de estas funciones multivariantes puede encontrarse en el libro *Probabilidad y Estadística* de Morris Degroot, (1988), Addison-Wesley Iberoamericana, Wilmington, EEUU., ya citado previamente, o también en *Estadística Matemática con Aplicaciones* de John Freund y Ronald Walpole, (1990), Prentice-Hall Hispanoamericana, México., entre otros.

$$f(y_1) = \frac{1}{\frac{\sigma_A \sqrt{2\pi}}{(1-\varphi_1^2)^{\frac{1}{2}}}} \cdot e^{-\frac{1}{2} \left[\frac{y_1 - \mu}{\frac{\sigma_A}{(1-\varphi_1^2)^{\frac{1}{2}}}} \right]^2}$$

Trabajando algebraicamente:

$$f(y_1) = \frac{(1-\varphi_1^2)^{\frac{1}{2}}}{\sqrt{2\pi\sigma_A^2}} \cdot e^{-\frac{1}{2} \left[\frac{(y_1 - \mu)(1-\varphi_1^2)^{\frac{1}{2}}}{\sigma_A} \right]^2}$$

Reacomodando factores queda:

$$f(y_1) = (2\pi\sigma_A^2)^{-\frac{1}{2}} \cdot (1-\varphi_1^2)^{\frac{1}{2}} \cdot e^{-\frac{1-\varphi_1^2}{2\sigma_A^2}(y_1 - \mu)^2} \quad (7.9)$$

Por otro lado, se mostró que f_A es una densidad normal con media 0 y varianza σ_A^2 , por lo que siguiendo la expresión que define a una normal con esas características:

$$f_A[(y_t - \mu) - \varphi_1(y_{t-1} - \mu)] = \frac{1}{\sqrt{2\pi\sigma_A^2}} \cdot e^{-\frac{1}{2} \left[\frac{(y_t - \mu) - \varphi_1(y_{t-1} - \mu)}{\sigma_A} \right]^2}$$

Calculando la productoria de (7.8):

$$\prod_{t=2}^n f_A[(y_t - \mu) - \varphi_1(y_{t-1} - \mu)] = \left(\frac{1}{\sqrt{2\pi\sigma_A^2}} \right)^{n-1} \cdot \left\{ e^{-\frac{1}{2} \left[\frac{(y_2 - \mu) - \varphi_1(y_1 - \mu)}{\sigma_A} \right]^2} \cdot e^{-\frac{1}{2} \left[\frac{(y_3 - \mu) - \varphi_1(y_2 - \mu)}{\sigma_A} \right]^2} \cdot \dots \cdot e^{-\frac{1}{2} \left[\frac{(y_n - \mu) - \varphi_1(y_{n-1} - \mu)}{\sigma_A} \right]^2} \right\}$$

Sumando exponentes en el segundo miembro de la igualdad:

$$\prod_{t=2}^n f_A[(y_t - \mu) - \varphi_1(y_{t-1} - \mu)] = (2\pi\sigma_A^2)^{-\frac{n-1}{2}} \cdot e^{-\frac{1}{2\sigma_A^2} \cdot \sum_{t=2}^n [(y_t - \mu) - \varphi_1(y_{t-1} - \mu)]^2} \quad (7.10)$$

En (7.8) por las expresiones obtenidas en (7.9) y (7.10) resulta:

$$f(y_1, y_2, \dots, y_n | \mu, \varphi_1, \sigma_A^2) = (2\pi\sigma_A^2)^{-\frac{n}{2}} \cdot (1-\varphi_1^2)^{\frac{1}{2}} \cdot e^{-\frac{S(\mu, \varphi_1)}{2\sigma_A^2}} \quad (7.11)$$

donde, claramente:

$$S(\mu, \varphi_1) = (1-\varphi_1^2) \cdot (y_1 - \mu)^2 + \sum_{t=2}^n [(y_t - \mu) - \varphi_1(y_{t-1} - \mu)]^2. \quad (7.12)$$

Esta expresión se conoce como suma de cuadrados no condicional.

Los valores de los parámetros que maximizan a la función de verosimilitud también maximizan al logaritmo natural de dicha función. Dado que la función de verosimilitud (7.11) está expresada como un producto, es conveniente considerar su logaritmo a los efectos de derivarla parcialmente y encontrar el valor que la maximiza. En efecto, tomando logaritmo natural en (7.11), se tiene:

$$\ln f(y_1, y_2, \dots, y_n | \mu, \varphi_1, \sigma_A^2) = \ln(2\pi\sigma_A^2)^{-\frac{n}{2}} + \ln(1 - \varphi_1)^{\frac{1}{2}} + \ln e^{-\frac{S(\mu, \varphi_1)}{2\sigma_A^2}}$$

Por las propiedades del logaritmo:

$$\ln f(y_1, y_2, \dots, y_n | \mu, \varphi_1, \sigma_A^2) = -\frac{n}{2} \ln(2\pi\sigma_A^2) + \ln(1 - \varphi_1)^{\frac{1}{2}} - \frac{S(\mu, \varphi_1)}{2\sigma_A^2} \quad (7.13)$$

Derivando parcialmente con respecto a σ_A^2 resulta:

$$\frac{\partial \ln f(y_1, y_2, \dots, y_n | \mu, \varphi_1, \sigma_A^2)}{\partial \sigma_A^2} = -\frac{n}{2} \cdot \frac{1}{2\pi\sigma_A^2} \cdot 2\pi + 0 + \frac{2S(\mu, \varphi_1)}{(2\sigma_A^2)^2} = -\frac{n}{2\sigma_A^2} + \frac{S(\mu, \varphi_1)}{2(\sigma_A^2)^2}$$

Igualando a cero para hallar el máximo:

$$-\frac{n}{2\sigma_A^2} + \frac{S(\mu, \varphi_1)}{2(\sigma_A^2)^2} = 0$$

Sumando los términos:

$$\frac{-n\sigma_A^2 + S(\mu, \varphi_1)}{2(\sigma_A^2)^2} = 0$$

Por lo que resulta:

$$\sigma_A^2 = \frac{S(\mu, \varphi_1)}{n} \quad (7.14)$$

que es el valor que maximiza la función de verosimilitud. De esta manera

$$\hat{\sigma}_A^2 = \frac{S(\hat{\mu}, \hat{\varphi}_1)}{n} \quad (7.15)$$

es el estimador máximo verosímil de σ_A^2 , donde $\hat{\mu}$ y $\hat{\varphi}_1$ son los estimadores máximo verosímiles de μ y de φ_1 . Si en (7.15) se sustituye n por $n - 2$ se obtiene la estimación por suma de cuadrados no condicional de σ_A^2 .

Por otro lado, si en (7.13) se sustituye σ_A^2 por su estimador máximo verosímil, queda:

$$\ln f(y_1, y_2, \dots, y_n | \mu, \varphi_1) = -\frac{n}{2} \ln \left[2\pi \cdot \frac{S(\mu, \varphi_1)}{n} \right] + \frac{1}{2} \cdot \ln(1 - \varphi_1) - \frac{S(\mu, \varphi_1)}{2 \cdot \frac{S(\mu, \varphi_1)}{n}}$$

Desarrollando cada término:

$$\ln f(y_1, y_2, \dots, y_n | \mu, \varphi_1) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \left[\frac{S(\mu, \varphi_1)}{n} \right] + \frac{1}{2} \cdot \ln(1 - \varphi_1) - \frac{n}{2}$$

De donde, factorizando se obtiene:

$$\ln f(y_1, y_2, \dots, y_n | \mu, \varphi_1) = -\frac{n}{2} \left\{ \ln(2\pi) + \ln \left[\frac{S(\mu, \varphi_1)}{n} \right] - \frac{\ln(1 - \varphi_1)}{n} + 1 \right\}$$

Esta igualdad puede escribirse como

$$\ln f(y_1, y_2, \dots, y_n | \mu, \varphi_1) = -\frac{n}{2} \{ \ln(2\pi) + 1 + l(\mu, \varphi_1) \} \quad (7.15)$$

Donde

$$l(\mu, \varphi_1) = \ln \left[\frac{S(\mu, \varphi_1)}{n} \right] - \frac{\ln(1 - \varphi_1)}{n} \quad (7.16)$$

Claramente los extremos de (7.15) son los extremos de (7.16). Puede emplearse el método de *Newton-Raphson* para hallar los valores $\hat{\mu}$ y $\hat{\varphi}_1$ que cumplen dicha condición. Los estimadores de máxima verosimilitud de μ y de φ_1 se encuentran resolviendo el sistema:

$$\begin{cases} \frac{\partial l(\mu, \varphi_1)}{\partial \mu} = 0 \\ \frac{\partial l(\mu, \varphi_1)}{\partial \varphi_1} = 0 \end{cases}$$

Si se tiene en cuenta el gradiente de (7.16):

$$\text{grad}l(\mu, \varphi_1) = \left(\frac{\partial l(\mu, \varphi_1)}{\partial \mu}, \frac{\partial l(\mu, \varphi_1)}{\partial \varphi_1} \right)_{1 \times 2},$$

puede observarse que $\text{grad}l(\hat{\mu}, \hat{\varphi}_1) = (0, 0)_{1 \times 2}$. Además, la matriz hessiana⁵¹ está dada por:

$$Hl(\mu, \varphi_1) = \begin{pmatrix} \frac{\partial^2 l(\mu, \varphi_1)}{\partial \mu \partial \mu} & \frac{\partial^2 l(\mu, \varphi_1)}{\partial \mu \partial \varphi_1} \\ \frac{\partial^2 l(\mu, \varphi_1)}{\partial \varphi_1 \partial \mu} & \frac{\partial^2 l(\mu, \varphi_1)}{\partial \varphi_1 \partial \varphi_1} \end{pmatrix}_{2 \times 2}$$

Sean $\mu^{(0)}$ y $\varphi_1^{(0)}$ estimadores iniciales de μ y de φ_1 . La aproximación de Taylor de $\text{grad}l(\hat{\mu}, \hat{\varphi}_1)$ alrededor de estas aproximaciones es:

$$\text{grad}l(\hat{\mu}, \hat{\varphi}_1) \approx \text{grad}l(\mu^{(0)}, \varphi_1^{(0)}) + Hl(\mu^{(0)}, \varphi_1^{(0)}) \cdot (\hat{h} - h^{(0)})$$

⁵¹ Si la matriz hessiana es definida positiva en un vector crítico de una función, entonces la función tiene un mínimo relativo en ese vector. Si la matriz es definida negativa, entonces existe un máximo relativo. Para mayor detalle puede consultarse Cálculo Vectorial de Jerrold Marsden y Anthony Tromba citado en la bibliografía de este documento.

donde $\hat{h} = (\hat{\mu}, \hat{\varphi}_1)$ y $h^{(0)} = (\mu^{(0)}, \varphi_1^{(0)})$. Si $Hl(\mu, \varphi_1)$ es no singular, una solución iterativa inmediata para

$$\text{gradl}(\mu^{(0)}, \varphi_1^{(0)}) + Hl(\mu^{(0)}, \varphi_1^{(0)}) \cdot (\hat{h} - h^{(0)}) = 0$$

está dada por la expresión:

$$h^{(1)} = h^{(0)} - [Hl(\mu^{(0)}, \varphi_1^{(0)})]^{-1} \cdot \text{gradl}(\mu^{(0)}, \varphi_1^{(0)})$$

Donde $h^{(1)} = (\mu^{(1)}, \varphi_1^{(1)})$. A su vez:

$$h^{(2)} = h^{(1)} - [Hl(\mu^{(1)}, \varphi_1^{(1)})]^{-1} \cdot \text{gradl}(\mu^{(1)}, \varphi_1^{(1)})$$

Donde $h^{(2)} = (\mu^{(2)}, \varphi_1^{(2)})$. En general, la continuación de las iteraciones genera una secuencia $h^{(0)}, h^{(1)}, h^{(2)}, \dots$ que converge a la solución \hat{h} . Se remite al lector a consultar los detalles de interés sobre este método a cualquier libro de Análisis Numérico como el de *Richard Burden y J. Douglas Faires*, por ejemplo.

Asociada a la matriz Hessiana de la función de verosimilitud aparece la matriz de información, a partir de la cual es posible obtener una aproximación asintótica de la varianza del vector máximo verosímil. La matriz de información del vector $h = (\mu, \varphi_1)$ se define:

$$I(h) = E \left\{ \left[\frac{\partial \ln f(y_1, y_2, \dots, y_n | h)}{\partial h} \right]^2 \right\}$$

y se puede demostrar que $I(h) = -E[H \ln f(y_1, y_2, \dots, y_n | h)]$, donde $H \ln f(y_1, y_2, \dots, y_n | h)$ es la matriz Hessiana del logaritmo de la función de verosimilitud⁵². Bajo condiciones apropiadas, la inversa de la matriz de información es un estimador asintótico de la varianza de \hat{h} . Esto es:

$$\text{Var}[\hat{h}] \approx I^{-1}(h).$$

En la siguiente sección se vuelve sobre este resultado. El lector interesado en un detalle más completo al respecto puede consultar *Time Series Analysis and its Applications de Robert Shumway y David Stoffer*, pág.133 o también pág. 243 del libro *Time Series Analysis: Forecasting and Control* de Box, Jenkins y Reinsel (1994), entre otros.

⁵² Esta es una generalización de la Información de Fisher respecto de la cual se da un elegante tratamiento en el libro *Probabilidad y Estadística de Morris DeGroot* citado oportunamente en este documento. Allí puede encontrarse una demostración del resultado que se menciona, como también que la información de Fisher en una muestra aleatoria de n observaciones es n veces la información de Fisher en una sola observación. Además se proporciona una deducción de la denominada desigualdad de la información o desigualdad de Cramér-Rao que establece una cota inferior para la varianza de un estimador máximo verosímil y un análisis de las propiedades de los estimadores máximo verosímiles para muestras grandes (pág. 409).

En el caso de los modelos autorregresivos, es posible condicionar la función de verosimilitud para lograr eliminar el término que genera la no linealidad de la misma. En efecto, condicionando en y_1 , la verosimilitud queda:

$$f(y_1, \dots, y_n; \mu, \varphi_1, \sigma_A^2 | y_1) = \prod_{t=2}^n f_A[(y_t - \mu) - \varphi_1(y_{t-1} - \mu)] = (2\pi\sigma_A^2)^{-\frac{n-1}{2}} \cdot e^{-\frac{S_c(\mu, \varphi_1)}{2\sigma_A^2}} \quad (7.17)$$

donde la suma de cuadrados condicional es:

$$S_c(\mu, \varphi_1) = \sum_{t=2}^n [(y_t - \mu) - \varphi_1(y_{t-1} - \mu)]^2. \quad (7.18)$$

Por lo que el estimador máximo verosímil condicional de σ_A^2 es:

$$\hat{\sigma}_A^2 = \frac{S_c(\hat{\mu}, \hat{\varphi}_1)}{n-1} \quad (7.19)$$

donde $\hat{\mu}$ y $\hat{\varphi}_1$ son los valores que minimizan la suma de cuadrados condicional (7.18)⁵³. Considerando, como al inicio, $\delta = \mu(1 - \varphi_1)$, la suma de cuadrados condicional puede escribirse como:

$$S_c(\mu, \varphi_1) = \sum_{t=2}^n [y_t - (\delta + \varphi_1 y_{t-1})]^2. \quad (7.20)$$

Si se utiliza la estimación mediante el método de mínimos cuadrados ordinarios, se tiene $\hat{\delta} = \bar{y}_{(2)} - \hat{\varphi}_1 \bar{y}_{(1)}$, donde $\bar{y}_{(1)} = \frac{1}{n-1} \sum_{t=1}^{n-1} y_t$, y $\bar{y}_{(2)} = \frac{1}{n-1} \sum_{t=2}^n y_t$, y por lo tanto, los estimadores condicionales son:

$$\hat{\mu} = \frac{\bar{y}_{(2)} - \hat{\varphi}_1 \bar{y}_{(1)}}{1 - \hat{\varphi}_1} \quad (7.21)$$

$$\hat{\varphi}_1 = \frac{\sum_{t=2}^n [y_t - \bar{y}_{(2)}][y_{t-1} - \bar{y}_{(1)}]}{\sum_{t=2}^n [y_{t-1} - \bar{y}_{(1)}]^2} \quad (7.22)$$

Puede apreciarse que $\hat{\mu} \approx \hat{\mu}_y$ y $\hat{\varphi}_1 \approx \hat{\rho}_1$. Es decir, que los estimadores de Yule-Walker y los estimadores por mínimos cuadrados condicionales son aproximadamente iguales (Shumway & Stoffer, 2011).

La verosimilitud para un $AR(1)$ puede generalizarse para un modelo $ARMA(p, q)$. Sea el vector paramétrico $h = (\mu, \varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q)'$, $(p + q + 1) - dimensional$. Puede expresarse la función de verosimilitud mediante la expresión:

⁵³ Nótese que los valores que minimizan $S_c(\mu, \varphi_1)$ son los que maximizan la verosimilitud.

$$L(h, \sigma_A^2) = \prod_{t=1}^n f(y_t | y_{t-1}, \dots, y_1) \quad (7.23)$$

Escribiendo la ecuación de un proceso $ARMA(p, q)$ en la forma

$$A_t = \theta_1 A_{t-1} + \theta_2 A_{t-2} + \dots + \theta_q A_{t-q} + w_t - \varphi_1 w_{t-1} - \varphi_2 w_{t-2} - \dots - \varphi_p w_{t-p}, \quad (7.24)$$

donde $w_t = y_t - \mu$ y $A_t \sim iidN(0; \sigma_A^2)$, y teniendo en cuenta que la densidad de probabilidad conjunta de $A_t = (A_1, \dots, A_n)'$ está dada por:

$$f_A(A_1, \dots, A_n | h, \sigma_A^2) = (2\pi\sigma_A^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_A^2} \sum_{t=1}^n A_t^2}, \quad (7.25)$$

puede escribirse el logaritmo natural de la función de verosimilitud condicional de la siguiente manera:

$$\ln L^*(h, \sigma_A^2) = -\frac{n}{2} \ln(2\pi\sigma_A^2) - \frac{S^*(h)}{2\sigma_A^2}, \quad (7.26)$$

donde

$$S^*(h) = \sum_{t=1}^n A_t^2(h | w^*, A^*, w). \quad (7.27)$$

A su vez $w = (w_1, \dots, w_n)'$, con las condiciones iniciales conocidas $w^* = (w_{1-p}, \dots, w_{-1}, w_0)'$ y $A^* = (A_{1-q}, \dots, A_{-1}, A_0)'$. Estas condiciones se basan comúnmente en el siguiente criterio: los valores iniciales son las p primeras observaciones de la serie, mientras que las innovaciones previas A_i son nulas. Es decir, se determinan a partir de la ecuación (7.24) desde $t = p + 1$ en adelante.

El vector $\hat{h} = (\hat{\mu}, \hat{\varphi}_1, \dots, \hat{\varphi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q)'$, que maximiza la función de verosimilitud, es llamado el vector estimador de máxima verosimilitud condicional. Una vez que se tiene el vector \hat{h} puede calcularse el estimador $\hat{\sigma}_A^2$ de σ_A^2 mediante la expresión

$$\hat{\sigma}_A^2 = \frac{S^*(h)}{g.l.}, \quad (7.28)$$

donde el número de grados de libertad $g.l.$ es igual al número de términos usados en la suma $S^*(h)$ menos el número de parámetros estimados. Si se emplea la suma (7.27) para calcular $\hat{\sigma}_A^2$, entonces $g.l. = (n - p) - (p + q + 1) = n - (2p + q + 1)$.

De la misma manera que se realiza para el caso condicional, siguiendo a Box, Jenkins y Reinsel (1994), se puede escribir la función de verosimilitud no condicional de un modelo $ARMA(p, q)$ como:

$$\ln L(h, \sigma_A^2) = -\frac{n}{2} \ln(2\pi\sigma_A^2) - \frac{S(h)}{2\sigma_A^2} \quad (7.29)$$

donde $S(h)$ dada por

$$S(h) = \sum_{-\infty}^n [E(A_t|h, w)]^2$$

se llama suma de cuadrados no condicional y $E(A_t|h, w)$ es la esperanza condicional de A_t dados h y w .

En el caso de tratar con un modelo estacional, el vector estimador de máxima verosimilitud será $\hat{h} = (\hat{\mu}, \hat{\Phi}_1, \dots, \hat{\Phi}_p, \hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\Theta}_1, \dots, \hat{\Theta}_q, \hat{\theta}_1, \dots, \hat{\theta}_q)'$.

Ejemplo 7.2.1: Tomando la serie de registros de lluvia procedente de Los Charrúas, que cuenta con 228 observaciones, como ya se hizo en el ejemplo 7.1.1, se obtiene:

$$\bar{y}_{(1)} = \frac{1}{n-1} \sum_{t=1}^{n-1} y_t = \frac{2851,064175}{228-1} = 12,55975407$$

$$\bar{y}_{(2)} = \frac{1}{n-1} \sum_{t=2}^n y_t = \frac{2850,47898}{228-1} = 12,55717613$$

Luego, se calcula las diferencias $y_t - \bar{y}_{(2)}$ (que es cada observación de la serie a partir del segundo registro menos 12,55717613), y $y_{t-1} - \bar{y}_{(1)}$ (que es cada observación de la serie hasta el anteúltimo registro menos 12,55975407), para obtener:

$$\sum_{t=2}^n [y_t - \bar{y}_{(2)}][y_{t-1} - \bar{y}_{(1)}] = 4570,836761$$

$$\sum_{t=2}^n [y_{t-1} - \bar{y}_{(1)}]^2 = 13520,94458$$

Por lo tanto, a partir de (7.22), el estimador máximo verosímil condicional de ϕ_1 es

$$\hat{\phi}_1 = \frac{4570,836761}{13520,94458} = 0,33805602.$$

Una vez que se tiene $\hat{\phi}_1$, empleando (7.21), se obtiene:

$$\hat{\mu} = \frac{12,55717613 - 0,33805602 \cdot 12,55975407}{1 - 0,33805602} = 12,5558596.$$

De (7.18): $S_c(\hat{\mu}, \hat{\phi}_1) = 11986,1185$. Por lo que, de (7.19):

$$\hat{\sigma}_A^2 = \frac{11986,1185}{228-1} = 52,8022843.$$

La constante α del modelo puede estimarse ahora como:

$$\hat{\delta} = \hat{\mu}(1 - \hat{\phi}_1) = 8,311275676.$$

Luego, el modelo $ARIMA(1,0,0) \times (0,0,0)_{12}$ para la serie de registros de lluvia de Los Charrúas, obtenido mediante máxima verosimilitud condicional, es:

$$g(y_t) - 12,5558596 = 0,33805602[g(y_{t-1}) - 12,5558596] + A_t$$

Si se emplea $\hat{\alpha}$:

$$g(y_t) = 8,311275676 + 0,33805602 \cdot g(y_{t-1}) + A_t.$$

7.3. Contrastes de Significación de los Parámetros Estimados

La significación de los parámetros que se estimaron en la fase anterior es fundamental para decidir su inclusión en los modelos propuestos. Si los parámetros que se estimaron no son significativamente diferentes de cero, entonces deben suprimirse del modelo para evitar la sobreparametrización. Puesto que las series se ajustan a modelos *ARMA* estacionarios, estos contrastes de significación pueden realizarse mediante un estadístico con distribución normal. Particularmente, si el vector a estimar es

$$h = (\delta, \Phi_1, \dots, \Phi_p, \varphi_1, \dots, \varphi_p, \Theta_1, \dots, \Theta_q, \theta_1, \dots, \theta_q)', \quad (7.30)$$

para determinar si el modelo está sobreestimado, se plantean los siguientes test de hipótesis:

$$H_0: \delta = 0 \text{ versus } H_1: \delta \neq 0 \quad (7.31)$$

$$H_0: \Phi_i = 0 \text{ versus } H_1: \Phi_i \neq 0 \quad (7.32)$$

$$H_0: \varphi_i = 0 \text{ versus } H_1: \varphi_i \neq 0 \quad (7.33)$$

$$H_0: \Theta_i = 0 \text{ versus } H_1: \Theta_i \neq 0 \quad (7.34)$$

$$H_0: \theta_i = 0 \text{ versus } H_1: \theta_i \neq 0 \quad (7.35)$$

Para muestras grandes, el estimador máximo verosímil \hat{h} de h tiene una distribución aproximadamente normal con media h y matriz de covarianzas $Var[h]$. Por lo que el estadístico z a utilizar en los contrastes es:

$$z = \frac{\hat{h}_i - 0}{\sqrt{Var[\hat{h}_i]}} \rightsquigarrow N(0,1), \quad (7.36)$$

A los efectos de obtener una estimación para la matriz de covarianzas de \hat{h} supóngase que, para simplificar, $h = (\boldsymbol{\varphi}, \boldsymbol{\theta})'$ donde $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_p)'$ y $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)'$ y $\ln f(\boldsymbol{\varphi}, \boldsymbol{\theta})$ es el logaritmo de la función de verosimilitud asociada a un proceso *ARMA*(p, q). Las derivadas parciales del logaritmo de la función de verosimilitud con respecto a φ_i y a θ_j en el vector inicial $h_0 = (\boldsymbol{\varphi}_0, \boldsymbol{\theta}_0)$ son:

$$u_{t-i} = \frac{\partial \ln f(h_0)}{\partial \varphi_i} \quad y \quad v_{t-j} = \frac{\partial \ln f(h_0)}{\partial \theta_j}$$

Resultan así los siguientes vectores:

$$U_t = (u_{t-1}, u_{t-2}, \dots, u_{t-p})_{1 \times p} = \left(\frac{\partial \ln f}{\partial \varphi_1}, \frac{\partial \ln f}{\partial \varphi_2}, \dots, \frac{\partial \ln f}{\partial \varphi_p} \right)_{1 \times p} \quad \text{y} \quad V_t = (v_{t-1}, v_{t-2}, \dots, v_{t-q})_{1 \times q} = \left(\frac{\partial \ln f}{\partial \theta_1}, \frac{\partial \ln f}{\partial \theta_2}, \dots, \frac{\partial \ln f}{\partial \theta_q} \right)_{1 \times q}.$$

De esta manera $\{U_t\}$ y $\{V_t\}$ son tales que conforman los procesos autorregresivos: $\varphi(B)U_t = A_t$ y $\theta(B)V_t = A_t$.

Ahora, considérese los productos:

$$U_t' U_t = \begin{pmatrix} u_{t-1} \\ u_{t-2} \\ \vdots \\ u_{t-p} \end{pmatrix}_{p \times 1} \cdot (u_{t-1}, u_{t-2}, \dots, u_{t-p})_{1 \times p} = \begin{pmatrix} u_{t-1}^2 & u_{t-1}u_{t-2} & \dots & u_{t-1}u_{t-p} \\ u_{t-2}u_{t-1} & u_{t-2}^2 & \dots & u_{t-2}u_{t-p} \\ \vdots & \vdots & \ddots & \vdots \\ u_{t-p}u_{t-1} & u_{t-p}u_{t-2} & \dots & u_{t-p}^2 \end{pmatrix}_{p \times p}$$

$$U_t' V_t = \begin{pmatrix} u_{t-1} \\ u_{t-2} \\ \vdots \\ u_{t-p} \end{pmatrix}_{p \times 1} \cdot (v_{t-1}, v_{t-2}, \dots, v_{t-p})_{1 \times q} = \begin{pmatrix} u_{t-1}v_{t-1} & u_{t-1}v_{t-2} & \dots & u_{t-1}v_{t-q} \\ u_{t-2}v_{t-1} & u_{t-2}v_{t-2} & \dots & u_{t-2}v_{t-q} \\ \vdots & \vdots & \ddots & \vdots \\ u_{t-p}v_{t-1} & u_{t-p}v_{t-2} & \dots & u_{t-p}v_{t-q} \end{pmatrix}_{p \times q}$$

$$V_t' U_t = \begin{pmatrix} v_{t-1} \\ v_{t-2} \\ \vdots \\ v_{t-p} \end{pmatrix}_{q \times 1} \cdot (u_{t-1}, u_{t-2}, \dots, u_{t-p})_{1 \times p} = \begin{pmatrix} v_{t-1}u_{t-1} & v_{t-1}u_{t-2} & \dots & v_{t-1}u_{t-p} \\ v_{t-2}u_{t-1} & v_{t-2}u_{t-2} & \dots & v_{t-2}u_{t-p} \\ \vdots & \vdots & \ddots & \vdots \\ v_{t-q}u_{t-1} & v_{t-q}u_{t-2} & \dots & v_{t-q}u_{t-p} \end{pmatrix}_{q \times p}$$

$$V_t' V_t = \begin{pmatrix} v_{t-1} \\ v_{t-2} \\ \vdots \\ v_{t-p} \end{pmatrix}_{q \times 1} \cdot (v_{t-1}, v_{t-2}, \dots, v_{t-p})_{1 \times q} = \begin{pmatrix} v_{t-1}^2 & v_{t-1}v_{t-2} & \dots & v_{t-1}v_{t-q} \\ v_{t-2}v_{t-1} & v_{t-2}^2 & \dots & v_{t-2}v_{t-q} \\ \vdots & \vdots & \ddots & \vdots \\ v_{t-q}v_{t-1} & v_{t-q}v_{t-2} & \dots & v_{t-q}^2 \end{pmatrix}_{q \times q}$$

Se define la *matriz de información* de $h = (\varphi, \theta)$ para un modelo ARMA mediante:

$$I(\varphi, \theta) = \frac{1}{\sigma_A^2} E \begin{bmatrix} U_t' U_t & U_t' V_t \\ V_t' U_t & V_t' V_t \end{bmatrix} \quad (7.37)$$

Esta expresión puede escribirse como:

$$I(\varphi, \theta) = \frac{1}{\sigma_A^2} E \begin{bmatrix} u_{t-1}^2 & u_{t-1}u_{t-2} & \dots & u_{t-1}u_{t-p} & & u_{t-1}v_{t-1} & u_{t-1}v_{t-2} & \dots & u_{t-1}v_{t-q} \\ u_{t-2}u_{t-1} & u_{t-2}^2 & \dots & u_{t-2}u_{t-p} & \vdots & u_{t-2}v_{t-1} & u_{t-2}v_{t-2} & \dots & u_{t-2}v_{t-q} \\ \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ u_{t-p}u_{t-1} & u_{t-p}u_{t-2} & \dots & u_{t-p}^2 & & u_{t-p}v_{t-1} & u_{t-p}v_{t-2} & \dots & u_{t-p}v_{t-q} \\ & & & & \dots & & & & \\ & & & & & & & & \\ v_{t-1}u_{t-1} & v_{t-1}u_{t-2} & \dots & v_{t-1}u_{t-p} & & v_{t-1}^2 & v_{t-1}v_{t-2} & \dots & v_{t-1}v_{t-q} \\ v_{t-2}u_{t-1} & v_{t-2}u_{t-2} & \dots & v_{t-2}u_{t-p} & \vdots & v_{t-2}v_{t-1} & v_{t-2}^2 & \dots & v_{t-2}v_{t-q} \\ \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ v_{t-q}u_{t-1} & v_{t-q}u_{t-2} & \dots & v_{t-q}u_{t-p} & & v_{t-q}v_{t-1} & v_{t-q}v_{t-2} & \dots & v_{t-q}^2 \end{bmatrix}$$

$$= \frac{n}{\sigma_A^2} \begin{bmatrix} \gamma_0(uu) & \gamma_1(uu) & \dots & \gamma_{p-1}(uu) & & \gamma_0(uv) & \gamma_{-1}(uv) & \dots & \gamma_{1-q}(uv) \\ \gamma_1(uu) & \gamma_0(uu) & \dots & \gamma_{p-2}(uu) & \vdots & \vdots & \gamma_1(uv) & \gamma_0(uv) & \dots & \gamma_{2-q}(uv) \\ \vdots & \vdots \\ \gamma_{p-1}(uu) & \gamma_{p-2}(uu) & \dots & \gamma_0(uu) & \dots & \dots & \gamma_{p-1}(uv) & \gamma_{p-2}(uv) & \dots & \gamma_{p-q}(uv) \\ \gamma_0(uv) & \gamma_1(uv) & \dots & \gamma_{p-1}(uv) & \dots & \dots & \gamma_0(vv) & \gamma_1(vv) & \dots & \gamma_{q-1}(vv) \\ \gamma_{-1}(uv) & \gamma_0(uv) & \dots & \gamma_{p-2}(uv) & \vdots & \vdots & \gamma_1(vv) & \gamma_0(vv) & \dots & \gamma_{q-2}(vv) \\ \vdots & \vdots \\ \gamma_{1-q}(uv) & \gamma_{2-q}(uv) & \dots & \gamma_{p-q}(uv) & & & \gamma_{q-1}(vv) & \gamma_{q-2}(vv) & \dots & \gamma_0(vv) \end{bmatrix} \quad (7.38)$$

Aquí $\gamma_k(uu)$ y $\gamma_k(vv)$ son las autocovarianzas de las variables u_t y v_t , mientras que las covarianzas cruzadas se definen como:

$$\gamma_k(uv) = \gamma_{-k}(vu) = E[u_t v_{t+k}] = E[v_t u_{t-k}].$$

Para muestras considerablemente grandes la matriz de covarianzas de los estimadores de máxima verosimilitud se obtiene mediante:

$$Var[\hat{h}] = Var[\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\theta}}] \approx I^{-1}(\boldsymbol{\varphi}, \boldsymbol{\theta}). \quad (7.39)$$

Cabe en este punto, realizar la siguiente aclaración: para observaciones independientes e idénticamente distribuidas el estimador máximo verosímil \hat{h} de h converge en probabilidad a h y la distribución de $\sqrt{n}(\hat{h} - h)$ es aproximadamente normal con media 0 y matriz de covarianza $I^{-1}(h)$. En el análisis de series temporales las observaciones no son *iid*. Citando a Brockwell y Davis: "*Probabilidad en el contexto de series de tiempo se utiliza casi siempre en el sentido de probabilidad gaussiana, es decir, la probabilidad calculada conforme a la (posiblemente falsa) suposición de que la serie es gaussiana. No obstante, para muestras grandes, los estimadores de los coeficientes de los procesos ARMA calculados por la maximización de la probabilidad gaussiana tienen buenas propiedades análogas a las descritas (...)*." (Brockwell & Davis, Introduction to Time Series and Forecasting, 2002), pág. 387.

Sea $\Gamma_p(\boldsymbol{\varphi})$ la matriz de covarianzas de orden $p \times p$ de las p sucesivas observaciones de un proceso $AR(p)$ con vector de parámetros $\boldsymbol{\varphi}' = (\varphi_1, \dots, \varphi_p)$. En consecuencia, empleando (7.46) la matriz de covarianzas de orden $p \times p$ del vector $\hat{\boldsymbol{\varphi}}$ está dada por:

$$Var[\hat{\boldsymbol{\varphi}}] = \frac{\sigma_A^2}{n} \Gamma_p^{-1}(\boldsymbol{\varphi}) \quad (7.40)$$

En el caso de un proceso $MA(q)$ con vector de parámetros $\boldsymbol{\theta}' = (\theta_1, \dots, \theta_q)$, la matriz de covarianza de orden $q \times q$, es $\Gamma_q(\boldsymbol{\theta})$. Mediante (7.46) la matriz de covarianzas de orden $q \times q$ del vector $\hat{\boldsymbol{\theta}}$ está dada por:

$$Var[\hat{\boldsymbol{\theta}}] = \frac{\sigma_A^2}{n} \Gamma_q^{-1}(\boldsymbol{\theta}) \quad (7.41)$$

Ejemplo 7.3.1: Para un proceso $AR(1)$, $y_t = \varphi_1 y_{t-1} + A_t$, la varianza de $\hat{\varphi}_1$ es $Var[\hat{\varphi}_1] = \frac{1-\varphi_1^2}{n}$. En el caso de un proceso $AR(2)$, $(1 - \varphi_1 B - \varphi_2 B^2)y_t = A_t$, la matriz de covarianzas de los estimadores $\hat{\varphi}_1$ y $\hat{\varphi}_2$ es:

$$Var[\hat{\varphi}_1, \hat{\varphi}_2] = \frac{1}{n} \begin{bmatrix} 1 - \varphi_1^2 & -\varphi_1(1 + \varphi_2) \\ -\varphi_1(1 + \varphi_2) & 1 - \varphi_1^2 \end{bmatrix}$$

Supóngase que se tiene el proceso $AR(2)$ del ejemplo 4.2.2, $(1 + 0,4B - 0,5B^2)y_t = A_t$ para $n = 250$. Claramente $\hat{\varphi}_1 = -0,4$ y $\hat{\varphi}_2 = 0,5$. Por lo tanto,

$$Var[\hat{\varphi}_1, \hat{\varphi}_2] = \frac{1}{250} \begin{bmatrix} 1 - (-0,4)^2 & -(-0,4)(1 + 0,5) \\ -(-0,4)(1 + 0,5) & 1 - (-0,4)^2 \end{bmatrix} = \frac{1}{250} \begin{bmatrix} 0,84 & 0,6 \\ 0,6 & 0,84 \end{bmatrix}$$

Resulta, entonces, que $Var[\hat{\varphi}_1] = \frac{0,84}{250} = 0,00336$. Por otro lado:

$$Var[\hat{\varphi}_2, \hat{\varphi}_1] = \frac{1}{250} \begin{bmatrix} 1 - (0,5)^2 & -(0,5)(1 - 0,4) \\ -(0,5)(1 - 0,4) & 1 - (0,5)^2 \end{bmatrix} = \frac{1}{250} \begin{bmatrix} 0,75 & -0,3 \\ -0,3 & 0,75 \end{bmatrix}$$

La varianza de $\hat{\varphi}_2$ será entonces $Var[\hat{\varphi}_2] = \frac{0,75}{250} = 0,003$.

Ejemplo 7.3.2: Teniendo en cuenta que $\theta(B)V_t = A_t$, la matriz de covarianzas del estimador máximo verosímil $\hat{\theta}_1$ en el caso de un proceso $MA(1)$ es $Var[\hat{\theta}_1] = \frac{1-\theta_1^2}{n}$. Para un $MA(2)$ se tendrá:

$$Var[\hat{\theta}_1, \hat{\theta}_2] = \frac{1}{n} \begin{bmatrix} 1 - \theta_1^2 & \theta_1(1 - \theta_2) \\ \theta_1(1 - \theta_2) & 1 - \theta_1^2 \end{bmatrix}.$$

En el libro *Time Series Analysis: Forecasting and Control* de Box, Jenkins y Reinsel (1994), se da una forma alternativa basada en que un proceso $ARMA(p, q)$ puede escribirse en términos de los ceros de $\varphi(B)$ y de $\theta(B)$ de la siguiente manera:

$$\prod_{i=1}^p (1 - o_i B) w_t = \prod_{j=1}^q (1 - l_j B) A_t, \quad (7.42)$$

donde las raíces se asumen reales. Por lo que, considerando los vectores $\mathbf{o}' = (o_1, \dots, o_p)$ y $\mathbf{l}' = (l_1, \dots, l_q)$, la matriz de información queda⁵⁴ (7.43):

$$I(\mathbf{o}, \mathbf{l}) = n \cdot \begin{bmatrix} (1 - o_1^2)^{-1} & (1 - o_1 o_2)^{-1} & \dots & (1 - o_1 o_p)^{-1} & & & -(1 - o_1 l_1)^{-1} & \dots & -(1 - o_1 l_q)^{-1} \\ \vdots & \vdots \\ (1 - o_1 o_p)^{-1} & (1 - o_2 o_p)^{-1} & \dots & (1 - o_p^2)^{-1} & & & -(1 - o_p l_1)^{-1} & \dots & -(1 - o_p l_q)^{-1} \\ & \dots & & & \dots & \dots & & & \\ -(1 - o_1 l_1)^{-1} & -(1 - o_2 l_1)^{-1} & \dots & -(1 - o_p l_1)^{-1} & & & (1 - l_1^2)^{-1} & \dots & (1 - l_1 l_q)^{-1} \\ \vdots & \vdots \\ -(1 - o_1 l_q)^{-1} & -(1 - o_2 l_q)^{-1} & \dots & -(1 - o_p l_q)^{-1} & & & (1 - l_1 l_q)^{-1} & \dots & (1 - l_q^2)^{-1} \end{bmatrix}$$

Ejemplo 7.3.3: Considérese un proceso $ARMA(1,1)$ dado por la siguiente ecuación:

⁵⁴ Los detalles pueden consultarse en la página 258 del libro *Time Series Analysis: Forecasting and Control* de Box, Jenkins y Reinsel (1994) citado en la bibliografía de este documento.

$$(1 - \varphi_1 B)w_t = (1 - \theta_1 B)A_t$$

Entonces, de (7.43) se obtiene:

$$\begin{aligned} \text{Var}[\hat{\varphi}_1, \hat{\theta}_1] &\approx \frac{1}{n} \begin{bmatrix} (1 - \varphi_1^2)^{-1} & -(1 - \varphi_1 \theta_1)^{-1} \\ -(1 - \varphi_1 \theta_1)^{-1} & (1 - \theta_1^2)^{-1} \end{bmatrix}^{-1} \\ &= \frac{1}{n} \frac{1 + \varphi_1 \theta_1}{(\varphi_1 + \theta_1)^2} \begin{bmatrix} (1 - \varphi_1^2)(1 + \varphi_1 \theta_1) & -(1 - \theta_1^2)(1 - \varphi_1^2) \\ -(1 - \theta_1^2)(1 - \varphi_1^2) & (1 - \theta_1^2)(1 + \varphi_1 \theta_1) \end{bmatrix} \end{aligned}$$

Si se considera el proceso $ARMA(1,1)$ del ejemplo 4.1.3, $(1 - 0,5B)y_t = (1 + 0,4B)A_t$, entonces $\varphi_1 = 0,5$ y $\theta_1 = 0,4$. Por lo tanto, para $n = 250$, la matriz de covarianzas es:

$$\begin{aligned} \text{Var}[\hat{\varphi}_1, \hat{\theta}_1] &= \frac{1}{250} \frac{1 + 0,5 \cdot 0,4}{(0,5 + 0,4)^2} \begin{bmatrix} (1 - 0,5^2)(1 + 0,5 \cdot 0,4) & -(1 - 0,4^2)(1 - 0,5^2) \\ -(1 - 0,4^2)(1 - 0,5^2) & (1 - 0,4^2)(1 + 0,5 \cdot 0,4) \end{bmatrix} \\ &= \frac{1}{250} \frac{1,2}{0,81} \begin{bmatrix} 0,9 & -0,63 \\ -0,63 & 1,008 \end{bmatrix} \end{aligned}$$

Luego, $\text{Var}[\hat{\varphi}_1] \approx 0,00533$ y $\text{Var}[\hat{\theta}_1] \approx 0,00597$.

Para un proceso $ARMA(2,2)$, con vectores paramétricos $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$ y $\boldsymbol{\theta} = (\theta_1, \theta_2)$, dado por la ecuación:

$$(1 - \varphi_1 B - \varphi_2 B^2)w_t = (1 - \theta_1 B - \theta_2 B^2)A_t,$$

resulta (Box, Jenkins, & Reinsel, 1994):

$$\begin{aligned} \text{Var}[\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\theta}}] &\cong \frac{1}{n} \begin{bmatrix} (1 - \varphi_1^2)^{-1} & (1 - \varphi_1 \varphi_2)^{-1} & \vdots & \vdots & -(1 - \varphi_1 \theta_1)^{-1} & -(1 - \varphi_1 \theta_2)^{-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (1 - \varphi_1 \varphi_2)^{-1} & (1 - \varphi_2^2)^{-1} & \dots & \dots & -(1 - \varphi_2 \theta_1)^{-1} & -(1 - \varphi_2 \theta_2)^{-1} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -(1 - \varphi_1 \theta_1)^{-1} & -(1 - \varphi_2 \theta_1)^{-1} & \dots & \dots & (1 - \theta_1^2)^{-1} & (1 - \theta_1 \theta_2)^{-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -(1 - \varphi_1 \theta_2)^{-1} & -(1 - \varphi_2 \theta_2)^{-1} & \dots & \dots & (1 - \theta_1 \theta_2)^{-1} & (1 - \theta_2^2)^{-1} \end{bmatrix}^{-1}. \end{aligned}$$

En el caso de los modelos estacionales multiplicativos, la matriz de información $I(\boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\Phi}, \boldsymbol{\Theta})$ se define de manera similar a (7.38) y las aproximaciones de las varianzas y covarianzas de los estimadores se obtienen calculando la inversa de dicha matriz de información.

Ejemplo 7.3.4: Sea el proceso $ARMA(0,1)x(0,1)_{12}$ definido como en el capítulo 5 mediante:

$$w_t = A_t - \theta_1 A_{t-1} - \theta_1 A_{t-12} + \theta_1 \theta_1 A_{t-13}$$

La matriz de información es (Box, Jenkins, & Reinsel, 1994):

$$I(\theta_1, \Theta_1) = n \begin{bmatrix} (1 - \theta_1^2)^{-1} & \theta_1^{11}(1 - \theta_1^{12}\Theta)^{-1} \\ \theta_1^{11}(1 - \theta_1^{12}\Theta)^{-1} & (1 - \Theta_1^2)^{-1} \end{bmatrix}.$$

Ejemplo 7.3.5: De acuerdo con el ejemplo 7.2.1, el modelo para la serie de registros de lluvia de Los Charrúas, estimado empleando máxima verosimilitud condicional, es

$$g(y_t) = 8,311275676 + 0,33805602 \cdot g(y_{t-1}) + A_t$$

Es claro que $\hat{\varphi}_1 = 0,33805602$. Por lo tanto,

$$Var[\hat{\varphi}_1] = \frac{1 - \varphi_1^2}{n} = \frac{1 - (0,33805602)^2}{228} = 0,003884728.$$

Este resultado permite estudiar la significancia del parámetro mediante el contraste (7.33):

$$H_0: \varphi_1 = 0 \text{ versus } H_1: \varphi_1 \neq 0$$

El estadístico de contraste es

$$z = \frac{0,33805602 - 0}{\sqrt{0,003884728}} = 5,4239$$

Teniendo en cuenta (7.36), obtener un valor tan extremo como $z = 5,4239$, considerando cierta la hipótesis nula, tiene probabilidad prácticamente nula (concretamente, $valor - p = 2,91563 \cdot 10^{-8}$). Es decir, que el valor- $p < 0,05$ (también se emplean 0,10 y 0,01). Por lo tanto, se rechaza la hipótesis nula. Para estudiar la significatividad de la constante, se debe testear el contraste (7.31). O bien, $H_0: \mu = 0 \text{ versus } H_1: \mu \neq 0$. Del ejemplo 7.2.1, $\hat{\mu} = 12,5558596$ y su varianza se calcula como en la sección 3.3:

$$Var[\hat{\mu}_y] \approx \frac{\sigma_y^2}{n} \left(1 + 2 \sum_{k=1}^{\infty} \rho_k \right) = \frac{59,931}{228} (1 + 2 \cdot 0,3360) = 0,439494,$$

donde del ejemplo 7.1.1, $\hat{\rho}_1 = 0,3360$ y $\hat{\sigma}_y^2 = \hat{\gamma}_0 = 59,931$. El estadístico de contraste es

$$z = \frac{12,5558596}{\sqrt{0,439494}} = 18,939564.$$

Nuevamente, la probabilidad de obtener un valor tan extremo como el observado considerando cierta la hipótesis nula es prácticamente cero, por lo que se rechaza la hipótesis nula. Luego, el modelo no presenta estructura adicional. Después de esta fase se puede estudiar las condiciones de causalidad e invertibilidad del modelo. Esto es, las raíces de los polinomios autorregresivo y de media móvil deben estar fuera del círculo unidad. Analizando el modelo propuesto, es sencillo escribirlo en términos del operador de retardos de la siguiente manera:

$$(1 - 0,33805602 \cdot B)g(y_t) = 8,311275676 + A_t$$

El polinomio autorregresivo en variable compleja $\varphi(z) = 1 - 0,33805602 \cdot z$ tiene raíz real $z = \frac{1}{0,33805602} = 2,958089609 > 1$.

7.4. Análisis de Residuos

En las secciones previas, se pudo observar que la construcción de un modelo para una serie temporal consiste en diferentes etapas. Puntualmente, identificación del tipo de modelo, estimación de los coeficientes y contrastes de significación de los parámetros estimados. Ahora, se procede a estudiar la adecuación del modelo formulado. Las innovaciones de un modelo *ARMA* pueden escribirse mediante la expresión:

$$A_t = \frac{\varphi(B)}{\theta(B)} w_t. \quad (7.44)$$

Por lo tanto, es claro que si el modelo *ARMA* propuesto para el proceso generador de la serie es el correcto, (7.44) debe comportarse exactamente como un ruido blanco, con autocorrelaciones nulas, media cero y varianza constante. Para comprobar esto, se emplean estimaciones de A_t a partir de las cuales se calculan los *residuos* del modelo⁵⁵:

$$\hat{A}_t = \frac{\hat{\varphi}(B)}{\hat{\theta}(B)} w_t. \quad (7.45)$$

En principio, es importante observar el gráfico de la serie (7.45) contra el tiempo para comprobar que la serie oscile alrededor de cero con una variabilidad constante. Se puede realizar el siguiente contraste a dos colas para probar que la media es cero:

$$H_0: E[A_t] = 0 \text{ versus } H_1: E[A_t] \neq 0$$

El estadístico es un t habitual:

$$t = \frac{E[\hat{A}_t]}{\sqrt{\frac{Var[\hat{A}_t]}{n}}} \sim N(0,1) \quad (7.46)$$

Para chequear si los residuos se distribuyen normalmente, se construye un histograma de los residuos estandarizados, esto es $\frac{\hat{A}_t}{\hat{\sigma}_A^2}$, y luego se realiza una comparación con la distribución normal empleando un test de bondad de ajuste Chi-cuadrado⁵⁶. La prueba de bondad de ajuste se usa para probar la hipótesis de que una distribución de frecuencias (en este caso la distribución de frecuencias de los residuos estandarizados), se ajusta a alguna distribución que se asevera (aquí se trata de la distribución normal). Se rechazará la hipótesis nula de normalidad cuando el valor del estadístico de contraste sea grande y la probabilidad de obtener un valor tan extremo como el que se observa sea pequeña. El rechazo de la hipótesis nula significa que no existe evidencia suficiente para concluir un buen ajuste con la distribución supuesta.

⁵⁵ La idea subyacente es que el modelo propuesto genera estimaciones de los valores reales observados de la serie. Así, la diferencia entre el valor observado y el estimado será un residuo.

⁵⁶ En el libro *Probabilidad y Estadística* de Morris DeGroot, (1988), Addison-Wesley Iberoamericana, Wilmington, EEUU., pág. 495 en adelante, se encuentra un riguroso y elegante detalle de los contrastes de bondad de ajuste.

Existen diversos test que permiten contrastar la hipótesis de normalidad de los residuos, uno de los cuales es el test de Doornik-Hansen⁵⁷. En su trabajo original, *An Omnibus Test for Univariate and Multivariate Normality*, (1994), los autores proponen un estadístico de contraste basado en una serie de transformaciones sobre los momentos de una variable aleatoria⁵⁸. Sea (x_1, \dots, x_n) una muestra de n observaciones de una variable aleatoria unidimensional con media μ y varianza σ^2 . Se define el momento central de orden i de la variable x como $\mu_i = E[(x - \mu)^i]$, (nótese que el momento de segundo orden es la varianza). Doornik y Hansen definen el coeficiente de asimetría, β_1 , y el coeficiente de Kurtosis, β_2 , mediante las siguientes expresiones⁵⁹:

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

Los estadísticos muestrales correspondientes son:

$$\sqrt{b_1} = \frac{m_3}{m_2^{3/2}}$$

$$b_2 = \frac{m_4}{m_2^2}$$

donde $m_i = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^i$ y \bar{x} es la media de la muestra. Posteriormente, definen dos transformaciones de simetría y kurtosis, z_1 y z_2 , respectivamente, y proponen un estadístico que sigue aproximadamente una distribución Chi-cuadrado con dos grados de libertad:

$$E_p = z_1^2 + z_2^2 \sim \chi_{(2)}^2 \quad (7.47)$$

Las transformaciones z_1 y z_2 se detallan en el Capítulo 9, Apéndice XXII, siguiendo lo expuesto por Doornik y Hansen. Si el *valor-p* asociado es inferior a un valor crítico convenientemente especificado, entonces se rechaza la hipótesis nula de normalidad y se concluye que la distribución de los residuos no es normal.

Siguiendo el análisis de adecuación del modelo, se calcula la *ACF* y la *PACF* de \hat{A}_t , con el propósito de comprobar que los residuos no siguen ningún patrón que revele correlación, es decir que los coeficientes de la función de autocorrelación y de la función de autocorrelación

⁵⁷ Otro comúnmente utilizado es el de Jarque-Bera.

⁵⁸ Este contraste de normalidad trabaja con los momentos de tercer y cuarto orden de la muestra porque éstos están asociados a la asimetría y a la kurtosis. La kurtosis es una medida del "apuntamiento" de la distribución con respecto al comportamiento de sus "colas" y se debe a K. Pearson (1906). Si la distribución de datos presenta algún tipo de asimetría y/o un pronunciado "apuntamiento" (exceso de kurtosis), o demasiado "aplanamiento", entonces es muy probable que no presente un buen ajuste con la distribución normal.

⁵⁹ Cuando la distribución es normal, $\beta_1 = 0$ y $\beta_2 = 3$. Cuando $\beta_2 = 3$ se dice que la distribución es mesokúrtica; si $\beta_2 < 3$, la distribución es platokúrtica y es leptokúrtica si $\beta_2 > 3$. Estos nombres provienen del griego y hacen referencia a los términos "medio", "ancho" y "delgado", respectivamente.

parcial no son estadísticamente significativos. Para probar esto, se contrasta la hipótesis ya planteada en la sección 3.3:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0 \quad (7.48)$$

versus la hipótesis alternativa de que alguno de estos coeficientes es diferente de cero. Este contraste se realiza mediante el estadístico de Ljung-Box (1978):

$$Q = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k} \quad (7.49)$$

que bajo la hipótesis nula tiene aproximadamente una distribución χ^2 con $h - (p + q)$ grados de libertad. (Wei, 2006)

Con respecto a la varianza de los residuos, si esta no es constante se dice que existe heterocedasticidad. Una forma sencilla de estudiar la heterocedasticidad consiste en realizar un gráfico rango-media. La construcción de este gráfico se hace seleccionando una cierta cantidad k de submuestras (por ejemplo $k = 12$ meses) y luego se calcula la media \bar{x}_k y la varianza σ_k^2 de cada submuestra. Si un ajuste por regresión del gráfico señala que a mayor media aumenta la varianza, entonces existe heterocedasticidad y el diagrama suele presentar forma de “embudo”. Por el contrario, un conjunto de datos homocedástico tendrá una regresión con pendiente aproximadamente cero, puesto que las variables \bar{x}_k y σ_k^2 no están correlacionadas linealmente y el diagrama no sigue ningún patrón que refleje alguna forma. En este sentido, puede emplearse el test habitual que permite contrastar $H_0: \alpha = 0$ versus $H_1: \alpha \neq 0$ donde α es la pendiente de la recta de regresión $\sigma_k^2 = \alpha \bar{x}_k + \text{constante}$, a saber:

$$t = \frac{\hat{\alpha} - 0}{SE(\hat{\alpha})} \quad (7.50)$$

En este caso, $SE(\hat{\alpha})$ es el error estándar de la estimación puntual del parámetro α , que se

calcula como $SE(\hat{\alpha}) = \frac{s}{\sqrt{(\bar{x}_k - E[\bar{x}_k])^2}}$ donde $s = \sqrt{\frac{\sum(\sigma_k^2 - \hat{\sigma}_{k_i}^2)^2}{n-2}}$. La suma $\sum(\sigma_k^2 - \hat{\sigma}_{k_i}^2)^2$ se llama suma de cuadrados residual, y $\hat{\sigma}_{k_i}^2$ es el valor estimado predicho por la recta. El estadístico (7.50) tiene una distribución $t - Student$ con $n - 2$ grados de libertad.⁶⁰

Ejemplo 7.4.1: En ejemplos anteriores se estudió el modelo de la serie de registros de lluvia de Los Charrúas que, para recordar, es un $ARIMA(1,0,0)x(0,0,0)_{12}$ dado por la ecuación:

$$g(y_t) = 8,311275676 + 0,33805602 \cdot g(y_{t-1}) + A_t.$$

A continuación se muestran las primeras diez observaciones de la serie (a modo ilustrativo), los valores estimados con el modelo propuesto y los residuos:

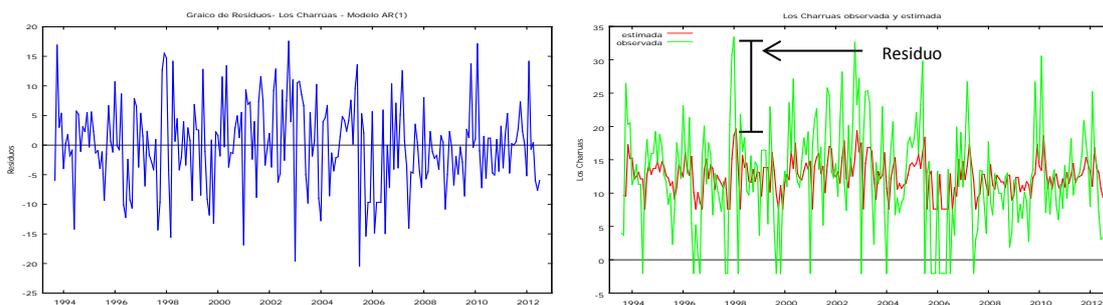
⁶⁰ A este estudio se puede agregar el coeficiente de determinación que representa la proporción de la variación total en los valores de la variable dependiente que puede ser explicada por su relación lineal con la variable independiente. Diversos libros de estadística inferencial abordan con mayor o menor grado de profundidad la teoría de regresión lineal.

Tabla 7.1

Fecha	Valores Observados	Valores Estimados	Residuos
sep-93	3,7048186	9,659543	-5,9547244
oct-93	26,4495121	9,563712	16,8858001
nov-93	20,273428	17,252692	3,02073598
dic-93	20,5409716	15,16483	5,37614156
ene-94	11,2939202	15,255275	-3,96135476
feb-94	12,371118	12,129253	0,24186495
mar-94	14,2673033	12,493407	1,77389631
abr-94	11,2939202	13,134423	-1,84050276
may-94	11,2939202	12,129253	-0,83533276
jun-94	-2,04937058	12,129253	-14,1786236

En el gráfico de los residuos del modelo, que se muestra en la siguiente figura (izquierda), puede verse que éstos oscilan alrededor de cero, lo que sugiere que la serie de residuos tiene media cero. A la derecha se ve la serie observada versus la estimada por el modelo.

Gráfico 7.1



El estadístico para testear la significatividad de la media es:

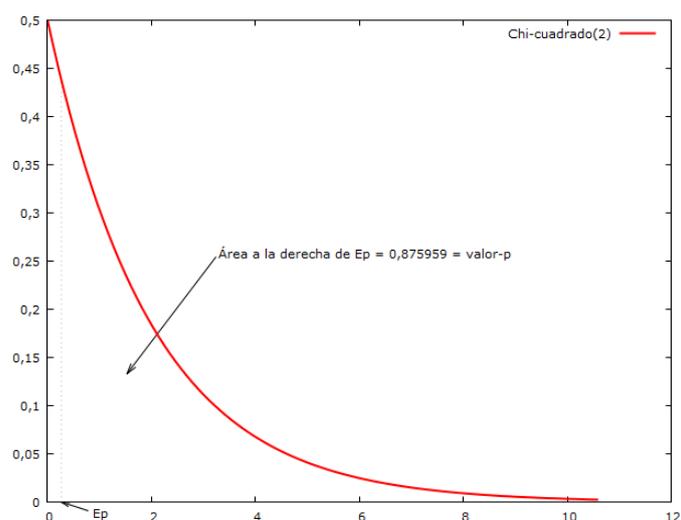
$$t = \frac{4,605 \cdot 10^{-16}}{\frac{7,2826}{\sqrt{228}}} = 9,22808 \cdot 10^{-16}$$

Es claro que la probabilidad de obtener un valor como el observado considerando cierta la hipótesis nula, $H_0: E[A_t] = 0$, es 1. Por lo que no existe evidencia suficiente para rechazar H_0 .

El siguiente paso consiste en contrastar la normalidad de los residuos. El estadístico de Doornik-Hansen observado es

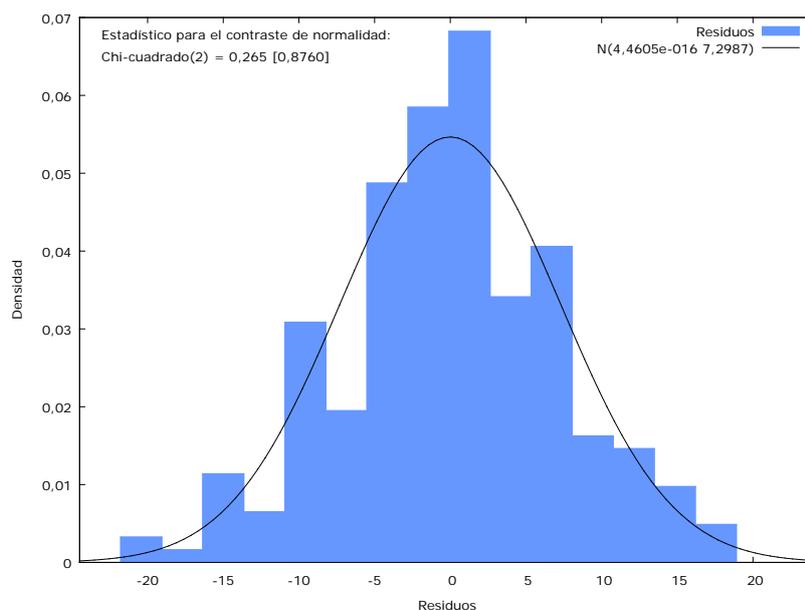
$$E_p = 0,264871$$

Gráfico 7.2



El mismo se obtuvo empleando las herramientas de Gretl aunque puede realizarse manualmente utilizando las fórmulas dadas en el Apéndice XXII. En la figura superior (Gráfico 7.2) se muestra el gráfico de la distribución $\chi^2_{(2)}$, la posición de E_p , muy cercano a cero, y el *valor-p* asociado que es el área bajo la curva a la derecha de E_p . Puesto que el *valor-p* = $0,875959 = P(\chi > E_p)$ es cercano a 1, se concluye que no existe evidencia suficiente para rechazar la hipótesis nula de normalidad. En otras palabras, la probabilidad de obtener un valor de E_p como el observado considerando cierta H_0 , es muy elevada, por lo que se concluye que los residuos se distribuyen normalmente. En la siguiente figura se observa el histograma de los residuos y el gráfico de la distribución normal superpuesta obtenido con Gretl.

Gráfico 7.3

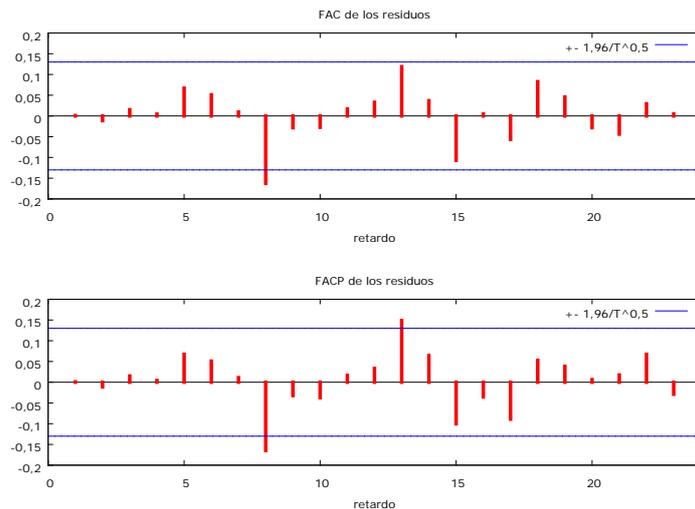


Seguidamente, se realiza el contraste que permite comprobar si la serie de los residuos se comporta como un ruido blanco. A continuación se muestra la salida de Gretl para la función de autocorrelación de los residuos y los correlogramas correspondientes. Las dos últimas columnas corresponden al valor del estadístico Q de Ljung-Box y el *valor-p* asociado⁶¹.

Tabla 7.2
Función de autocorrelación para Residuos

RETARDO	FAC	FACP	Estad-Q. [valor p]
1	0,0011	0,0011	0,0003 [0,987]
2	-0,0113	-0,0113	0,0299 [0,985]
3	0,0151	0,0151	0,0827 [0,994]
4	0,0045	0,0043	0,0874 [0,999]
5	0,0669	0,0673	1,1366 [0,951]
6	0,0510	0,0510	1,7488 [0,941]
7	0,0095	0,0111	1,7699 [0,971]
8	-0,1626 **	-0,1646 **	8,0419 [0,429]
9	-0,0285	-0,0322	8,2350 [0,511]
10	-0,0274	-0,0372	8,4149 [0,588]
11	0,0169	0,0165	8,4835 [0,669]
12	0,0330	0,0334	8,7472 [0,724]
13	0,1188 *	0,1489 **	12,1779 [0,513]
14	0,0370	0,0643	12,5117 [0,565]
15	-0,1071	-0,1003	15,3245 [0,428]
16	0,0049	-0,0354	15,3304 [0,501]
17	-0,0565	-0,0892	16,1209 [0,515]
18	0,0828	0,0528	17,8264 [0,467]
19	0,0459	0,0385	18,3525 [0,499]
20	-0,0282	0,0063	18,5525 [0,551]
21	-0,0438	0,0175	19,0373 [0,583]
22	0,0295	0,0673	19,2581 [0,629]
23	0,0048	-0,0290	19,2639 [0,686]

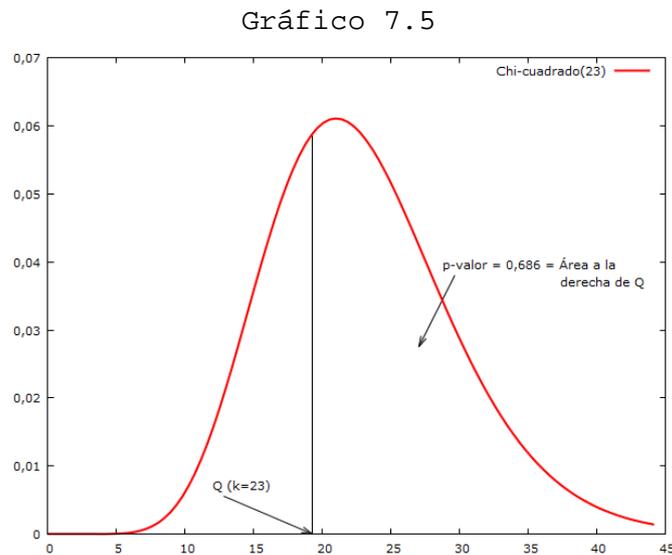
Gráfico 7.4



En todos los casos, la probabilidad de obtener un valor de Q como el observado considerando cierta la hipótesis nula permite concluir que no existe evidencia suficiente para rechazar la

⁶¹ Gretl utiliza la distribución Chi-Cuadrado con h grados de libertad. Es decir, no emplea $h - (p + q)$ grados de libertad como fue propuesto en (7.49).

hipótesis nula. Por lo que la serie de los residuos se comporta como un ruido blanco. El siguiente gráfico muestra la distribución Chi-cuadrado con 23 grados de libertad, la posición del estadístico Q para el retardo $k = 23$ y el valor- p asociado.



Por último, el gráfico rango-media de los residuos no revela que en la medida en que aumenta la media, también lo haga la varianza. Gretl ofrece una salida con la selección automática de las submuestras, el rango y la media de cada submuestra y el contraste de la pendiente:

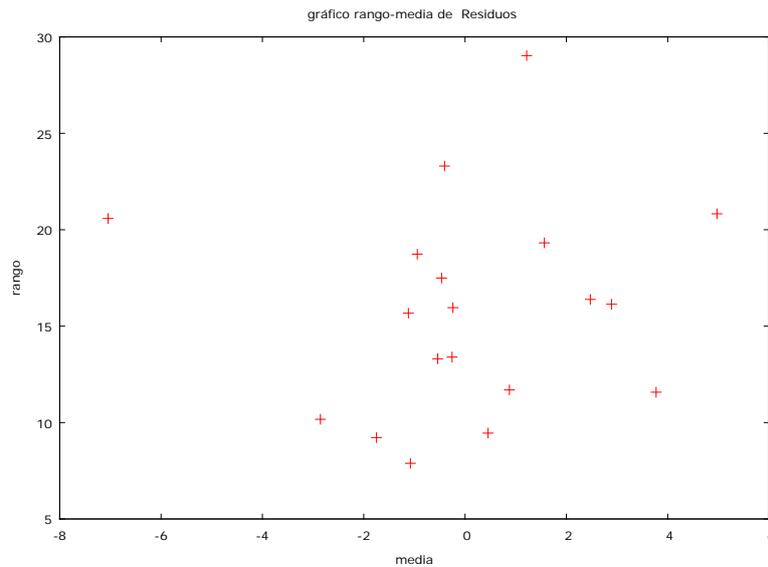
Tabla 7.3

Estadísticos de rango-media para Residuos utilizando 19 submuestras de tamaño 12

	rango	media
1993:09 - 1994:08	11,6921	0,876187
1994:09 - 1995:08	9,45332	0,454139
1995:09 - 1996:08	18,7350	-0,942992
1996:09 - 1997:08	13,3956	-0,256309
1997:09 - 1998:08	29,0305	1,21641
1998:09 - 1999:08	15,9594	-0,240091
1999:09 - 2000:08	23,3050	-0,404493
2000:09 - 2001:08	16,1360	2,89071
2001:09 - 2002:08	16,3852	2,47239
2002:09 - 2003:08	20,8260	4,97241
2003:09 - 2004:08	15,6712	-1,11844
2004:09 - 2005:08	11,5746	3,77084
2005:09 - 2006:08	20,5845	-7,04514
2006:09 - 2007:08	17,4954	-0,462352
2007:09 - 2008:08	9,22095	-1,74822
2008:09 - 2009:08	10,1696	-2,85487
2009:09 - 2010:08	19,3195	1,56462
2010:09 - 2011:08	7,88920	-1,07826
2011:09 - 2012:07	13,3084	-0,539442

Pendiente de 'rango' con respecto a 'media' = 0,198782 el valor p para H_0 : pendiente = 0 es 0,695983.

Gráfico 7.5



El estadístico (7.50) es:

$$t = \frac{0,198792 - 0}{0,5} = 0,397584$$

donde $\hat{\alpha} = 0,198792$ y $SE(\hat{\alpha}) = 0,5$ y sigue una distribución t con 225 grados de libertad. El *valor-p* a dos colas señala que no existe evidencia suficiente para rechazar la hipótesis nula.

7.5. Criterios de Información

En el análisis de series temporales empleando la metodología de modelización *ARIMA* es común hallar más de un modelo que recoge adecuadamente la estructura del proceso generador de la serie. Para seleccionar de un conjunto de modelos propuestos el que mejor represente la dinámica de los datos se han propuesto los llamados criterios de selección de modelos⁶².

El criterio más conocido es, tal vez, el de Akaike. Se asume, en principio, que un modelo estadístico de M parámetros es apropiado para explicar la estructura del conjunto de datos que conforman la serie. Con el propósito de evaluar la calidad del modelo adecuado, Akaike (1973) introduce un criterio de información conocido como criterio AIC (Akaike's information criterion) que se define como

$$AIC(M) = -2 \ln L(\hat{h}, \hat{\sigma}_A^2) + 2M, \quad (7.51)$$

donde $L(\hat{h}, \hat{\sigma}_A^2)$ es la función de verosimilitud evaluada en las estimaciones de los parámetros y M es el número de parámetros en el modelo. De (7.29), para un modelo *ARMA* que ha sido estimado con n observaciones, resulta:

⁶² Estos criterios ofrecen una estimación de la "cantidad" de información que se pierde al emplear un modelo en particular, por lo que cuanto menor valor alcancen mejor será el modelo.

$$\ln L(h, \sigma_A^2) = -\frac{n}{2} \ln(2\pi\sigma_A^2) - \frac{S(h)}{2\sigma_A^2}$$

Recordando que $\hat{\sigma}_A^2 = \frac{S(\hat{h})}{n}$ se tiene:

$$\ln L(\hat{h}, \hat{\sigma}_A^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \hat{\sigma}_A^2 - \frac{n\hat{\sigma}_A^2}{\hat{\sigma}_A^2}$$

Es decir:

$$\ln L(\hat{h}, \hat{\sigma}_A^2) = -\frac{n}{2} \ln \hat{\sigma}_A^2 - \frac{n}{2} (1 + \ln 2\pi) \quad (7.52)$$

En (7.51):

$$AIC(M) = -2 \cdot \left[-\frac{n}{2} \ln \hat{\sigma}_A^2 - \frac{n}{2} (1 + \ln 2\pi) \right] + 2M$$

Como el segundo miembro de (7.52) es una constante, el criterio se reduce a:

$$AIC(M) = n \ln \hat{\sigma}_A^2 + 2M \quad (7.53)$$

Puesto que el valor de M depende de p y de q , se concluye que un modelo es tanto más adecuado conforme al valor de M que haga más pequeña (7.53).

Estudios posteriores, mostraron que el criterio AIC tiende a sobreestimar el orden autorregresivo, por lo que aparecieron otros criterios con análoga finalidad al de Akaike. Schwartz (1978) sugiere un criterio bayesiano para seleccionar un modelo. Este criterio se conoce con el nombre SBC (Schwartz's Bayesian Criterion) y se define mediante la expresión:

$$SBC(M) = n \ln \hat{\sigma}_A^2 + M \ln n. \quad (7.54)$$

En (7.54), M es el número de parámetros en el modelo, $\hat{\sigma}_A^2$ es el estimador máximo verosímil de σ_A^2 y n es el número de observaciones que es igual al número de residuos que pueden calcularse con la serie. (Wei, 2006)

Otro criterio empleado frecuentemente es el de Hannan y Quinn (1979) el cual se define mediante⁶³:

$$HQ(M) = n \ln \hat{\sigma}_A^2 + 2Mc \ln(\ln n) \quad \text{con } c > 1 \quad (7.55)$$

Estudios de simulación mostraron que estos criterios son más apropiados que el criterio AIC cuando se tienen muestras grandes. Es importante remarcar que un modelo es tanto más adecuado cuanto más reducidos sean los valores de estos criterios, y señalar que la selección

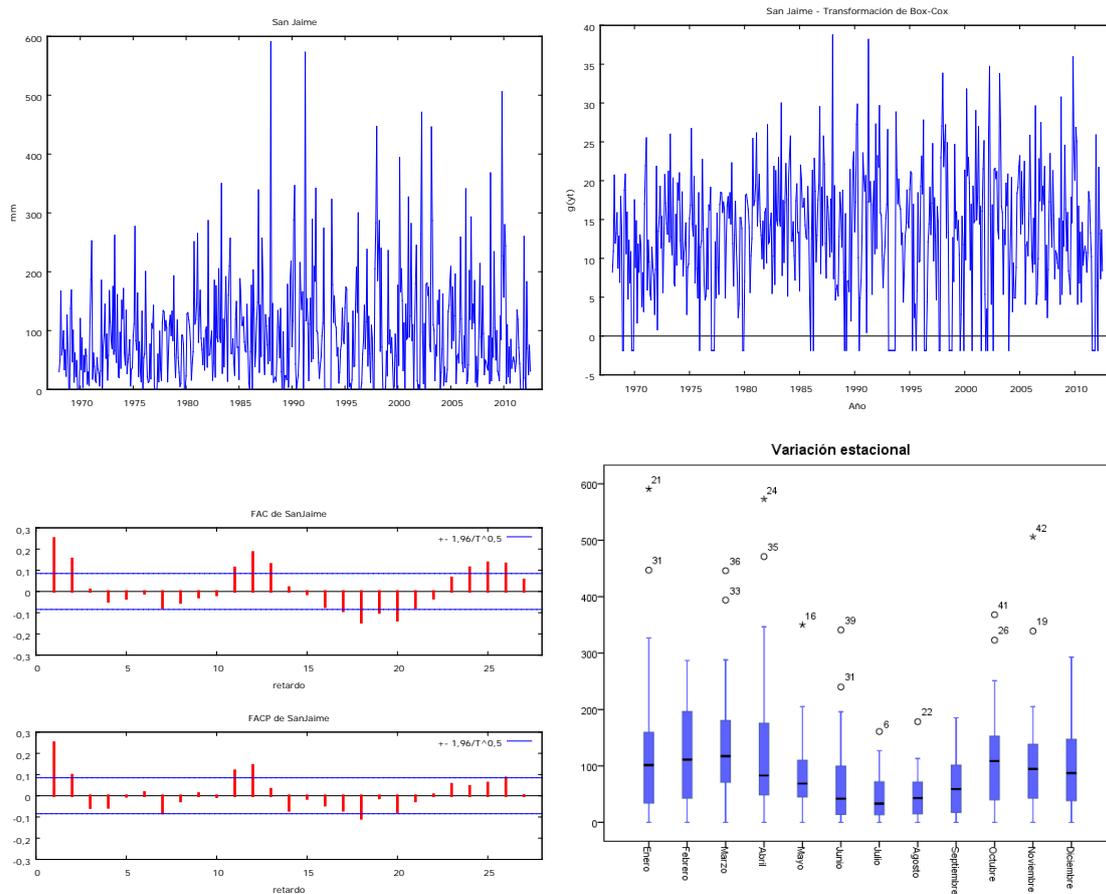
⁶³ La propuesta se basa en que para muestras finitas el criterio AIC tiende a sobrestimar el modelo mientras que el SBC tiende a subestimarlos por lo que la variante $2Mc \ln(\ln n)$ busca corregir estas irregularidades.

de un modelo no debe basarse exclusivamente en estos criterios. No obstante, son medidas muy útiles para acompañar el procedimiento de selección de órdenes y validación.

Ejemplo 7.5.1: Seguidamente se estudia cuál de los modelos propuestos para la serie procedente de la Localidad de San Jaime es el que mejor explica la estructura del proceso generador de los datos.

1. *Representaciones gráficas.* A continuación se muestra el gráfico de la serie y de la serie modificada por la transformación de Box-Cox. Además, se observa el correlograma y el diagrama de cajas para el estudio de la variación estacional.

Gráfico 7.6



2. *Modelos.* De acuerdo a lo discutido en el Capítulo 6, se propuso los siguientes modelos candidatos a explicar la estructura del proceso generador de los datos:

Modelo 1: $ARIMA(2,0,1)x(1,1,1)_{12}$

Modelo 2: $ARIMA(2,0,1)x(1,0,1)_{12}$

Modelo 3: $ARIMA(0,0,2)x(1,0,1)_{12}$

3. *Estimación de los parámetros.* A continuación, las salidas de Gretl para cada modelo propuesto, donde se muestran las estimaciones de los coeficientes, la desviación estándar, el estadístico para los contrastes de significación y el valor-p asociado:

Tabla 7.4

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	0,0457706	0,0505492	0,9055	0,36522	
phi_1	0,403661	0,247573	1,6305	0,10300	
phi_2	0,0646739	0,0649717	0,9954	0,31953	
Phi_1	0,00265931	0,0460862	0,0577	0,95399	
theta_1	-0,224068	0,247163	-0,9066	0,36464	
Theta_1	-0,840634	0,0264737	-31,7535	<0,00001	***

Tabla 7.5

Modelo 2					
	Coeficiente	Desv. Típica	z	Valor p	
const	7,24863	2,84241	2,5502	0,01077	**
phi_1	0,074321	0,299795	0,2479	0,80421	
phi_2	0,154501	0,0756368	2,0427	0,04109	**
Phi_1	0,291909	0,172396	1,6932	0,09041	*
theta_1	0,121272	0,304132	0,3987	0,69008	
Theta_1	-0,122814	0,181344	-0,6772	0,49825	

Tabla 7.6

Modelo 3					
	Coeficiente	Desv. Típica	z	Valor p	
const	6,87485	2,2544	3,0495	0,00229	***
Phi_1	0,48298	0,168087	2,8734	0,00406	***
theta_1	0,186048	0,0395847	4,7000	<0,00001	***
theta_2	0,179255	0,0411064	4,3608	0,00001	***
Theta_1	-0,323666	0,183584	-1,7630	0,07789	*

El Modelo 1 tiene sólo el parámetro de media móvil estacional significativo, el Modelo 2 también aparece sobreparametrizado y el 3 tiene el parámetro de media móvil estacional apenas significativo.

4. *Causalidad e Invertibilidad.* En el capítulo 4 se mostró que las raíces de los polinomios autorregresivo y de media móvil deben estar fuera del círculo unidad para que se cumplan las condiciones de causalidad e invertibilidad. A continuación, las salidas de Gretl:

Tabla 7.7

Modelo 1					
		Real	Imaginaria	Módulo	Frecuencia
AR	Raíz 1	1,8993	0,0000	1,8993	0,0000
	Raíz 2	-8,1408	0,0000	8,1408	0,5000
AR (estacional)					
	Raíz 1	376,0369	0,0000	376,0369	0,0000
MA					
	Raíz 1	4,4629	0,0000	4,4629	0,0000
MA (estacional)					
	Raíz 1	1,1896	0,0000	1,1896	0,0000

Tabla 7.8

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	2,3149	0,0000	2,3149	0,0000
	Raíz 2	-2,7960	0,0000	2,7960	0,5000
AR (estacional)					
	Raíz 1	3,4257	0,0000	3,4257	0,0000
MA					
	Raíz 1	-8,2459	0,0000	8,2459	0,5000
MA (estacional)					
	Raíz 1	8,1424	0,0000	8,1424	0,0000

Tabla 7.9

Modelo 3					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	2,0705	0,0000	2,0705	0,0000
MA					
	Raíz 1	-0,5189	-2,3042	2,3619	-0,2853
	Raíz 2	-0,5189	2,3042	2,3619	0,2853
MA (estacional)					
	Raíz 1	3,0896	0,0000	3,0896	0,0000

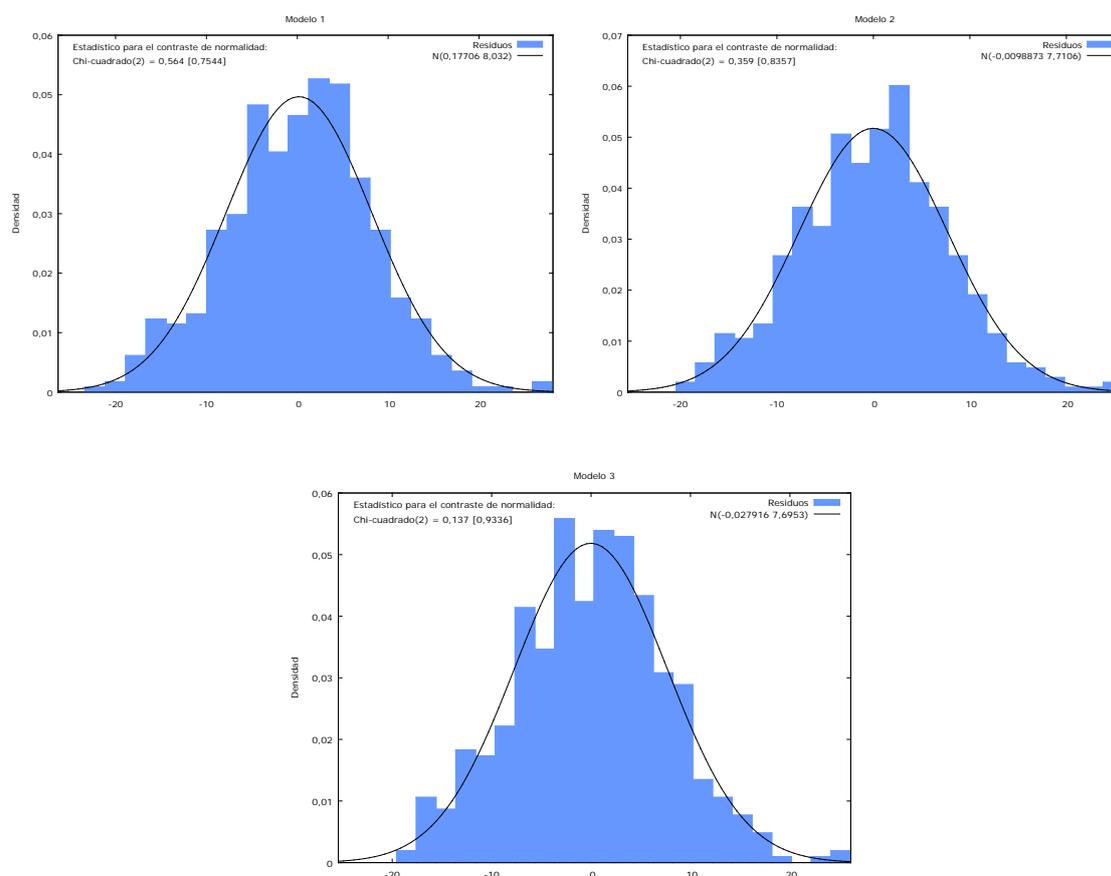
Observación: Los tres modelos cumplen con las condiciones de causalidad e invertibilidad (ver la columna *Módulo*).

5. *Análisis de Residuos.* A continuación se muestra el ajuste de los residuos de cada modelo con una distribución normal que tiene por media la media de los residuos y por varianza la varianza de los residuos, acompañada por el valor del estadístico de Doornik-Hansen y el valor-p asociado.

Tabla 7.10

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	0,564	0,75443
Modelo 2	0,359	0,83568
Modelo 3	0,137	0,93359

Gráfico 7.7



Los residuos de los modelos deben comportarse como un ruido blanco, es decir, deben tener media cero, varianza constante y covarianza nula.

Tabla 7.11

	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,17706	7,9924	0,499807	0,6174
Modelo 2	-0,0098873	7,6734	-0,0290702	0,9768
Modelo 3	-0,027916	7,6657	-0,08216	0,9346

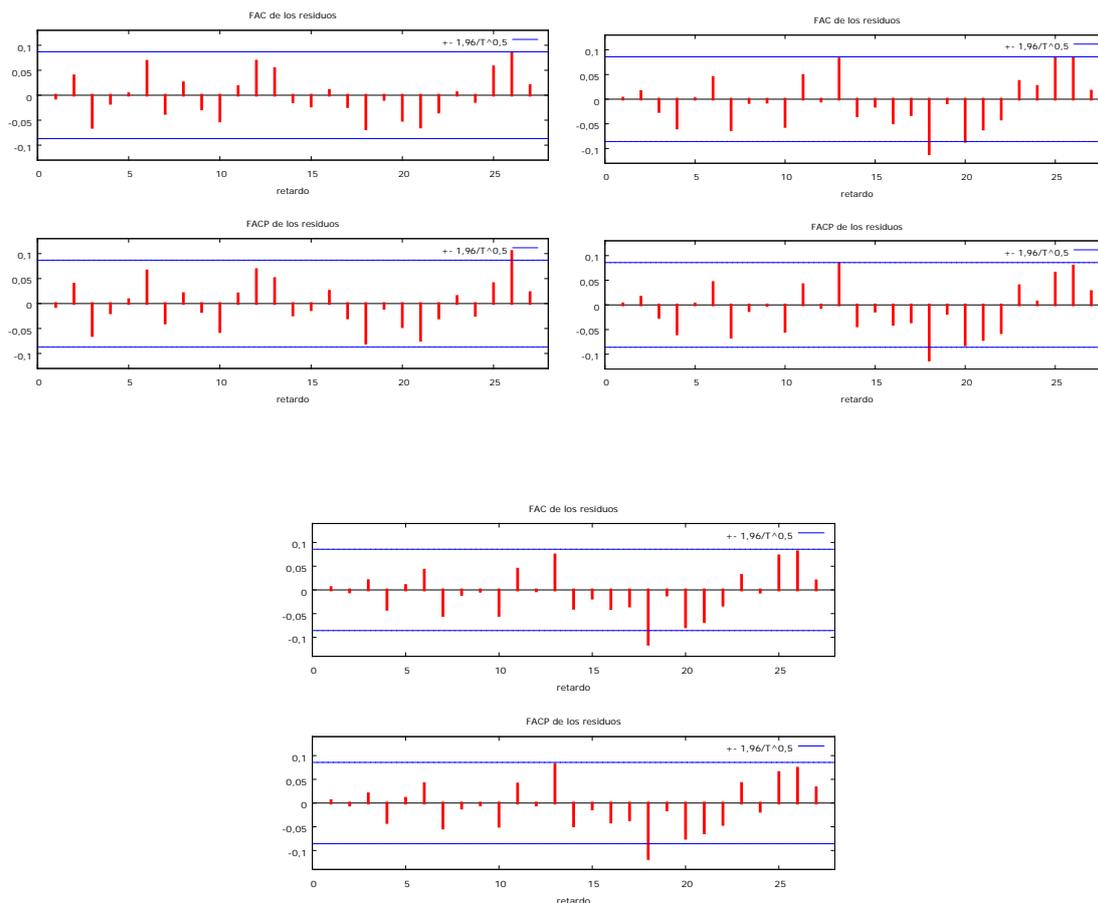
En la primera columna de la Tabla 7.11 se observa la media de la serie de los residuos, seguida de la desviación típica, el estadístico de contraste y el valor-p asociado. En los tres modelos, el valor-p indica que no existe evidencia suficiente para rechazar la hipótesis nula $\mu_A = 0$.

Comprobar la nulidad de la covarianza de los residuos es equivalente a comprobar que la ACF es nula. La figura inferior muestra los correlogramas de la ACF residual asociada a cada modelo. En ninguno de los tres casos puede rechazarse la hipótesis de correlaciones nulas ni en los retardos regulares ni en los estacionales. También se anexa una tabla con el estadístico de Ljung-Box para el retardo estacional $k = 36$:

Tabla 7.12

	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0032	48,7495	0,076
Modelo 2	0,0559	58,4947	0,01
Modelo 3	0,0299	50,946	0,051

Gráfico 7.8

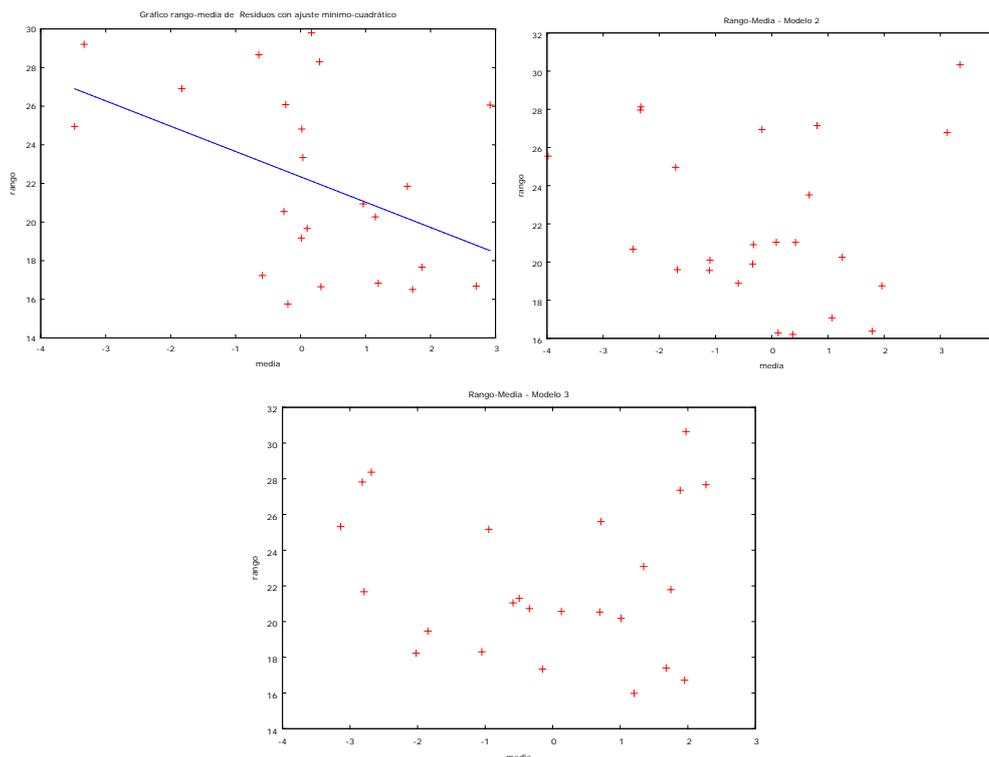


Para el estudio de la homocedasticidad de los residuos, se muestra a continuación los gráficos rango-media y una tabla que contiene la pendiente de la recta de regresión que mejor ajusta a los puntos del diagrama y el valor-p asociado para la hipótesis nula de pendiente cero:

Tabla 7.13

	Pendiente estimada	Valor-p
Modelo 1	-1,31061	0,0344695
Modelo 2	-0,167943	0,74144
Modelo 3	-0,171226	0,737523

Gráfico 7.9



En el Modelo 1 se observa que el gráfico no presenta un comportamiento aleatorio como en los otros casos. Además, para este modelo también se rechaza la hipótesis nula de pendiente cero para la recta de regresión.

6. *Correlación entre los coeficientes.* Un análisis importante en la validación de un modelo es el cálculo de la matriz de correlación entre los parámetros a los efectos de detectar posible colinealidad⁶⁴. Valores altos para el coeficiente de correlación puede significar la existencia de factores comunes en el modelo. A continuación se muestra la matriz de covarianza de cada modelo (obtenida con Gretl) y el cálculo del coeficiente de correlación entre los estimadores de los parámetros:

Tabla 7.14

Matriz de Covarianzas del Modelo 1						
const	phi_1	phi_2	Phi_1	theta_1	Theta_1	
0,00255522	-0,00279668	0,000399626	-5,76885e-005	0,00302163	-8,69880e-005	const
	0,0612925	-0,0131990	-0,000970048	-0,0603688	9,33708e-005	phi_1
		0,00422133	0,000477603	0,0125796	2,98037e-005	phi_2
			0,00212394	0,000753801	-0,000426087	Phi_1
				0,0610897	3,19277e-005	theta_1
					0,000700858	Theta_1

$$Corr(\varphi_1, \varphi_2) = \frac{Cov(\varphi_1, \varphi_2)}{\sqrt{Var[\varphi_1] \cdot Var[\varphi_2]}} = \frac{-0,013199}{\sqrt{(0,247573)^2 \cdot (0,0649717)^2}} = -0,820566$$

⁶⁴ La colinealidad aparece cuando hay correlación lineal alta entre los parámetros de un modelo. Esto significa básicamente que los parámetros no son independientes y, por lo tanto, hay redundancia entre ellos. Una posible solución es eliminar del modelo los parámetros no significativos.

$$\text{Corr}(\varphi_1, \Phi_1) = \frac{\text{Cov}(\varphi_1, \Phi_1)}{\sqrt{\text{Var}[\varphi_1] \cdot \text{Var}[\Phi_1]}} = \frac{-0,000970048}{\sqrt{(0,247573)^2 \cdot (0,0460862)^2}} = -0,0850196$$

$$\text{Corr}(\varphi_1, \theta_1) = \frac{\text{Cov}(\varphi_1, \theta_1)}{\sqrt{\text{Var}[\varphi_1] \cdot \text{Var}[\theta_1]}} = \frac{-0,0603688}{\sqrt{(0,247573)^2 \cdot (0,247163)^2}} = -0,98656523$$

$$\text{Corr}(\varphi_1, \Theta_1) = \frac{\text{Cov}(\varphi_1, \Theta_1)}{\sqrt{\text{Var}[\varphi_1] \cdot \text{Var}[\Theta_1]}} = \frac{0,0000933708}{\sqrt{(0,247573)^2 \cdot (0,0264737)^2}} = 0,01424601$$

$$\text{Corr}(\varphi_2, \Phi_1) = \frac{\text{Cov}(\varphi_2, \Phi_1)}{\sqrt{\text{Var}[\varphi_2] \cdot \text{Var}[\Phi_1]}} = \frac{0,000477603}{\sqrt{(0,0649717)^2 \cdot (0,0460862)^2}} = 0,15950412$$

$$\text{Corr}(\varphi_2, \theta_1) = \frac{\text{Cov}(\varphi_2, \theta_1)}{\sqrt{\text{Var}[\varphi_2] \cdot \text{Var}[\theta_1]}} = \frac{0,000477603}{\sqrt{(0,0649717)^2 \cdot (0,247163)^2}} = 0,78335594$$

$$\text{Corr}(\varphi_2, \Theta_1) = \frac{\text{Cov}(\varphi_2, \Theta_1)}{\sqrt{\text{Var}[\varphi_2] \cdot \text{Var}[\Theta_1]}} = \frac{0,0000298037}{\sqrt{(0,0649717)^2 \cdot (0,0264737)^2}} = 0,01732732$$

$$\text{Corr}(\Phi_1, \theta_1) = \frac{\text{Cov}(\Phi_1, \theta_1)}{\sqrt{\text{Var}[\Phi_1] \cdot \text{Var}[\theta_1]}} = \frac{0,000753801}{\sqrt{(0,0460862)^2 \cdot (0,247163)^2}} = 0,06617628$$

$$\text{Corr}(\Phi_1, \Theta_1) = \frac{\text{Cov}(\Phi_1, \Theta_1)}{\sqrt{\text{Var}[\Phi_1] \cdot \text{Var}[\Theta_1]}} = \frac{-0,000426087}{\sqrt{(0,0460862)^2 \cdot (0,0264737)^2}} = -0,34923096$$

$$\text{Corr}(\theta_1, \Theta_1) = \frac{\text{Cov}(\theta_1, \Theta_1)}{\sqrt{\text{Var}[\theta_1] \cdot \text{Var}[\Theta_1]}} = \frac{0,000700858}{\sqrt{(0,247163)^2 \cdot (0,0264737)^2}} = 0,10711047$$

Tabla 7.15

Matriz de Covarianzas del Modelo 2						
const	phi_1	phi_2	Phi_1	theta_1	Theta_1	
8,07932	-0,648533	0,120756	-0,296506	0,644739	0,300651	const
	0,0898770	-0,0190966	-0,00128399	-0,0903801	0,00100732	phi_1
		0,00572093	0,000449505	0,0193655	-0,000131488	phi_2
			0,0297205	0,00193714	-0,0302941	Phi_1
				0,0924960	-0,00183264	theta_1
					0,0328857	Theta_1

Las correlaciones entre los parámetros estimados del modelo pueden organizarse en una matriz:

Tabla 7.16

Matriz de Correlaciones del Modelo 2					
	φ_1	φ_2	Φ_1	θ_1	Θ_1
φ_1	1	-0,84216758	-0,02484335	-0,99125711	0,01852848
φ_2		1	0,03447261	0,84184755	-0,00571598
Φ_1			1	0,03694636	-0,96900871
θ_1				1	-0,03322859
Θ_1					1

Tabla 7.17

Matriz de Covarianzas del Modelo 3					
const	Phi_1	theta_1	theta_2	Theta_1	
5,08232	-0,375223	-0,0102386	0,00366854	0,397462	const
	0,0282534	0,000819123	-0,000287650	-0,0300143	Phi_1
		0,00156695	0,000239132	-0,000943766	theta_1
			0,00168974	0,000539536	theta_2
				0,0337033	Theta_1

Tabla 7.18

Matriz de correlaciones del Modelo 3				
	Φ_1	θ_1	θ_2	Θ_1
Φ_1	1	0,12310839	-0,04163138	-0,97265599
θ_1		1	0,14696059	-0,129868
θ_2			1	0,07149508
Θ_1				1

Los tres modelos presentan correlaciones altas aunque en muchos casos estas corresponden a estimaciones paramétricas no significativas. Si se escriben los modelos de acuerdo con sus respectivas ecuaciones, el modelo 3 tiene muy alta la correlación entre sus parámetros estacionales, aunque Θ_1 no es significativo al 5% (lo es al 10%) y podría suprimirse del modelo:

Modelo 1: $(1 - B^{12})g(y_t) = (1 - 0,840634B^{12})A_t$

Modelo 2: $(1 - 0,154501B^2)(1 - 0,291909B^{12})g(y_t) = 7,24863 + A_t$

Modelo 3:

$$(1 - 0,48298B^{12})g(y_t) = 6,87485 + (1 + 0,186048B + 0,179255B^2)(1 - 0,323666B^{12})A_t$$

7. *Criterios de selección.* El estudio precedente muestra que los modelos satisfacen las condiciones de causalidad e invertibilidad. También superan las pruebas del análisis residual excepto la prueba de homocedasticidad que no está claro que la cumpla el Modelo 1 y, en general, las correlaciones entre los parámetros significativos no son altas. En este punto, es donde los criterios de información pueden traer claridad a las decisiones. A continuación se muestran los resultados obtenidos con Gretl, de los tres criterios para cada modelo:

Tabla 7.19

	AIC	SBC	HQ
Modelo 1	3573,635	3603,262	3585,252
Modelo 2	3614,884	3644,674	3626,553
Modelo 3	3625,668	3651,226	3635,678

De los modelos propuestos el que tiene los valores más bajos es el Modelo 1, por lo que sería el que mejor recoge la estructura del proceso, aunque como ya se mencionó no está claro que los residuos sean homocedásticos. Este modelo se propuso teniendo en cuenta los resultados del test de Canova y Hansen y las características del correlograma de la ACF. En realidad, lo conveniente es elegir aquel modelo que supere todas las pruebas detalladas precedentemente. Los modelos 2 y 3 no incluyen orden de diferenciación estacional porque el test HEGY no detecta la presencia de raíces unitarias. De éstos, el ajuste residual tiene diferencias sutiles y el análisis de colinealidad revela correlación alta entre parámetros que pueden suprimirse. Por lo tanto, la decisión final pesa sobre la información que aportan los criterios de selección que en este caso muestran los valores más bajos para el Modelo 2. Puede concluirse que tanto el Modelo 1 como el 2 explican la estructura del proceso, al margen de las diferencias en los resultados de los test de raíces unitarias.

7.6. Modelos propuestos para las series de registros de lluvia mensual acumulada de Entre Ríos

En la Tabla 6.3 se propuso los posibles modelos candidatos a explicar la estructura de los procesos que generan los datos. Como en el ejemplo 7.5.1, se estudió el ajuste de los mismos y a continuación se muestra los modelos seleccionados para cada estación de registro. El análisis detallado de cada modelo y la fundamentación para su elección puede consultarse en los respectivos apéndices del Capítulo 9.

Los Charrúas

$$ARIMA(1,0,0)x(0,0,0)_{12}$$

$$(1 - 0,33805602B)g(y_t) = 8,311275676 + A_t$$

Lucas Sur

$$ARIMA(0,0,2)x(1,0,0)_{12}$$

$$(1 - 0,174254 \cdot B^{12})g(y_t) = 9,9792 + (1 + 0,215065 \cdot B + 0,235373 \cdot B^2)A_t$$

Octavo Distrito

Modelo 2:

$$ARIMA(0,0,2)x(1,1,1)_{12}$$

$$(1 + 0,140214B^{12})(1 - B^{12})g(y_t) = 0,135715 + (1 + 0,00518022B + 0,253185B^2)(1 - 0,701589B^{12})A_t$$

Eliminando los parámetros no significativos:

$$(1 + 0,140214B^{12})(1 - B^{12})g(y_t) = (1 + 0,253185B^2)(1 - 0,701589B^{12})A_t$$

Modelo 4:

$$\begin{aligned} & ARIMA(2,0,0)x(1,1,1)_{12} \\ & (1 + 0,111677B^{12})(1 - 0,0136424B - 0,25624B^2)(1 - B^{12})g(y_t) \\ & = 0,0773948 + (1 - 0,709892B^{12})A_t \end{aligned}$$

Si se eliminan los parámetros no significativos:

$$(1 - 0,25624B^2)(1 - B^{12})g(y_t) = (1 - 0,709892B^{12})A_t$$

Puente de Hierro

$$\begin{aligned} & ARIMA(0,0,5)x(1,0,1)_{12} \\ & (1 - 0,535312B^{12})g(y_t) \\ & = 5,86861 \\ & + (1 + 0,112918B + 0,208105B^2 - 0,0644063B^3 - 0,136021B^4 \\ & - 0,191963B^5)(1 - 0,410155B^{12})A_t \end{aligned}$$

Si se eliminan los parámetros no significativos (que a su vez puede ser una solución al problema de la colinealidad entre los parámetros estacionales), resulta:

$$\begin{aligned} & (1 - 0,535312B^{12})g(y_t) \\ & = 5,86861 \\ & + (1 + 0,112918B + 0,208105B^2 - 0,136021B^4 - 0,191963B^5)A_t \end{aligned}$$

San Víctor

$$\begin{aligned} & ARIMA(1,0,0)x(1,1,1)_{12} \\ & (1 - 0,192611B)(1 + 0,0155332B^{12})(1 - B^{12})g(y_t) \\ & = 0,0949959 + (1 - 0,817495B^{12})A_t \end{aligned}$$

Se eliminan las estimaciones que no son significativas y queda:

$$(1 - 0,192611B)(1 - B^{12})g(y_t) = (1 - 0,817495B^{12})A_t$$

Santa Anita

Modelo 2:

$$\begin{aligned} & ARIMA(1,1,1)x(1,1,1)_{12} \\ & (1 - 0,0809936.B)(1 + 0,100335B^{12})(1 - B^{12})(1 - B)g(y_t) \\ & = -0,0176411 + (1 - 0,899614B)(1 - 0,767096B^{12})A_t \end{aligned}$$

Si se elimina del modelo los parámetros no significativos resulta:

$$(1 - B^{12})(1 - B)g(y_t) = (1 - 0,899614B)(1 - 0,767096B^{12})A_t$$

Modelo 4:

$$\begin{aligned} & ARIMA(1,0,1)x(1,1,1)_{12} \\ & (1 - 0,0,600822B)(1 - 0,0849492B^{12})(1 - B^{12})g(y_t) \\ & = 0,027795 + (1 - 0,489336B)(1 - 0,782526B^{12})A_t \end{aligned}$$

Eliminando las estimaciones que no son significativamente distintas de cero:

$$(1 - 0,0,600822B)(1 - B^{12})g(y_t) = (1 - 0,782526B^{12})A_t$$

Séptimo Distrito

$$\begin{aligned} & ARIMA(1,0,2)x(1,1,1)_{12} \\ & (1 - 0,325231B)(1 - 0,0837906B^{12})(1 - B^{12})g(y_t) \\ & = (1 + 0,348893B + 0,142392B^2)(1 - 0,752449B^{12})A_t \end{aligned}$$

Eliminando los parámetros no significativos:

$$(1 - B^{12})g(y_t) = (1 + 0,142392B^2)(1 - 0,752449B^{12})A_t$$

Viale

$$\begin{aligned} & ARIMA(1,0,0)x(0,1,1)_{12} \\ & (1 - 0,222247B)(1 - B^{12})g(y_t) = 0,265992 + (1 - 0,687592B^{12})A_t \end{aligned}$$

La constante no es significativa por lo que puede quitarse del modelo:

$$(1 - 0,222247B)(1 - B^{12})g(y_t) = (1 - 0,687592B^{12})A_t$$

Villa Paranacito

$$\begin{aligned} & ARIMA(1,0,0)x(1,0,0)_{12} \\ & (1 - 0,197151B)(1 - 0,185279B^{12})g(y_t) = 9,1651 + A_t \end{aligned}$$

Colón

$$\begin{aligned} & ARIMA(1,0,0)x(1,1,1)_{12} \\ & (1 - 0,161077B)(1 - 0,0154802B^{12})(1 - B^{12})g(y_t) \\ & = -0,0421835 + (1 - 0,836472B^{12})A_t \end{aligned}$$

Eliminando las estimaciones no significativas:

$$(1 - 0,161077B)(1 - B^{12})g(y_t) = (1 - 0,836472B^{12})A_t$$

La Lila

$$ARIMA(1,0,2)x(1,1,1)_{12}$$

$$(1 - 0,108857B)(1 - 0,028278B^{12})(1 - B^{12})g(y_t) \\ = (1 - 0,00681973B + 0,110075B^2)(1 - 0,86073B^{12})A_t$$

Eliminando los parámetros no significativos:

$$(1 - B^{12})g(y_t) = (1 + 0,110075B^2)(1 - 0,86073B^{12})A_t$$

Lucas González

$$ARIMA(0,0,0)x(0,1,1)_{12}$$

$$(1 - B^{12})g(y_t) = -0,0047778 + (1 - 0,854797B^{12})A_t$$

La constante no es significativa por lo que puede quitarse del modelo:

$$(1 - B^{12})g(y_t) = (1 - 0,854797B^{12})A_t$$

Santa María del Tatutí

Modelo 1:

$$ARIMA(1,0,0)x(1,0,1)_{12}$$

$$(1 - 0,136604B)(1 - 0,96121B^{12})g(y_t) = 0,481199 + (1 - 0,88506B^{12})A_t$$

Suprimiendo la constante por no ser significativa:

$$(1 - 0,136604B)(1 - 0,96121B^{12})g(y_t) = (1 - 0,88506B^{12})A_t$$

Modelo 2:

$$ARIMA(1,0,0)x(0,1,1)_{12}$$

$$(1 - 0,127983B)(1 - B^{12})g(y_t) = 0,0132856 + (1 - 0,927642B^{12})A_t$$

Sin la constante:

$$(1 - 0,127983B)(1 - B^{12})g(y_t) = (1 - 0,927642B^{12})A_t$$

Villa Elisa

$$ARIMA(1,0,0)x(0,0,0)_{12}$$

$$(1 - 0,296586B)g(y_t) = 9,02309 + A_t$$

Antelo

$$ARIMA(1,0,0)x(1,0,1)_{12}$$

$$(1 - 0,0800517B)(1 - 0,900118B^{12})g(y_t) = 1,01489 + (1 - 0,758945B^{12})A_t$$

Febre

Modelo 1:

$$\begin{aligned} & ARIMA(1,0,1)x(1,0,1)_{12} \\ (1 - 0,558112B)(1 - 0,938242B^{12})g(y_t) \\ & = 0,31656 + (1 - 0,550027B)(1 - 0,827128B^{12})A_t \end{aligned}$$

Modelo 2:

$$\begin{aligned} & ARIMA(1,0,1)x(0,1,1)_{12} \\ (1 + 0,533195B)(1 - B)g(y_t) & = -0,00523827 + (1 - 0,540941B)(1 - 0,868914B^{12})A_t \end{aligned}$$

Suprimiendo la constante que no es significativa:

$$(1 + 0,533195B)(1 - B)g(y_t) = (1 - 0,540941B)(1 - 0,868914B^{12})A_t$$

Feliciano

$$\begin{aligned} & ARIMA(1,0,2)x(1,1,1)_{12} \\ (1 - 0,282252B)(1 - 0,0122436B^{12})(1 - B^{12})g(y_t) \\ & = 0,10535 + (1 + 0,458383B + 0,173515B^2)(1 - 0,827634B^{12})A_t \end{aligned}$$

Eliminando los parámetros no significativos:

$$(1 - B^{12})g(y_t) = (1 + 0,458383B + 0,173515B^2)(1 - 0,827634B^{12})A_t$$

San Jaime

Modelo 1:

$$\begin{aligned} & ARIMA(2,0,1)x(1,1,1)_{12} \\ (1 - B^{12})g(y_t) & = (1 - 0,840634B^{12})A_t \end{aligned}$$

Modelo 2:

$$\begin{aligned} & ARIMA(2,0,1)x(1,0,1)_{12} \\ (1 - 0,154501B^2)(1 - 0,291909B^{12})g(y_t) & = 7,24863 + A_t \end{aligned}$$

San Salvador

$$\begin{aligned} & ARIMA(1,0,0)x(2,0,0)_{12} \\ (1 - 0,206252B)(1 - 0,1483B^{12} - 0,177606B^{24})g(y_t) & = 5,97247 + A_t \end{aligned}$$

La Paz

$$ARIMA(0,0,5)x(2,0,0)_{12}$$

$$\begin{aligned}(1 - 0,144538B^{12} - 0,0810902B^{24})g(y_t) \\ = 8,45036 \\ + (1 + 0,140548B + 0,15108B^2 + 0,012178B^3 - 0,0209091B^4 \\ - 0,109353B^5)A_t\end{aligned}$$

Eliminando los parámetros que no son significativos, resulta:

$$\begin{aligned}(1 - 0,144538B^{12} - 0,0810902B^{24})g(y_t) \\ = 8,45036 + (1 + 0,140548B + 0,15108B^2 - 0,109353B^5)A_t\end{aligned}$$

San Gustavo

$$ARIMA(1,0,1)x(2,0,0)_{12}$$

$$\begin{aligned}(1 - 0,435584B)(1 - 0,149665B^{12} - 0,109613B^{24})g(y_t) \\ = 4,74973 + (1 - 0,162094B)A_t\end{aligned}$$

Se suprime el parámetro de media móvil por no ser significativo:

$$(1 - 0,435584B)(1 - 0,149665B^{12} - 0,109613B^{24})g(y_t) = 4,74973 + A_t$$

Paraná

$$ARIMA(0,0,6)x(1,0,1)_{12}$$

$$\begin{aligned}(1 - 0,974067B^{12})g(y_t) \\ = 0,274851 \\ + (1 + 0,0592301B + 0,0491378B^2 - 0,0246789B^3 - 0,0188897B^4 \\ - 0,0105108B^5 - 0,0999834B^6)(1 - 0,887638B^{12})A_t\end{aligned}$$

Eliminando los parámetros que no son significativos:

$$(1 - 0,974067B^{12})g(y_t) = 0,274851 + (1 - 0,0999834B^6)(1 - 0,887638B^{12})A_t$$

Capítulo 8

Conclusiones

El estudio de la metodología de modelización ARIMA y su aplicación para el modelado de las series de registros de lluvia, permitió obtener las conclusiones que se presentan a continuación.

Los modelos ARIMA son modelos paramétricos donde cada observación en un instante particular es modelada en función de los valores pasados. Puntualmente, permiten describir un valor como una función lineal de observaciones anteriores y errores aleatorios. A lo largo del trabajo se puede observar que los instrumentos claves en la identificación y análisis de estos modelos son las funciones de autocorrelación simple y parcial y los contrastes de raíces unitarias. Hallar el modelo con mejor ajuste no es una tarea sencilla; a la identificación de los posibles modelos candidatos a recoger la estructura del proceso sobreviene la validación del mismo y esta etapa incluye una secuencia de pruebas rigurosas que los modelos propuestos deben satisfacer. Si las pruebas de validación no son satisfechas, entonces se procede a reformular el modelo hasta obtener uno que sea el adecuado (si es que existe).

El trabajo realizado sobre las series de registros históricos de lluvia no estuvo exento de esta dificultad. De las veintidós series seleccionadas, en ocho de ellas se detectó evidencia sobre la presencia de raíz unitaria en alguna frecuencia. Pero en otros casos, no es posible llegar a resultados concluyentes a partir de la información que aportan los test. Por ejemplo, en el caso del test HEGY el no rechazo de la hipótesis nula para algunas series se basa en un valor- p que supera a 0,01 pero no a 0,05, por lo que no puede concluirse presencia de raíz unitaria a un 5% de significación. Además, muchos de estos resultados dependieron del criterio de información utilizado (AIC, SBC y/o HQ) y de la elección o no de términos trigonométricos como componentes estacionales. El contraste de Canova-Hansen parece menos ambiguo, aunque hubo casos en los que no detectó inestabilidad y el modelo con mejor ajuste sí incluyó un orden de diferenciación estacional sugerida por el test HEGY o por el comportamiento de la ACF (el modelo propuesto para la serie que proviene del registro de La Lila tiene un orden de diferenciación que no es detectada por el contraste de Canova y Hansen y apenas lo es con el test HEGY bajo ciertas restricciones, ver Apéndice X, pág. 198). Aunque estos contrastes son sumamente valiosos para detectar la presencia de raíces unitarias no muestran, en este estudio, resultados unánimes. La decisión de incluir un orden de diferenciación en alguna frecuencia estuvo en todos los casos sustentada en el conjunto de informaciones que revelaron los test, el correlograma de la función de autocorrelación y el análisis de validación.

Todas las series reflejaron un comportamiento sujeto a variaciones estacionales. Los meses de invierno se caracterizan por tener los registros acumulados más bajos, mientras que primavera, verano y parte de otoño exhiben los valores acumulados más altos. Los modelos que presentaron el mejor ajuste cuentan con parámetros estacionales que recogen esa estructura. Sólo las series procedentes de Villa Elisa y Los Charrúas tienen como mejor ajuste un modelo autorregresivo sin órdenes estacionales.

Un problema recurrente en muchos casos fue el de la alta correlación particularmente entre los parámetros estacionales. Esto representa una complicación cuando las estimaciones de los parámetros no pueden suprimirse del modelo, es decir, cuando son significativamente distintos de cero. Éste fue el caso de los modelos propuestos para las series procedentes de Febre y La Lila.

El análisis residual no mostró en todos los casos que los residuos se comporten perfectamente como un ruido blanco. Incrementar el orden del polinomio de media móvil regular permitió alcanzar residuos con covarianzas nulas, pero a riesgo de la pérdida de homocedasticidad o de normalidad.

Es importante señalar que fue resuelto el problema de investigación planteado, ya que se encontró un modelo ARIMA particular para cada serie de registro de lluvia, aunque éste en algunos casos no haya sido del todo satisfactorio (por ejemplo, no se encontró un modelo con un buen ajuste para la serie procedente de La Lila o de Lucas González, ver Apéndice XI, pág. 204). Por otro lado, respecto de las hipótesis de investigación planteadas, en el caso de la Hipótesis I, es cierto que los modelos que explican los procesos generadores de las series son modelos ARIMA estacionales multiplicativos, pero hay registros tal que los modelos no incluyen órdenes estacionales. Con respecto a la Hipótesis II, existe evidencia de que ésta no se corroboró en ocho series de las veintidós estudiadas.

Los objetivos fijados en un inicio se lograron concretar. El trabajo muestra los aspectos básicos de la metodología de modelización ARIMA y cómo a partir de ella se logra obtener modelos que describen las características de las series de registros de lluvia. Esto a su vez puede ser de utilidad para otras investigaciones relacionadas.

La primera línea de investigación que continúa inmediatamente de este trabajo es estudiar la capacidad predictiva de estos modelos. Se trata puntualmente, de obtener predicciones óptimas de la serie en algún tiempo futuro, basándose para ello en la información que la misma revela a partir de su propio pasado. Básicamente, si se conoce los valores de la serie temporal hasta un instante k , se intenta obtener un valor puntual o un intervalo que contenga el 95% de las veces el valor que se espera que tome la serie en un instante futuro $k + l$.

La segunda línea de investigación es analizar la presencia de outliers en los datos observados y estudiar modelos que recojan esos valores extremos. Las causas de valores atípicos pueden encontrarse en cambios climáticos bruscos o algunos prolongados como por ejemplo, el fenómeno conocido como “*la corriente del niño*”. La idea básica consiste en considerar la serie observada x_t y la serie libre de outliers z_t . Se supone que z_t sigue un modelo $ARMA(p, q)$:

$$\varphi(B)z_t = \theta(B)A_t,$$

donde los operadores φ y θ son causales e invertibles y $A_t \sim N(0; \sigma_A^2)$. Un modelo de *outliers aditivo* se define como:

$$x_t = \begin{cases} z_t & \text{si } t \neq T \\ z_t + \omega & \text{si } t = T \end{cases}$$

$$x_t = z_t + \omega I_t^T = \frac{\theta(B)}{\varphi(B)} A_t + \omega I_t^T$$

donde

$$I_t^T = \begin{cases} 1 & \text{si } t = T \\ 0 & \text{si } t \neq T \end{cases}$$

es una variable que indica la presencia o ausencia de un outliers en el instante T . Un modelo de *outliers innovacional* actúa como adición al error A_t en un punto particular de la serie y en series estacionarias afecta a varias observaciones. Se define como:

$$x_t = \frac{\theta(B)}{\varphi(B)} (A_t + \omega I_t^T).$$

Estos modelos recogen la presencia de outliers aditivos e innovacionales aunque existen otras clasificaciones y formas de outliers como los de cambio de nivel y aditivos estacionales, entre otros. La manera de detectar la presencia de outliers requiere en general de métodos iterativos y existen programas de cómputo como SPSS y/o TRAMO que identifican los outliers y los corrigen con técnicas de interpolación.

Otra línea de investigación es considerar la existencia de otros modelos de series temporales, diferentes a los modelos ARIMA, que puedan presentar mejores ajustes y recojan con más precisión las características de las series de registros de lluvia. Por ejemplo, los modelos de *suavizado exponencial* donde se encuentran los conocidos como estacionales simples y los aditivos de Winters, entre otros. Un modelo *estacional simple* es similar a un ARIMA con orden cero en la estructura autorregresiva, órdenes $1, s$ y $s + 1$ de media móvil, donde $s = 12$ en el caso de las series mensuales, y un orden de diferenciación tanto en la parte regular como en la estacional. Esta forma de modelo se adecúa muy bien a series que presentan tendencia y estacionalidad. Por otro lado, un modelo *aditivo de Winters* se parece a un modelo ARIMA con orden cero en los polinomios autorregresivos y $s + 1$ órdenes de media móvil, un orden de diferenciación regular y un orden de diferenciación estacional. Son útiles para representar series temporales que presentan una forma de estacionalidad que no depende del nivel de la serie y que además tienen tendencia lineal.

Por último cabría explorar la posibilidad de que las series de registros de lluvia se modelen mejor empleando modelos *GARCH (Generalized Autoregressive Conditional Heteroscedasticity)*. Estos modelos permiten estudiar series que tienen desviaciones grandes (pequeñas) con respecto a un nivel constante, seguidas inmediatamente de variaciones pequeñas (grandes). En otras palabras, la varianza de la serie está afectada por un fenómeno de volatilidad. En general, cuando un modelo incorpora la posibilidad de que los errores tengan varianza no constante, se dice que el modelo es heterocedástico. Básicamente se parte de un modelo de regresión estándar:

$$Y_t = X_t' \beta + n_t$$

donde Y_t es la variable dependiente, X_t' es un vector de variables independientes de orden $1 \times k$, β es un vector de parámetros de dimensión $k \times 1$ y n_t es un error no correlacionado pero con varianza cambiante en el tiempo. Engle (1982) define el término de error de la siguiente manera:

$$n_t = \sigma_t \cdot e_t$$

donde $e_t \sim iid(0,1)$ y es independiente de las realizaciones pasadas n_{t-i} y además:

$$\sigma_t^2 = \theta_0 + \theta_1 n_{t-1}^2 + \theta_2 n_{t-2}^2 + \dots + \theta_s n_{t-s}^2.$$

Teniendo en cuenta toda la información hasta el momento $t - 1$, la varianza condicional de n_t es:

$$Var_{t-1}(n_t) = E_{t-1}(n_t^2) = E(n_t^2 | n_{t-1}, n_{t-2}, \dots) = \sigma_t^2$$

que está relacionada con los cuadrados de los errores pasados y cambia con el tiempo (*volatility clustering*). El modelo $n_t^2 = \theta_0 + \theta_1 n_{t-1}^2 + \theta_2 n_{t-2}^2 + \dots + \theta_s n_{t-s}^2 + \varepsilon_t$ es un $AR(s)$ con $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. Engle llamó a este modelo *ARCH (Autoregressive Conditional Heteroscedasticity)*. Si se considera que la varianza condicional del proceso de error se relaciona además con las últimas varianzas condicionales, entonces se tiene una generalización del modelo *ARCH*:

$$\sigma_t^2 = \theta_0 + \varphi_1 \sigma_{t-1}^2 + \dots + \varphi_r \sigma_{t-r}^2 + \theta_1 n_{t-1}^2 + \dots + \theta_s n_{t-s}^2$$

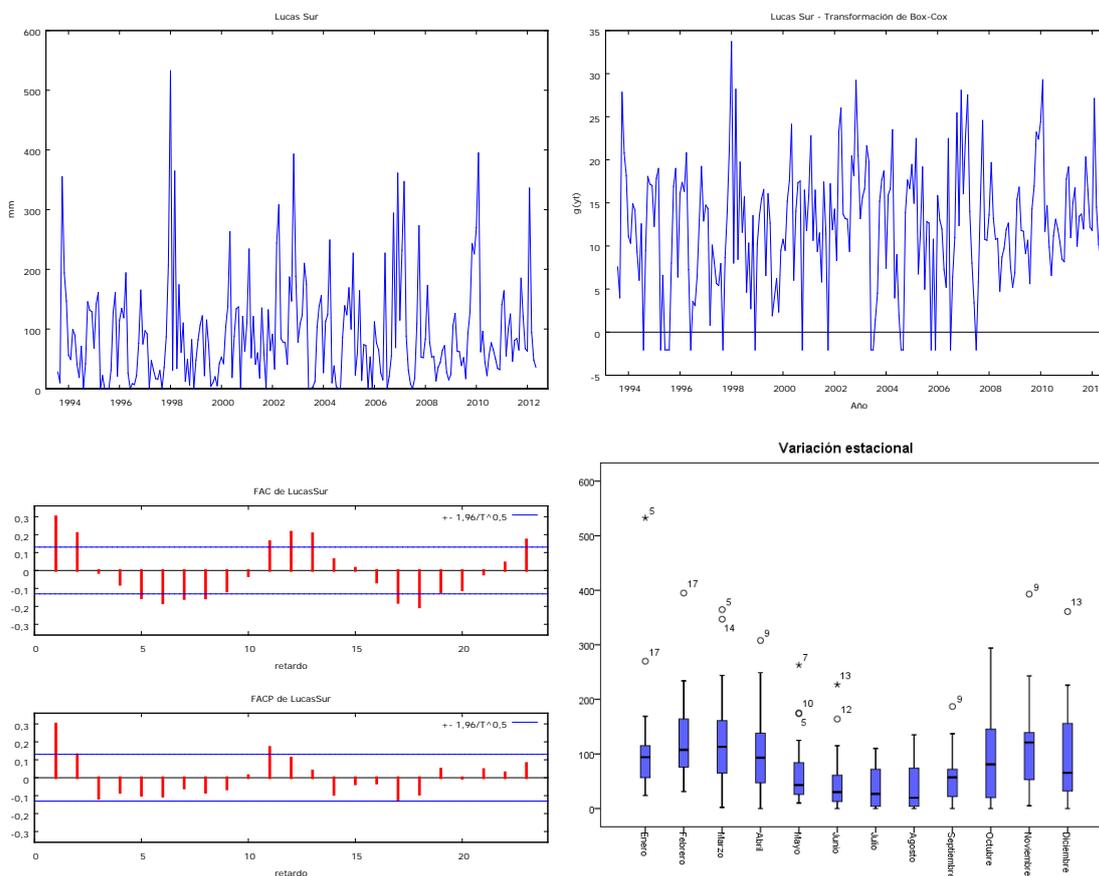
donde el polinomio $\varphi(B)$ tiene sus raíces fuera del círculo unitario y se asume $\theta_0 > 0$ y los parámetros φ_i y θ_j también mayores que cero para garantizar que $\sigma_t^2 > 0$. El modelo así definido es un *GARCH(r, s)*. Es de esperar que algunas de las series analizadas en este trabajo puedan presentar características que se modelen mejor con estos modelos que con los ARIMA.

Se continuará estas líneas de investigación con el propósito de obtener más y mejores resultados en la modelación estadística de las series de registros de lluvia, que puedan servir como aporte metodológico a otras investigaciones relacionadas.

Capítulo 9

Apéndice I: Lucas Sur

1. Representaciones gráficas.



2. Modelo propuesto: $ARIMA(0,0,2) \times (1,0,0)_{12}$

3. Estimación de los parámetros:

	Coefficiente	Desv. Típica	z	Valor p	
const	9,9792	0,987233	10,1083	<0,00001	***
Phi_1	0,174254	0,0660593	2,6378	0,00834	***
theta_1	0,215065	0,0682543	3,1509	0,00163	***
theta_2	0,235373	0,0678049	3,4713	0,00052	***

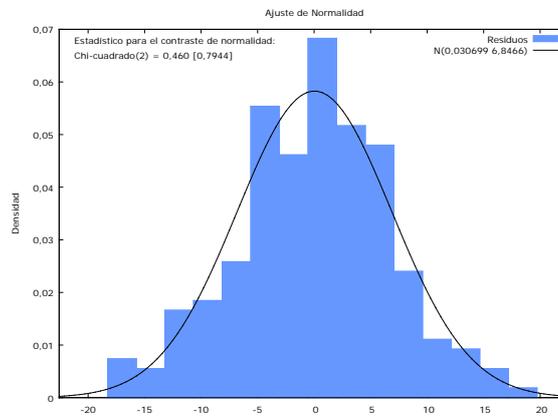
4. Causalidad e invertibilidad:

		Real	Imaginaria	Módulo	Frecuencia
AR (estacional)	Raíz 1	5,7387	0,0000	5,7387	0,0000
MA	Raíz 1	-0,4569	-2,0099	2,0612	-0,2856
	Raíz 2	-0,4569	2,0099	2,0612	0,2856

5. Análisis de los residuos.

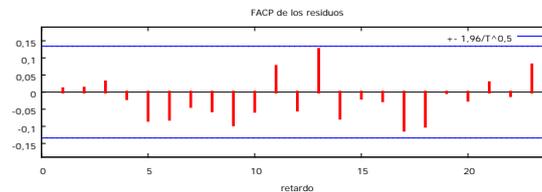
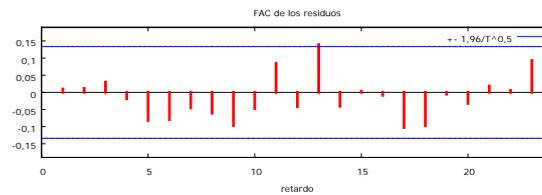
5.a. Normalidad:

Estadístico de Doornik-Hansen	Valor-p
0,460	0,79443



5.b. Media y Correlograma residual:

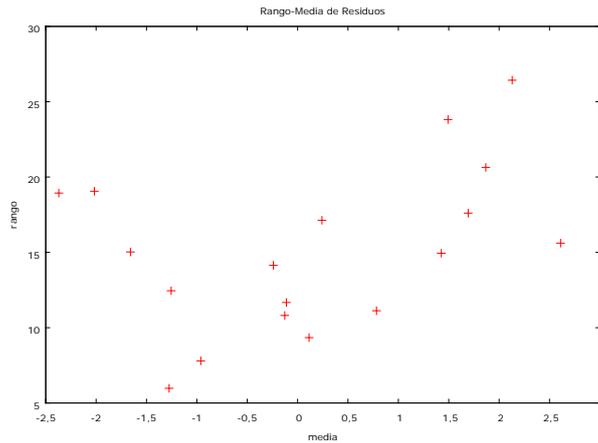
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,030699	6,7982	0,0660598	0,9474



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	0,0934	21,2283	0,567

5.c. Homocedasticidad:

Pendiente estimada	Valor-p
1,42794	0,101101



6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas				
const	Phi_1	theta_1	theta_2	
0,974629	-0,0461485	-0,000171397	-0,0154609	const
	0,00436383	-0,000801045	0,000451731	Phi_1
		0,00465865	0,000894810	theta_1
			0,00459751	theta_2

Matriz de correlación entre los parámetros			
	Φ_1	θ_1	θ_2
Φ_1	1	-0,17766134	0,10085208
θ_1		1	0,19334802
θ_2			1

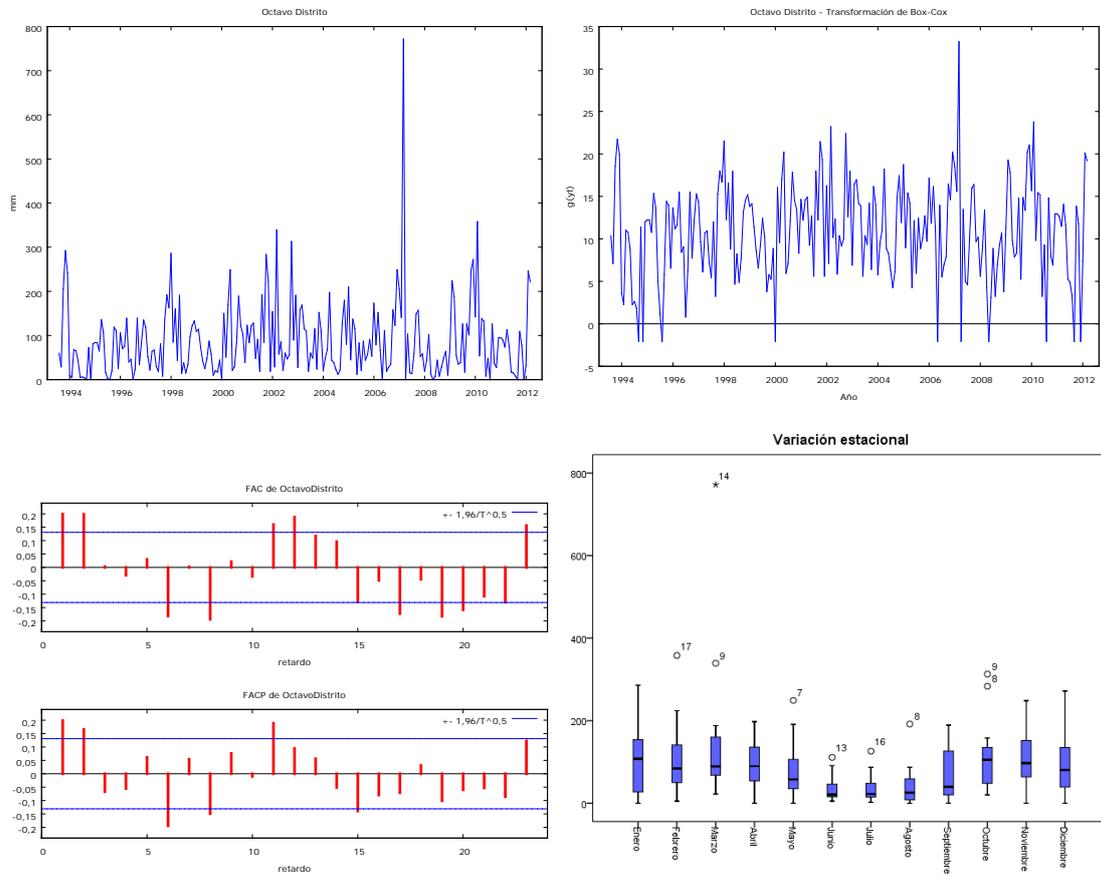
7. Criterios de selección:

AIC	SBC	HQ
1436,636	1453,466	1443,437

Conclusión: El modelo propuesto supera las pruebas de validación.

Apéndice II: Octavo Distrito

1. Representaciones gráficas.



2. Modelos Propuestos:

Modelo 1: $ARIMA(0,0,2)x(0,1,1)_{12}$

Modelo 2: $ARIMA(0,0,2)x(1,1,1)_{12}$

Modelo 3: $ARIMA(0,0,2)x(1,0,1)_{12}$

Modelo 4: $ARIMA(2,0,0)x(1,1,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	0,0436733	0,156018	0,2799	0,77954	
theta_1	0,0715247	0,0648652	1,1027	0,27017	
theta_2	0,174989	0,0673569	2,5979	0,00938	***
Theta_1	-0,744906	0,0519658	-14,3345	<0,00001	***

Modelo 2					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,135715	0,172019	0,7890	0,43014	
Phi_1	-0,140214	0,0797801	-1,7575	0,07883	*
theta_1	0,00518022	0,0710742	0,0729	0,94190	
theta_2	0,253185	0,0807789	3,1343	0,00172	***
Theta_1	-0,701589	0,0647893	-10,8288	<0,00001	***

Modelo 3					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	10,0875	1,95803	5,1519	<0,00001	***
Phi_1	0,0692682	0,173262	0,3998	0,68931	
theta_1	0,10884	0,0635301	1,7132	0,08668	*
theta_2	0,220816	0,0698453	3,1615	0,00157	***
Theta_1	0,118984	0,191442	0,6215	0,53426	

Modelo 4					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,0773948	0,136268	0,5680	0,57006	
phi_1	0,0136424	0,06397	0,2133	0,83112	
phi_2	0,25624	0,0817952	3,1327	0,00173	***
Phi_1	-0,111677	0,0782949	-1,4264	0,15376	
Theta_1	-0,709892	0,0628497	-11,2951	<0,00001	***

4. Causalidad e invertibilidad:

Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
MA					
	Raíz 1	-0,2044	-2,3818	2,3905	-0,2636
	Raíz 2	-0,2044	2,3818	2,3905	0,2636
MA (estacional)					
	Raíz 1	1,3425	0,0000	1,3425	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	-7,1320	0,0000	7,1320	0,5000
MA					
	Raíz 1	-0,0102	-1,9874	1,9874	-0,2508
	Raíz 2	-0,0102	1,9874	1,9874	0,2508
MA (estacional)					
	Raíz 1	1,4253	0,0000	1,4253	0,0000

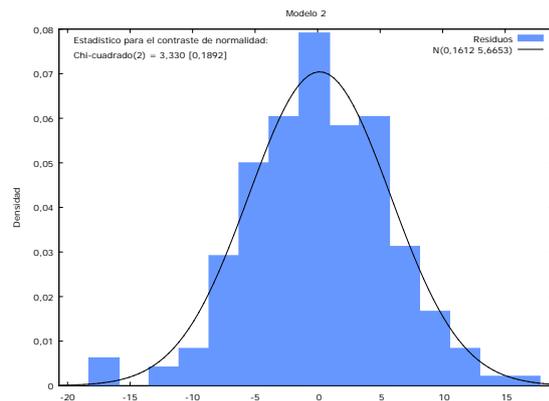
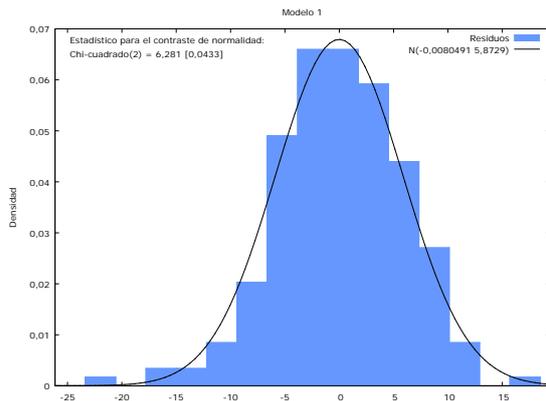
Modelo 3					
		Real	Imaginaria	Módulo	Frecuencia
AR (estacional)					
	Raíz 1	14,4366	0,0000	14,4366	0,0000
MA					
	Raíz 1	-0,2464	-2,1137	2,1281	-0,2685
	Raíz 2	-0,2464	2,1137	2,1281	0,2685
MA (estacional)					
	Raíz 1	-8,4045	0,0000	8,4045	0,5000

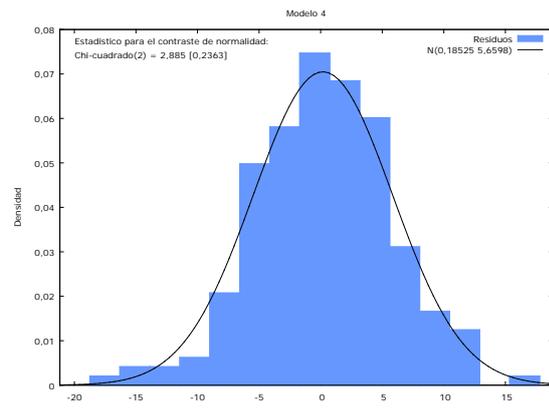
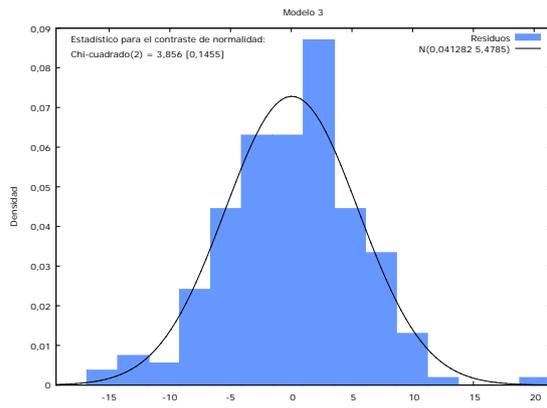
Modelo 4					
		Real	Imaginaria	Módulo	Frecuencia
AR					
	Raíz 1	-2,0023	0,0000	2,0023	0,5000
	Raíz 2	1,9491	0,0000	1,9491	0,0000
AR (estacional)					
	Raíz 1	-8,9544	0,0000	8,9544	0,5000
MA (estacional)					
	Raíz 1	1,4087	0,0000	1,4087	0,0000

5. Análisis de los residuos

5.a. Normalidad:

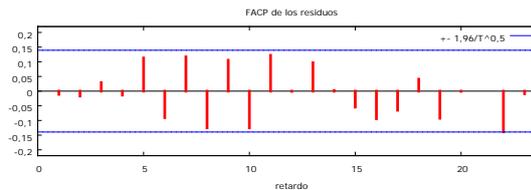
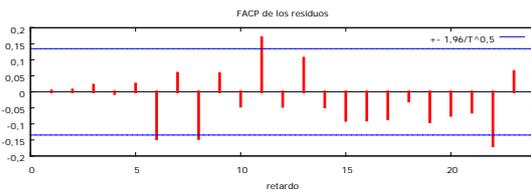
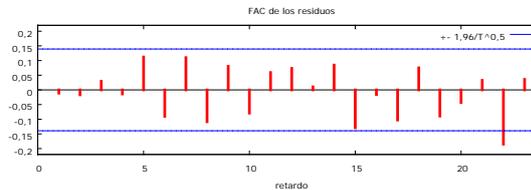
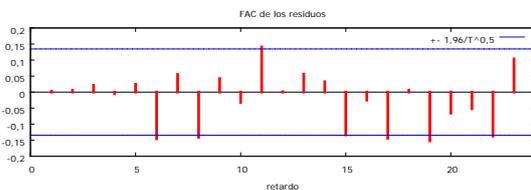
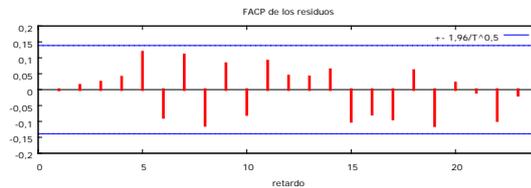
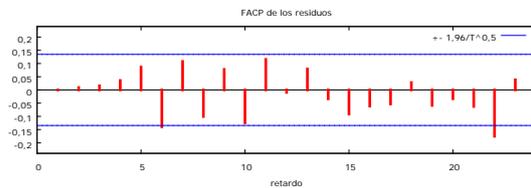
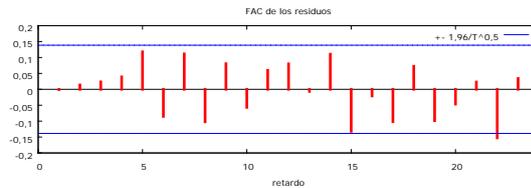
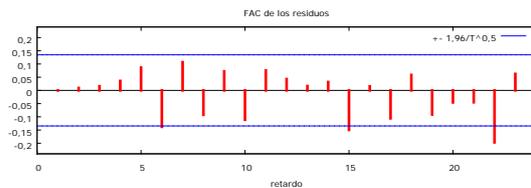
	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	6,281	0,04325
Modelo 2	3,330	0,18920
Modelo 3	3,856	0,1455
Modelo 4	2,885	0,23632





5.b. Media y Correlograma residual:

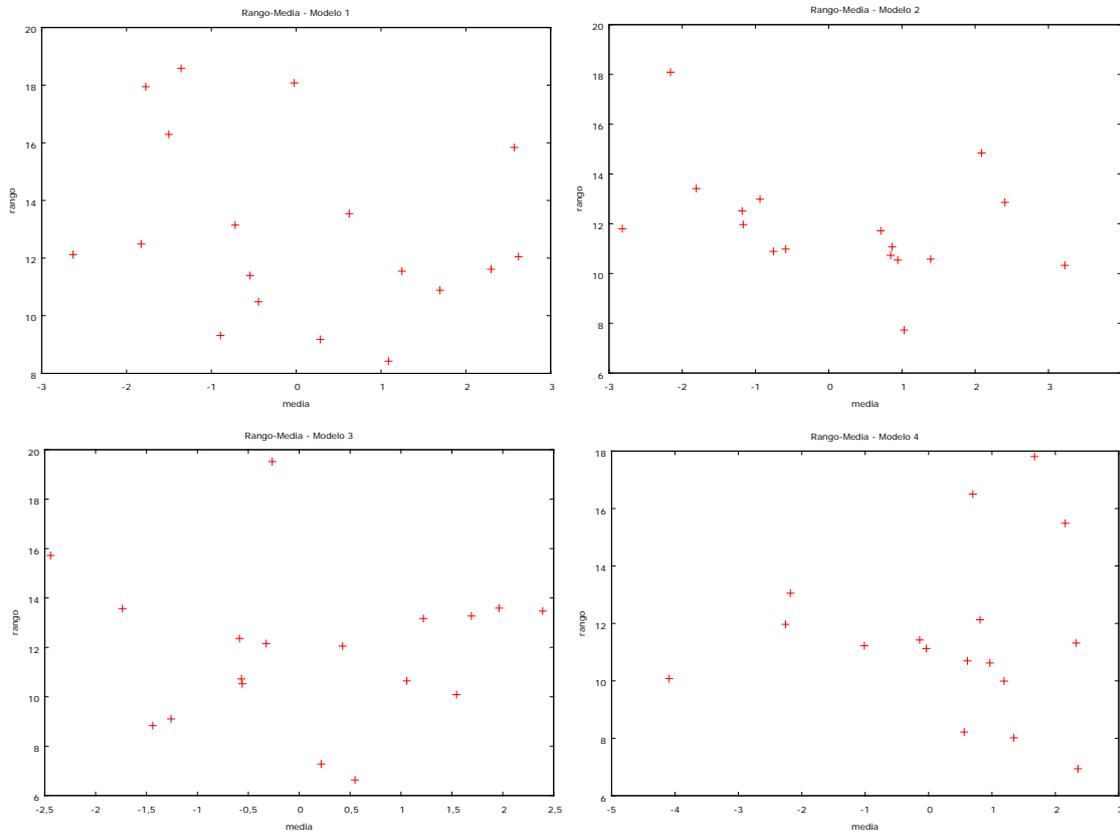
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	-0,00804914	5,83101	-0,020099	0,984
Modelo 2	0,1612	5,60809	0,406504	0,6848
Modelo 3	0,041282	5,42628	0,110771	0,9119
Modelo 4	0,185253	5,60205	0,465318	0,6422



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	0,0055	46,9926	0,104
Modelo 2	-0,0068	51,4435	0,046
Modelo 3	0,1000	52,7302	0,036
Modelo 4	-0,0095	51,7566	0,043

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,508463	0,29785
Modelo 2	-0,479112	0,146698
Modelo 3	-0,0664356	0,907159
Modelo 4	0,0719738	0,865854



6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1				
const	theta_1	theta_2	Theta_1	
0,0243415	-0,000901776	0,000633924	-0,000913868	const
	0,00420749	-0,000427602	3,06450e-005	theta_1
		0,00453695	0,000483196	theta_2
			0,00270045	Theta_1

Correlación entre los parámetros estimados del Modelo 1			
	θ_1	θ_2	Θ_1
θ_1	1	-0,09786916	0,00909139
θ_2		1	0,13804593
Θ_1			1

Matriz de Covarianzas del Modelo 2					
const	Phi_1	theta_1	theta_2	Theta_1	
0,0295905	0,000514741	-0,00199105	0,00114391	-0,00111007	const
	0,00636487	0,000983011	0,000529552	-0,00209520	Phi_1
		0,00505154	-0,000765481	-0,000250164	theta_1
			0,00652523	-0,000232900	theta_2
				0,00419765	Theta_1

Correlación entre los parámetros estimados del Modelo 2				
	Φ_1	θ_1	θ_2	Θ_1
Φ_1	1	0,17336117	0,08217053	-0,40534761
θ_1		1	-0,13332896	-0,05432623
θ_2			1	-0,04450085
Θ_1				1

Matriz de Covarianzas del Modelo 3					
const	Phi_1	theta_1	theta_2	Theta_1	
3,83390	-0,325357	0,0219373	-0,0363124	0,322559	const
	0,0300198	-0,00248742	0,00340560	-0,0297129	Phi_1
		0,00403607	-0,000484226	0,00254851	theta_1
			0,00487836	-0,00420496	theta_2
				0,0366499	Theta_1

Correlación entre los parámetros estimados del Modelo 3				
	Φ_1	θ_1	θ_2	Θ_1
Φ_1	1	-0,22597805	0,28141881	-0,89578646
θ_1		1	-0,10912679	0,20954127
θ_2			1	-0,31447597
Θ_1				1

Matriz de Covarianzas del Modelo 4					
const	phi_1	phi_2	Phi_1	Theta_1	
0,0185689	-0,00178387	-0,000292856	8,62092e-005	-0,000758746	const
	0,00409216	-0,00148256	0,000349807	1,49518e-005	phi_1
		0,00669046	0,000910512	-0,000505723	phi_2
			0,00613009	-0,00209586	Phi_1
				0,00395009	Theta_1

Correlación entre los parámetros del Modelo 4				
	φ_1	φ_2	Φ_1	Θ_1
φ_1	1	-0,28334014	0,06984232	0,0037189
φ_2		1	0,14217537	-0,09837431
Φ_1			1	-0,4259176
Θ_1				1

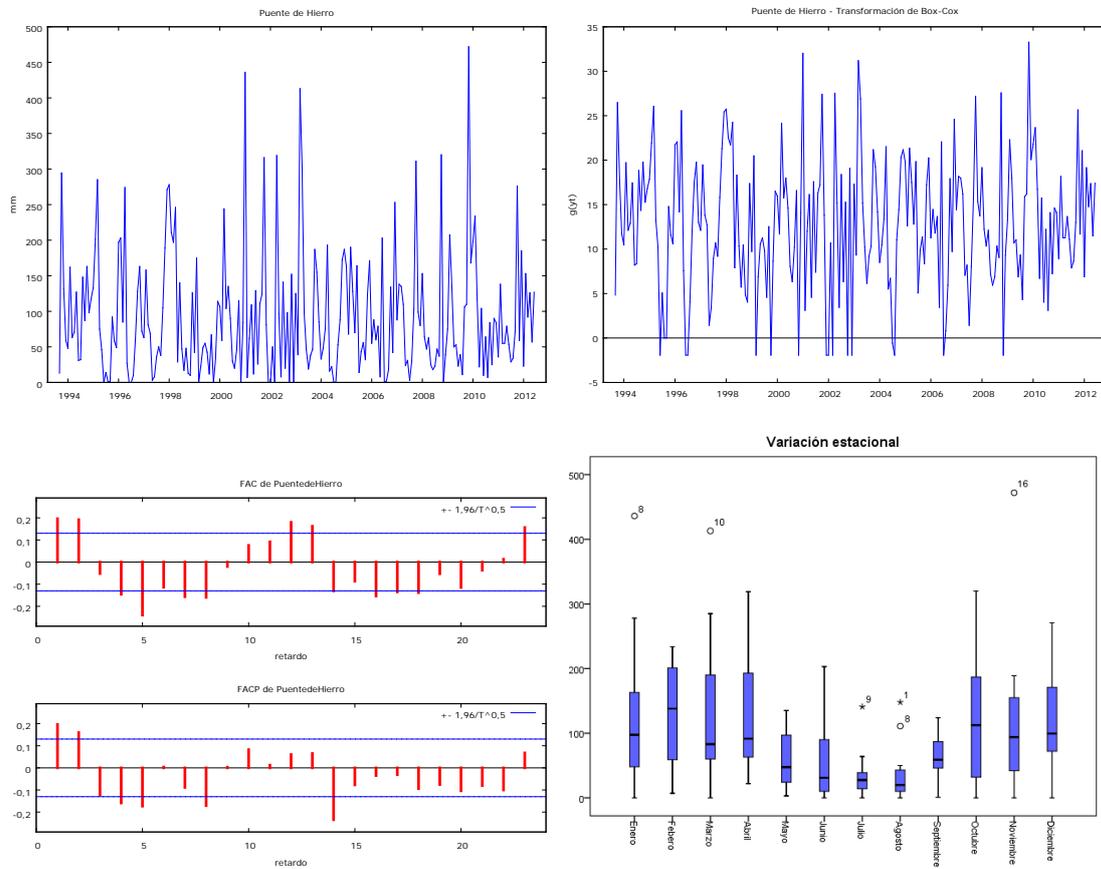
7. Criterios de selección:

	AIC	SBC	HQ
Modelo 1	1358,221	1375,004	1365,004
Modelo 2	1268,423	1288,213	1276,432
Modelo 3	1329,732	1349,871	1337,872
Modelo 4	1255,475	1275,205	1263,461

Conclusión: Los modelos que mejor explican el proceso son el 2 y el 4, tanto por el análisis de validación como por los resultados obtenidos con los criterios de información. No obstante, puede decirse que ambos recogen la estructura del proceso.

Apéndice III: Puente de Hierro

1. Representaciones gráficas.



2. Modelos Propuestos:

Modelo 1: $ARIMA(0,0,5)x(1,0,1)_{12}$

Modelo 2: $ARIMA(0,0,2)x(1,0,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	5,86861	2,92105	2,0091	0,04453	**
Phi_1	0,535312	0,22895	2,3381	0,01938	**
theta_1	0,112918	0,0668079	1,6902	0,09099	*
theta_2	0,208105	0,0680419	3,0585	0,00222	***
theta_3	-0,0644063	0,0678719	-0,9489	0,34265	
theta_4	-0,136021	0,068708	-1,9797	0,04774	**
theta_5	-0,191963	0,0712073	-2,6958	0,00702	***
Theta_1	-0,410155	0,24331	-1,6857	0,09185	*

Modelo 2					
	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	5,79322	2,69188	2,1521	0,03139	**
Phi_1	0,54403	0,205932	2,6418	0,00825	***
theta_1	0,106249	0,0620737	1,7117	0,08696	*
theta_2	0,26892	0,0638355	4,2127	0,00003	***
Theta_1	-0,376233	0,223699	-1,6819	0,09259	*

4. Causalidad e invertibilidad:

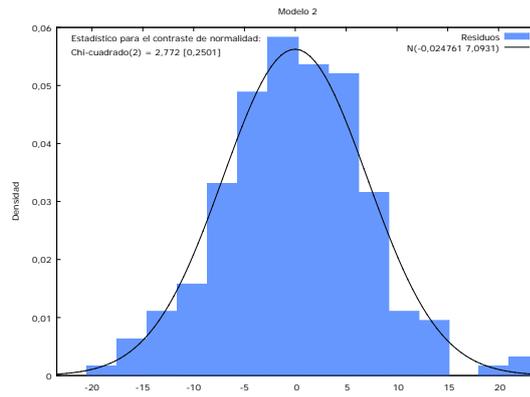
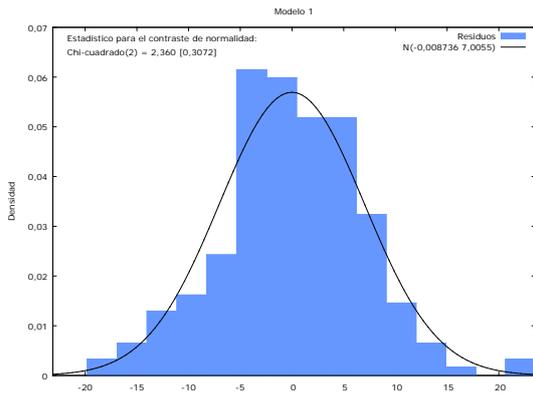
Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	1,8681	0,0000	1,8681	0,0000
MA					
	Raíz 1	-1,2021	-0,9242	1,5164	-0,3957
	Raíz 2	-1,2021	0,9242	1,5164	0,3957
	Raíz 3	1,3641	0,0000	1,3641	0,0000
	Raíz 4	0,1658	-1,2780	1,2887	-0,2295
	Raíz 5	0,1658	1,2780	1,2887	0,2295
MA (estacional)					
	Raíz 1	2,4381	0,0000	2,4381	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	1,8381	0,0000	1,8381	0,0000
MA					
	Raíz 1	-0,1975	-1,9182	1,9284	-0,2663
	Raíz 2	-0,1975	1,9182	1,9284	0,2663
MA (estacional)					
	Raíz 1	2,6579	0,0000	2,6579	0,0000

5. Análisis de los residuos.

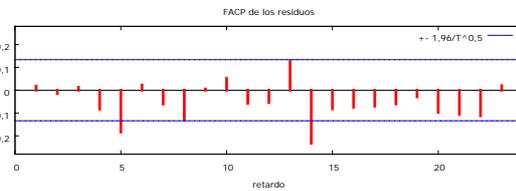
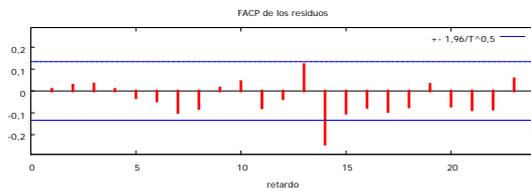
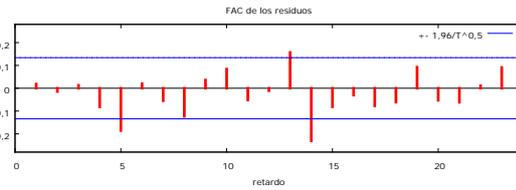
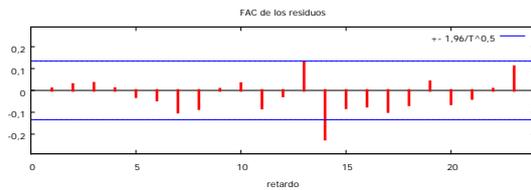
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	2,360	0,3072
Modelo 2	2,772	0,2501



5.b. Media y Correlograma residual:

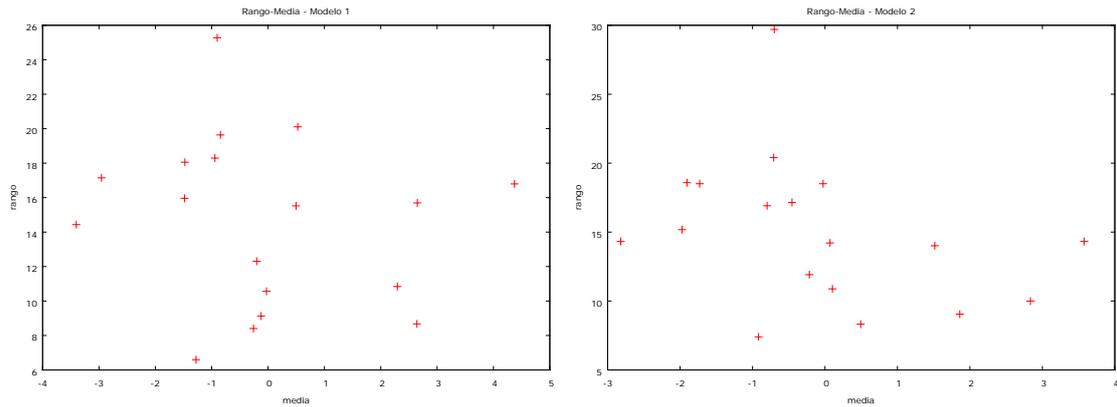
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	-0,00873602	6,88947	-0,0185496	0,9852
Modelo 2	-0,0247615	7,02615	-0,0515544	0,9589



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	0,0328	42,5861	0,209
Modelo 2	0,0278	56,2686	0,017

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,395301	0,526071
Modelo 2	-1,11634	0,152676



6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1								
const	Phi_1	theta_1	theta_2	theta_3	theta_4	theta_5	Theta_1	
8,53256	-0,665760	0,0452190	-0,00972466	0,0127609	0,0110031	4,67072e-007	0,686557	const
	0,0524181	-0,00364830	0,000942125	-0,000978397	-0,000804051	0,000181083	-0,0540074	Phi_1
		0,00446330	0,000723060	0,00142301	0,000333570	-0,000457842	0,00410042	theta_1
			0,00462970	0,000969612	0,000866711	0,00103593	-0,00106796	theta_2
				0,00460659	0,00121323	0,00107172	5,93245e-005	theta_3
					0,00472078	0,000806169	0,000760481	theta_4
						0,00507047	-0,000272341	theta_5
							0,0591999	Theta_1

Correlaciones entre los parámetros estimados del Modelo 1							
	Φ_1	θ_1	θ_2	θ_3	θ_4	θ_5	Θ_1
Φ_1	1	-0,2385185	0,06047717	-0,06296286	-0,0511135	0,0111074	-0,96951083
θ_1		1	0,15906334	0,31382687	0,07266945	-0,0962417	0,25225545
θ_2			1	0,20995757	0,18539182	0,21381071	-0,06450875
θ_3				1	0,26016318	0,22175162	0,0035924
θ_4					1	0,16477612	0,04549054
θ_5						1	-0,01571913
Θ_1							1

Matriz de Covarianzas del Modelo 2					
const	Phi_1	theta_1	theta_2	Theta_1	
7,24623	-0,547274	0,0423147	-0,0184947	0,575881	const
	0,0424079	-0,00328798	0,00137049	-0,0445656	Phi_1
		0,00385314	0,000431212	0,00420039	theta_1
			0,00407497	-0,00142761	theta_2
				0,0500411	Theta_1

Correlaciones entre los parámetros estimados del Modelo 2				
	Φ_1	θ_1	θ_2	Θ_1
Φ_1	1	-0,25721583	0,10425329	-0,96741291
θ_1		1	0,10882306	0,30249482
θ_2			1	-0,09997313
Θ_1				1

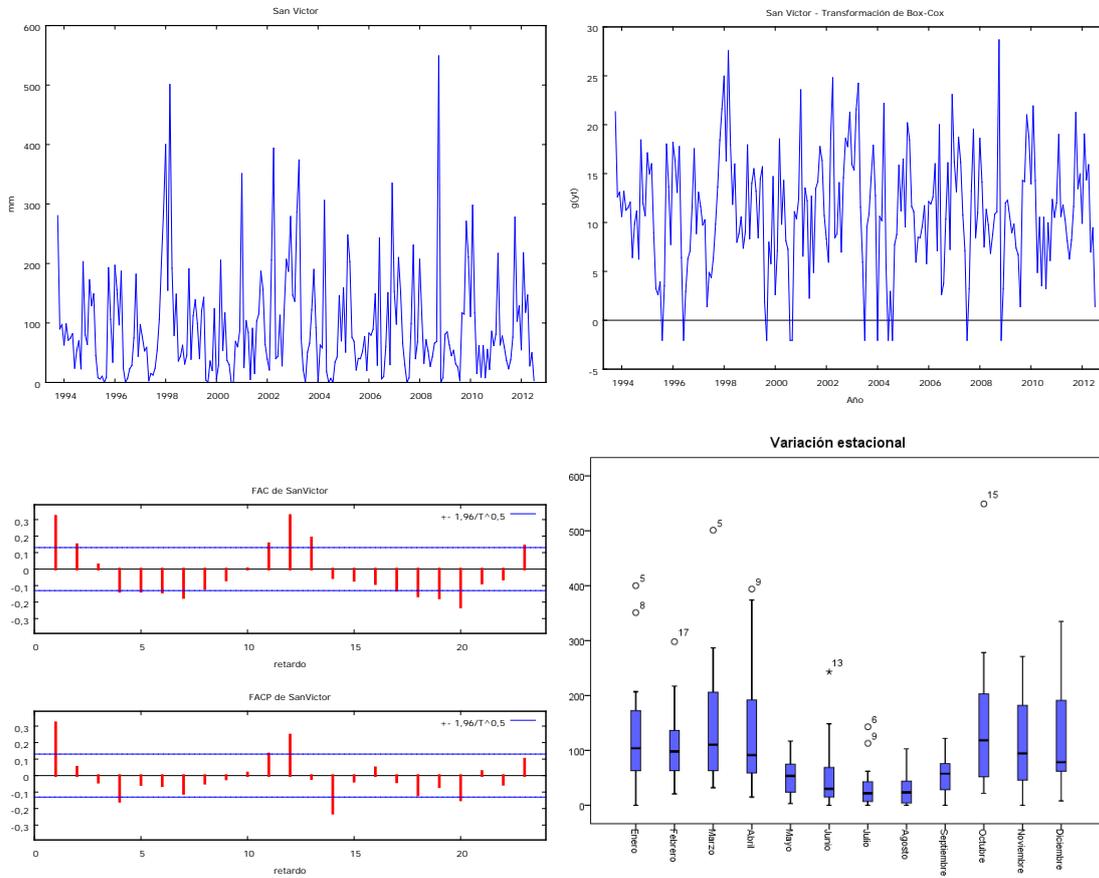
7. Criterios de selección:

	AIC	SBC	HQ
Modelo 1	1450,341	1480,635	1462,582
Modelo 2	1452,751	1472,947	1460,912

Conclusión: Ambos modelos recogen la estructura del proceso, no obstante el Modelo 1 presenta mejor ajuste residual. En los dos casos, los parámetros estacionales tienen correlaciones altas (aunque puede eliminarse el parámetro de media móvil estacional en ambos modelos), mientras que el criterio AIC favorece al modelo 1 y los criterios SBC y HQ al modelo 2. A pesar de las diferencias en los criterios de información, el Modelo 1 parece explicar mejor las características del proceso (ver que para el Modelo 2 no está claro que los residuos se comporten como un ruido blanco).

Apéndice IV: San Víctor

1. Representaciones gráficas.



2. Modelos Propuestos:

Modelo 1: $ARIMA(1,0,0)x(1,1,1)_{12}$

Modelo 2: $ARIMA(1,0,1)x(1,1,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	0,0949959	0,0990595	0,9590	0,33757	
phi_1	0,192611	0,0720418	2,6736	0,00750	***
Phi_1	-0,0155332	0,0798525	-0,1945	0,84577	
Theta_1	-0,817495	0,05449	-15,0027	<0,00001	***

Modelo 2					
	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,0424905	0,0624137	0,6808	0,49601	
phi_1	0,637133	0,204649	3,1133	0,00185	***
Phi_1	-0,00271228	0,0821904	-0,0330	0,97367	
theta_1	-0,461702	0,237559	-1,9435	0,05195	*
Theta_1	-0,809014	0,0562599	-14,3799	<0,00001	***

4. Causalidad e invertibilidad:

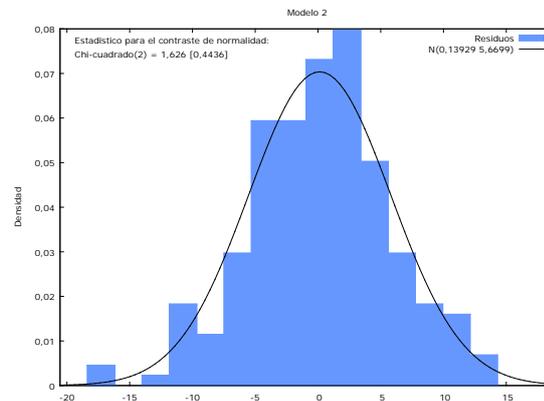
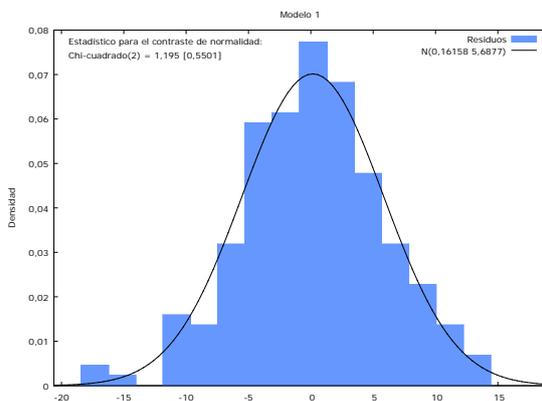
Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	5,1918	0,0000	5,1918	0,0000
AR (estacional)					
	Raíz 1	-64,3782	0,0000	64,3782	0,5000
MA (estacional)					
	Raíz 1	1,2232	0,0000	1,2232	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	1,5695	0,0000	1,5695	0,0000
AR (estacional)					
	Raíz 1	-368,6941	0,0000	368,6941	0,5000
MA					
	Raíz 1	2,1659	0,0000	2,1659	0,0000
MA (estacional)					
	Raíz 1	1,2361	0,0000	1,2361	0,0000

5. Análisis de los residuos

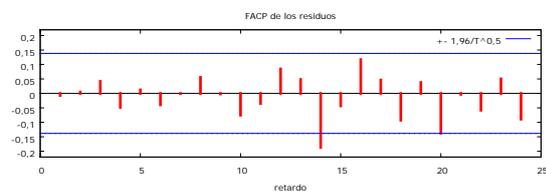
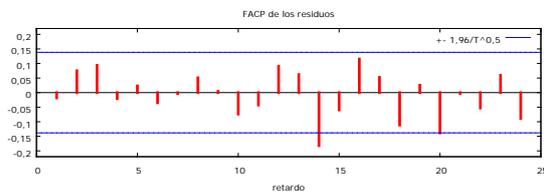
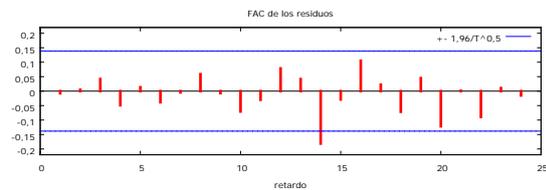
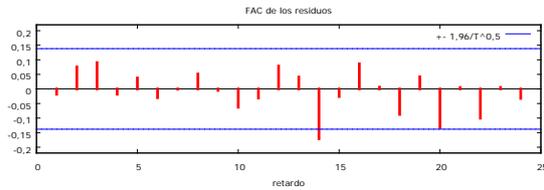
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	1,195	0,55007
Modelo 2	1,626	0,44356



5.b. Media y Correlograma residual:

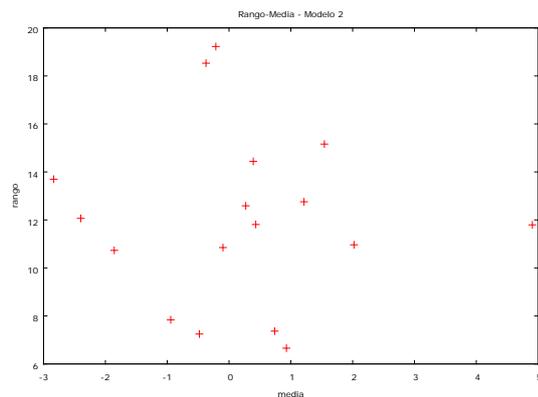
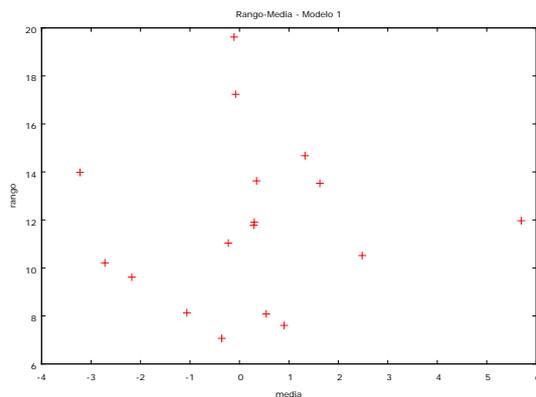
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,161583	5,64489	0,405824	0,6853
Modelo 2	0,139291	5,61293	0,351829	0,7253



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0743	42,7794	0,203
Modelo 2	-0,0746	39,7184	0,308

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	0,0856444	0,84443
Modelo 2	-0,0943264	0,857855



6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de covarianzas del Modelo 1				
const	phi_1	Phi_1	Theta_1	
0,00981278	-0,00126873	0,00131036	-9,84013e-005	const
	0,00519003	-0,000459944	0,000644896	phi_1
		0,00637643	-0,00244866	Phi_1
			0,00296916	Theta_1

Matriz de covarianzas del Modelo 1			
	φ_1	Φ_1	Θ_1
φ_1	1	-0,07995247	0,16428137
Φ_1		1	-0,56275992
Θ_1			1

Matriz de correlaciones del Modelo 2					
const	phi_1	Phi_1	theta_1	Theta_1	
0,00389547	-0,00630179	0,000856401	0,00666307	-2,68498e-006	const
	0,0418811	-0,00182288	-0,0466171	0,000667186	phi_1
		0,00675526	0,00158042	-0,00274032	Phi_1
			0,0564344	-1,64545e-005	theta_1
				0,00316518	Theta_1

Matriz de correlaciones del Modelo 2				
	φ_1	Φ_1	θ_1	Θ_1
φ_1	1	-0,10837456	-0,95887973	0,05794799
Φ_1		1	0,08094312	-0,59262672
θ_1			1	-0,00123116
Θ_1				1

7. Criterios de selección:

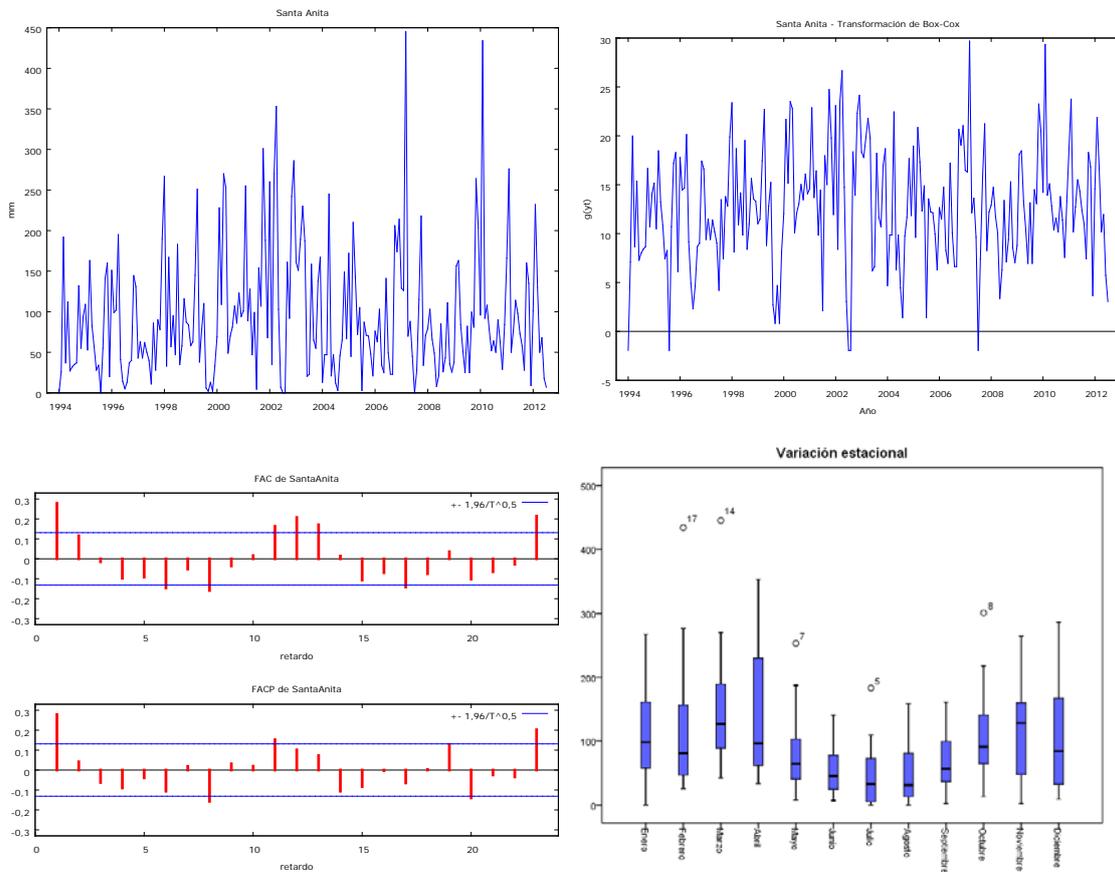
	AIC	SBC	HQ
Modelo 1	1275,338	1291,855	1282,021
Modelo 2	1275,014	1294,834	1283,034

Conclusión: Los dos modelos propuestos incluyen un orden de diferenciación estacional sugerido por los resultados de los contrastes HEGY y CH. Además, ambos presentan sobreparametrización, es decir, hay parámetros estimados que no son significativamente distintos de cero. Las condiciones de causalidad e invertibilidad están satisfechas, y el mejor ajuste de normalidad residual lo tiene el Modelo 1. Los correlogramas residuales no contradicen la hipótesis de ruido blanco en los retardos regulares y tampoco en los

estacionales y no hay presencia de heterocedasticidad. Las correlaciones entre los coeficientes estimados no son altas a excepción de la correlación entre el parámetro autorregresivo y el de media móvil del Modelo 2, lo que podría ser un indicio de factores comunes, sin embargo θ_1 no es significativo al 5% y puede suprimirse. En este punto resulta útil la información que ofrecen los criterios. El Modelo 1 es el que presenta los valores más bajos de los criterios SBC y de HQC, mientras que el criterio de Akaike en ambos modelos tiene diferencia mínima. El análisis integrado sugiere al Modelo 1 como el que mejor explica la estructura del proceso generador de la serie. Se han probado otros modelos sin mejores resultados.

Apéndice V: Santa Anita

1. Representaciones gráficas.



2. Modelos Propuestos:

Modelo 1: $ARIMA(0,1,1) \times (0,1,1)_{12}$

Modelo 2: $ARIMA(1,1,1) \times (1,1,1)_{12}$

Modelo 3: $ARIMA(1,0,0) \times (0,1,1)_{12}$

Modelo 4: $ARIMA(1,0,1) \times (1,1,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	-0,0633445	0,0465571	-1,3606	0,17365	
theta_1	-0,684303	0,0495033	-13,8234	<0,00001	***
Theta_1	-0,741231	0,0491371	-15,0850	<0,00001	***

Modelo 2					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	-0,0176411	0,0138096	-1,2775	0,20144	
phi_1	0,0809936	0,0869522	0,9315	0,35161	
Phi_1	-0,100335	0,0861088	-1,1652	0,24393	
theta_1	-0,899614	0,0354449	-25,3806	<0,00001	***
Theta_1	-0,767096	0,0643552	-11,9197	<0,00001	***

Modelo 3					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,0640494	0,0861283	0,7437	0,45709	
phi_1	0,112078	0,0661722	1,6937	0,09032	*
Theta_1	-0,842313	0,0402163	-20,9446	<0,00001	***

Modelo 4					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,027795	0,0720822	0,3856	0,69979	
phi_1	0,600822	0,331067	1,8148	0,06955	*
Phi_1	-0,0849492	0,0817888	-1,0386	0,29897	
theta_1	-0,489336	0,358083	-1,3665	0,17177	
Theta_1	-0,782526	0,0539199	-14,5127	<0,00001	***

4. Causalidad e invertibilidad:

Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
MA					
	Raíz 1	1,4613	0,0000	1,4613	0,0000
MA (estacional)					
	Raíz 1	1,3491	0,0000	1,3491	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	12,3467	0,0000	12,3467	0,0000
AR (estacional)					
	Raíz 1	-9,9666	0,0000	9,9666	0,5000
MA					
	Raíz 1	1,1116	0,0000	1,1116	0,0000
MA (estacional)					
	Raíz 1	1,3036	0,0000	1,3036	0,0000

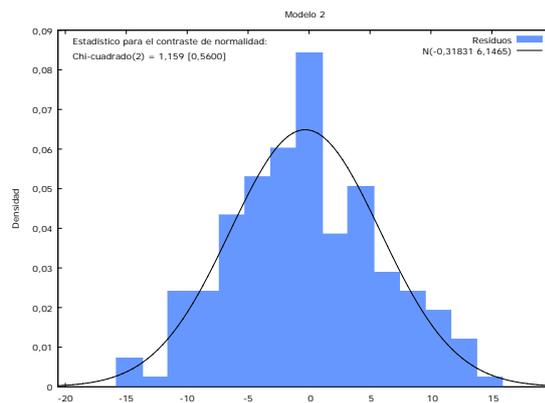
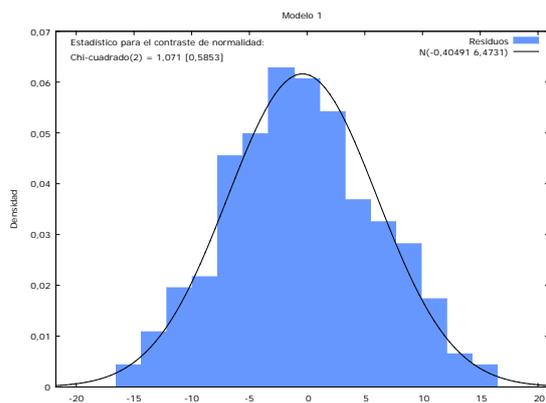
Modelo 3					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	8,9224	0,0000	8,9224	0,0000
MA (estacional)					
	Raíz 1	1,1872	0,0000	1,1872	0,0000

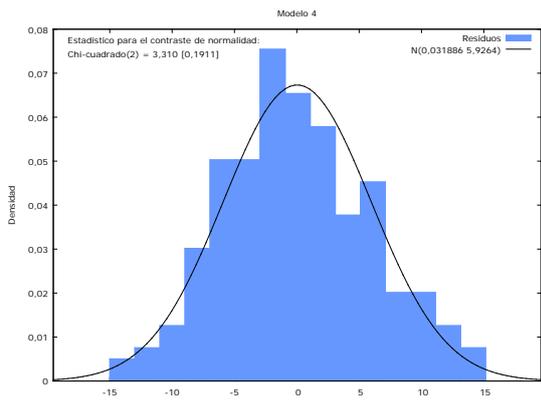
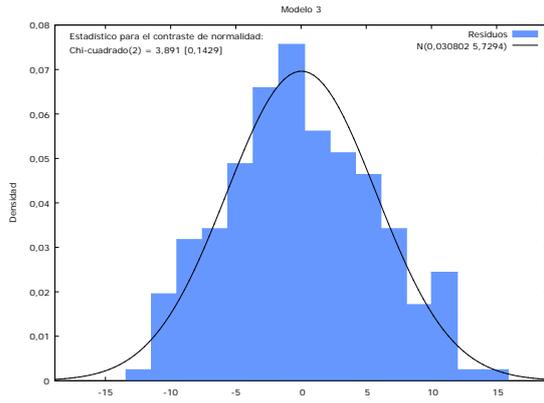
Modelo 4					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	1,6644	0,0000	1,6644	0,0000
AR (estacional)					
	Raíz 1	-11,7717	0,0000	11,7717	0,5000
MA					
	Raíz 1	2,0436	0,0000	2,0436	0,0000
MA (estacional)					
	Raíz 1	1,2779	0,0000	1,2779	0,0000

5. Análisis de los residuos.

4.a. Normalidad:

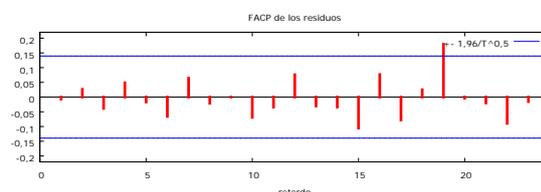
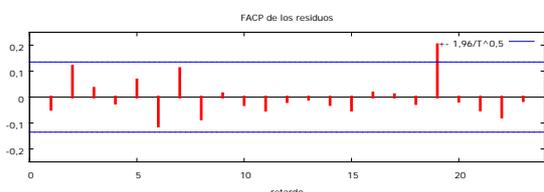
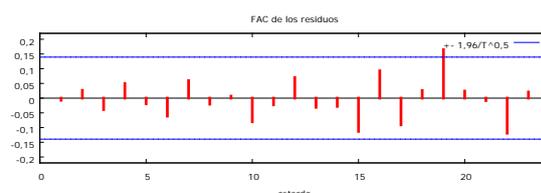
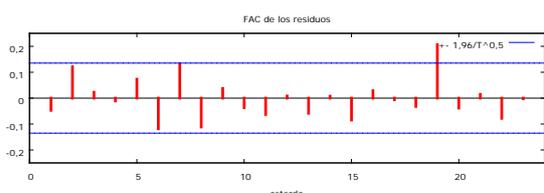
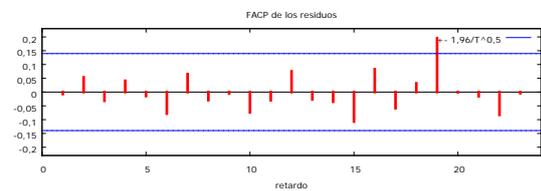
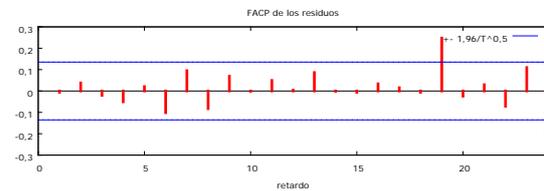
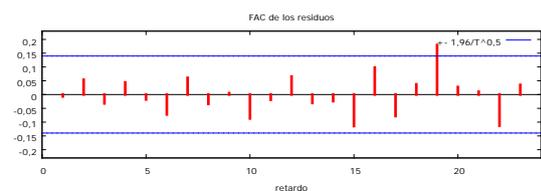
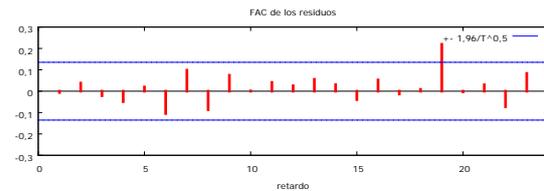
	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	1,071	0,5853
Modelo 2	1,159	0,56004
Modelo 3	3,891	0,14291
Modelo 4	3,310	0,19112





5.b. Media y Correlograma residual:

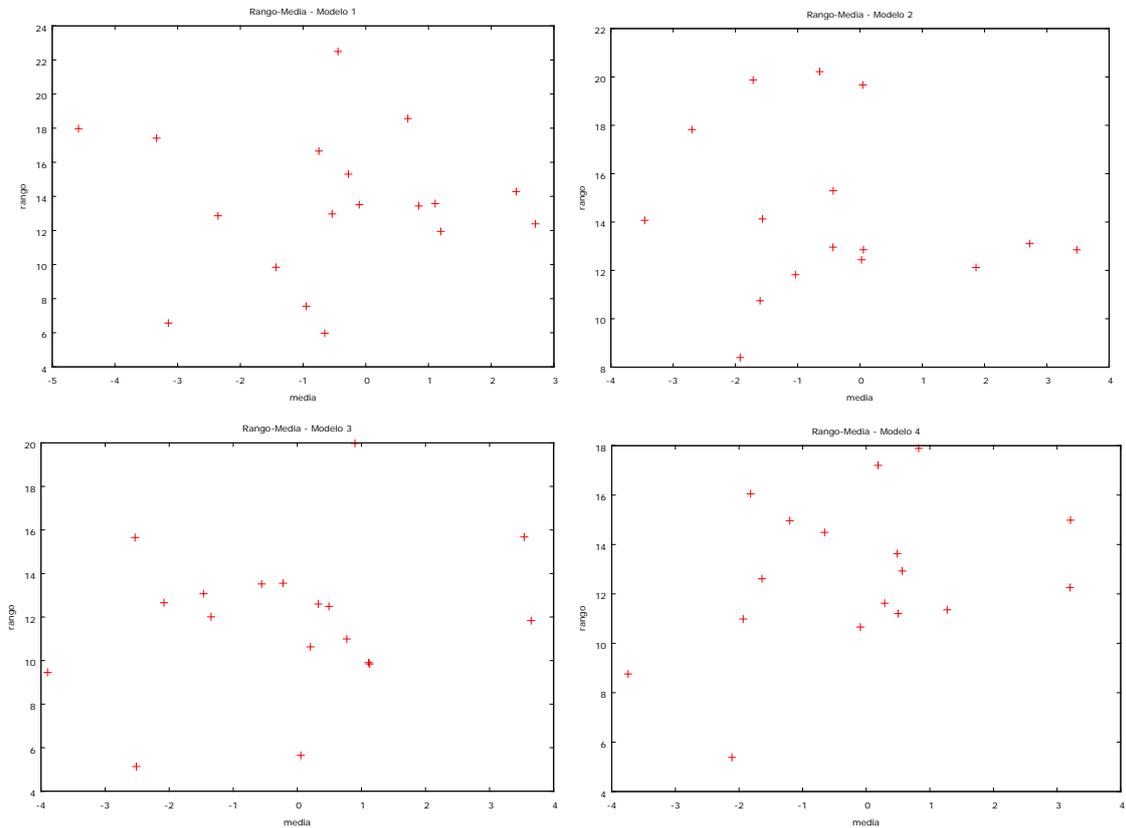
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	-0,404912	6,44207	-0,910845	0,3634
Modelo 2	-0,318313	6,0835	-0,734402	0,4636
Modelo 3	0,0308015	5,70191	0,0782819	0,9377
Modelo 4	0,0318865	5,86595	0,0764893	0,9391



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0246	39,8621	0,302
Modelo 2	-0,0492	42,7488	0,204
Modelo 3	-0,0134	41,3612	0,248
Modelo 4	-0,0425	39,9142	0,300

4.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,00554642	0,992139
Modelo 2	-0,276171	0,578985
Modelo 3	0,413648	0,344841
Modelo 4	0,61518	0,153498



6. Matriz de Covarianzas y Correlación entre coeficientes:

const	theta_1	Theta_1	
0,00216756	0,000179313	0,000578946	const
	0,00245058	-0,000143604	theta_1
		0,00241445	Theta_1

	θ_1	Θ_1
θ_1	1	-0,05903681
Θ_1		1

Matriz de covarianzas del Modelo 2					
const	phi_1	Phi_1	theta_1	Theta_1	
0,000190705	0,000218190	-6,14956e-005	-2,00033e-005	0,000132055	const
	0,00756068	0,000242322	-0,00178637	0,000564816	phi_1
		0,00741473	4,71224e-005	-0,00292396	Phi_1
			0,00125634	-0,000491716	theta_1
				0,00414159	Theta_1

Matriz de correlaciones del Modelo 2				
	φ_1	Φ_1	θ_1	Θ_1
φ_1	1	0,03236419	-0,57961162	0,10093523
Φ_1		1	0,01543925	-0,52764308
θ_1			1	-0,21556435
Θ_1				1

Matriz de Covarianzas del Modelo 3			
const	phi_1	Theta_1	
0,00741808	-0,000485548	-0,000640945	const
	0,00437876	0,000185012	phi_1
		0,00161735	Theta_1

Correlación entre los parámetros estimados del Modelo 3		
	φ_1	Θ_1
φ_1	1	0,06952199
Θ_1		1

Matriz de covarianzas del Modelo 4					
const	phi_1	Phi_1	theta_1	Theta_1	
0,00519585	-0,0131364	-9,17905e-005	0,0142052	-0,000304383	const
	0,109606	0,00265331	-0,116749	-0,00182926	phi_1
		0,00668941	-0,00259969	-0,00187275	Phi_1
			0,128224	0,00209239	theta_1
				0,00290736	Theta_1

Matriz de correlaciones del Modelo 4				
	φ_1	Φ_1	θ_1	Θ_1
φ_1	1	0,09798921	-0,98481255	-0,10247324
Φ_1		1	-0,08876546	-0,42465563
θ_1			1	0,10837019
Θ_1				1

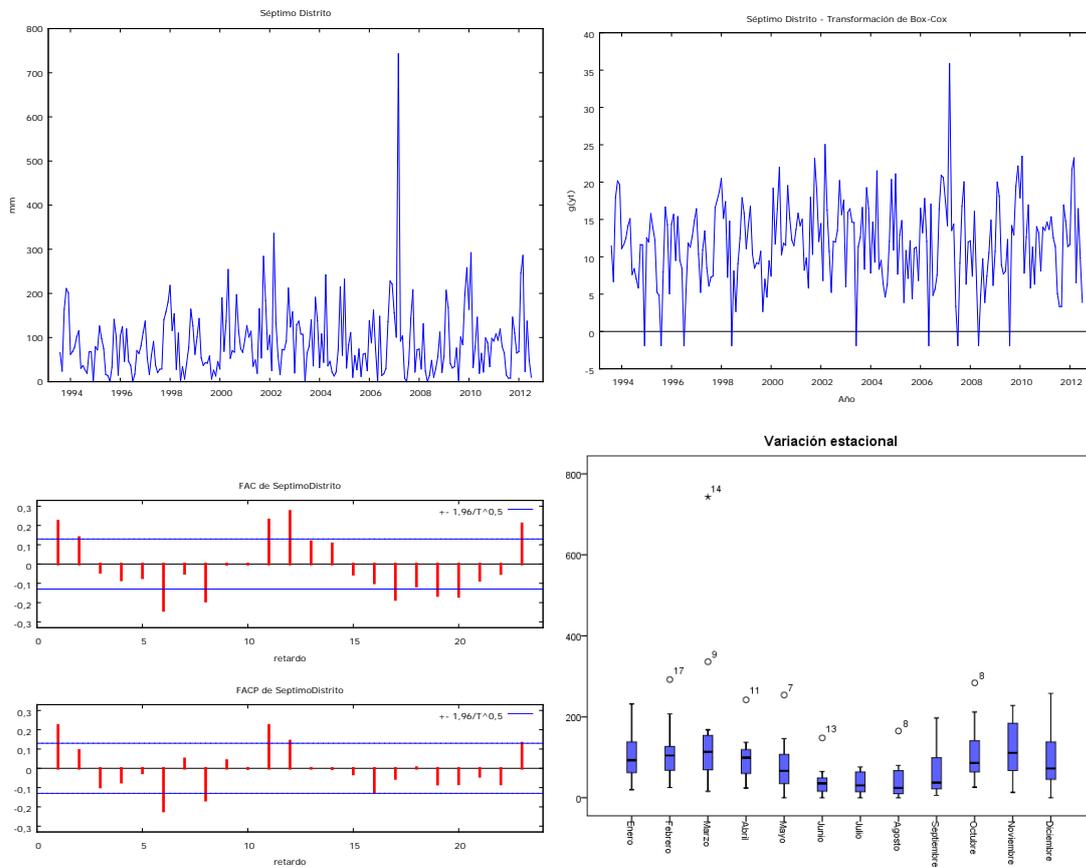
7. Criterios de selección:

	AIC	SBC	HQ
Modelo 1	1386,181	1399,569	1391,593
Modelo 2	1281,999	1301,698	1289,973
Modelo 3	1334,095	1347,483	1339,507
Modelo 4	1273,492	1293,222	1281,478

Conclusión: Los cuatro modelos propuestos incluyen un orden de diferenciación estacional por lo que sugiere el test de Canova y Hansen y el test HEGY bajo criterio AIC y términos trigonométricos, mientras que los modelos 1 y 2 contienen, además, un orden de diferenciación regular sugerida por el contraste HEGY con criterio AIC y constante (aunque el correlograma de la ACF no da indicios de presencia de raíces unitarias). Los cuatro modelos superan las pruebas residuales y de colinealidad. Los criterios de información arrojan los valores más bajos para el Modelo 4. Como puede apreciarse es dudosa la inclusión de un orden de diferenciación regular. En caso de optarse por un modelo con un orden de diferenciación regular, entonces el Modelo 2 es el adecuado entre los propuestos (tener en cuenta, en este punto las consecuencias de la sobrediferenciación).

Apéndice VI: Séptimo Distrito

1. Representaciones gráficas.



2. Modelos propuestos:

Modelo 1: $ARIMA(1,0,2)x(1,0,1)_{12}$

Modelo 2: $ARIMA(1,0,2)x(1,1,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	1,25373	0,726771	1,7251	0,08451	*
phi_1	-0,342006	0,409384	-0,8354	0,40348	
Phi_1	0,920475	0,0427953	21,5088	<0,00001	***
theta_1	0,418898	0,421808	0,9931	0,32066	
theta_2	0,155406	0,0780203	1,9919	0,04639	**
Theta_1	-0,804388	0,0662642	-12,1391	<0,00001	***

Modelo 2					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,22378	0,179616	1,2459	0,21281	
phi_1	-0,325231	0,345044	-0,9426	0,34590	
Phi_1	-0,0837906	0,0888778	-0,9428	0,34580	
theta_1	0,348893	0,351507	0,9926	0,32092	
theta_2	0,142392	0,075422	1,8879	0,05903	*
Theta_1	-0,752449	0,0521087	-14,4400	<0,00001	***

4. Causalidad e Invertibilidad:

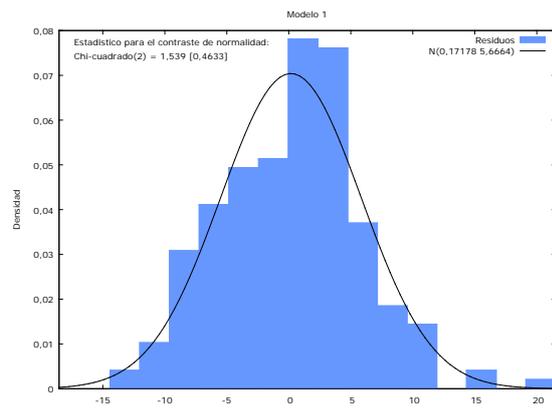
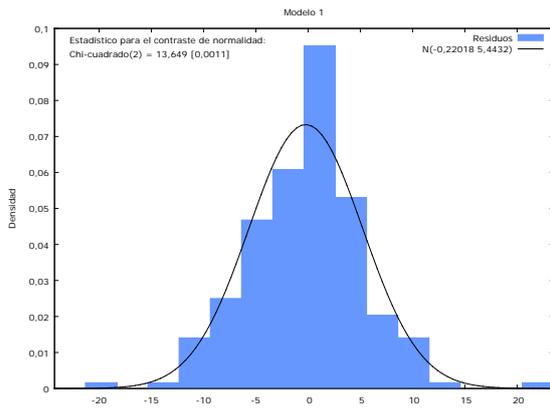
Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	-2,9239	0,0000	2,9239	0,5000
AR (estacional)					
	Raíz 1	1,0864	0,0000	1,0864	0,0000
MA					
	Raíz 1	-1,3478	-2,1490	2,5367	-0,3391
	Raíz 2	-1,3478	2,1490	2,5367	0,3391
MA (estacional)					
	Raíz 1	1,2432	0,0000	1,2432	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	-3,0747	0,0000	3,0747	0,5000
AR (estacional)					
	Raíz 1	-11,9345	0,0000	11,9345	0,5000
MA					
	Raíz 1	-1,2251	-2,3499	2,6501	-0,3265
	Raíz 2	-1,2251	2,3499	2,6501	0,3265
MA (estacional)					
	Raíz 1	1,3290	0,0000	1,3290	0,0000

5. Análisis de los residuos

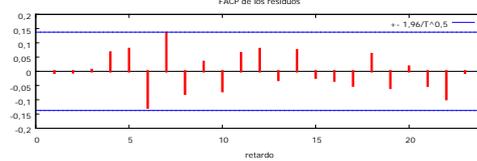
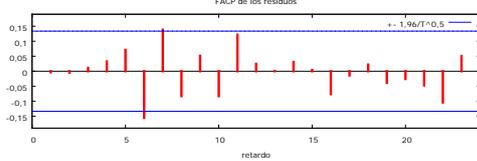
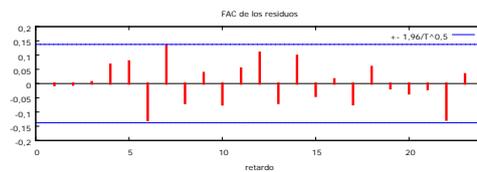
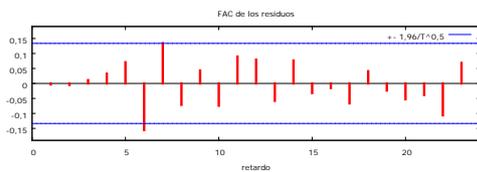
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	13,649	0,00109
Modelo 2	1,539	0,46329



5.b. Media y Correlograma residual:

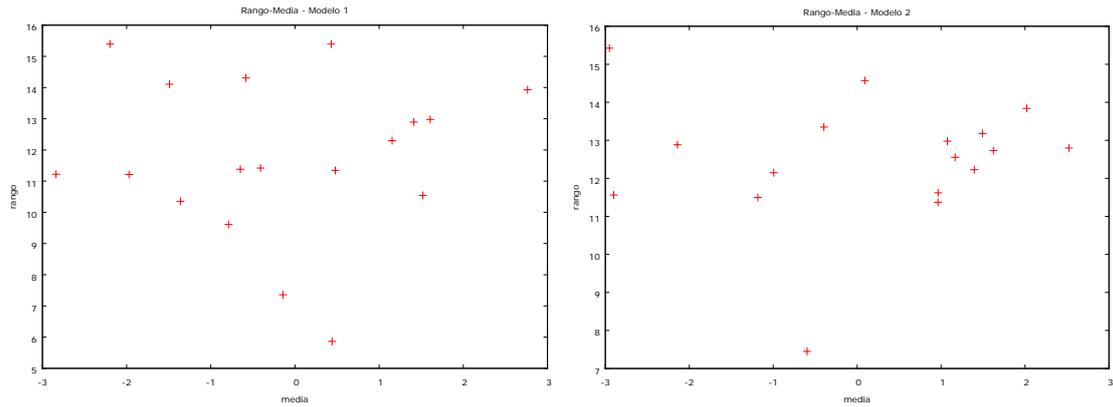
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	-0,220181	5,37926	-0,600173	0,549
Modelo 2	0,171784	5,59581	0,437389	0,6623



	\hat{p}_{36}	Q_{36}	Valor-p
Modelo 1	0,0038	40,2524	0,288
Modelo 2	-0,0025	43,6744	0,178

4.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	0,0348065	0,935178
Modelo 2	0,0175925	0,946565



4.d. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1						
const	phi_1	Phi_1	theta_1	theta_2	Theta_1	
0,528196	-0,0876963	-0,0261179	0,0902492	0,00806202	0,0300557	const
	0,167596	-0,00412461	-0,169540	-0,00515651	0,00351323	phi_1
		0,00183144	0,00407260	-0,000261466	-0,00197797	Phi_1
			0,177922	0,00686144	-0,00435390	theta_1
				0,00608717	4,04304e-005	theta_2
					0,00439095	Theta_1

Correlaciones entre los parámetros del Modelo 1					
	φ_1	Φ_1	θ_1	θ_2	Θ_1
φ_1	1	-0,23542683	-0,98180785	-0,16144232	0,12950805
Φ_1		1	0,22561131	-0,07830897	-0,69750073
θ_1			1	0,20849365	-0,15577031
θ_2				1	0,00782026
Θ_1					1

Matriz de Covarianzas del Modelo 2						
const	phi_1	Phi_1	theta_1	theta_2	Theta_1	
0,0322618	-0,0198551	0,000773492	0,0195835	-0,000364722	-0,000528464	const
	0,119056	-0,00253035	-0,117869	0,00163015	0,000129345	phi_1
		0,00789926	0,00304593	6,29985e-005	-0,00192151	Phi_1
			0,123557	-0,000314155	-0,000969919	theta_1
				0,00568847	-0,000579762	theta_2
					0,00271531	Theta_1

Correlaciones entre los parámetros del Modelo 2					
	φ_1	Φ_1	θ_1	θ_2	Θ_1
φ_1	1	-0,08251118	-0,97183188	0,06264048	0,00719391
Φ_1		1	0,09749731	0,00939807	-0,41489587
θ_1			1	-0,01184983	-0,05295308
θ_2				1	-0,14751679
Θ_1					1

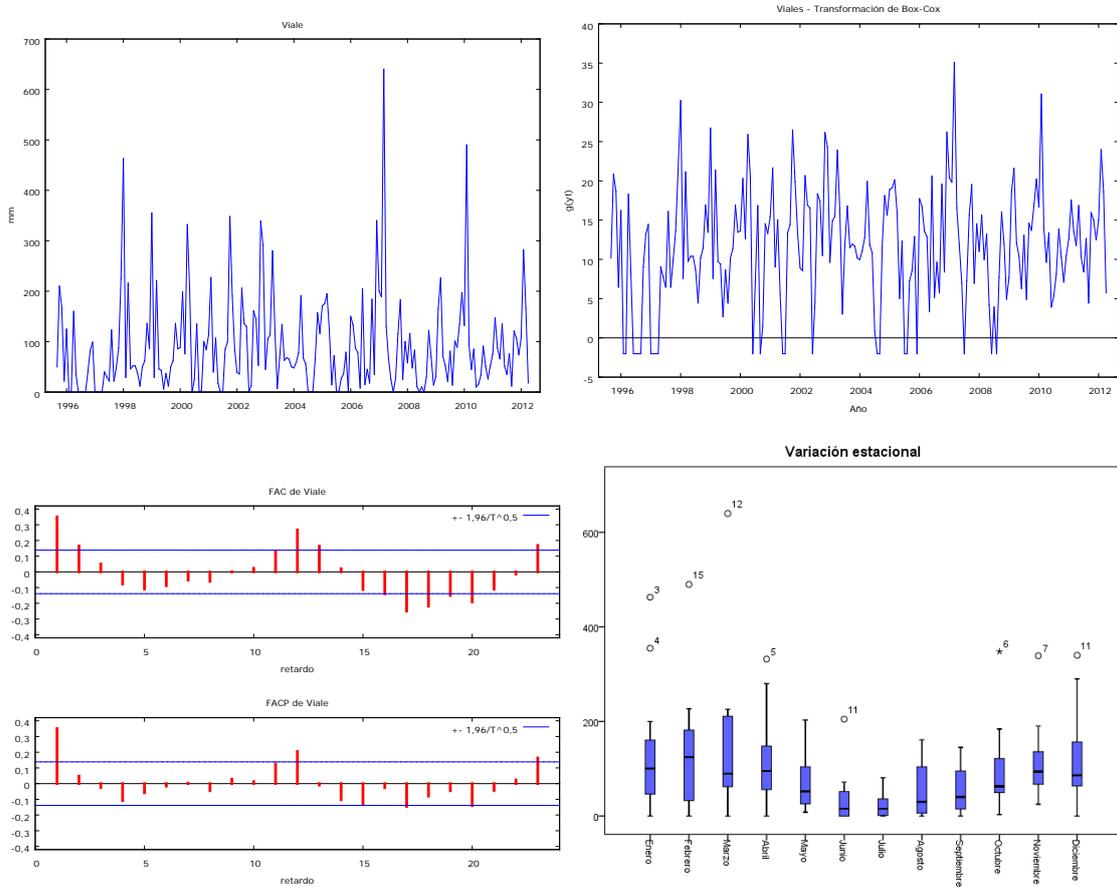
4.c. Criterios de selección:

	AIC	SBC	HQ
Modelo 1	1346,999	1370,594	1356,533
Modelo 2	1288,418	1311,611	1297,801

Conclusión: Claramente el Modelo 2 explica mejor la estructura del proceso generador de los datos, no solo por el análisis residual sino también por los criterios de información. Nótese que existe colinealidad entre los parámetros autorregresivo y de media móvil de la parte regular, sin embargo estos parámetros no son significativos por lo que se pueden eliminar del modelo.

Apéndice VII: Viale

1. Representaciones gráficas.



2. Modelos Propuestos:

Modelo 1: $ARIMA(1,0,0)x(0,1,1)_{12}$

Modelo 2: $ARIMA(1,0,0)x(1,0,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	0,265992	0,209049	1,2724	0,20324	
phi_1	0,222247	0,0747073	2,9749	0,00293	***
Theta_1	-0,687592	0,0523783	-13,1274	<0,00001	***

Modelo 2					
	Coeficiente	Desv. Típica	z	Valor p	
const	5,81185	1,565	3,7136	0,00020	***
phi_1	0,288408	0,0743478	3,8792	0,00010	***
Phi_1	0,324758	0,158168	2,0532	0,04005	**
Theta_1	0,0954281	0,170849	-0,5586	0,57647	

4. Causalidad e invertibilidad:

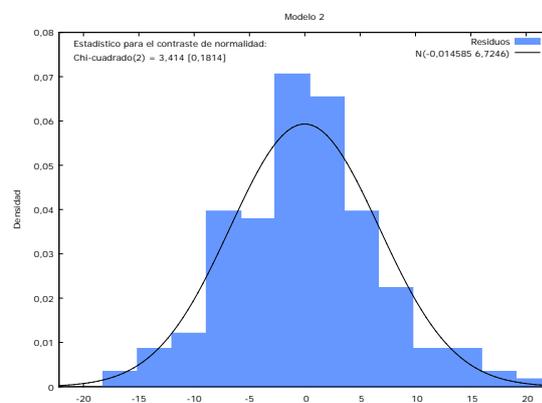
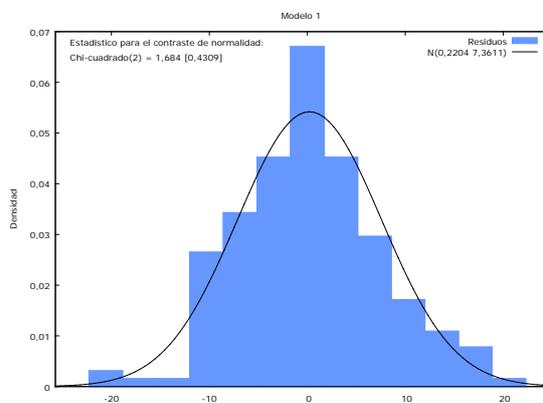
Modelo 1					
		Real	Imaginaria	Módulo	Frecuencia
AR					
	Raíz 1	4,4995	0,0000	4,4995	0,0000
MA (estacional)					
	Raíz 1	1,4544	0,0000	1,4544	0,0000

Modelo 2					
		Real	Imaginaria	Módulo	Frecuencia
AR					
	Raíz 1	3,4673	0,0000	3,4673	0,0000
AR (estacional)					
	Raíz 1	3,0792	0,0000	3,0792	0,0000
MA (estacional)					
	Raíz 1	10,4791	0,0000	10,4791	0,0000

5. Análisis de los residuos

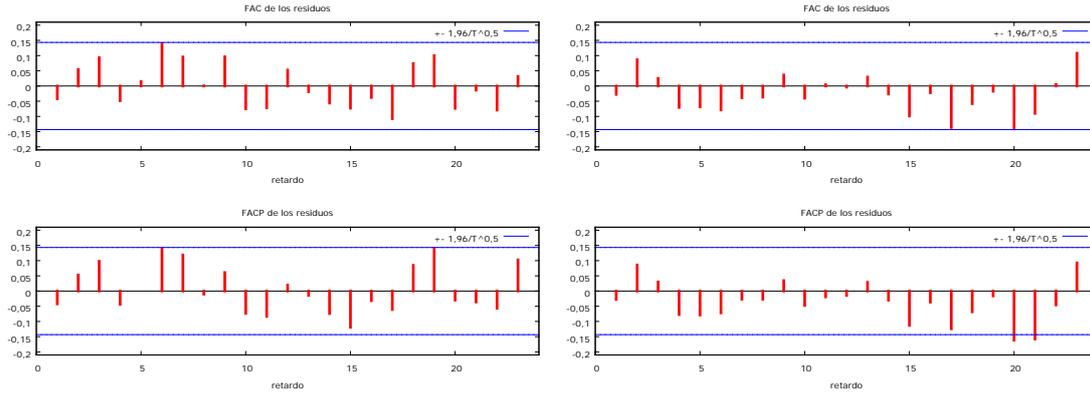
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	1,684	0,43086
Modelo 2	3,414	0,18143



5.b. Media y Correlograma residual:

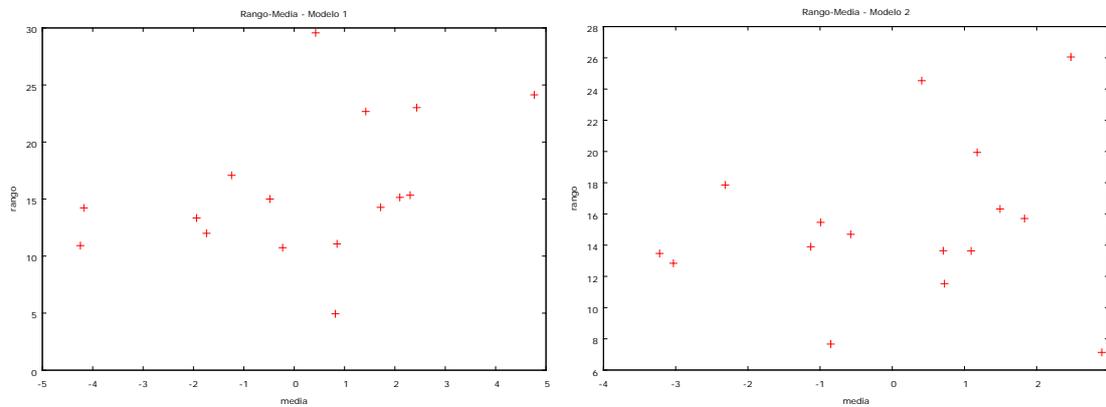
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,220399	7,3214	0,411658	0,6811
Modelo 2	-0,0145851	6,67018	-0,0299015	0,9762



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0249	36,5625	0,443
Modelo 2	0,0707	40,7329	0,270

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	1,05671	0,110901
Modelo 2	0,468738	0,524051



6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de covarianzas del Modelo 1			
const	phi_1	Theta_1	
0,0437017	-0,00228236	-0,00142656	const
	0,00558118	0,000721025	phi_1
		0,00274348	Theta_1

Matriz de correlaciones del Modelo 2		
	φ_1	θ_1
φ_1	1	0,18426204
θ_1		1

Matriz de covarianzas del Modelo 2				
const	phi_1	Phi_1	Theta_1	
2,44922	-0,0448452	-0,214895	0,216532	const
	0,00552760	-0,000174867	0,000712074	phi_1
		0,0250172	-0,0248285	Phi_1
			0,0291895	Theta_1

Matriz de correlaciones del Modelo 2			
	φ_1	Φ_1	θ_1
φ_1	1	-0,01487035	0,0560589
Φ_1		1	-0,91879668
θ_1			1

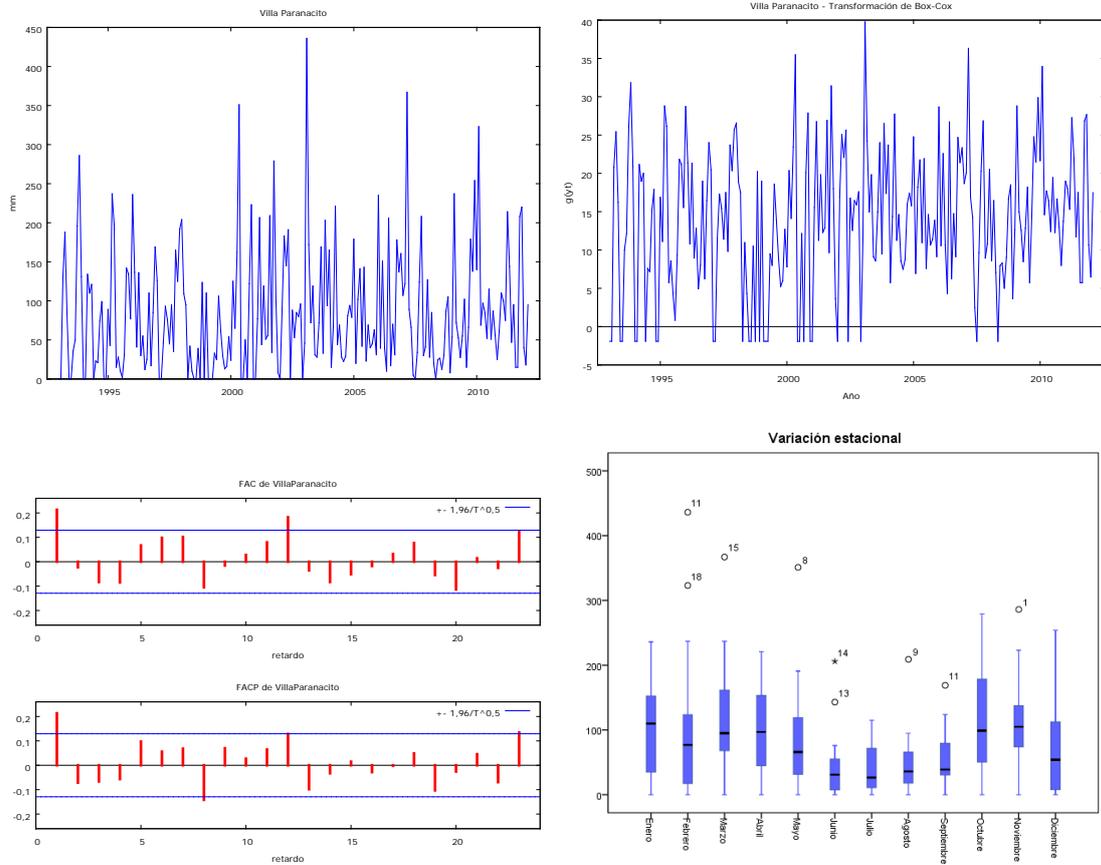
7. Criterios de selección:

	AIC	SBC	HQ
Modelo 1	1282,410	1295,335	1287,647
Modelo 2	1249,401	1265,557	1255,947

Conclusión: El hecho de que el test HEGY y el de Canova y Hansen detectan presencia de raíces unitarias estacionales sugiere que el modelo adecuado es el 1, aunque el Modelo 2 tiene los valores más bajos de los criterios de información. También se consideraron los modelos $ARIMA(1,1,1)x(0,1,1)_{12}$ y $ARIMA(1,1,0)x(0,1,1)_{12}$ pero no mostraron buenos ajustes residuales. En el primer caso los residuos son heterocedásticos y en el segundo no se ajustan a un ruido blanco.

Apéndice VIII: Villa Paranacito

1. Representaciones gráficas.



2. Modelos propuestos.

Modelo 1: $ARIMA(0,0,1)x(1,0,0)_{12}$

Modelo 2: $ARIMA(0,1,1)x(1,0,0)_{12}$

Modelo 3: $ARIMA(1,0,0)x(1,0,0)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	11,1939	1,24498	8,9912	<0,00001	***
Phi_1	0,195294	0,0680065	2,8717	0,00408	***
theta_1	0,193261	0,0666016	2,9018	0,00371	***

Modelo 2					
	Coeficiente	Desv. Típica	z	Valor p	
const	0,0449031	0,0707095	0,6350	0,52540	
Phi_1	0,196819	0,0657344	2,9942	0,00275	***
theta_1	-0,893847	0,0312074	-28,6421	<0,00001	***

Modelo 3					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	9,1651	1,1808	7,7618	<0,00001	***
phi_1	0,197151	0,0667793	2,9523	0,00315	***
Phi_1	0,185279	0,0658586	2,8133	0,00490	***

4. Causalidad e invertibilidad:

Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	5,1205	0,0000	5,1205	0,0000
MA					
	Raíz 1	-5,1743	0,0000	5,1743	0,5000

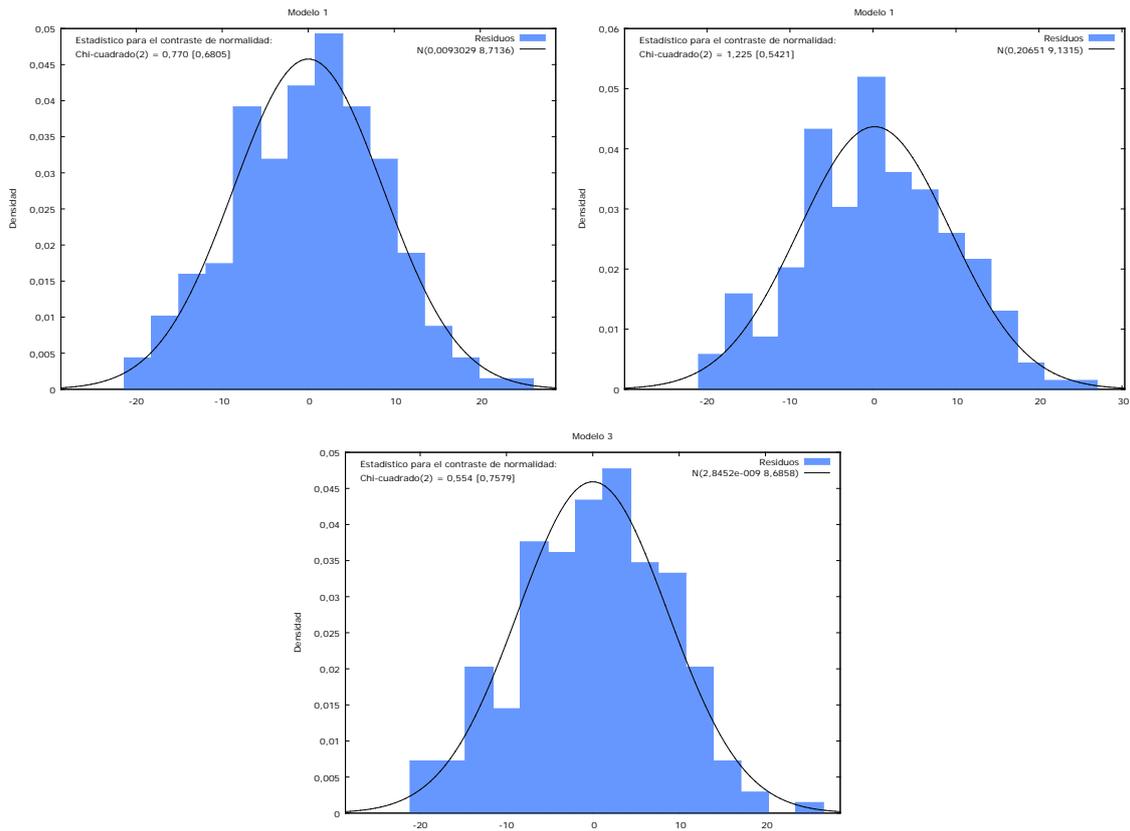
Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	5,0808	0,0000	5,0808	0,0000
MA					
	Raíz 1	1,1188	0,0000	1,1188	0,0000

Modelo 3					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	5,0723	0,0000	5,0723	0,0000
AR (estacional)					
	Raíz 1	5,3973	0,0000	5,3973	0,0000

5. Análisis de los residuos

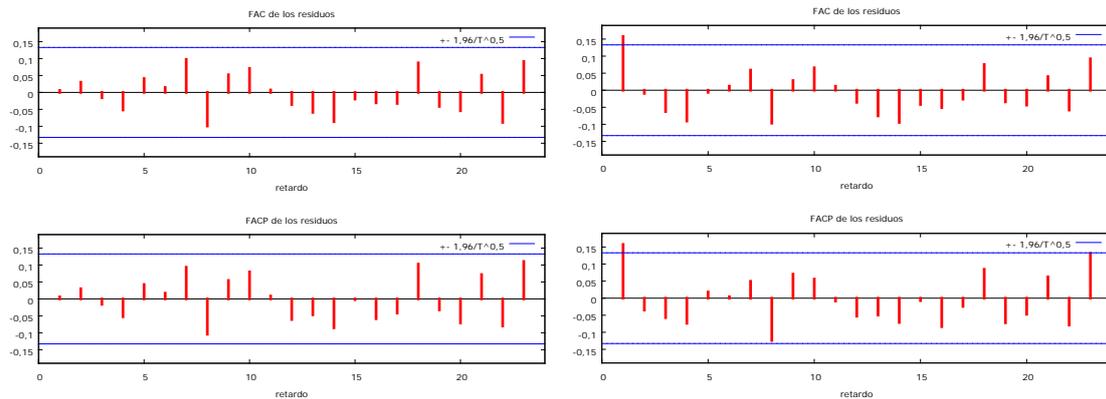
5.a. Normalidad:

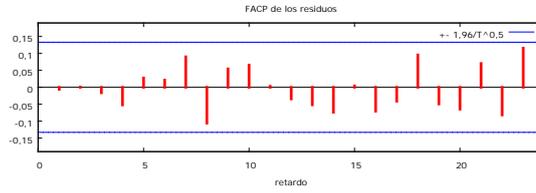
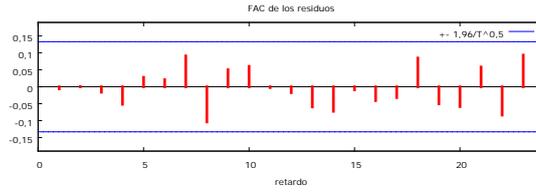
	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	0,769795	0,68052
Modelo 2	1,22462	0,542098
Modelo 3	0,554	0,75789



5.b. Media y Correlograma residual:

	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,00930294	8,6733	0,0158367	0,9874
Modelo 2	0,206507	9,08909	0,334692	0,7382
Modelo 3	0,00000000284525	8,64548	0,00000000484798	1

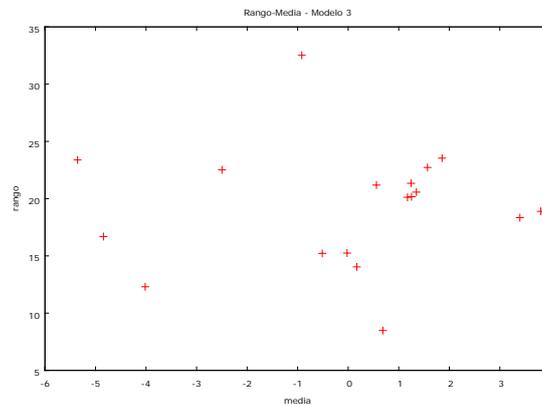
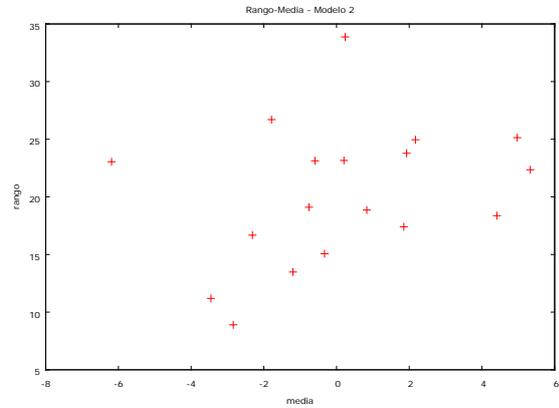
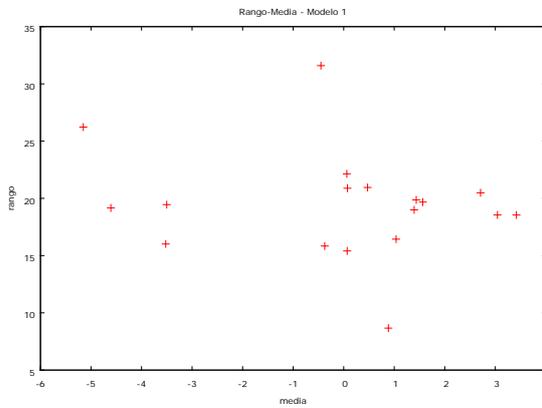




	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0249	35,8992	0,473
Modelo 2	-0,0298	41,5603	0,241
Modelo 3	-0,0184	33,9316	0,567

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,368635	0,427956
Modelo 2	0,606581	0,224753
Modelo 3	0,0730593	0,887784



6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1			
const	Phi_1	theta_1	
1,54998	-0,0696529	-2,57010e-005	const
	0,00462488	-0,000366275	Phi_1
		0,00443577	theta_1

Correlación entre los parámetros estimados del Modelo 1		
	Φ_1	θ_1
Φ_1	1	-0,08086716
θ_1		1

Matriz de Covarianzas del Modelo 2			
const	Phi_1	theta_1	
0,00499983	-0,000621011	-0,000326184	const
	0,00432101	-0,000570302	Phi_1
		0,000973904	theta_1

Correlación entre los parámetros estimados del Modelo 2		
	Φ_1	θ_1
Φ_1	1	-0,27800628
θ_1		1

Matriz de Covarianzas del Modelo 3			
const	phi_1	Phi_1	
1,39429	-0,0479770	-0,0454454	const
	0,00445948	-0,000228360	phi_1
		0,00433735	Phi_1

Correlación entre los parámetros estimados del Modelo 3		
	Φ_1	φ_1
Φ_1	1	-0,05192371
φ_1		1

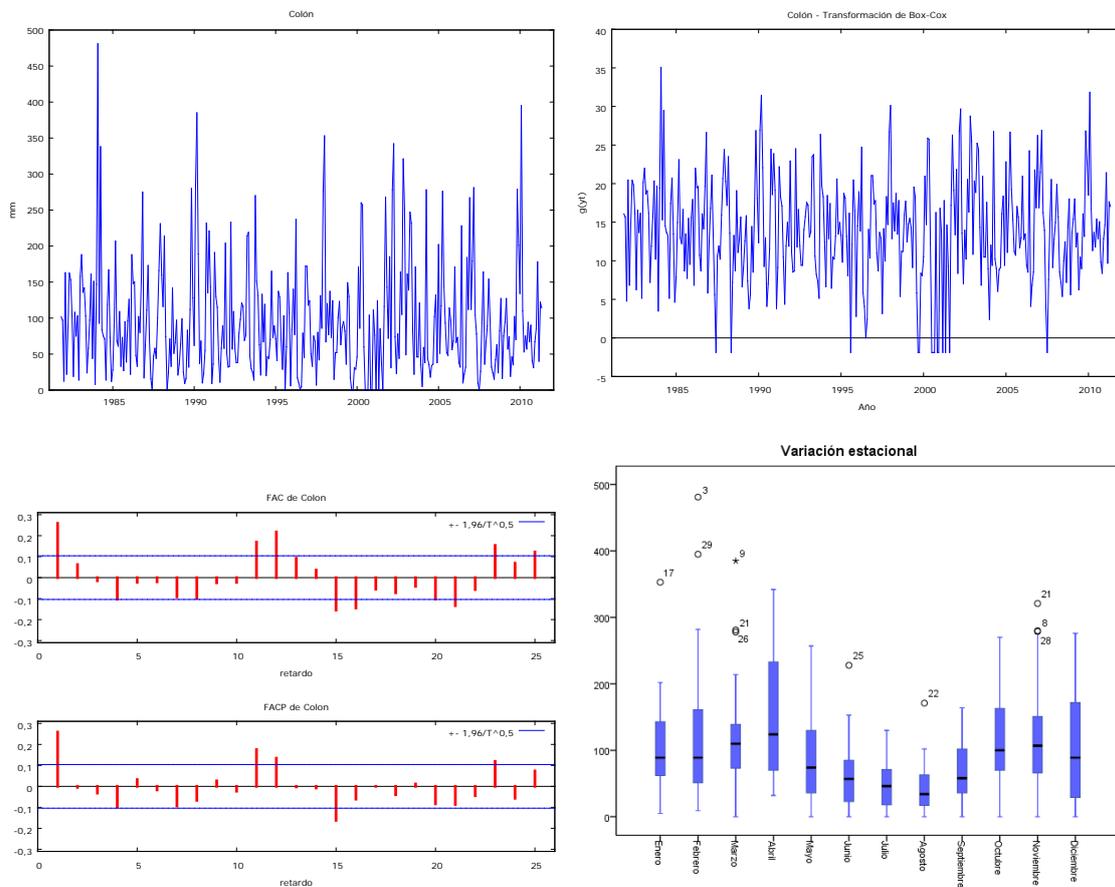
7. Criterios de selección:

	AIC	SBC	HQ
Modelo 1	1567,524	1581,062	1572,992
Modelo 2	1580,800	1594,319	1586,261
Modelo 3	1556,971	1567,110	1561,067

Conclusión: El Modelo 2 se propuso con un orden de diferenciación regular por la posible presencia de raíz unitaria en la frecuencia cero sugerida por los resultados del contraste HEGY. Sin embargo, $0,01 < \text{valor} - p < 0,05$ por lo que la hipótesis nula de raíz unitaria no se rechaza al 1% de significación pero sí al 5%, lo que implica que no es en modo alguno contundente el resultado que ofrece el test. Aunque el modelo supera la mayoría de las pruebas, los residuos no se ajustan a un ruido blanco. Mejor adaptación presentaron los modelos 1 y 3 pero los contrastes favorecen al Modelo 3, por lo que es el que mejor explica el proceso.

Apéndice IX: Colón

1. Representaciones gráficas.



2. Modelos propuestos.

Modelo 1: $ARIMA(1,0,0)x(0,1,1)_{12}$

Modelo 2: $ARIMA(1,0,0)x(1,0,1)_{12}$

Modelo 3: $ARIMA(1,0,0)x(1,1,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	-0,0159831	0,0788842	-0,2026	0,83944	
phi_1	0,158235	0,0526634	3,0047	0,00266	***
Theta_1	-0,821793	0,0316758	-25,9439	<0,00001	***

Modelo 2					
	Coeficiente	Desv. Típica	z	Valor p	
const	14,9032	1,63209	9,1314	<0,00001	***
phi_1	0,27496	0,0520132	5,2863	<0,00001	***
Phi_1	-0,484323	0,0990559	-4,8894	<0,00001	***
Theta_1	0,698817	0,0891395	7,8396	<0,00001	***

Modelo 3					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	-0,0421835	0,073381	-0,5749	0,56539	
phi_1	0,161077	0,0592854	2,7170	0,00659	***
Phi_1	-0,0154802	0,0599386	-0,2583	0,79620	
Theta_1	-0,836472	0,0338442	-24,7154	<0,00001	***

4. Causalidad e invertibilidad:

Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	6,3197	0,0000	6,3197	0,0000
MA (estacional)					
	Raíz 1	1,2169	0,0000	1,2169	0,0000

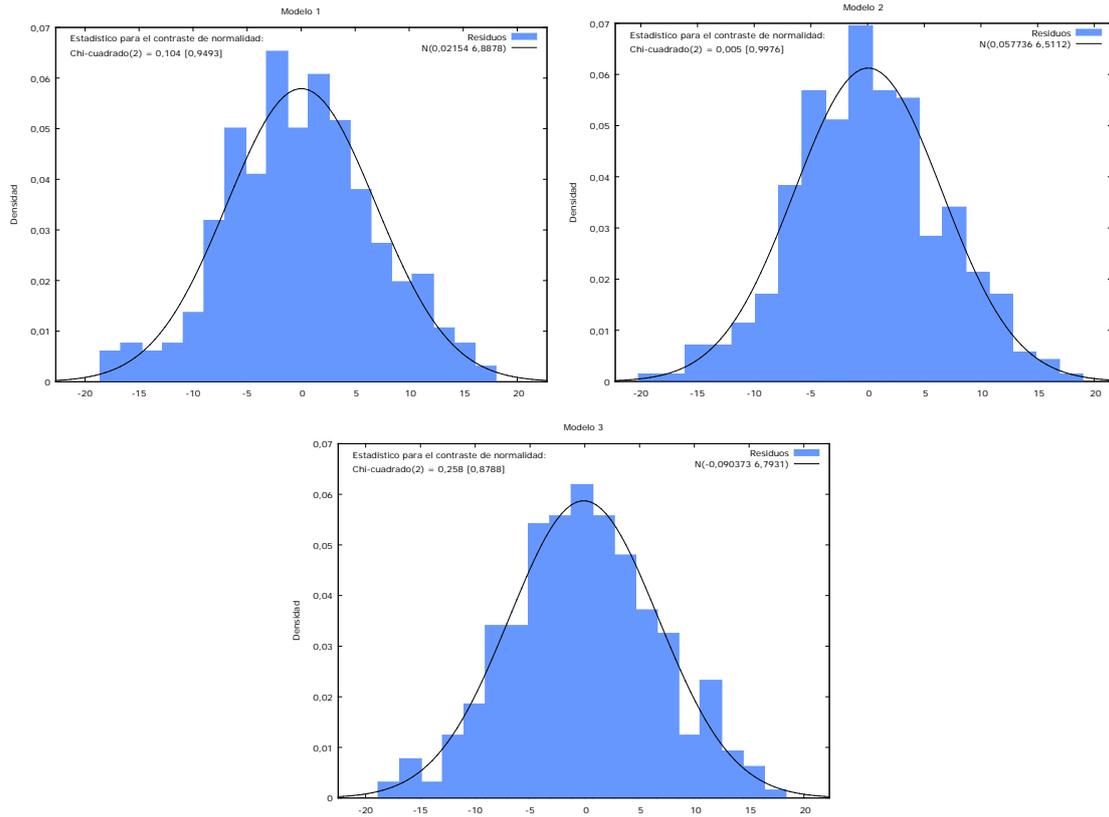
Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	3,6369	0,0000	3,6369	0,0000
AR (estacional)					
	Raíz 1	-2,0647	0,0000	2,0647	0,5000
MA (estacional)					
	Raíz 1	-1,4310	0,0000	1,4310	0,5000

Modelo 3					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	6,2082	0,0000	6,2082	0,0000
AR (estacional)					
	Raíz 1	-64,5985	0,0000	64,5985	0,5000
MA (estacional)					
	Raíz 1	1,1955	0,0000	1,1955	0,0000

5. Análisis de los residuos.

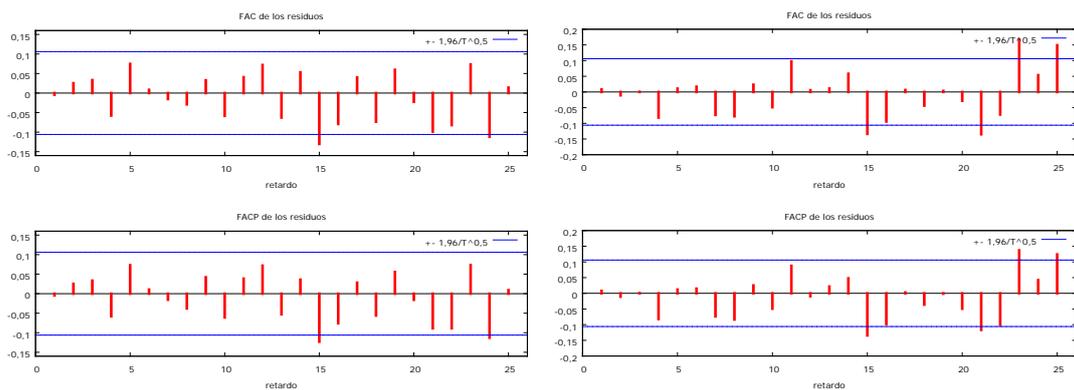
5.a. Normalidad:

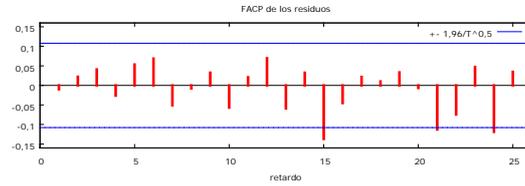
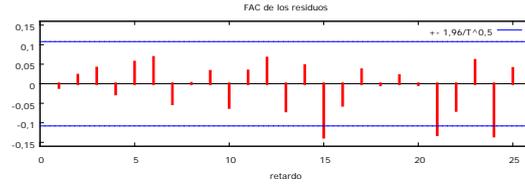
	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	0,104	0,94929
Modelo 2	0,005	0,99760
Modelo 3	0,258	0,87877



5.b. Media y Correlograma residual:

	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,0215401	6,86759	0,0580038	0,9538
Modelo 2	0,0577357	6,48245	0,164709	0,8693
Modelo 3	-0,0903732	6,76208	-0,242782	0,8083

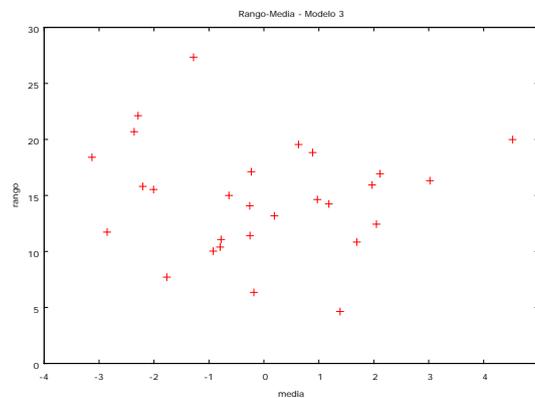
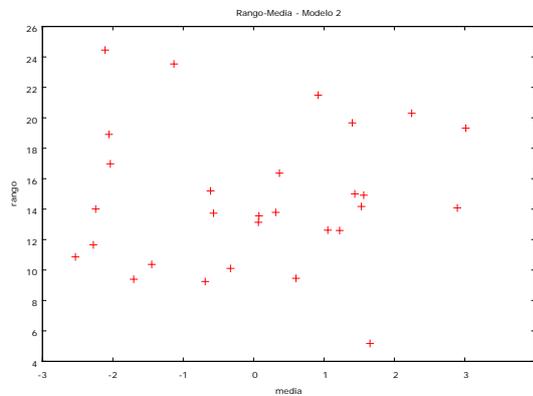
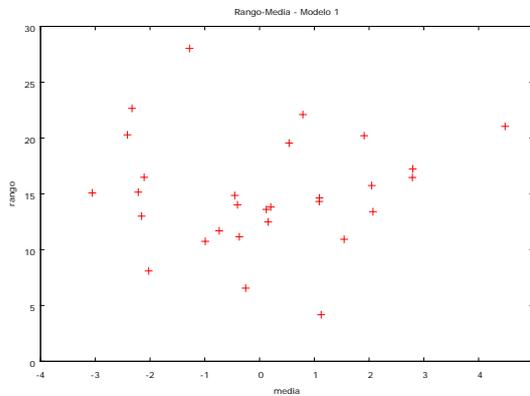




	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	0,0441	55,1957	0,021
Modelo 2	0,0782	69,5856	0,001
Modelo 3	0,0277	54,2286	0,026

4.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	0,113836	0,828889
Modelo 2	0,126661	0,814056
Modelo 3	-0,137144	0,79287



4.d. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1			
const	phi_1	Theta_1	
0,00622272	0,000264662	0,000219124	const
	0,00277344	-2,05824e-006	phi_1
		0,00100336	Theta_1

Correlación entre los parámetros estimados del Modelo 1		
	φ_1	θ_1
φ_1	1	-0,00123384
θ_1		1

Matriz de Covarianzas del Modelo 2				
const	phi_1	Phi_1	Theta_1	
2,66371	-0,0581854	-0,104142	0,0854662	const
	0,00270538	9,21106e-005	-0,000277109	phi_1
		0,00981207	-0,00775073	Phi_1
			0,00794585	Theta_1

Correlación entre los parámetros estimados del Modelo 2			
	φ_1	Φ_1	θ_1
φ_1	1	0,01787787	-0,05976774
Φ_1		1	-0,87779291
θ_1			1

Matriz de Covarianzas del Modelo 3				
const	phi_1	Phi_1	Theta_1	
0,00538478	0,000461752	4,08835e-005	0,000147532	const
	0,00351476	-0,000729632	0,000347379	phi_1
		0,00359264	-0,000811844	Phi_1
			0,00114543	Theta_1

Correlación entre los parámetros estimados del Modelo 3			
	φ_1	Φ_1	θ_1
φ_1	1	-0,20532864	0,17312969
Φ_1		1	-0,40020429
θ_1			1

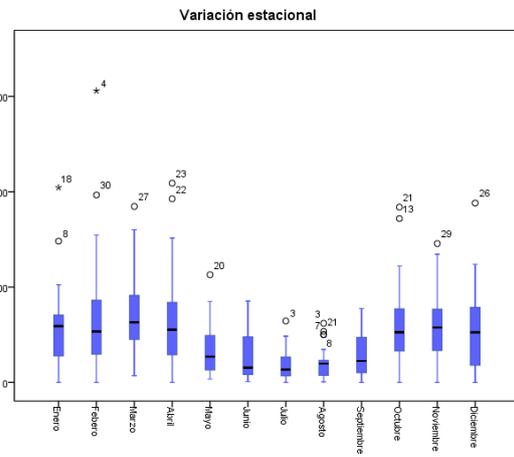
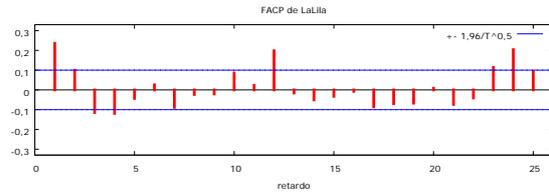
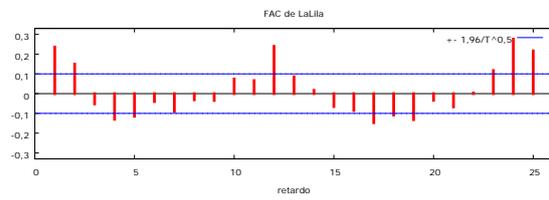
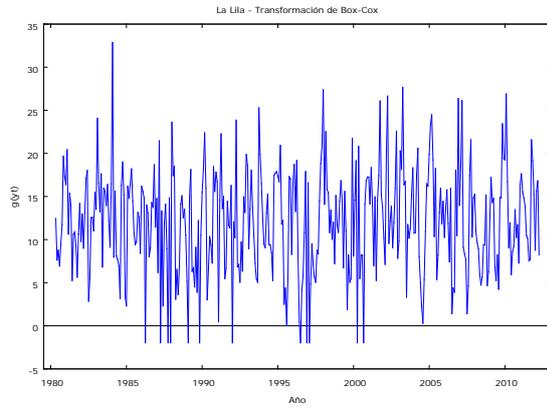
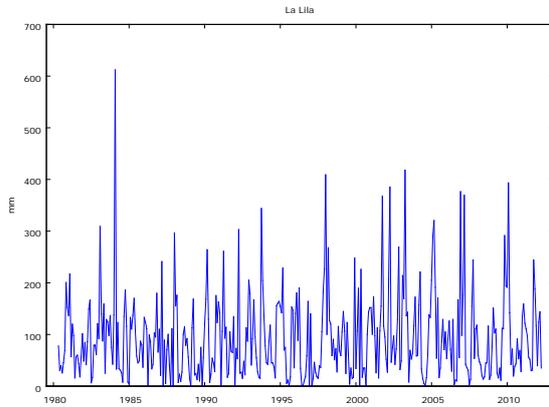
4.c. Criterios de selección:

	AIC	SBC	HQ
Modelo 1	2295,496	2310,835	2301,607
Modelo 2	2258,043	2277,217	2265,682
Modelo 3	2207,035	2226,031	2214,612

Conclusión: Los tres modelos parecen recoger la estructura del proceso generador de datos. Cumplen con las condiciones de causalidad e invertibilidad y el análisis residual muestra buena adaptación aunque los correlogramas de los residuos no se ajustan a un ruido blanco. La matriz de correlaciones tiene alta correlación entre los parámetros estacionales estimados en el Modelo 2, no así en los modelos 1 y 3. Por último, los criterios de información sugieren que el modelo que mejor explica el proceso es el Modelo 3. Los contrastes HEGY y CH no dan evidencias claras de presencia de raíces unitarias (aunque bajo ciertas restricciones sí lo hace el test HEGY).

Apéndice X: La Lila

1. Representaciones gráficas.



2. Modelos propuestos.

Modelo 1: $ARIMA(0,0,2)x(0,1,1)_{12}$

Modelo 2: $ARIMA(1,0,2)x(1,0,1)_{12}$

Modelo 3: $ARIMA(1,0,2)x(1,1,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coefficiente	Desv. Típica	z	Valor p	
const	-0,00821116	0,0604521	-0,1358	0,89196	
theta_1	0,100659	0,0510562	1,9715	0,04866	**
theta_2	0,159253	0,0503317	3,1641	0,00156	***
Theta_1	-0,879052	0,0275341	-31,9259	<0,00001	***

Modelo 2					
	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	2,46492	1,17011	2,1066	0,03516	**
phi_1	0,0194667	0,289878	0,0672	0,94646	
Phi_1	0,789857	0,0762361	10,3607	<0,00001	***
theta_1	0,131086	0,286398	0,4577	0,64716	
theta_2	0,179976	0,0626089	2,8746	0,00405	***
Theta_1	-0,607607	0,0969834	-6,2651	<0,00001	***

Modelo 3					
	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,043819	0,0655296	0,6687	0,50369	
phi_1	0,108857	0,385306	0,2825	0,77754	
Phi_1	0,028278	0,0624611	0,4527	0,65074	
theta_1	-0,00681973	0,387265	-0,0176	0,98595	
theta_2	0,110075	0,0635756	1,7314	0,08338	*
Theta_1	-0,86073	0,0319638	-26,9283	<0,00001	***

4. Causalidad e invertibilidad:

Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
MA					
	Raíz 1	-0,3160	-2,4858	2,5059	-0,2701
	Raíz 2	-0,3160	2,4858	2,5059	0,2701
MA (estacional)					
	Raíz 1	1,1376	0,0000	1,1376	0,0000

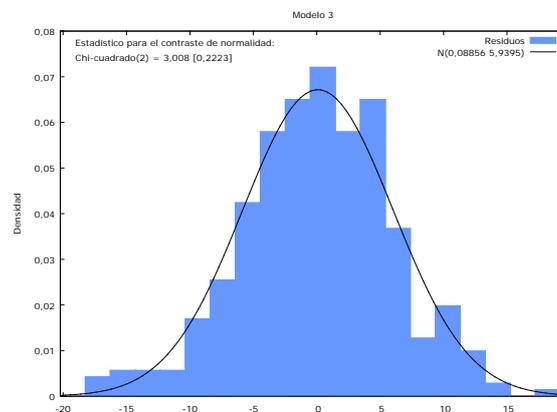
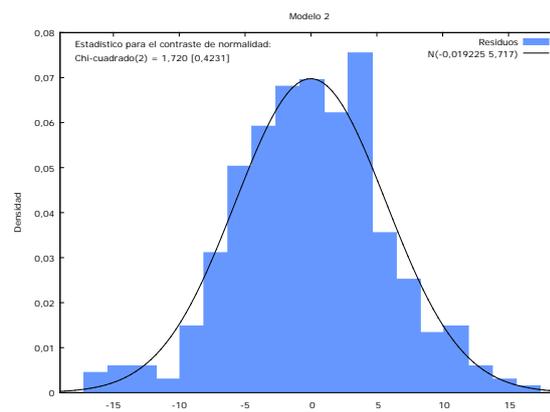
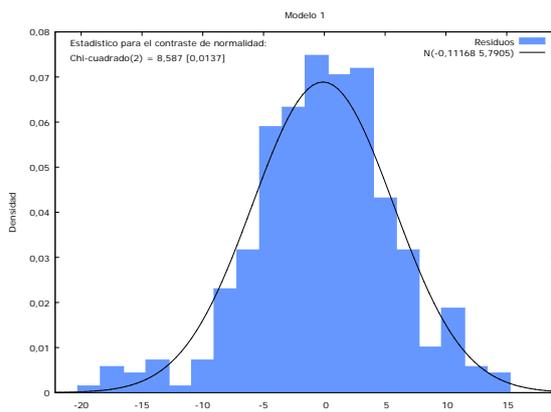
Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	51,3696	0,0000	51,3696	0,0000
AR (estacional)					
	Raíz 1	1,2661	0,0000	1,2661	0,0000
MA					
	Raíz 1	-0,3642	-2,3289	2,3572	-0,2747
	Raíz 2	-0,3642	2,3289	2,3572	0,2747
MA (estacional)					
	Raíz 1	1,6458	0,0000	1,6458	0,0000

Modelo 3					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	9,1863	0,0000	9,1863	0,0000
AR (estacional)					
	Raíz 1	35,3632	0,0000	35,3632	0,0000
MA					
	Raíz 1	0,0310	-3,0139	3,0141	-0,2484
	Raíz 2	0,0310	3,0139	3,0141	0,2484
MA (estacional)					
	Raíz 1	1,1618	0,0000	1,1618	0,0000

5. Análisis de los residuos.

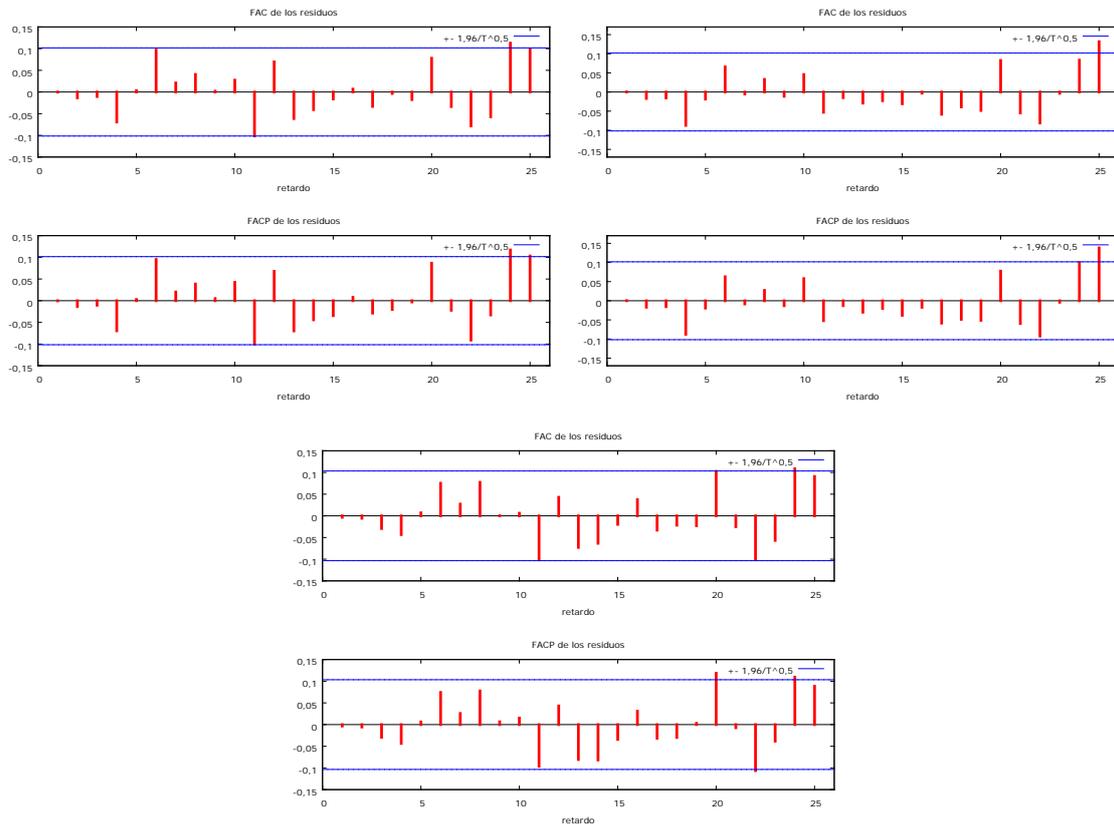
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	8,587	0,01365
Modelo 2	1,720	0,42306
Modelo 3	3,008	0,22229



5.b. Media y Correlograma residual:

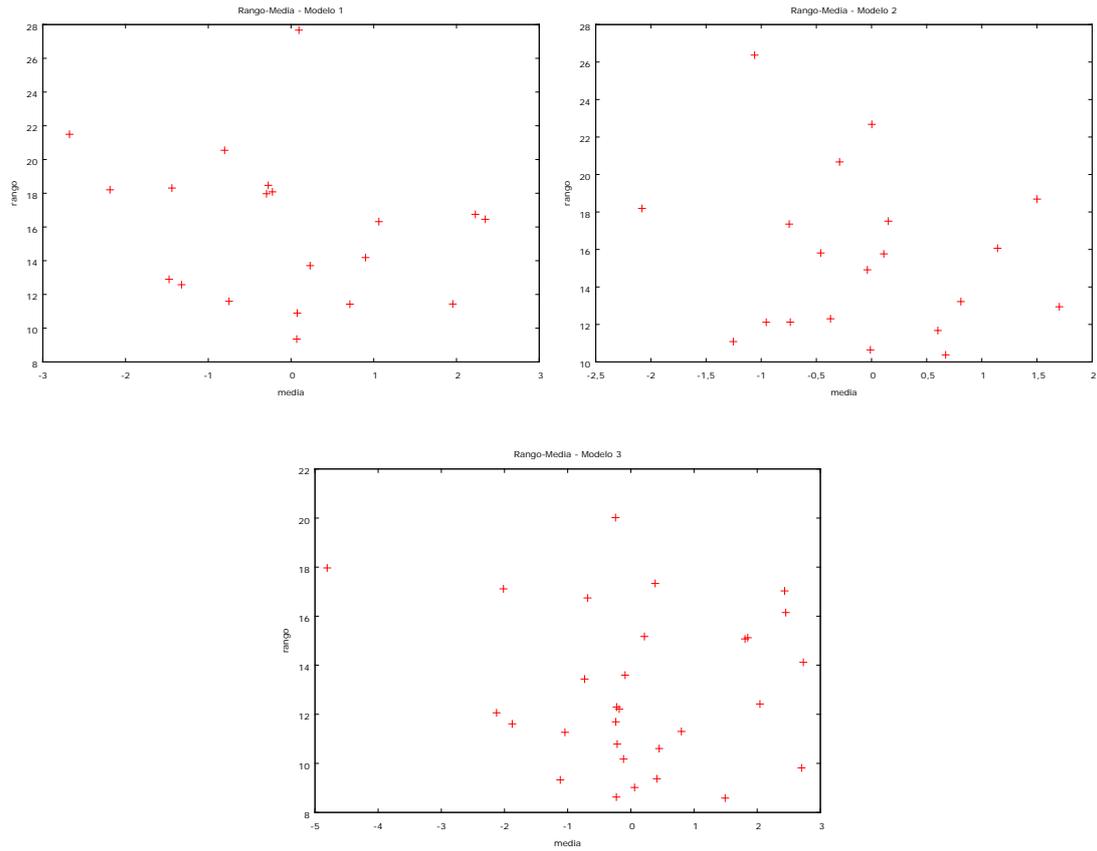
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	-0,111677	5,767	-0,373496	0,709
Modelo 2	-0,0192253	5,67827	-0,0652146	0,948
Modelo 3	0,0885603	5,89787	0,284506	0,7762



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0701	39,4713	0,318
Modelo 2	-0,0841	43,2886	0,188
Modelo 3	-0,0802	41,0444	0,259

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,748904	0,323746
Modelo 2	-0,793091	0,457124
Modelo 3	-0,176651	0,625223



4.d. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1				
const	theta_1	theta_2	Theta_1	
0,00365446	-4,83946e-005	0,000191169	0,000323811	const
	0,00260674	0,000426421	0,000139160	theta_1
		0,00253328	-0,000387333	theta_2
			0,000758129	Theta_1

Correlación entre los parámetros estimados del Modelo 1			
	θ_1	θ_2	θ_1
θ_1	1	0,16593901	0,09899085
θ_2		1	-0,27949369
θ_1			1

Matriz de Covarianzas del Modelo 2						
const	phi_1	Phi_1	theta_1	theta_2	Theta_1	
1,36917	-0,199795	-0,0690432	0,198525	0,0313687	0,0786155	const
	0,0840290	-0,000819254	-0,0817312	-0,0102261	0,000786189	phi_1
		0,00581194	0,000389010	-0,000404254	-0,00658522	Phi_1
			0,0820240	0,0105636	-0,000234734	theta_1
				0,00391988	0,000255532	theta_2
					0,00940577	Theta_1

Correlación entre los parámetros estimados del Modelo 2					
	φ_1	Φ_1	θ_1	θ_2	Θ_1
φ_1	1	-0,03707171	-0,98447034	-0,56345431	0,02796497
Φ_1		1	0,01781682	-0,08469496	-0,89066054
θ_1			1	0,58912288	-0,00845101
θ_2				1	0,0420835
Θ_1					1

Matriz de Covarianzas del Modelo 3						
const	phi_1	Phi_1	theta_1	theta_2	Theta_1	
0,00429412	-0,00757094	0,000320165	0,00732353	0,00106083	-0,000245818	const
	0,148461	0,00290322	-0,147804	-0,0143544	-0,000984317	phi_1
		0,00390139	-0,00267025	-0,000321995	-0,000907639	Phi_1
			0,149974	0,0145754	0,00111921	theta_1
				0,00404185	1,64285e-005	theta_2
					0,00102169	Theta_1

Correlación entre los parámetros estimados del Modelo 3					
	φ_1	Φ_1	θ_1	θ_2	Θ_1
φ_1	1	0,12063256	-0,99054034	-0,58598811	-0,07992282
Φ_1		1	-0,1103911	-0,08108659	-0,45461642
θ_1			1	0,59200008	0,09041592
θ_2				1	0,00808442
Θ_1					1

4.c. Criterios de selección:

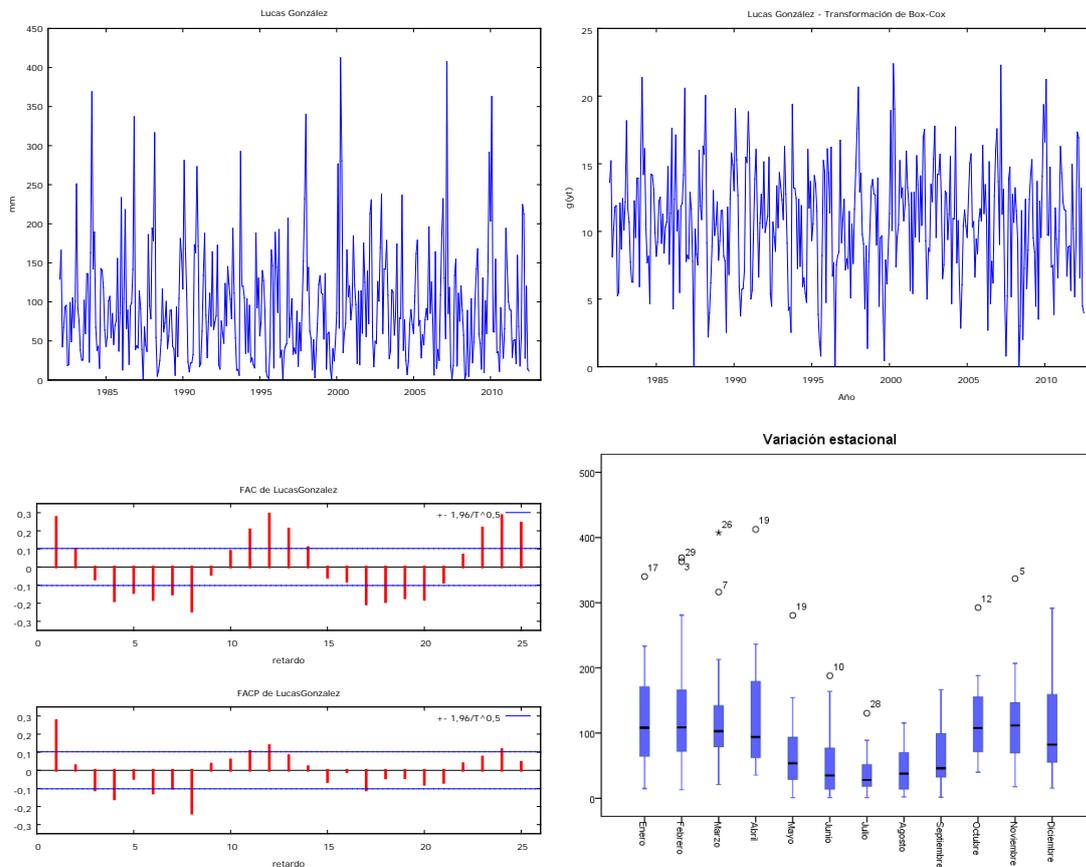
	AIC	SBC	HQ
Modelo 1	2368,430	2388,025	2376,212
Modelo 2	2354,447	2381,860	2365,334
Modelo 3	2306,035	2333,218	2316,844

Conclusión: Los criterios de selección toman los valores más bajos para el Modelo 3. Este modelo tiene correlación cercana a 1 entre los parámetros autorregresivo y de media móvil de la parte regular, sin embargo éstos no son significativos. En el análisis residual, existen algunos valores de la ACF de los residuos que alcanzan las barras de confianza en el correlograma, aunque esto no es suficiente para decir que no se trata de un ruido blanco. Por otro lado, el Modelo 2 también muestra buen ajuste residual y cumplimiento de las condiciones de causalidad e invertibilidad. Sin embargo, tiene alta correlación entre sus parámetros estacionales y éstos son significativos y no pueden eliminarse simplemente del modelo. El Modelo 3 incluye un orden de diferenciación aunque el test de Canova-Hansen no detecta presencia de raíces unitarias y el test HEGY sólo lo hace con constante y criterio AIC. Por lo que

ambos modelos, 2 y 3, son portadores de un riesgo, el de colinealidad y el de sobrediferenciación, respectivamente.

Apéndice XI: Lucas González

1. Representaciones gráficas.



2. Modelos propuestos.

Modelo 1: $ARIMA(0,0,0)x(0,1,1)_{12}$

Modelo 2: $ARIMA(0,0,1)x(0,0,2)_{12}$

Modelo 3⁶⁵: $ARIMA(0,0,8)x(1,0,1)_{12}$

2. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	-0,0047778	0,0364311	-0,1311	0,89566	
Theta_1	-0,854797	0,0291063	-29,3681	<0,00001	***

⁶⁵ Este modelo surge de una reformulación basada en que el correlograma de la ACF residual de los modelos 1 y 2 tiene el retardo 8 diferente de cero. Esto se repitió en otros modelos que no se incluyeron en este análisis por no presentar mejor ajuste que los sugeridos.

Modelo 2					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	10,624	0,33125	32,0725	<0,00001	***
theta_1	0,178726	0,0511014	3,4975	0,00047	***
Theta_1	0,197083	0,0564779	3,4896	0,00048	***
Theta_2	0,162234	0,0556352	2,9160	0,00355	***

Modelo 3					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	1,34772	0,590283	2,2832	0,02242	**
Phi_1	0,872628	0,0551739	15,8159	<0,00001	***
theta_1	0,122501	0,0547769	2,2364	0,02533	**
theta_2	0,0233794	0,0546147	0,4281	0,66859	
theta_3	-0,0157237	0,0548431	-0,2867	0,77434	
theta_4	-0,113309	0,0563288	-2,0116	0,04427	**
theta_5	0,0302438	0,0555947	0,5440	0,58644	
theta_6	-0,0223494	0,0551121	-0,4055	0,68509	
theta_7	0,0346446	0,054569	0,6349	0,52551	
theta_8	-0,143762	0,0560141	-2,5665	0,01027	**
Theta_1	-0,683227	0,0872796	-7,8280	<0,00001	***

3. Causalidad e invertibilidad:

Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
MA (estacional)					
	Raíz 1	1,1699	0,0000	1,1699	0,0000

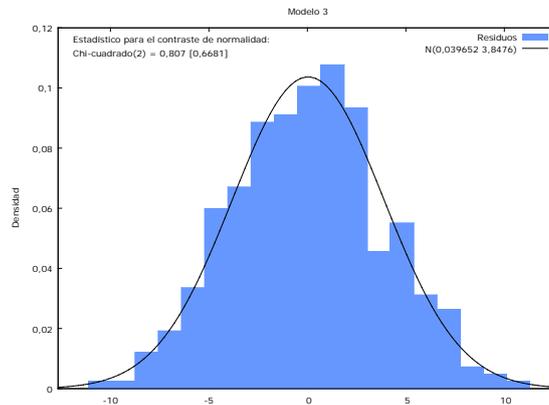
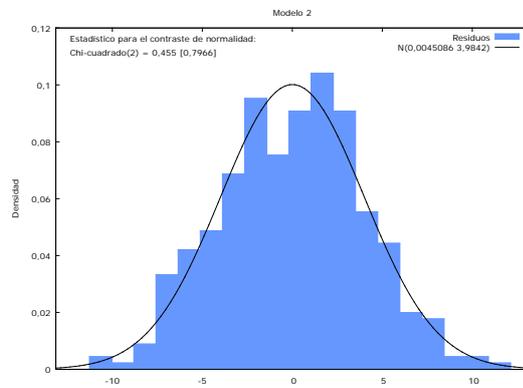
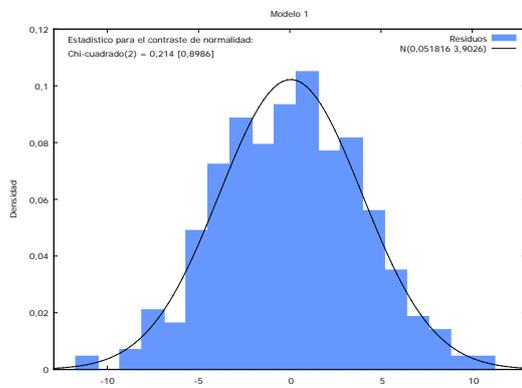
Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
MA					
	Raíz 1	-5,5952	0,0000	5,5952	0,5000
MA (estacional)					
	Raíz 1	-0,6074	-2,4073	2,4827	-0,2893
	Raíz 2	-0,6074	2,4073	2,4827	0,2893

Modelo 3					
		Real	Imaginaria	Módulo	Frecuencia
AR (estacional)					
	Raíz 1	1,1460	0,0000	1,1460	0,0000
MA					
	Raíz 1	-1,1642	0,0000	1,1642	0,5000
	Raíz 2	1,2836	0,0000	1,2836	0,0000
	Raíz 3	0,9609	-0,9550	1,3548	-0,1245
	Raíz 4	0,9609	0,9550	1,3548	0,1245
	Raíz 5	-0,8796	-0,9389	1,2865	-0,3698
	Raíz 6	-0,8796	0,9389	1,2865	0,3698
	Raíz 7	-0,0205	-1,2377	1,2378	-0,2526
	Raíz 8	-0,0205	1,2377	1,2378	0,2526
MA (estacional)					
	Raíz 1	1,4636	0,0000	1,4636	0,0000

4. Análisis de los residuos

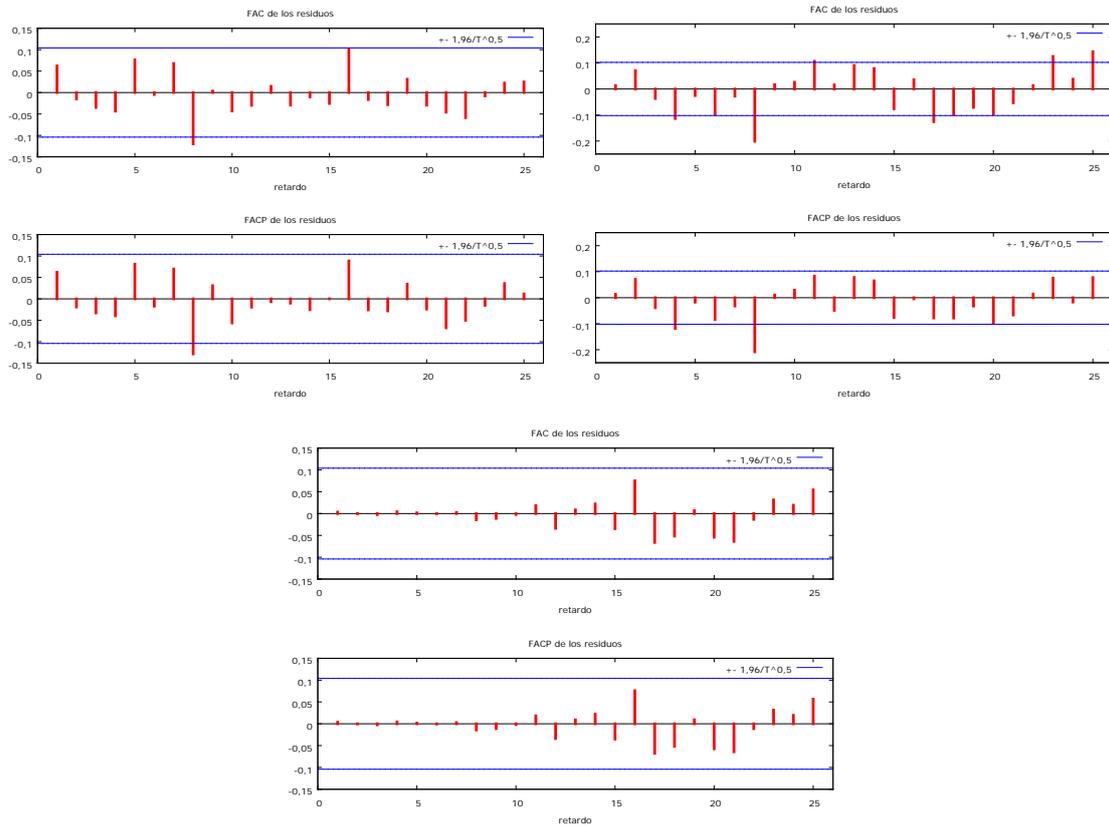
4.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	0,214	0,89860
Modelo 2	0,455	0,79657
Modelo 3	0,807	0,66807



5.b. Media y Correlograma residual:

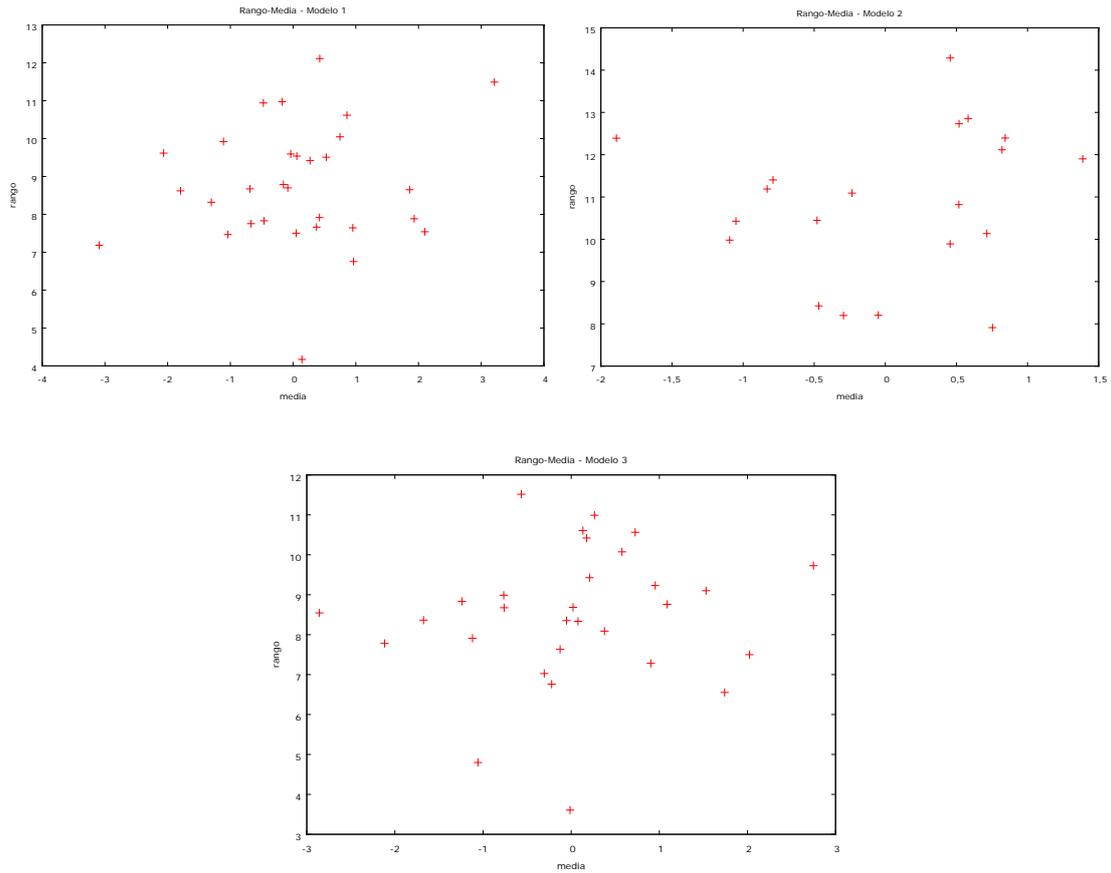
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,0518158	3,89711	0,250515	0,8023
Modelo 2	0,0396516	3,79291	0,196971	0,844
Modelo 3	0,00450859	3,96787	0,0217679	0,9826



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0073	27,3864	0,848
Modelo 2	0,1338	98,0695	0,000
Modelo 3	0,0172	16,1180	0,998

4.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	0,168344	0,48018
Modelo 2	0,317605	0,521006
Modelo 3	0,168189	0,533062



4.d. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1		
const	Theta_1	
0,00132722	2,47558e-005	const
	0,000847175	Theta_1

Matriz de Covarianzas del Modelo 2				
const	theta_1	Theta_1	Theta_2	
0,109727	-0,000921400	-0,000819977	0,000114960	const
	0,00261135	0,000136936	-0,000450984	theta_1
		0,00318975	0,000657931	Theta_1
			0,00309528	Theta_2

Correlación entre los parámetros estimados del Modelo 2			
	θ_1	θ_1	θ_2
θ_1	1	0,04744673	-0,15862757
θ_1		1	0,21255976
θ_2			1

Matriz de Covarianzas del Modelo 3					
const	Phi_1	theta_1	theta_2	theta_3	
0,348434	-0,0323771	0,00500630	0,00514973	-0,00199094	const
	0,00304416	-0,000465041	-0,000502401	0,000187331	Phi_1
		0,00300051	0,000404163	-0,000169317	theta_1
			0,00298276	0,000342515	theta_2
				0,00300776	theta_3

Matriz de Covarianzas del Modelo 3 (continuación)						
theta_4	theta_5	theta_6	theta_7	theta_8	Theta_1	
-0,00278712	-0,00689862	0,00130243	0,00276470	-0,0100654	0,0438174	const
0,000240556	0,000654365	-0,000113424	-0,000239163	0,000927348	-0,00410505	Phi_1
0,000328870	-0,000414006	5,57521e-006	-0,000119756	8,11329e-005	0,000643758	theta_1
0,000141390	0,000238969	-8,59220e-005	0,000390447	-4,42409e-005	0,000567328	theta_2
0,000360588	-0,000177883	6,26126e-005	-0,000405303	0,000327812	-0,000327180	theta_3
0,00317293	0,000169332	-4,39351e-005	-0,000168126	-0,000107063	-0,000510675	theta_4
	0,00309077	0,000305325	-0,000101640	9,70975e-005	-0,000667862	theta_5
		0,00303735	0,000648605	-0,000223917	0,000652755	theta_6
			0,00297778	-6,52536e-005	0,000680000	theta_7
				0,00313758	-0,00172174	theta_8
					0,00761772	Theta_1

Matriz de Correlación del Modelo 3										
	Φ_1	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8	θ_1
Φ_1	1	-0,15387218	-0,16672751	0,06190903	0,07740197	0,21333052	-0,03730133	-0,07943544	0,30006248	-0,8524561
θ_1		1	0,1350982	-0,05636132	0,10658505	-0,13594893	0,00184679	-0,04006396	0,02644248	0,13465189
θ_2			1	0,11435313	0,04595985	0,07870433	-0,02854617	0,13101063	-0,01446162	0,1190178
θ_3				1	0,11672363	-0,0583417	0,02071537	-0,13542905	0,10671009	-0,06835212
θ_4					1	0,05407234	-0,01415253	-0,05469636	-0,03393216	-0,10387268
θ_5						1	0,09965108	-0,03350312	0,03118008	-0,13763871
θ_6							1	0,21566881	-0,07253419	0,13570333
θ_7								1	-0,02134819	0,14277434
θ_8									1	-0,35217414
θ_1										1

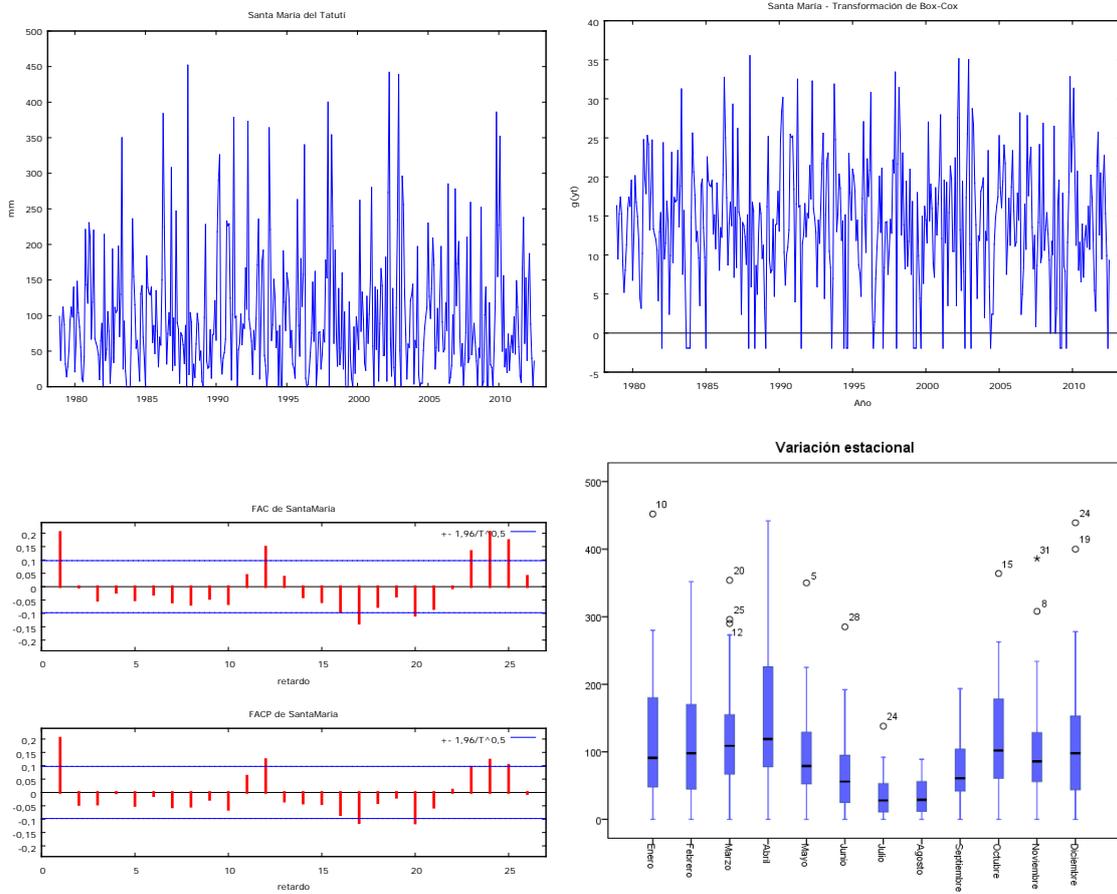
4.c. Criterios de selección:

	AIC	SBC	HQ
Modelo 1	1978,275	1989,891	1982,896
Modelo 2	2062,121	2081,647	2069,879
Modelo 3	1977,009	2023,475	1995,495

Conclusión: De los modelos propuestos, el 1 y el 3 son los que presentan mejor ajuste. Obsérvese que el Modelo 1 tiene el mejor ajuste de normalidad mientras que el 3 tiene residuos que se adaptan mejor a un ruido blanco que en los otros dos casos (de hecho, el Modelo 2 tiene residuos que no se ajustan a un ruido blanco). Los tres modelos pasan la prueba de homocedasticidad, aunque existe colinealidad entre los parámetros estacionales del Modelo 3 (además, ambos son significativos). Los criterios de información muestran valores más bajos para el Modelo 1 y, si a esto se agrega el hecho de que este modelo cumple mejor con el *principio de parsimonia* (recoge la estructura del proceso con muchos menos parámetros que el modelo 3), resulta el adecuado para explicar la serie. No obstante, el orden de diferenciación estacional incluido en el Modelo 1 surge sólo de la observación del correlograma de la serie ya que los test HEGY y CH no detectan presencia de raíz unitaria. Sobre este modelo recae también la posibilidad de que $\hat{\rho}_8 \neq 0$ no se deba sólo al azar y en ese caso los residuos no se ajusten a un ruido blanco. No se encontró un buen modelo.

Apéndice XII: Santa María del Tatutí

1. Representaciones gráficas.



2. Modelos propuestos.

Modelo 1: $ARIMA(1,0,0)x(1,0,1)_{12}$

Modelo 2⁶⁶: $ARIMA(1,0,0)x(0,1,1)_{12}$

3. Estimación de los parámetros.

Modelo 1					
	Coefficiente	Desv. Típica	z	Valor p	
const	0,481199	0,299596	1,6062	0,10824	
phi_1	0,136604	0,0428524	3,1878	0,00143	***
Phi_1	0,96121	0,0236173	40,6994	<0,00001	***
Theta_1	-0,88506	0,0349402	-25,3307	<0,00001	***

⁶⁶ Los contrastes HEGY y CH no revelan presencia de raíces unitarias estacionales, sin embargo el correlograma de la ACF de la serie muestra que los coeficientes estacionales no tienden a decrecer y es por ello que se propone el Modelo 2.

Modelo 2					
	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,0132856	0,0427963	0,3104	0,75623	
phi_1	0,127983	0,0420325	3,0449	0,00233	***
Theta_1	-0,927642	0,0193586	-47,9189	<0,00001	***

4. Causalidad e Invertibilidad.

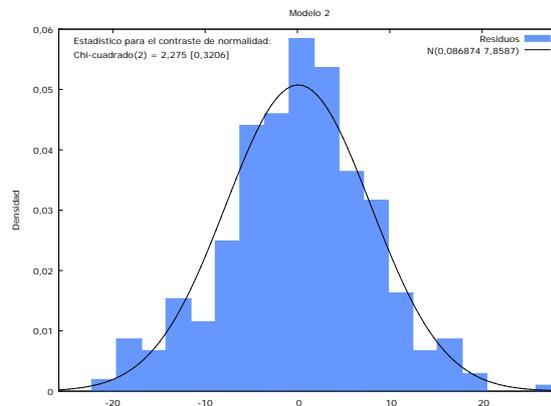
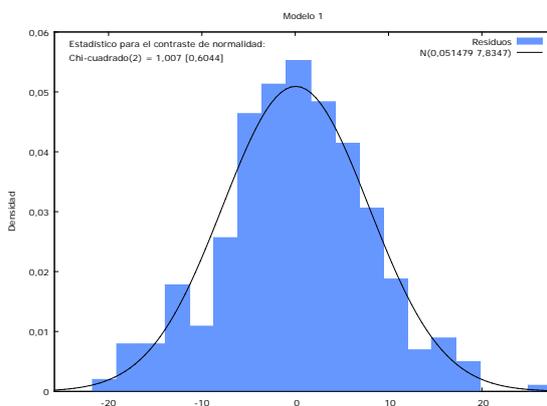
Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	7,3204	0,0000	7,3204	0,0000
AR (estacional)					
	Raíz 1	1,0404	0,0000	1,0404	0,0000
MA (estacional)					
	Raíz 1	1,1299	0,0000	1,1299	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	7,8136	0,0000	7,8136	0,0000
MA (estacional)					
	Raíz 1	1,0780	0,0000	1,0780	0,0000

5. Análisis de los residuos.

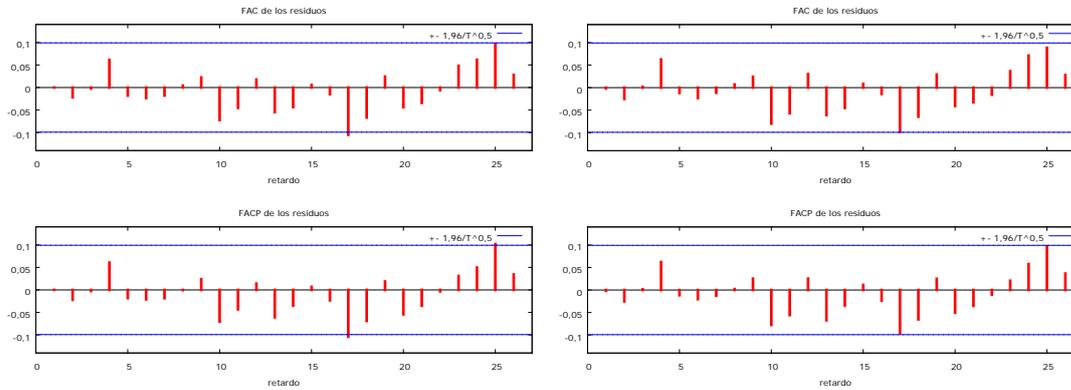
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	1,007	0,60443
Modelo 2	2,275	0,32058



5.b. Media y Correlograma residual:

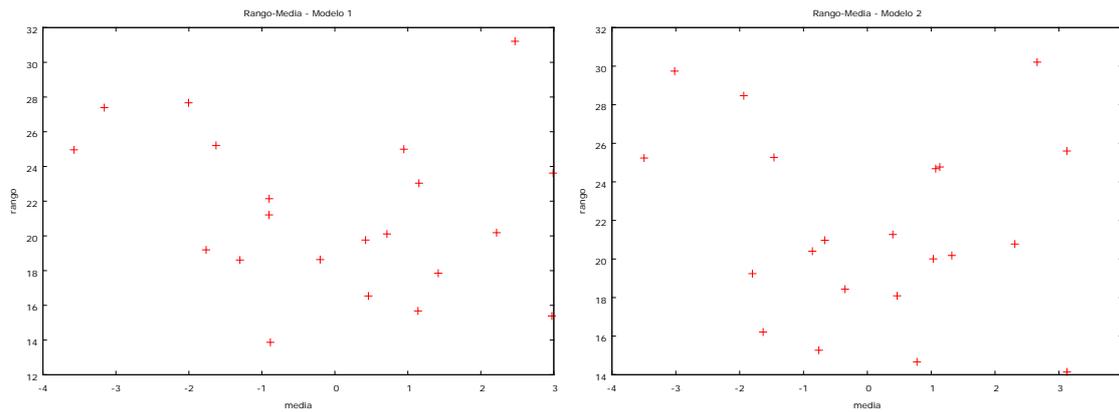
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,086874	7,83854	0,219151	0,8266
Modelo 2	0,0514792	7,80451	0,130429	0,8963



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0672	35,7735	0,479
Modelo 2	-0,0753	36,8262	0,430

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,50025	0,360621
Modelo 2	-0,398989	0,489969



6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1				
const	phi_1	Phi_1	Theta_1	
0,0897575	-0,00103360	-0,00692789	0,00711472	const
	0,00183633	9,82658e-006	-9,83473e-005	phi_1
		0,000557776	-0,000585299	Phi_1
			0,00122082	Theta_1

Correlaciones entre los parámetros estimados del Modelo 1			
	φ_1	Φ_1	Θ_1
φ_1	1	0,0097095	-0,06568435
Φ_1		1	-0,70928724
Θ_1			1

Matriz de Covarianzas del Modelo 2			
const	phi_1	Theta_1	
0,00183153	8,10093e-005	-0,000138815	const
	0,00176673	-2,67187e-005	phi_1
		0,000374755	Theta_1

Correlaciones entre los parámetros estimados del Modelo 2		
	φ_1	Θ_1
φ_1	1	-0,03283645
Θ_1		1

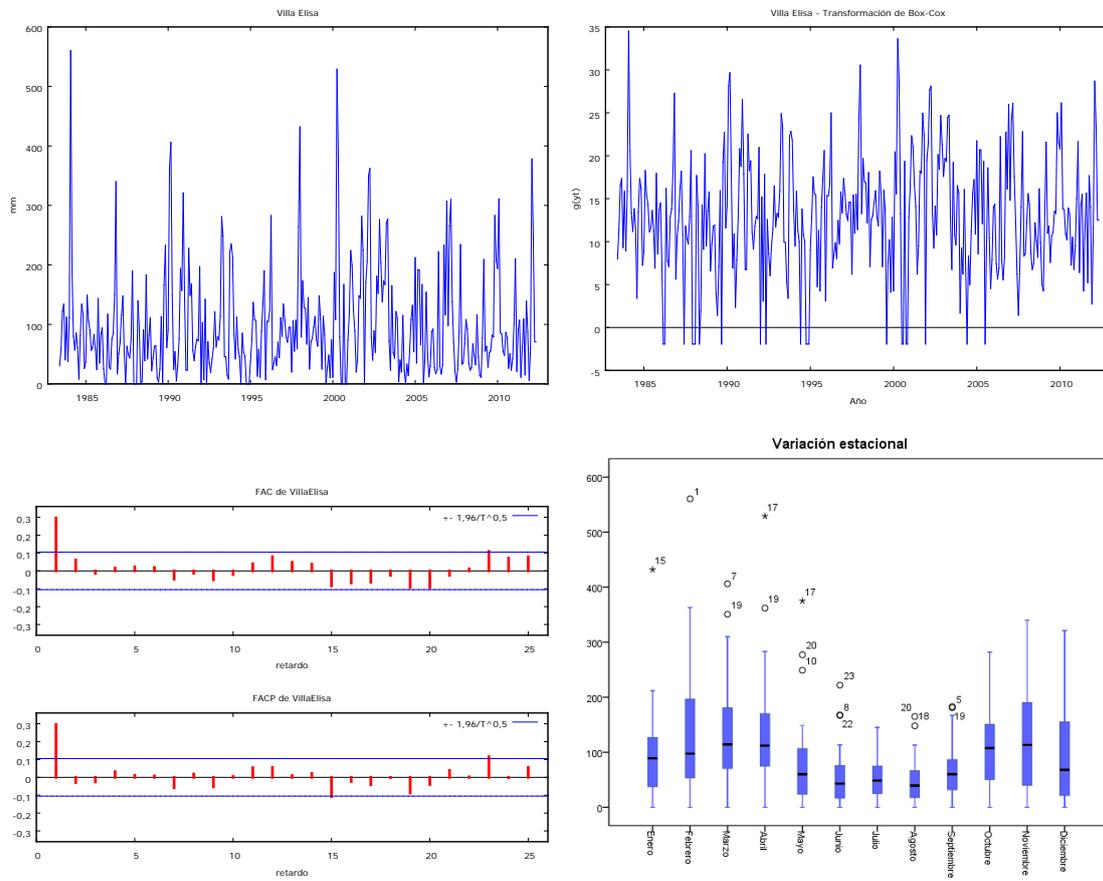
7. Criterios de selección:

	AIC	SBC	HQ
Modelo 1	2725,402	2745,246	2733,268
Modelo 2	2726,835	2742,710	2733,128

Conclusión: El análisis residual muestra que ambos modelos presentan buen ajuste y los criterios de información no tienen diferencias importantes, por lo que los dos modelos recogen la estructura del proceso generador de los datos. Sin embargo, para concordar con los resultados de los test de raíces unitarias puede elegirse el Modelo 1.

Apéndice XIII: Villa Elisa

1. Representaciones gráficas.



2. Modelos propuestos.

Modelo 1: $ARIMA(1,0,0)x(0,0,0)_{12}$

3. Estimación de los parámetros:

	Coefficiente	Desv. Típica	z	Valor p	
const	9,02309	0,75148	12,0071	<0,00001	***
phi_1	0,296586	0,0513775	5,7727	<0,00001	***

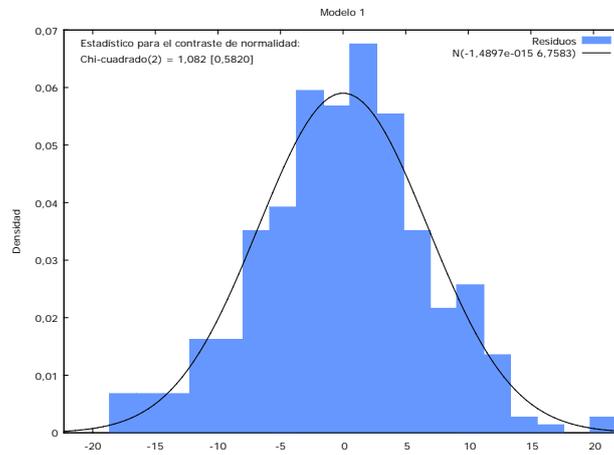
4. Causalidad e Invertibilidad

		Real	Imaginaria	Módulo	Frecuencia
AR					
	Raíz 1	3,3717	0,0000	3,3717	0,0000

5. Análisis de los residuos

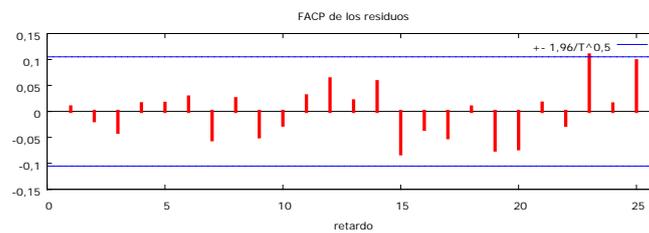
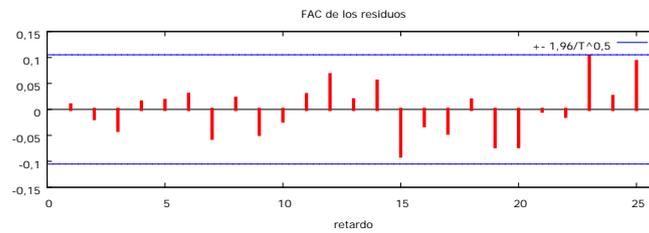
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	1,082	0,58203



5.b. Media y Correlograma residual:

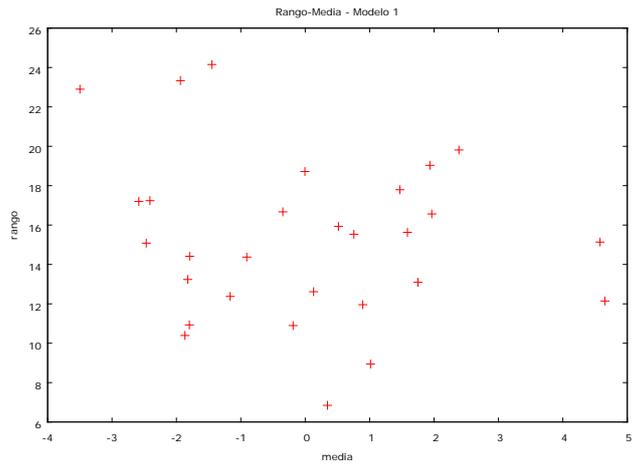
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	-1,55319E-15	6,74856	-4,28725e-015	1



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	0,0848	38,8753	0,341

4.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,369841	0,342711



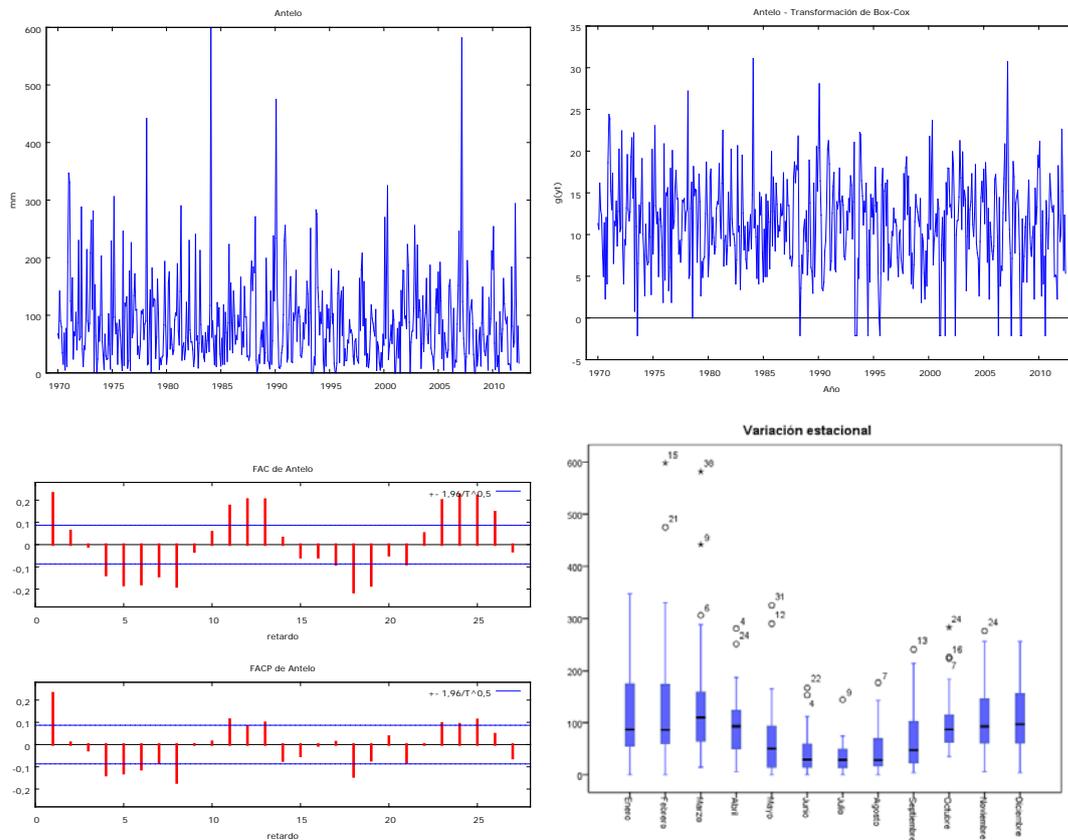
4.c. Criterios de información:

	AIC	SBC	HQ
Modelo 1	2312,816	2320,515	2315,882

Conclusión: el modelo seleccionado presenta buen ajuste.

Apéndice XIV: Antelo

1. Representaciones gráficas.



2. Modelos propuestos:

Modelo 1: $ARIMA(1,0,0)x(1,0,1)_{12}$

Modelo 2: $ARIMA(1,0,1)x(1,0,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	1,01489	0,310348	3,2702	0,00107	***
phi_1	0,0800517	0,0411713	1,9444	0,05185	*
Phi_1	0,900118	0,029459	30,5550	<0,00001	***
Theta_1	-0,758945	0,0457787	-16,5786	<0,00001	***

Modelo 2					
	Coeficiente	Desv. Típica	z	Valor p	
const	0,649181	0,251635	2,5799	0,00988	***
phi_1	0,403933	0,17044	2,3699	0,01779	**
Phi_1	0,900463	0,0286677	31,4104	<0,00001	***
theta_1	-0,345541	0,174384	-1,9815	0,04754	**
Theta_1	-0,76015	0,0453562	-16,7596	<0,00001	***

4. Causalidad e Invertibilidad:

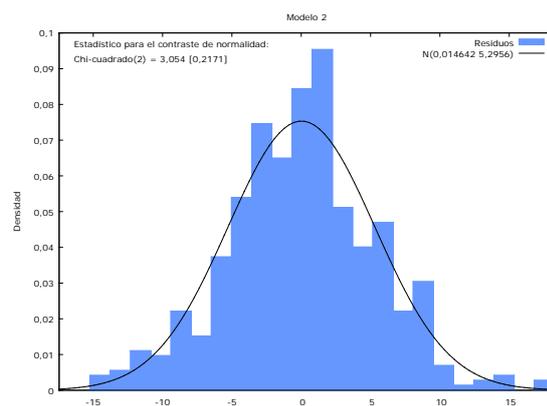
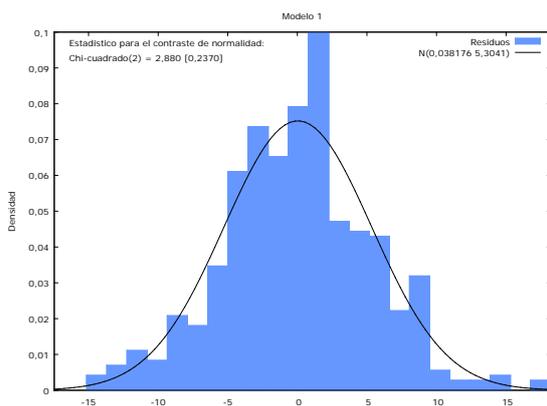
Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	12,4919	0,0000	12,4919	0,0000
AR (estacional)					
	Raíz 1	1,1110	0,0000	1,1110	0,0000
MA (estacional)					
	Raíz 1	1,3176	0,0000	1,3176	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	2,4757	0,0000	2,4757	0,0000
AR (estacional)					
	Raíz 1	1,1105	0,0000	1,1105	0,0000
MA					
	Raíz 1	2,8940	0,0000	2,8940	0,0000
MA (estacional)					
	Raíz 1	1,3155	0,0000	1,3155	0,0000

5. Análisis de los residuos

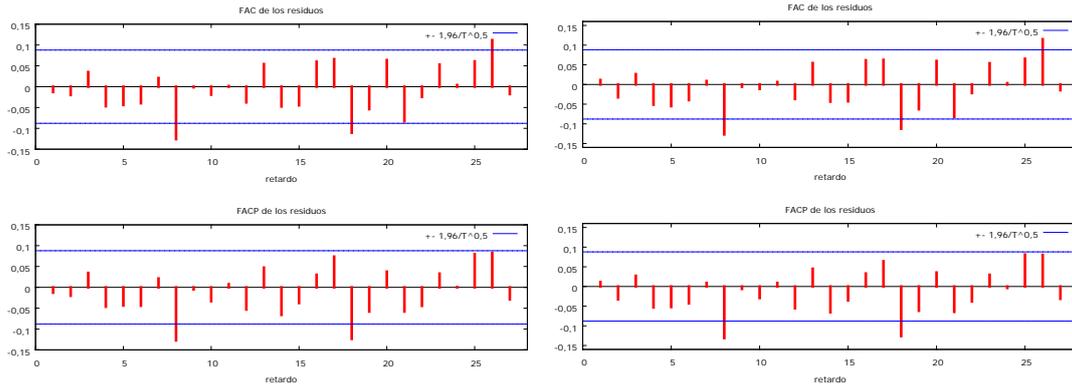
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	2,880	0,23696
Modelo 2	3,054	0,21715



5.b. Media y Correlograma residual:

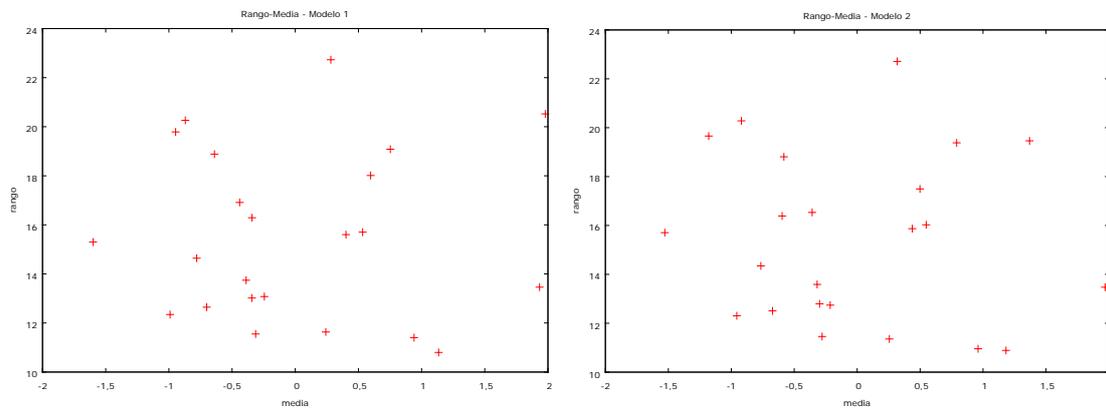
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,0381762	5,28803	0,160945	0,8722
Modelo 2	0,0146417	5,27416	0,0618893	0,9507



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	0,0006	48,5670	0,079
Modelo 2	0,0022	50,4713	0,055

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	0,0252303	0,975478
Modelo 2	-0,34851	0,682911



6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1				
const	phi_1	Phi_1	Theta_1	
0,0963160	-0,000463246	-0,00885001	0,00991035	const
	0,00169507	-0,000140852	5,50998e-005	phi_1
		0,000867832	-0,000964542	Phi_1
			0,00209569	Theta_1

Correlaciones entre los parámetros estimados del Modelo 1			
	φ_1	Φ_1	Θ_1
φ_1	1	-0,11613161	0,02923425
Φ_1		1	-0,71522006
Θ_1			1

Matriz de Covarianzas del Modelo 2					
const	phi_1	Phi_1	theta_1	Theta_1	
0,0633201	-0,0271103	-0,00458454	0,0278459	0,00507988	const
	0,0290500	-0,000855748	-0,0289245	0,00101005	phi_1
		0,000821836	0,000723751	-0,000921491	Phi_1
			0,0304098	-0,000927368	theta_1
				0,00205718	Theta_1

Correlaciones entre los parámetros estimados del Modelo 2				
	φ_1	Φ_1	θ_1	Θ_1
φ_1	1	-0,17513846	-0,97316773	0,1306576
Φ_1		1	0,14477371	-0,70869862
θ_1			1	-0,11724891
Θ_1				1

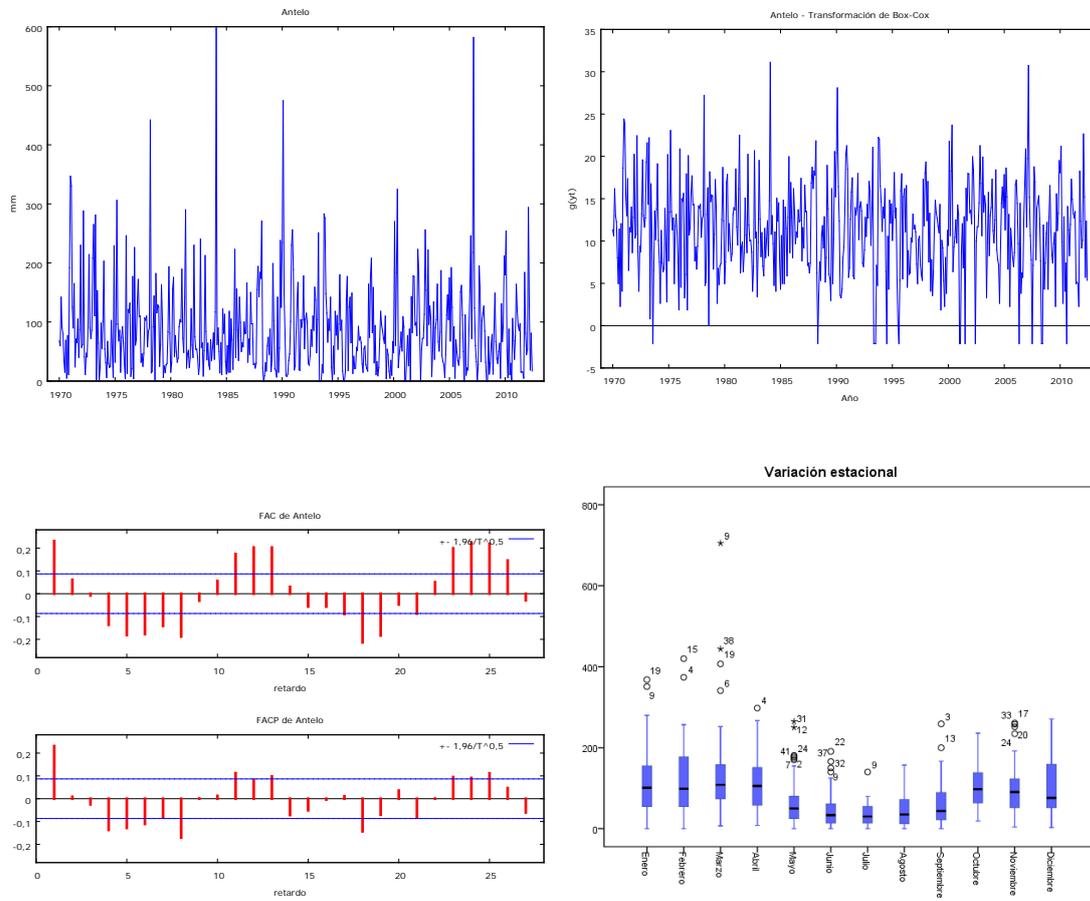
7. Criterios de información:

	AIC	SBC	HQ
Modelo 1	3074,903	3095,946	3083,162
Modelo 2	3074,270	3099,522	3084,182

Conclusión: Los dos modelos propuestos superan las pruebas de normalidad y homocedasticidad (el modelo 1 muestra mejor ajuste) y las condiciones de causalidad e invertibilidad. Sin embargo, no está claro que los residuos de ambos modelos tengan estructura de ruido blanco, ya que existen retardos para los que la ACF residual supera los límites de confianza. En lo que respecta al análisis de correlaciones, obsérvese que el Modelo 2 tiene alta correlación entre los parámetros estimados autorregresivo y de media móvil de la estructura regular. Los criterios de información SBC y HQ alcanzan menor valor para el Modelo 1, mientras que el AIC es casi igual en ambos casos. Luego, el Modelo 1 presenta mejor ajuste que el 2.

Apéndice XV: Febre

1. Representaciones gráficas.



2. Modelos propuestos:

Modelo 1: $ARIMA(1,0,1)x(1,0,1)_{12}$

Modelo 2: $ARIMA(1,0,1)x(0,1,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	0,31656	0,105812	2,9917	0,00277	***
phi_1	0,558112	0,0535391	10,4244	<0,00001	***
Phi_1	0,938242	0,0187316	50,0887	<0,00001	***
theta_1	-0,550027	0,0664849	-8,2730	<0,00001	***
Theta_1	-0,827128	0,0314841	-26,2713	<0,00001	***

Modelo 2					
	Coeficiente	Desv. Típica	z	Valor p	
const	-0,00523827	0,019084	-0,2745	0,78371	
phi_1	0,533195	0,0478893	11,1339	<0,00001	***
theta_1	-0,540941	0,0615757	-8,7850	<0,00001	***
Theta_1	-0,868914	0,0225612	-38,5137	<0,00001	***

4. Causalidad e Invertibilidad:

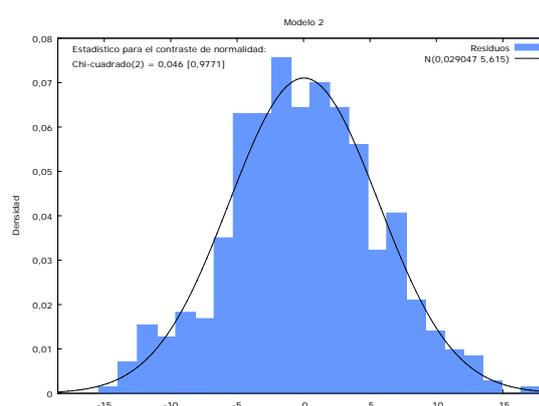
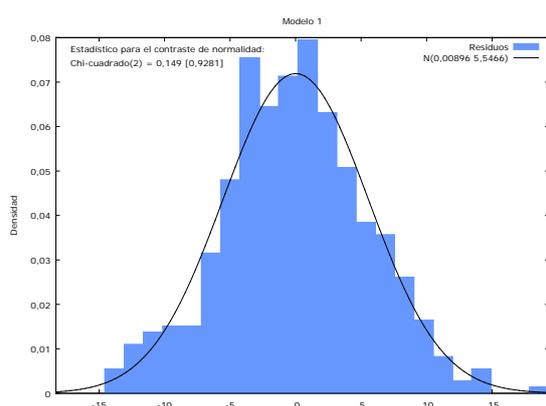
Modelo 1					
		Real	Imaginaria	Módulo	Frecuencia
AR					
	Raíz 1	1,7918	0,0000	1,7918	0,0000
AR (estacional)					
	Raíz 1	1,0658	0,0000	1,0658	0,0000
MA					
	Raíz 1	1,8181	0,0000	1,8181	0,0000
MA (estacional)					
	Raíz 1	1,2090	0,0000	1,2090	0,0000

Modelo 2					
		Real	Imaginaria	Módulo	Frecuencia
AR					
	Raíz 1	1,8755	0,0000	1,8755	0,0000
MA					
	Raíz 1	1,8486	0,0000	1,8486	0,0000
MA (estacional)					
	Raíz 1	1,1509	0,0000	1,1509	0,0000

5. Análisis de los residuos

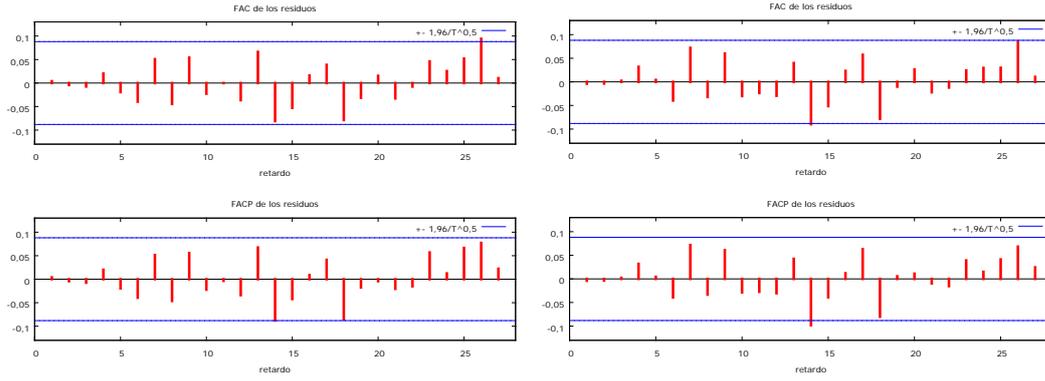
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	0,149	0,92810
Modelo 2	0,046	0,97713



5.b. Media y Correlograma residual:

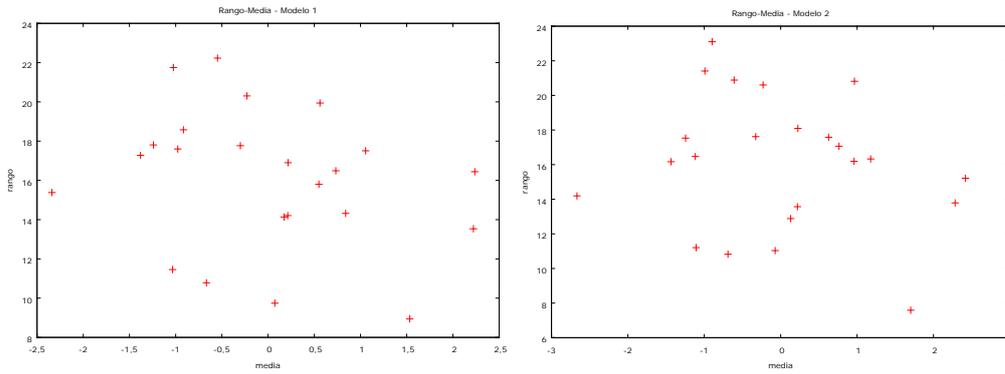
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,00896005	5,52415	0,0360868	0,9712
Modelo 2	0,0290472	5,59788	0,115447	0,9081



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0333	34,4885	0,541
Modelo 2	-0,0339	33,4567	0,590

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,781182	0,245816
Modelo 2	-0,573237	0,399221



6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1					
const	phi_1	Phi_1	theta_1	Theta_1	
0,0111962	-0,00178095	-0,00175653	0,00249906	0,00129891	const
	0,00286643	-0,000107819	-0,00290704	0,000196072	phi_1
		0,000350873	-4,15445e-005	-0,000280473	Phi_1
			0,00442024	-7,18402e-005	theta_1
				0,000991249	Theta_1

Correlaciones entre los parámetros estimados del Modelo 1				
	φ_1	Φ_1	θ_1	Θ_1
φ_1	1	-0,10751012	-0,8166894	0,11631969
Φ_1		1	-0,0333592	-0,47558144
θ_1			1	-0,03432047
Θ_1				1

Matriz de Covarianzas del Modelo 2				
const	phi_1	theta_1	Theta_1	
0,000364200	-0,000220776	0,000165626	-2,76803e-005	const
	0,00229338	-0,00229115	7,85850e-005	phi_1
		0,00379156	-6,15552e-006	theta_1
			0,000509007	Theta_1

Correlaciones entre los parámetros estimados del Modelo 2			
	φ_1	θ_1	Θ_1
φ_1	1	-0,77697255	0,07273425
θ_1		1	-0,00443091
Θ_1			1

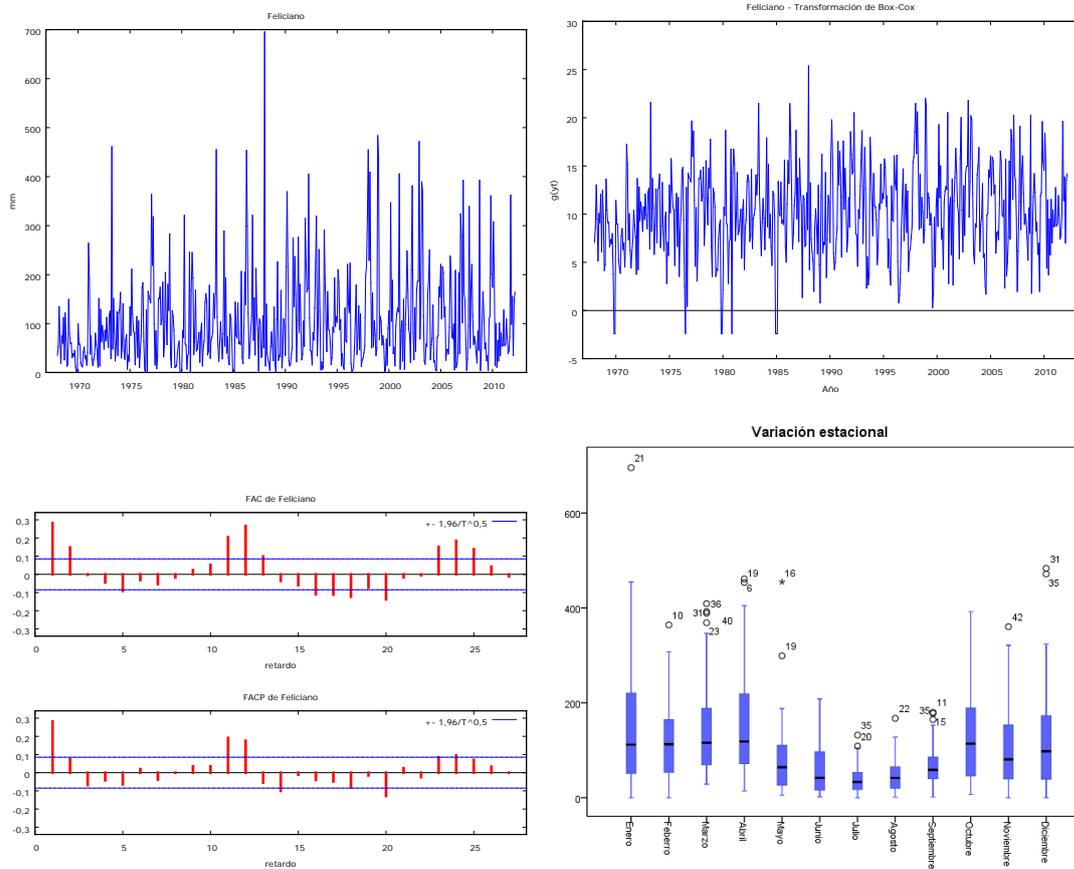
7. Criterios de información:

	AIC	SBC	HQ
Modelo 1	3107,787	3133,014	3117,690
Modelo 2	3118,925	3139,948	3127,178

Conclusión: Puede observarse, en principio, que los dos modelos propuestos reúnen la estructura del proceso generador de los datos. El Modelo 2 se propuso de acuerdo con el resultado del contraste HEGY que sugiere presencia de raíz unitaria estacional y del correlograma de la ACF y de la PACF de la serie. El análisis residual no refleja grandes discrepancias y los criterios de información son menores para el Modelo 1, que parece ser el que mejor ajuste presenta. Sin embargo persiste en ambos modelos el problema de la colinealidad entre los parámetros autorregresivo y de media móvil de la parte regular, más notable en el Modelo 1 que en el 2.

Apéndice XVI: Feliciano

1. Representaciones gráficas.



2. Modelos propuestos:

Modelo 1: $ARIMA(0,0,2)x(1,1,1)_{12}$

Modelo 2: $ARIMA(1,0,2)x(1,1,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coefficiente	Desv. Típica	z	Valor p	
const	0,0883605	0,0515502	1,7141	0,08652	*
Phi_1	-0,00269797	0,0451866	-0,0597	0,95239	
theta_1	0,18244	0,0400622	4,5539	<0,00001	***
theta_2	0,14259	0,042108	3,3863	0,00071	***
Theta_1	-0,827231	0,0239382	-34,5570	<0,00001	***

Modelo 2					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,10535	0,0674746	1,5613	0,11845	
phi_1	-0,282252	0,233162	-1,2105	0,22607	
Phi_1	-0,0122436	0,0443257	-0,2762	0,78238	
theta_1	0,458383	0,235014	1,9505	0,05112	*
theta_2	0,173515	0,0524364	3,3091	0,00094	***
Theta_1	-0,827634	0,0228133	-36,2786	<0,00001	***

4. Causalidad e Invertibilidad:

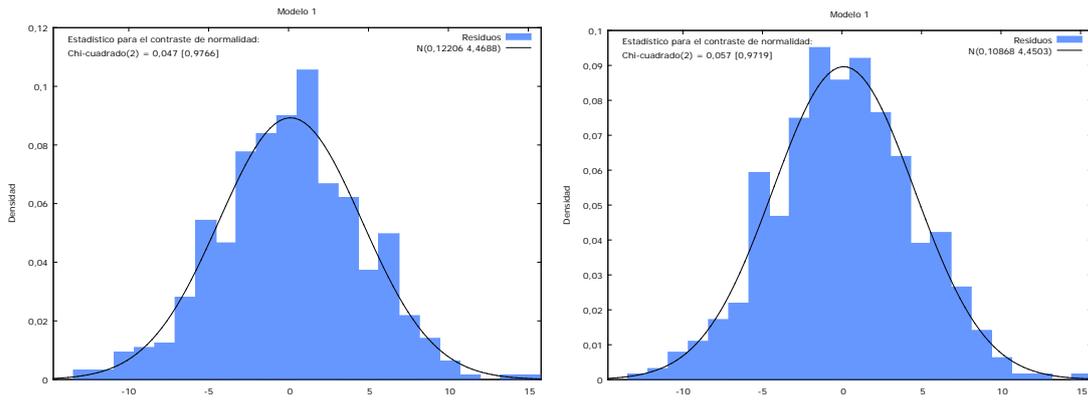
Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	-370,6485	0,0000	370,6485	0,5000
MA					
	Raíz 1	-0,6397	-2,5698	2,6482	-0,2888
	Raíz 2	-0,6397	2,5698	2,6482	0,2888
MA (estacional)					
	Raíz 1	1,2089	0,0000	1,2089	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	-3,5429	0,0000	3,5429	0,5000
AR (estacional)					
	Raíz 1	-81,6755	0,0000	81,6755	0,5000
MA					
	Raíz 1	-1,3209	-2,0046	2,4007	-0,3427
	Raíz 2	-1,3209	2,0046	2,4007	0,3427
MA (estacional)					
	Raíz 1	1,2083	0,0000	1,2083	0,0000

5. Análisis de los residuos.

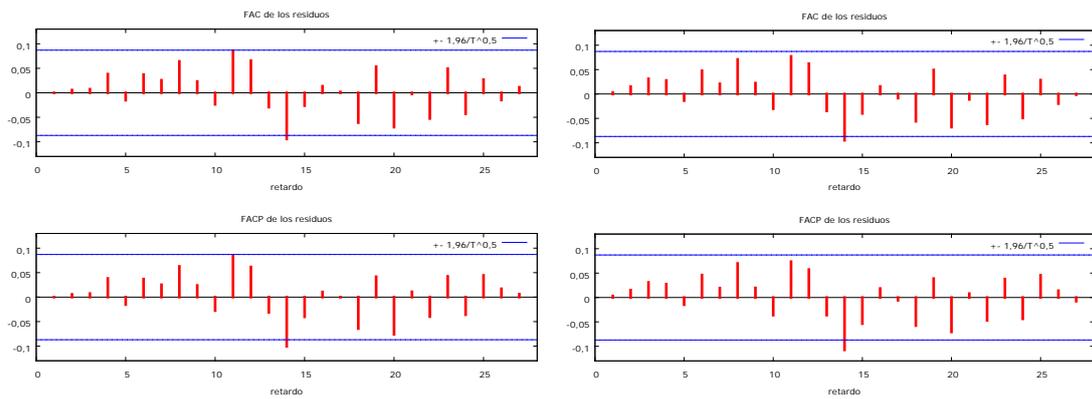
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	0,047	0,97656
Modelo 2	0,057	0,97190



5.b. Media y Correlograma residual:

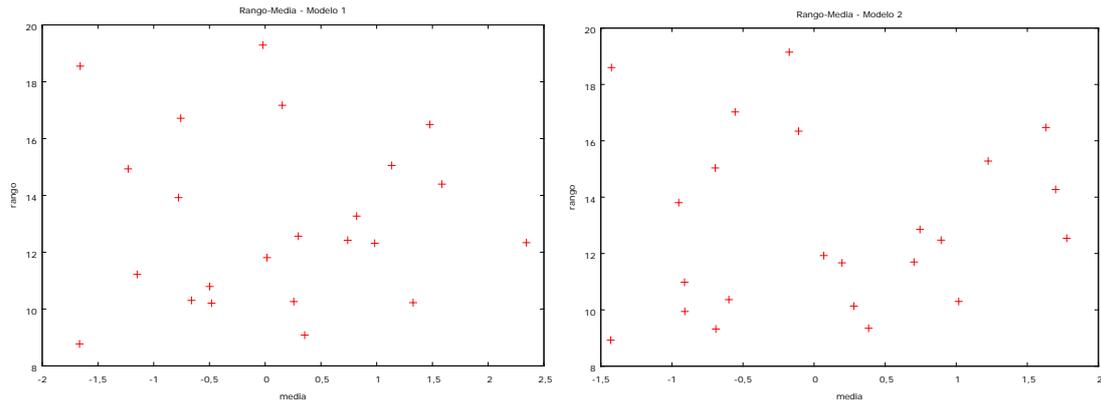
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,122061	4,45111	0,617466	0,5372
Modelo 2	0,108682	4,42821	0,552086	0,5811



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0216	36,0904	0,464
Modelo 2	-0,0191	36,9745	0,424

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,0121139	0,984199
Modelo 2	0,177505	0,794436



6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1					
const	Phi_1	theta_1	theta_2	Theta_1	
0,00265743	0,000195106	0,000108653	-1,73697e-006	-0,000227837	const
	0,00204183	-1,46443e-005	0,000251452	-0,000459811	Phi_1
		0,00160498	0,000202038	-0,000126994	theta_1
			0,00177308	-6,90287e-005	theta_2
				0,000573037	Theta_1

Correlación entre los parámetros estimados del Modelo 1				
	Φ_1	θ_1	θ_2	Θ_1
Φ_1	1	-0,00808955	0,13215415	-0,42508737
θ_1		1	0,11976603	-0,13242102
θ_2			1	-0,06848155
Θ_1				1

Matriz de Covarianzas del Modelo 2						
const	phi_1	Phi_1	theta_1	theta_2	Theta_1	
0,00455282	-0,00587439	0,000181926	0,00600573	0,000877334	-0,000214054	const
	0,0543646	8,01870e-005	-0,0539777	-0,00737648	-0,000158618	phi_1
		0,00196476	-6,38341e-005	0,000159271	-0,000394791	Phi_1
			0,0552315	0,00799853	3,17749e-005	theta_1
				0,00274958	-4,05634e-005	theta_2
					0,000520445	Theta_1

Correlación entre los parámetros estimados del Modelo 2					
	φ_1	Φ_1	θ_1	θ_2	Θ_1
φ_1	1	0,00775873	-0,98506042	-0,60333501	-0,02981993
Φ_1		1	-0,00612778	0,06852487	-0,39041233
θ_1			1	0,6490581	0,00592656
θ_2				1	-0,03390887
Θ_1					1

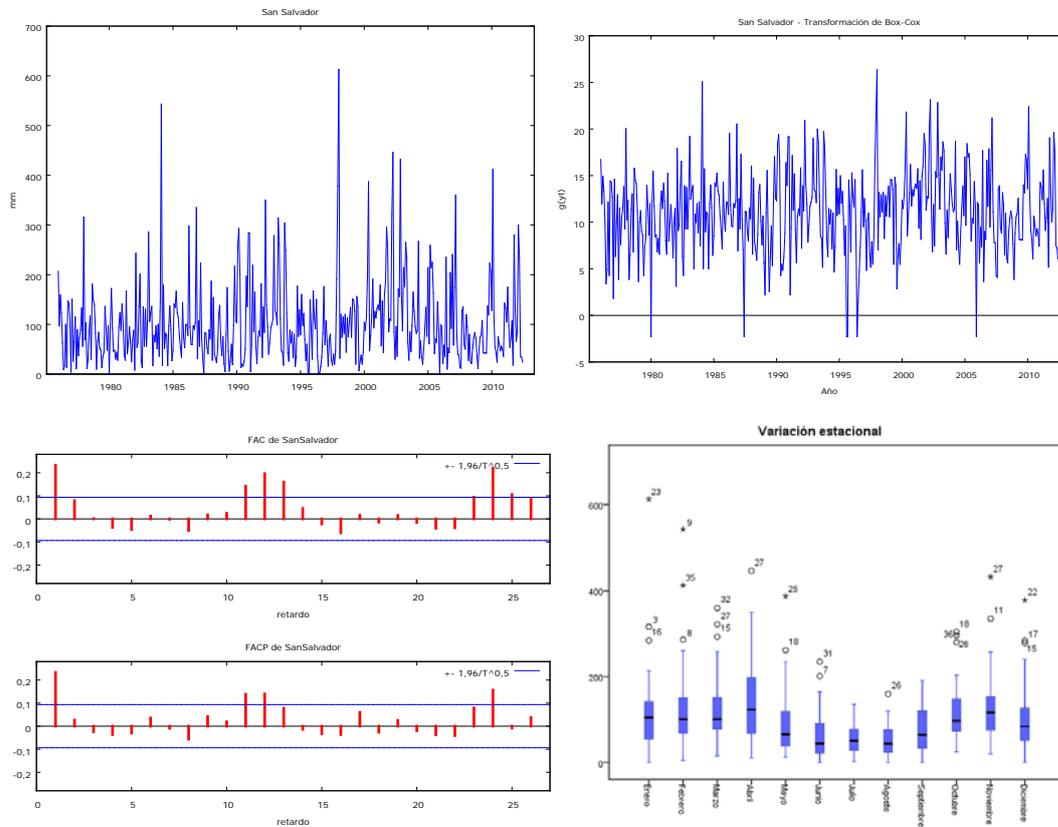
7. Criterios de información:

	AIC	SBC	HQ
Modelo 1	2964,243	2989,614	2974,193
Modelo 2	2955,121	2984,707	2966,725

Conclusión: Los dos modelos presentan buen ajuste residual y satisfacen las condiciones de causalidad e invertibilidad. Los criterios de información toman valores más bajos para el Modelo 2 aunque éste tiene una correlación alta en las estimaciones de los parámetros autorregresivo y de media móvil de la parte regular. No obstante, la no significatividad de estos parámetros permite suprimirlos del modelo.

Apéndice XVII: San Salvador

1. Representaciones gráficas.



2. Modelos propuestos:

Modelo 1: $ARIMA(0,0,1)x(1,0,1)_{12}$

Modelo 2: $ARIMA(1,0,0)x(2,0,0)_{12}$

Modelo 3: $ARIMA(1,0,1)x(2,0,0)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	4,40053	1,19948	3,6687	0,00024	***
Phi_1	0,605851	0,105986	5,7163	<0,00001	***
theta_1	0,157524	0,0452522	3,4810	0,00050	***
Theta_1	-0,423067	0,115412	-3,6657	0,00025	***

Modelo 2					
	Coeficiente	Desv. Típica	z	Valor p	
const	5,97247	0,672737	8,8779	<0,00001	***
phi_1	0,206252	0,0482971	4,2705	0,00002	***
Phi_1	0,1483	0,0487694	3,0408	0,00236	***
Phi_2	0,177606	0,0480171	3,6988	0,00022	***

Modelo 3					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	5,98293	1,94648	3,0737	0,00211	***
phi_1	0,20487	0,234725	0,8728	0,38277	
Phi_1	0,14831	0,0460211	3,2226	0,00127	***
Phi_2	0,177594	0,0467659	3,7975	0,00015	***
theta_1	0,00143609	0,244339	0,0059	0,99531	

4. Causalidad e Invertibilidad:

Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	1,6506	0,0000	1,6506	0,0000
MA					
	Raíz 1	-6,3482	0,0000	6,3482	0,5000
MA (estacional)					
	Raíz 1	2,3637	0,0000	2,3637	0,0000

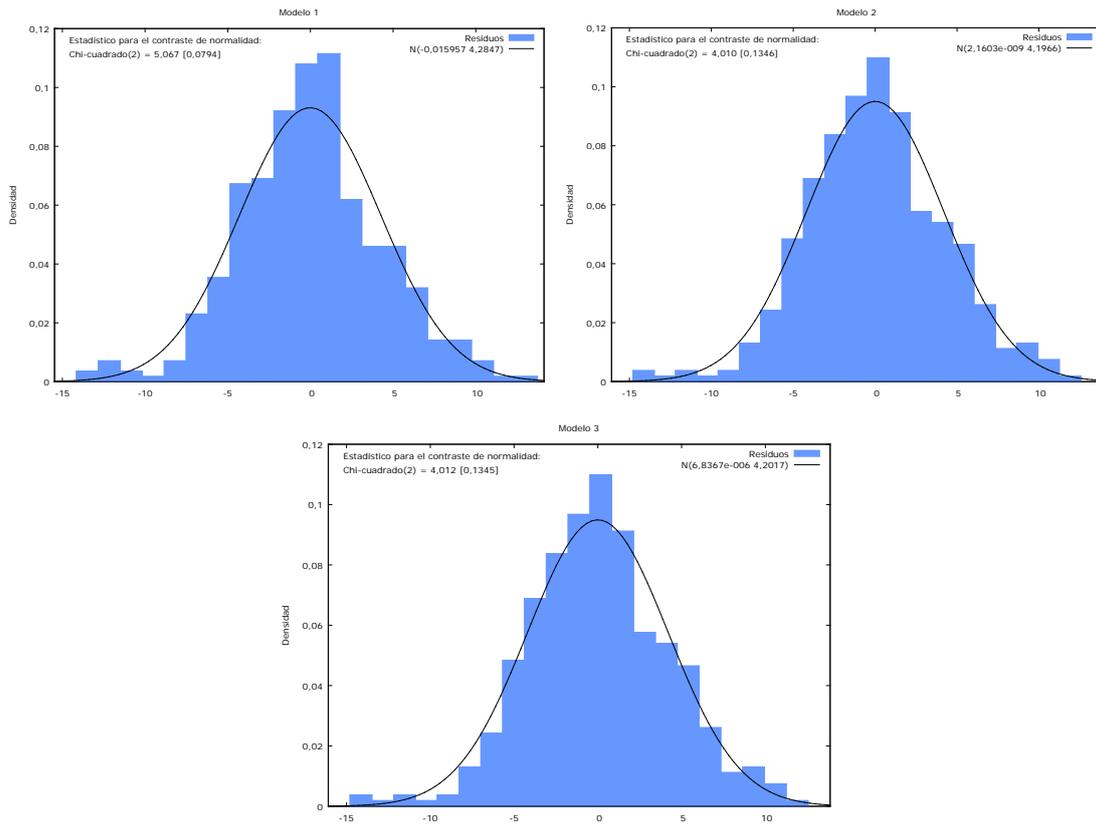
Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	4,8484	0,0000	4,8484	0,0000
AR (estacional)					
	Raíz 1	1,9918	0,0000	1,9918	0,0000
	Raíz 2	-2,8268	0,0000	2,8268	0,5000

Modelo 3					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	4,8811	0,0000	4,8811	0,0000
AR (estacional)					
	Raíz 1	1,9918	0,0000	1,9918	0,0000
	Raíz 2	-2,8269	0,0000	2,8269	0,5000
MA					
	Raíz 1	-696,3373	0,0000	696,3373	0,5000

5. Análisis de los residuos

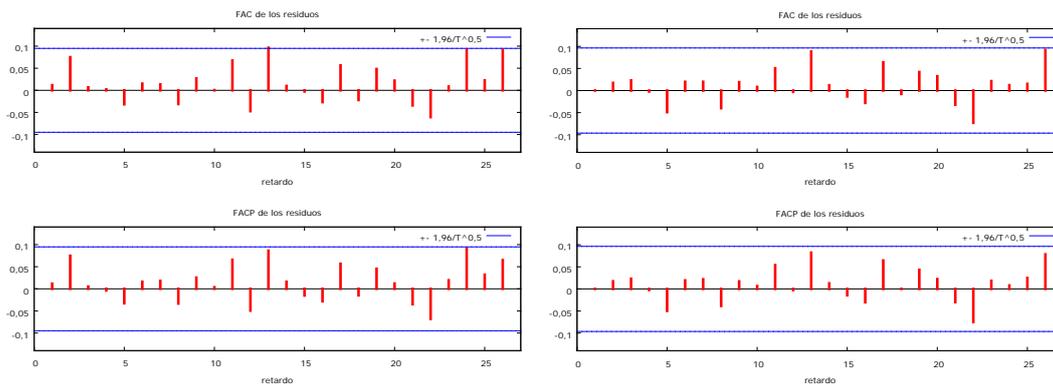
5.a. Normalidad:

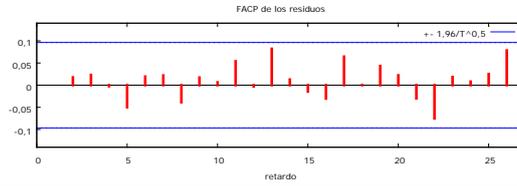
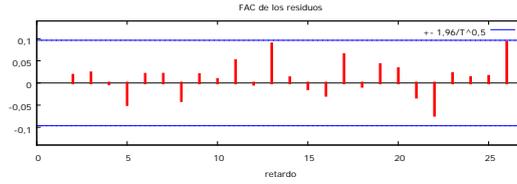
	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	5,067	0,07936
Modelo 2	4,010	0,13465
Modelo 3	4,012	0,13452



5.b. Media y Correlograma residual:

	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	-0,0159574	4,26957	-0,0771407	0,9385
Modelo 2	2,16034e-009	4,18127	1,05e-008	1
Modelo 3	6,83668e-006	4,18127	3,32286e-005	1

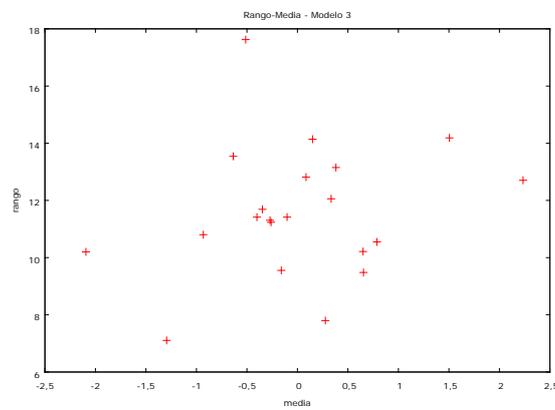
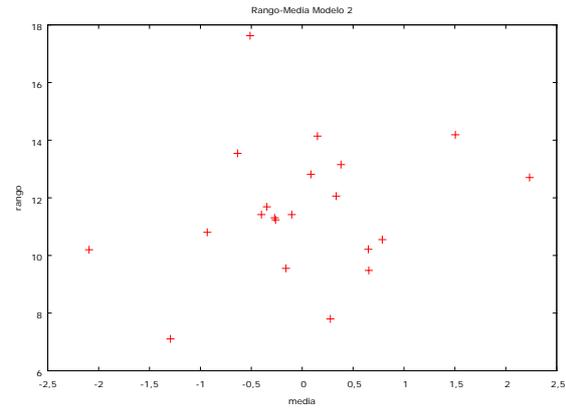
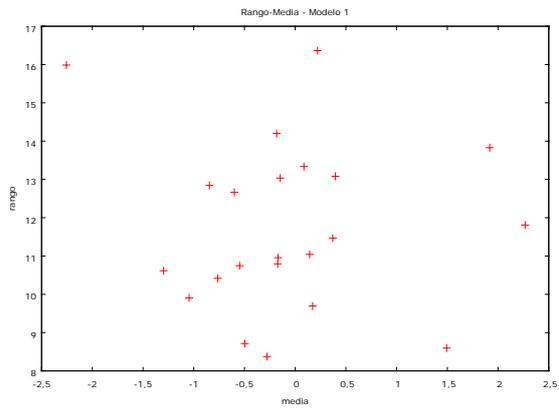




	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	0,0603	33,9408	0,567
Modelo 2	0,0579	29,8523	0,755
Modelo 3	0,0579	29,8545	0,755

4.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,158819	0,746593
Modelo 2	0,534636	0,353387
Modelo 3	0,534014	0,353836



4.d. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1				
const	Phi_1	theta_1	Theta_1	
1,43874	-0,126229	0,00613035	0,128372	const
	0,0112331	-0,000585137	-0,0113852	Phi_1
		0,00204776	0,000307206	theta_1
			0,0133200	Theta_1

Correlaciones entre los parámetros estimados del Modelo 1			
	Φ_1	θ_1	Θ_1
Φ_1	1	-0,12200268	-0,93076747
θ_1		1	0,0588219
Θ_1			1

Matriz de Covarianzas del Modelo 2				
const	phi_1	Phi_1	Phi_2	
0,452575	-0,0157268	-0,0164291	-0,0165406	const
	0,00233261	-0,000164689	-5,24478e-005	phi_1
		0,00237846	-0,000381151	Phi_1
			0,00230564	Phi_2

Correlaciones entre los parámetros estimados del Modelo 2			
	φ_1	Φ_1	Φ_2
φ_1	1	-0,06991915	-0,02261571
Φ_1		1	-0,16276226
Φ_2			1

Matriz de Covarianzas del Modelo 3					
const	phi_1	Phi_1	Phi_2	theta_1	
3,78878	-0,436053	-0,0108473	-0,0359818	0,449007	const
	0,0550956	-0,000713998	0,00237519	-0,0564002	phi_1
		0,00211794	-0,000199766	0,000516701	Phi_1
			0,00218705	-0,00261387	Phi_2
				0,0597015	theta_1

Correlaciones entre los parámetros estimados del Modelo 3				
	φ_1	θ_1	Φ_1	Φ_2
φ_1	1	-0,98339615	-0,06609683	0,21637631
θ_1		1	0,04595042	-0,22875042
Φ_1			1	-0,09281866
Φ_2				1

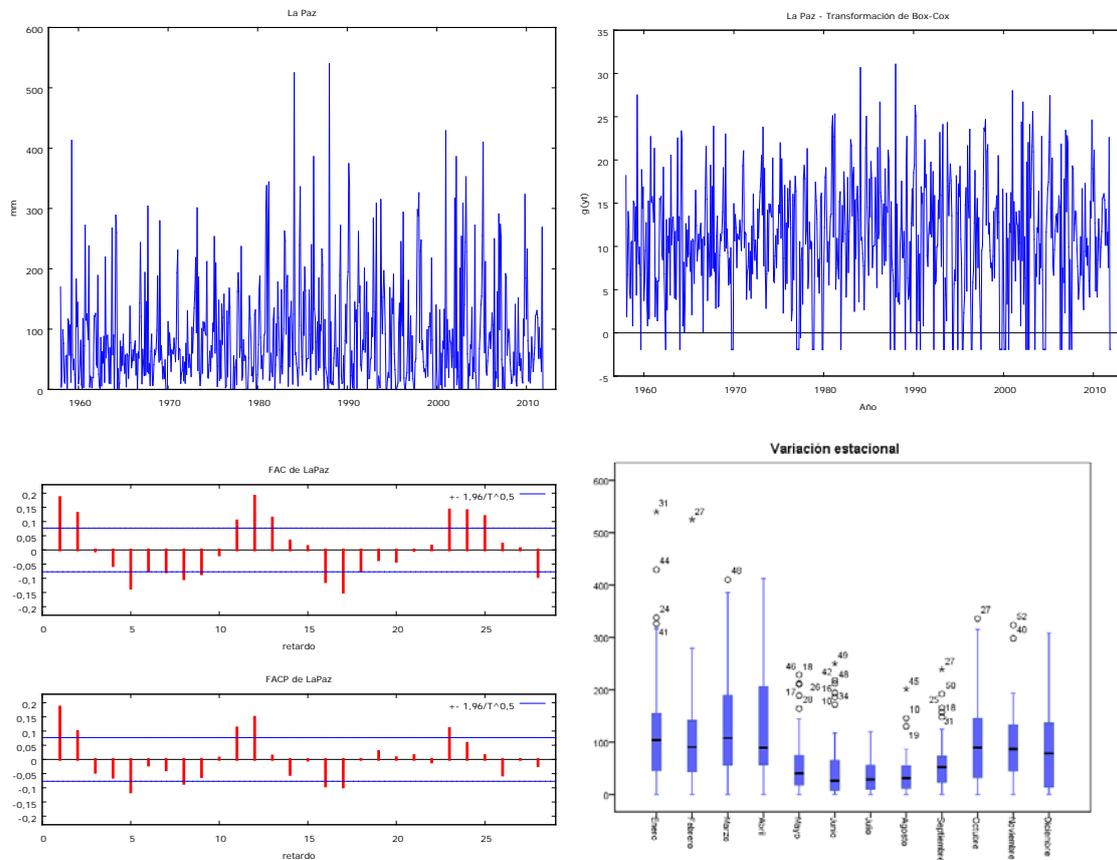
4.c. Criterios de información:

	AIC	SBC	HQ
Modelo 1	2454,629	2474,901	2462,637
Modelo 2	2360,729	2376,823	2367,095
Modelo 3	2364,729	2388,870	2374,277

Conclusión: Los modelos 2 y 3 tienen residuos que se ajustan mejor a un ruido blanco y superan más claramente la prueba de normalidad de Doornik-Hansen. No se evidencia problemas de heterocedasticidad residual y las condiciones de causalidad e invertibilidad están satisfechas. El Modelo 1 presenta los valores más bajos de los criterios de información pero tiene fuerte correlación entre los parámetros autorregresivo y de media móvil de la parte estacional. Además, estos coeficientes son significativos por lo que no pueden suprimirse del modelo. Los dos órdenes de autorregresividad en los modelos 2 y 3 surgen de la información procedente de la PACF de la serie. No aparecen correlaciones entre los parámetros del Modelo 2 y tiene los criterios de información más bajos.

Apéndice XVIII: La Paz

1. Representaciones gráficas.



2. Modelos propuestos:

Modelo 1: $ARIMA(1,0,2)x(1,0,1)_{12}$

Modelo 2: $ARIMA(0,0,2)x(1,0,1)_{12}$

Modelo 3⁶⁷: $ARIMA(0,0,5)x(1,0,1)_{12}$

Modelo 4⁶⁸: $ARIMA(0,0,5)x(2,0,0)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coefficiente	Dev. Típica	z	Valor p	
const	2,80049	1,14733	2,4409	0,01465	**
phi_1	0,0990559	0,274601	0,3607	0,71830	
Phi_1	0,717717	0,0786199	9,1290	<0,00001	***
theta_1	0,0158705	0,275109	0,0577	0,95400	
theta_2	0,121594	0,0503251	2,4162	0,01568	**
Theta_1	-0,563062	0,0935504	-6,0188	<0,00001	***

⁶⁷ Reformulado para obtener residuos con estructura de ruido blanco.

⁶⁸ Reformulado del Modelo 3 para obtener mejor ajuste y parámetros no correlacionados.

Modelo 2					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	4,02648	1,04989	3,8351	0,00013	***
Phi_1	0,632675	0,093769	6,7472	<0,00001	***
theta_1	0,118201	0,0347078	3,4056	0,00066	***
theta_2	0,133279	0,0384218	3,4688	0,00052	***
Theta_1	-0,469003	0,106761	-4,3930	0,00001	***

Modelo 3					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	5,25214	1,30481	4,0252	0,00006	***
Phi_1	0,520398	0,117254	4,4382	<0,00001	***
theta_1	0,128565	0,0348977	3,6841	0,00023	***
theta_2	0,143626	0,0370954	3,8718	0,00011	***
theta_3	0,00777455	0,0397416	0,1956	0,84490	
theta_4	-0,0191999	0,0409197	-0,4692	0,63892	
theta_5	-0,106431	0,0410851	-2,5905	0,00958	***
Theta_1	-0,364877	0,127401	-2,8640	0,00418	***

Modelo 4					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	8,45036	0,676637	12,4888	<0,00001	***
Phi_1	0,144538	0,0392053	3,6867	0,00023	***
Phi_2	0,0810902	0,0379696	2,1357	0,03271	**
theta_1	0,140548	0,0347181	4,0483	0,00005	***
theta_2	0,15108	0,0379761	3,9783	0,00007	***
theta_3	0,012178	0,039814	0,3059	0,75970	
theta_4	-0,0209091	0,0410786	-0,5090	0,61075	
theta_5	-0,109353	0,0413495	-2,6446	0,00818	***

4. Causalidad e Invertibilidad:

Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	10,0953	0,0000	10,0953	0,0000
AR (estacional)					
	Raíz 1	1,3933	0,0000	1,3933	0,0000
MA					
	Raíz 1	-0,0653	-2,8670	2,8678	-0,2536
	Raíz 2	-0,0653	2,8670	2,8678	0,2536
MA (estacional)					
	Raíz 1	1,7760	0,0000	1,7760	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	1,5806	0,0000	1,5806	0,0000
MA					
	Raíz 1	-0,4434	-2,7030	2,7392	-0,2759
	Raíz 2	-0,4434	2,7030	2,7392	0,2759
MA (estacional)					
	Raíz 1	2,1322	0,0000	2,1322	0,0000

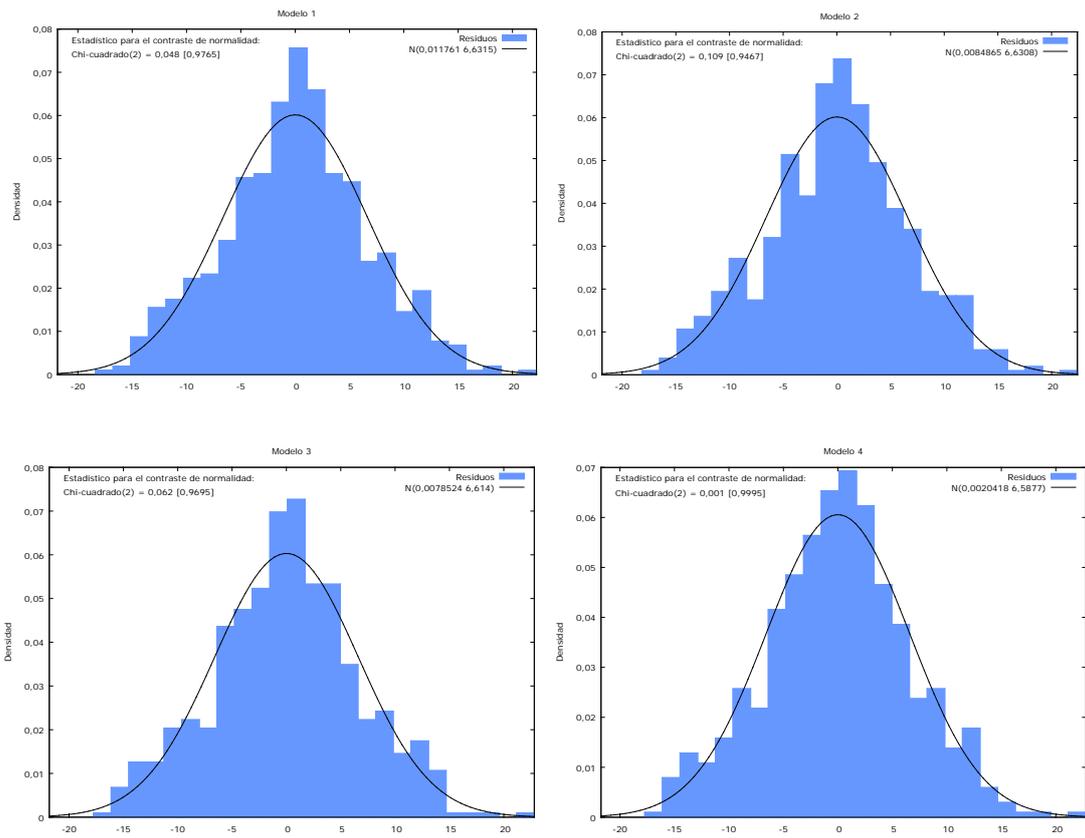
Modelo 3					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	1,9216	0,0000	1,9216	0,0000
MA					
	Raíz 1	-1,2561	-0,9626	1,5826	-0,3959
	Raíz 2	-1,2561	0,9626	1,5826	0,3959
	Raíz 3	1,7001	0,0000	1,7001	0,0000
	Raíz 4	0,3159	-1,4515	1,4855	-0,2159
	Raíz 5	0,3159	1,4515	1,4855	0,2159
MA (estacional)					
	Raíz 1	2,7406	0,0000	2,7406	0,0000

Modelo 4					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	2,7318	0,0000	2,7318	0,0000
	Raíz 2	-4,5142	0,0000	4,5142	0,5000
MA					
	Raíz 1	-1,2530	-0,9534	1,5745	-0,3965
	Raíz 2	-1,2530	0,9534	1,5745	0,3965
	Raíz 3	1,7019	0,0000	1,7019	0,0000
	Raíz 4	0,3064	-1,4400	1,4722	-0,2166
	Raíz 5	0,3064	1,4400	1,4722	0,2166

5. Análisis de los residuos

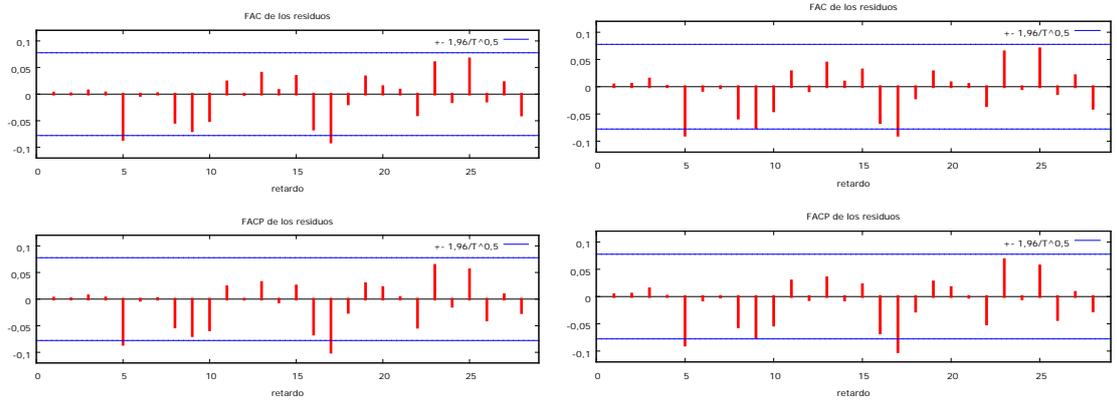
5.a. Normalidad:

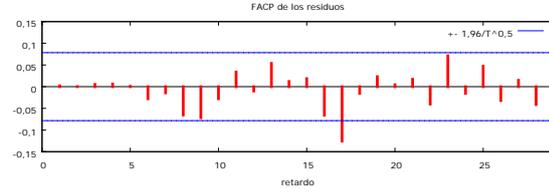
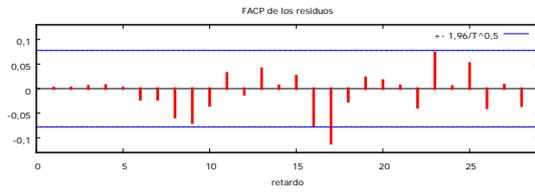
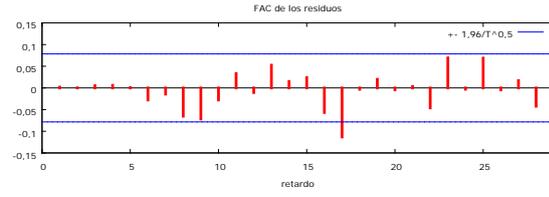
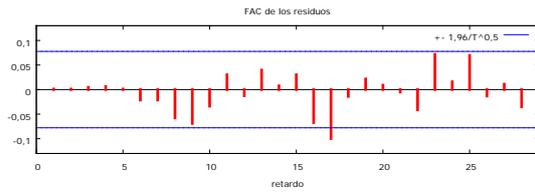
	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	0,048	0,97648
Modelo 2	0,109	0,94674
Modelo 3	0,062	0,96946
Modelo 4	0,001	0,99951



5.b. Media y Correlograma residual:

	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,0117612	6,60533	0,0448688	0,9642
Modelo 2	0,00848647	6,60986	0,032379	0,9742
Modelo 3	0,00785241	6,57748	0,0301073	0,976
Modelo 4	0,00204179	6,55056	0,00778618	0,9938

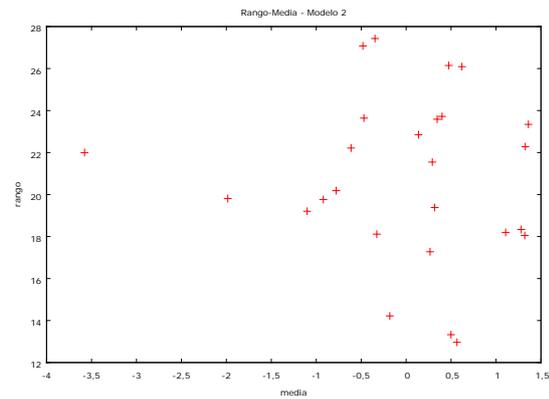
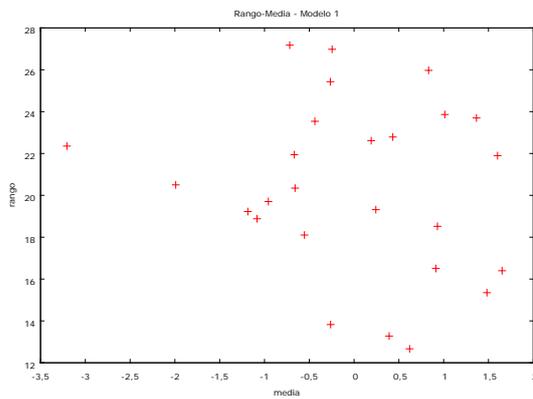


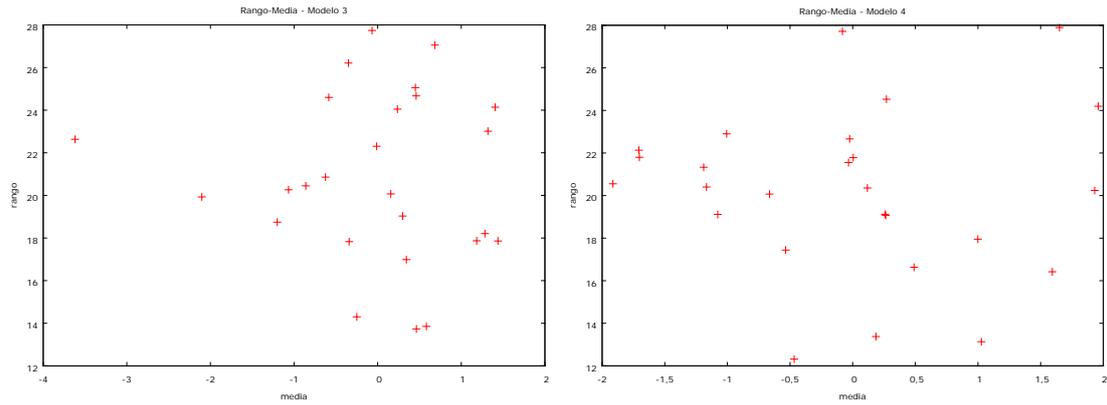


	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0172	38,7157	0,348
Modelo 2	-0,0013	40,2796	0,287
Modelo 3	0,0162	35,5476	0,490
Modelo 4	0,0389	38,9409	0,339

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,619364	0,394495
Modelo 2	-0,325992	0,657681
Modelo 3	-0,227949	0,750702
Modelo 4	-0,0858718	0,90546





6. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1						
const	phi_1	Phi_1	theta_1	theta_2	Theta_1	
1,31636	-0,227716	-0,0591977	0,229239	0,0292316	0,0645171	const
	0,0754056	-0,000727379	-0,0749411	-0,00876390	0,00101934	phi_1
		0,00618109	0,000467030	-0,000205400	-0,00684421	Phi_1
			0,0756847	0,00864399	-0,000795144	theta_1
				0,00253262	0,000456177	theta_2
					0,00875167	Theta_1

Correlación entre los parámetros estimados del Modelo 1					
	φ_1	Φ_1	θ_1	θ_2	Θ_1
φ_1	1	-0,03369195	-0,99200338	-0,63417724	0,03967997
Φ_1		1	0,02159273	-0,05191386	-0,93056174
θ_1			1	0,62434525	-0,03089551
θ_2				1	0,09689538
Θ_1					1

Matriz de Covarianzas del Modelo 2					
const	Phi_1	theta_1	theta_2	Theta_1	
1,10227	-0,0970138	0,00392445	0,00293640	0,103466	const
	0,00879262	-0,000316031	-0,000273431	-0,00941620	Phi_1
		0,00120463	5,75701e-005	0,000264874	theta_1
			0,00147624	0,000518269	theta_2
				0,0113978	Theta_1

Correlación entre los parámetros estimados del Modelo 2				
	Φ_1	θ_1	θ_2	Θ_1
Φ_1	1	-0,09710539	-0,07589458	-0,94059736
θ_1		1	0,04317101	0,07148248
θ_2			1	0,12634699
Θ_1				1

Matriz de Covarianzas del Modelo 3								
const	Phi_1	theta_1	theta_2	theta_3	theta_4	theta_5	Theta_1	
1,70252	-0,151262	0,00554860	0,000418726	-0,00224619	0,00350288	0,00137032	0,156506	const
	0,0137486	-0,000469867	-4,07991e-005	0,000159783	-0,000348506	-0,000119437	-0,0142726	Phi_1
		0,00121785	0,000109371	8,19339e-005	4,59709e-006	-0,000131573	0,000416396	theta_1
			0,00137607	0,000136082	6,57020e-005	-7,43913e-005	0,000217998	theta_2
				0,00157939	0,000277530	0,000225906	-5,90300e-005	theta_3
					0,00167442	0,000222378	0,000456636	theta_4
						0,00168799	0,000128770	theta_5
							0,0162310	Theta_1

Correlación entre los parámetros estimados del Modelo 3							
	Φ_1	θ_1	θ_2	θ_3	θ_4	θ_5	Θ_1
Φ_1	1	-0,11482871	-0,00938	0,03428922	-0,0726357	-0,02479287	-0,95543818
θ_1		1	0,08448611	0,05907741	0,00321924	-0,09176681	0,09365628
θ_2			1	0,09230714	0,04328387	-0,048811	0,04612747
θ_3				1	0,17066016	0,13835602	-0,01165882
θ_4					1	0,13227417	0,08759209
θ_5						1	0,02460127
Θ_1							1

Matriz de Covarianzas del Modelo 4								
const	Phi_1	Phi_2	theta_1	theta_2	theta_3	theta_4	theta_5	
0,457837	-0,0159580	-0,0154510	0,00239778	-0,00447715	-0,00230981	-0,000916284	0,000431763	const
	0,00153706	-0,000124193	-4,78491e-005	8,19321e-005	0,000100190	0,000128283	-5,83326e-006	Phi_1
		0,00144169	-0,000126134	0,000308832	3,99910e-005	-0,000102836	-2,57205e-005	Phi_2
			0,00120534	0,000125145	9,50651e-005	2,14527e-005	-0,000140561	theta_1
				0,00144219	0,000161705	3,43543e-005	-8,27046e-005	theta_2
					0,00158516	0,000297839	0,000202374	theta_3
						0,00168745	0,000245251	theta_4
							0,00170978	theta_5

Correlación entre los parámetros estimados del Modelo 4							
	Φ_1	Φ_2	θ_1	θ_2	θ_3	θ_4	θ_5
Φ_1	1	-0,08342886	-0,03515386	0,05502993	0,06418651	0,0796542	-0,00359829
Φ_2		1	-0,09568419	0,21417852	0,02645394	-0,06593159	-0,01638223
θ_1			1	0,0949177	0,0687748	0,01504217	-0,09791261
θ_2				1	0,10694913	0,02202192	-0,05266827
θ_3					1	0,18210846	0,12292738
θ_4						1	0,14438594
θ_5							1

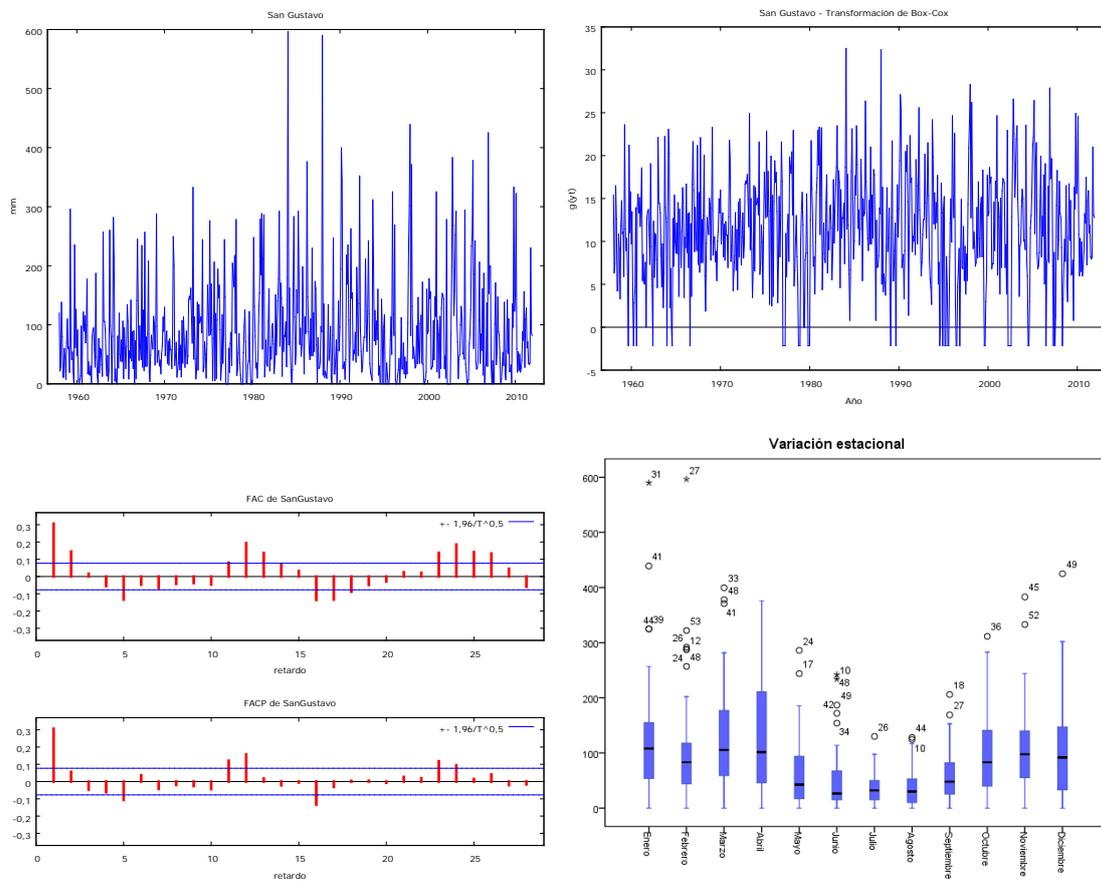
7. Criterios de información:

	AIC	SBC	HQ
Modelo 1	4212,657	4243,832	4224,762
Modelo 2	4218,141	4244,872	4228,520
Modelo 3	4217,896	4257,993	4233,464
Modelo 4	4133,514	4173,440	4149,029

Conclusión: Claramente el mejor ajuste residual lo muestra el Modelo 4. No está clara la homocedasticidad en el caso de los otros modelos y puntualmente en el 1 y el 2 los correlogramas residuales tienen retardos en los que la ACF está fuera de las bandas de confianza. Las condiciones de causalidad e invertibilidad están satisfechas y sólo el Modelo 4 no tiene parámetros con altas correlaciones. Además, los criterios de información alcanzan los valores más bajos para este modelo.

Apéndice XIX: San Gustavo

1. Representaciones gráficas.



2. Modelos propuestos:

Modelo 1: $ARIMA(1,0,1) \times (2,0,0)_{12}$

Modelo 2: $ARIMA(1,0,10) \times (1,0,1)_{12}$

3. Estimación de los parámetros:

Modelo 1					
	Coficiente	Desv. Típica	z	Valor p	
const	4,74973	1,11621	4,2552	0,00002	***
phi_1	0,435584	0,120808	3,6056	0,00031	***
Phi_1	0,149665	0,0380447	3,9339	0,00008	***
Phi_2	0,109613	0,0414099	2,6470	0,00812	***
theta_1	-0,162094	0,131881	-1,2291	0,21904	

Modelo 2					
	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,499134	0,293221	1,7022	0,08871	*
phi_1	0,374342	0,295048	1,2687	0,20453	
Phi_1	0,932982	0,0300563	31,0411	<0,00001	***
theta_1	-0,167003	0,2959	-0,5644	0,57249	
theta_2	0,0214262	0,070112	0,3056	0,75991	
theta_3	0,00100291	0,050317	0,0199	0,98410	
theta_4	0,0470373	0,0436023	1,0788	0,28069	
theta_5	-0,0756945	0,0497806	-1,5206	0,12837	
theta_6	0,0312166	0,0426081	0,7326	0,46378	
theta_7	0,000217146	0,0387302	0,0056	0,99553	
theta_8	0,0761299	0,0442873	1,7190	0,08561	*
theta_9	-0,0223114	0,0503172	-0,4434	0,65747	
theta_10	-0,121764	0,0406633	-2,9945	0,00275	***
Theta_1	-0,809876	0,0471394	-17,1804	<0,00001	***

4. Causalidad e Invertibilidad:

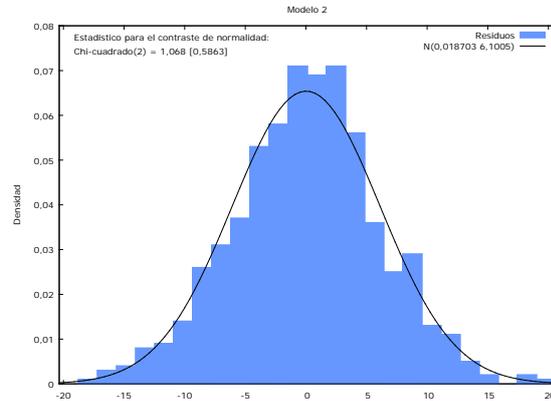
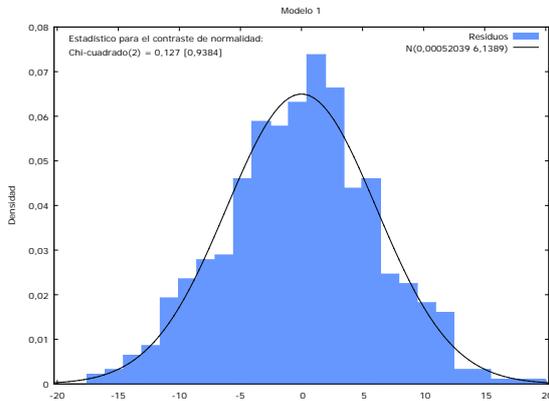
Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	2,2958	0,0000	2,2958	0,0000
AR (estacional)					
	Raíz 1	2,4139	0,0000	2,4139	0,0000
	Raíz 2	-3,7793	0,0000	3,7793	0,5000
MA					
	Raíz 1	6,1693	0,0000	6,1693	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	2,6714	0,0000	2,6714	0,0000
AR (estacional)					
	Raíz 1	1,0718	0,0000	1,0718	0,0000
MA					
	Raíz 1	1,2479	0,0000	1,2479	0,0000
	Raíz 2	-1,4118	0,0000	1,4118	0,5000
	Raíz 3	1,0131	-0,6747	1,2172	-0,0935
	Raíz 4	1,0131	0,6747	1,2172	0,0935
	Raíz 5	-1,0168	-0,6900	1,2288	-0,4051
	Raíz 6	-1,0168	0,6900	1,2288	0,4051
	Raíz 7	0,3937	-1,1001	1,1684	-0,1953
	Raíz 8	0,3937	1,1001	1,1684	0,1953
	Raíz 9	-0,3996	-1,1690	1,2354	-0,3024
	Raíz 10	-0,3996	1,1690	1,2354	0,3024
MA (estacional)					
	Raíz 1	1,2348	0,0000	1,2348	0,0000

5. Análisis de los residuos.

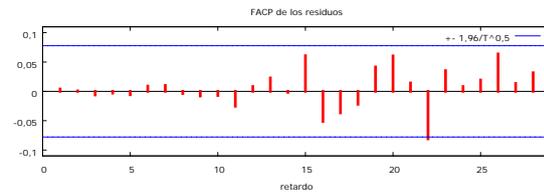
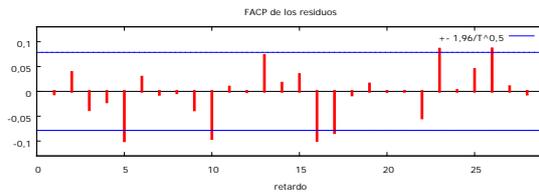
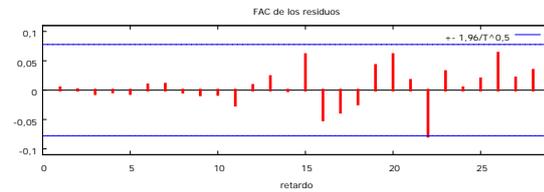
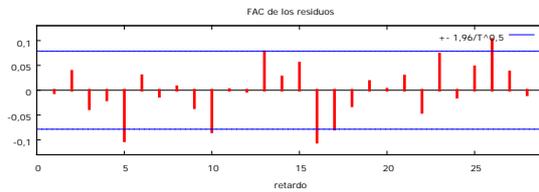
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	0,127	0,93839
Modelo 2	1,068	0,58627



5.b. Media y Correlograma residual:

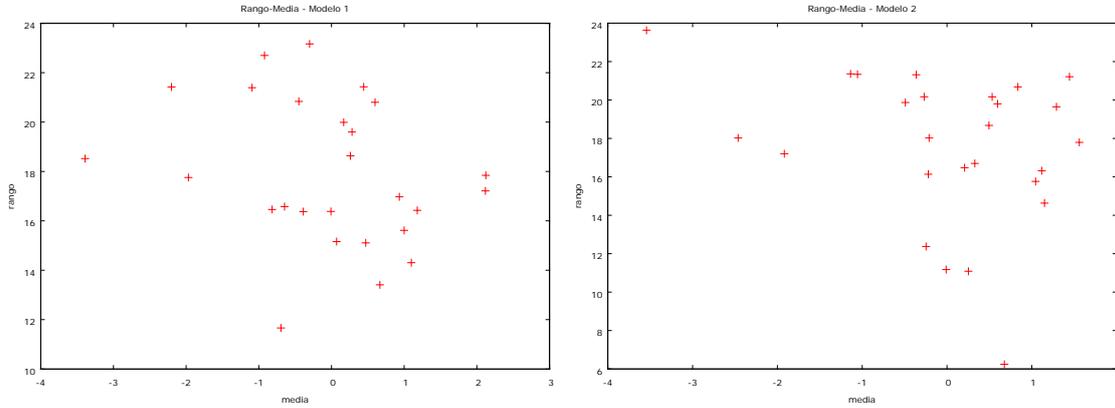
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,000520393	6,11916	0,00212268	0,9983
Modelo 2	0,018703	6,0376	0,0780611	0,9378



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	0,0407	61,5028	0,005
Modelo 2	-0,0437	26,7810	0,868

5.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	-0,62283	0,19266
Modelo 2	-0,91258	0,160599



6. Matriz de Covarianzas y Correlación entre coeficientes:

const	phi_1	Phi_1	Phi_2	theta_1	
1,24593	-0,126009	-0,0117139	-0,0101563	0,132289	const
	0,0145946	0,000460227	7,39013e-005	-0,0153198	phi_1
		0,00144740	-0,000280949	-0,000497647	Phi_1
			0,00171478	-0,000116821	Phi_2
				0,0173926	theta_1

Correlación entre los parámetros estimados del Modelo 1				
	φ_1	Φ_1	Φ_2	θ_1
φ_1	1	0,10013415	0,01477244	-0,9615573
Φ_1		1	-0,17833194	-0,09918477
Φ_2			1	-0,02139117
θ_1				1

Matriz de Covarianzas del Modelo 2					
const	Phi_1	theta_8	theta_10	Theta_1	
0,0859788	-0,00519074	4,54337e-04	4,84854e-04	0,00472438	const
	9,03383e-04	1,01507e-04	-1,28376e-04	-0,00119225	Phi_1
		0,00196136	1,34044e-04	-5,71799e-05	theta_8
			0,0016535	1,90057e-04	theta_10
				0,00222213	Theta_1

Correlación entre los parámetros estimados del Modelo 2				
	Φ_1	θ_8	θ_{10}	θ_1
Φ_1	1	0,07625727	-0,10503782	-0,84148768
θ_8		1	0,07443301	-0,02738925
θ_{10}			1	0,09915102
θ_1				1

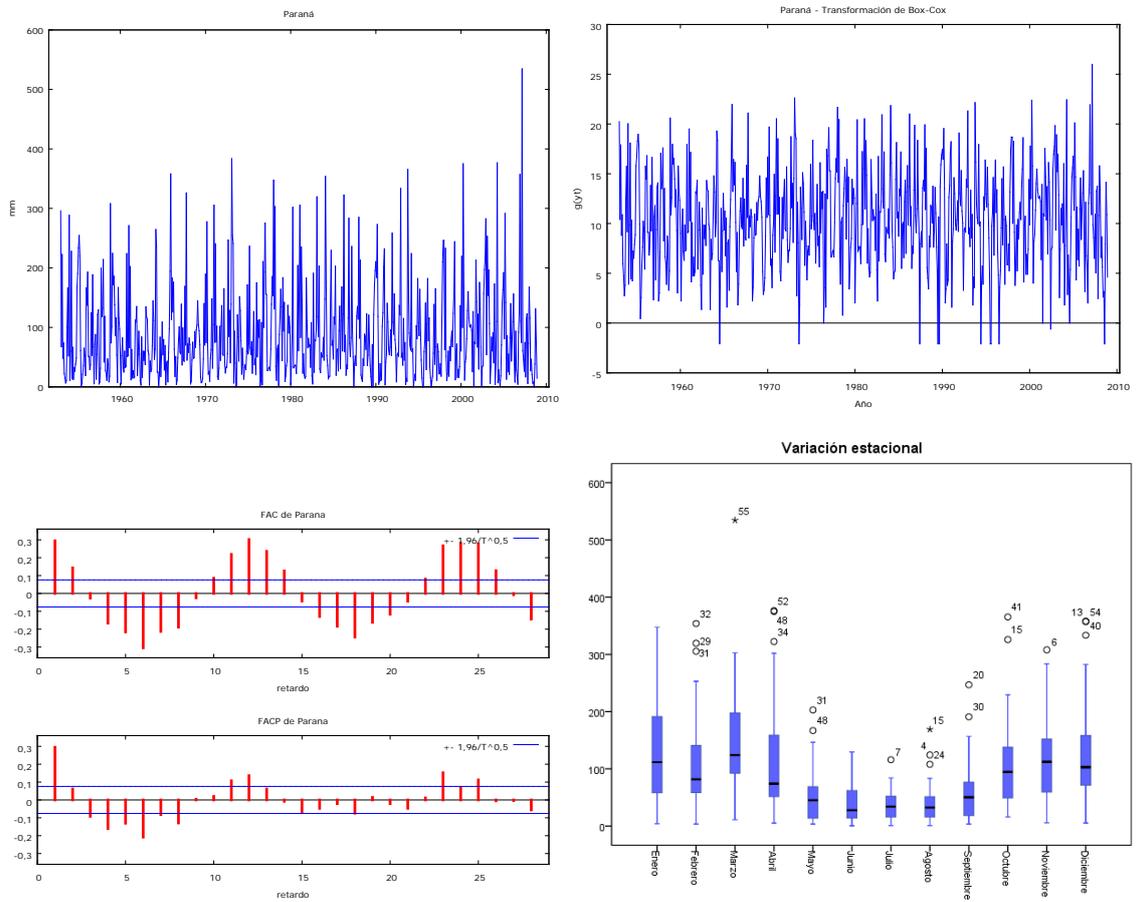
7. Criterios de información:

	AIC	SBC	HQ
Modelo 1	4036,032	4062,640	4046,373
Modelo 2	4114,525	4181,330	4140,464

Conclusión: El Modelo 1 se sugiere de la observación del correlograma. A excepción de que no presenta residuos con estructura de ruido blanco, el resto del ajuste es bueno. Existe alta correlación entre los parámetros autorregresivo y de media móvil de la parte regular pero este último no es significativo y, por lo tanto, puede suprimirse. El Modelo 2 se propone con un orden 10 de media móvil para alcanzar residuos con ajuste a un ruido blanco. Aunque se logra el objetivo, se pierde homocedasticidad (ver que el gráfico rango-media tiene forma de embudo con pendiente negativa) y *pasimonia*, los parámetros autorregresivo y de media móvil estacionales tienen elevada correlación y los criterios de información toman valores más altos que los del Modelo 1. Otras opciones que no se incluyen fueron exploradas sin obtener mejores resultados.

Apéndice XX: Paraná

1. Representaciones gráficas.



2. Modelos propuestos:

Modelo 1: $ARIMA(2,0,2)x(1,0,1)_{12}$

Modelo 2: $ARIMA(0,0,6)x(1,0,1)_{12}$

3. Estimación de parámetros:

Modelo 1					
	Coeficiente	Desv. Típica	z	Valor p	
const	0,397554	0,164086	2,4228	0,01540	**
phi_1	0,108337	0,195497	0,5542	0,57947	
phi_2	-0,389634	0,182569	-2,1342	0,03283	**
Phi_1	0,970953	0,0101659	95,5112	<0,00001	***
theta_1	-0,0551719	0,191662	-0,2879	0,77345	
theta_2	0,431453	0,177685	2,4282	0,01517	**
Theta_1	-0,866282	0,022681	-38,1941	<0,00001	***

Modelo 2					
	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Valor p</i>	
const	0,274851	0,0969164	2,8360	0,00457	***
Phi_1	0,974067	0,00898099	108,4588	<0,00001	***
theta_1	0,0592301	0,0415522	1,4254	0,15403	
theta_2	0,0491378	0,0394127	1,2468	0,21249	
theta_3	-0,0246789	0,0373785	-0,6602	0,50910	
theta_4	-0,0188897	0,0403399	-0,4683	0,63960	
theta_5	-0,0105108	0,0422774	-0,2486	0,80366	
theta_6	-0,0999834	0,0413079	-2,4204	0,01550	**
Theta_1	-0,887638	0,0209949	-42,2787	<0,00001	***

4. Causalidad e Invertibilidad:

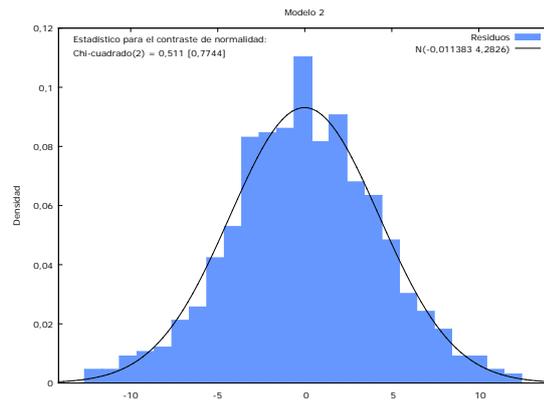
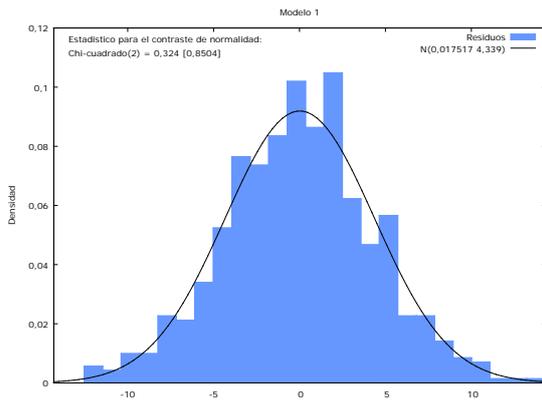
Modelo 1					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR					
	Raíz 1	0,1390	-1,5960	1,6020	-0,2362
	Raíz 2	0,1390	1,5960	1,6020	0,2362
AR (estacional)					
	Raíz 1	1,0299	0,0000	1,0299	0,0000
MA					
	Raíz 1	0,0639	-1,5211	1,5224	-0,2433
	Raíz 2	0,0639	1,5211	1,5224	0,2433
MA (estacional)					
	Raíz 1	1,1544	0,0000	1,1544	0,0000

Modelo 2					
		<i>Real</i>	<i>Imaginaria</i>	<i>Módulo</i>	<i>Frecuencia</i>
AR (estacional)					
	Raíz 1	1,0266	0,0000	1,0266	0,0000
MA					
	Raíz 1	-1,4882	0,0000	1,4882	0,5000
	Raíz 2	1,4578	0,0000	1,4578	0,0000
	Raíz 3	0,6780	-1,3236	1,4871	-0,1747
	Raíz 4	0,6780	1,3236	1,4871	0,1747
	Raíz 5	-0,7153	-1,2541	1,4438	-0,3325
	Raíz 6	-0,7153	1,2541	1,4438	0,3325
MA (estacional)					
	Raíz 1	1,1266	0,0000	1,1266	0,0000

5. Análisis de los residuos

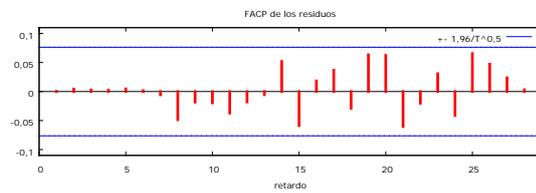
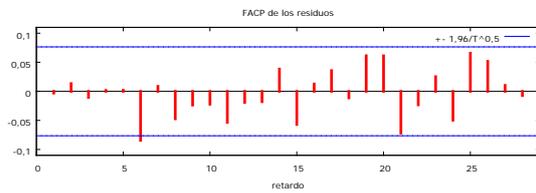
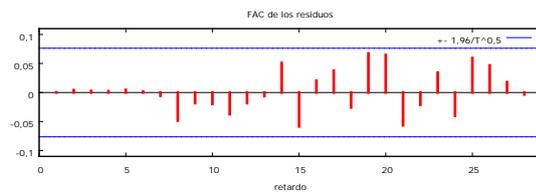
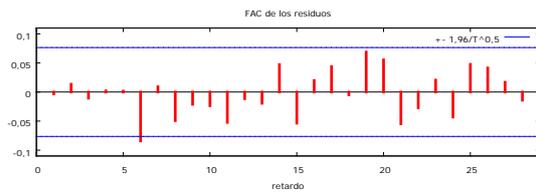
5.a. Normalidad:

	Estadístico de Doornik-Hansen	Valor-p
Modelo 1	0,324	0,85043
Modelo 2	0,511	0,77445



5.b. Media y Correlograma residual:

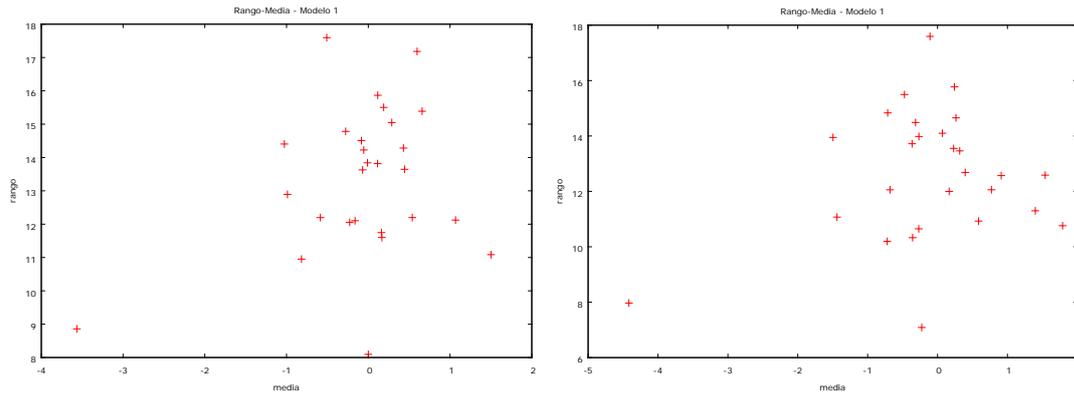
	Media	Desviación Típica	Estadístico t	Valor-p a dos colas
Modelo 1	0,0175166	4,31916	0,104031	0,9172
Modelo 2	-0,0113828	4,25648	-0,0687024	0,9452



	$\hat{\rho}_{36}$	Q_{36}	Valor-p
Modelo 1	-0,0047	45,0221	0,144
Modelo 2	-0,0033	40,6910	0,271

4.c. Homocedasticidad:

	Pendiente estimada	Valor-p
Modelo 1	0,723538	0,139407
Modelo 2	0,441801	0,269524



4.d. Matriz de Covarianzas y Correlación entre coeficientes:

Matriz de Covarianzas del Modelo 1							
const	phi_1	phi_2	Phi_1	theta_1	theta_2	Theta_1	
0,0269241	-0,0126713	-0,00768716	-0,00144376	0,0124576	0,00854746	0,00137284	const
	0,0382192	-0,00591245	0,000212567	-0,0367671	0,00285299	-0,000211846	phi_1
		0,0333315	-4,91687e-005	0,00668349	-0,0316501	-1,55851e-005	phi_2
			0,000103345	-0,000254188	1,50419e-005	-0,000104661	Phi_1
				0,0367344	-0,00375789	0,000298315	theta_1
					0,0315718	-2,03309e-005	theta_2
						0,000514429	Theta_1

Correlación entre los parámetros estimados del Modelo 1				
	φ_2	Φ_1	θ_2	Θ_1
φ_2	1	-0,02649207	-0,97565727	-0,00376375
Φ_1		1	0,00832734	-0,45391742
θ_2			1	-0,0050448
Θ_1				1

Matriz de Covarianzas del Modelo 2					
const	Phi_1	theta_1	theta_2	theta_3	
0,00939279	-0,000850306	0,000249120	0,000188881	-6,48788e-005	const
	8,06582e-005	-3,01068e-005	-1,78880e-005	5,54830e-006	Phi_1
		0,00172658	1,11934e-005	0,000161369	theta_1
			0,00155336	2,22617e-005	theta_2
				0,00139715	theta_3
	theta_4	theta_5	theta_6	Theta_1	
	5,77069e-005	-9,45781e-005	-0,000245217	0,000756126	const
	-4,96928e-006	5,50692e-006	2,39976e-005	-7,65566e-005	Phi_1
	0,000102573	2,58907e-005	-2,04344e-005	4,18162e-005	theta_1
	-2,58968e-005	-8,03083e-005	3,81339e-005	-5,44251e-005	theta_2
	0,000180286	9,45743e-005	9,88637e-005	1,57522e-005	theta_3
	0,00162731	0,000175655	6,75258e-005	2,04040e-005	theta_4
		0,00178738	8,73210e-005	8,47505e-005	theta_5
			0,00170634	-9,19856e-005	theta_6
				0,000440786	Theta_1

Correlación entre los parámetros estimados del Modelo 2			
	Φ_1	θ_6	θ_1
Φ_1	1	0,06468603	-0,40601737
θ_2		1	-0,1060652
θ_1			1

4.c. Criterios de información:

	AIC	SBC	HQ
Modelo 1	3807,722	3843,636	3821,644
Modelo 2	3803,946	3848,868	3821,358

Conclusión: El Modelo 1, propuesto a partir de la observación del correlograma, no tiene buen ajuste a un ruido blanco en los residuos y los parámetros autorregresivo y de media móvil de orden 2 tienen alta correlación y no pueden suprimirse porque ambos son significativos. Por otro lado, el segundo modelo se sugiere mediante la introducción de términos de media móvil hasta lograr que los residuos tengan comportamiento de ruido blanco. En general, este modelo muestra mejor ajuste y no aparecen altas correlaciones entre sus parámetros significativos.

Apéndice XXI: Series de precipitaciones de lluvia con evidencia de raíces unitarias en alguna frecuencia. Resultados de los contrastes HEGY y Canova-Hansen

1. Octavo Distrito

Canova-Hansen:

Degrees of freedom (T-k) = 211, lag order = 5

Statistic	p-value	Ang. Frequency	Period
L_1 = 0,1487	0,88779	$\pm\pi/6$	12
L_2 = 0,6411	0,08357 *	$\pm\pi/3$	6
L_3 = 0,6857	0,06523 *	$\pm\pi/2$	4
L_4 = 0,1905	0,77241	$\pm 2\pi/3$	3
L_5 = 0,2572	0,58519	$\pm 5\pi/6$	2,40
L_pi= 0,1280	0,49424	π	2
L_f = 1,9028	0,27570	Joint test	

2. San Víctor

a. Test HEGY:

Deterministic component: constant - Dof (T-k) = 197

Statistic	p-value	Ang. Frequency	Period
t1= -3,94	0,00091 ***	zero	infinity
F1= 4,99	0,00871 ***	$\pm\pi/6$	12
F2= 4,62	0,01232 **	$\pm\pi/3$	6
F3= 6,17	0,00283 ***	$\pm\pi/2$	4
F4= 8,34	0,00036 ***	$\pm 2\pi/3$	3
F5= 16,22	0,00000 ***	$\pm 5\pi/6$	2,40
t2= -4,09	0,00000 ***	π	2
Fs= 9,45	0,00000	All the seasonal cycles	
Ft= 10,31	0,00000	Delta_s (all the seas. + zero freq.)	

b. Canova-Hansen:

Degrees of freedom (T-k) = 213, lag order = 5

Statistic	p-value	Ang. Frequency	Period
L_1 = 0,1027	0,97659	$+-\pi/6$	12
L_2 = 0,2664	0,56102	$+-\pi/3$	6
L_3 = 0,7655	0,04140 **	$+-\pi/2$	4
L_4 = 0,2784	0,53123	$+2\pi/3$	3
L_5 = 0,3436	0,38915	$+5\pi/6$	2,40
L_pi= 0,0969	0,63650	π	2
L_f = 2,0134	0,17774	Joint test	

3. Santa Anita

Canova-Hansen:

Degrees of freedom (T-k) = 210, lag order = 5

Statistic	p-value	Ang. Frequency	Period
L_1 = 0,1266	0,93806	$+-\pi/6$	12
L_2 = 0,3978	0,29776	$+-\pi/3$	6
L_3 = 0,6129	0,09743 *	$+-\pi/2$	4
L_4 = 0,3022	0,47579	$+2\pi/3$	3
L_5 = 0,2687	0,55559	$+5\pi/6$	2,40
L_pi= 0,1292	0,48959	π	2
L_f = 1,4161	0,82463	Joint test	

4. Séptimo Distrito

Test HEGY

Deterministic component: constant - Dof (T-k) = 203

Statistic	p-value	Ang. Frequency	Period
t1= -3,35	0,00573 ***	zero	infinity
F1= 4,59	0,01413 **	$+-\pi/6$	12
F2= 9,76	0,00011 ***	$+-\pi/3$	6
F3= 10,27	0,00000 ***	$+-\pi/2$	4
F4= 15,89	0,00000 ***	$+2*\pi/3$	3
F5= 14,09	0,00000 ***	$+5*\pi/6$	2,40
t2= -5,41	0,00000 ***	π	2
Fs= 13,47	0,00000	All the seasonal cycles	
Ft= 13,27	0,00000	Delta_s (all the seas. + zero freq.)	

5. Viale

a. Test HEGY

Deterministic component: constant - Dof (T-k) = 175

Statistic	p-value	Ang. Frequency	Period
t1= -3,03	0,01439 **	zero	infinity
F1= 5,93	0,00411 ***	$\pm\pi/6$	12
F2= 12,20	0,00000 ***	$\pm\pi/3$	6
F3= 9,77	0,00011 ***	$\pm\pi/2$	4
F4= 11,05	0,00000 ***	$\pm 2\pi/3$	3
F5= 10,81	0,00000 ***	$\pm 5\pi/6$	2,40
t2= -3,66	0,00000 ***	π	2
Fs= 13,00	0,00000	All the seasonal cycles	
Ft= 12,52	0,00000	Delta_s (all the seas. + zero freq.)	

b. Canova-Hansen

Degrees of freedom (T-k) = 187, lag order = 5

Statistic	p-value	Ang. Frequency	Period
L_1 = 0,2691	0,55876	$\pm\pi/6$	12
L_2 = 0,7819	0,03715 **	$\pm\pi/3$	6
L_3 = 0,8762	0,02092 **	$\pm\pi/2$	4
L_4 = 0,4199	0,26822	$\pm 2\pi/3$	3
L_5 = 0,4156	0,27417	$\pm 5\pi/6$	2,40
L_pi= 0,2022	0,28291	π	2
L_f = 2,4252	0,02046 **	Joint test	

6. Villa Paranacito

Test HEGY: Deterministic component: constant - Dof (T-k) = 205

Statistic	p-value	Ang. Frequency	Period
t1= -3,07	0,01387 **	zero	infinity
F1= 12,40	0,00000 ***	$\pm\pi/6$	12
F2= 10,68	0,00000 ***	$\pm\pi/3$	6
F3= 18,51	0,00000 ***	$\pm\pi/2$	4
F4= 12,75	0,00000 ***	$\pm 2\pi/3$	3
F5= 11,27	0,00000 ***	$\pm 5\pi/6$	2,40
t2= -4,54	0,00000 ***	π	2
Fs= 15,02	0,00000	All the seasonal cycles	
Ft= 14,52	0,00000	Delta_s (all the seas. + zero freq.)	

Test HEGY: Deterministic component: constant + (s-1) trigonometric terms - Dof
(T-k) = 194

Statistic	p-value	Ang. Frequency	Period
t1= -3,08	0,01114 **	zero	infinity
F1= 16,95	0,00000 ***	+ $-\pi/6$	12
F2= 12,71	0,00000 ***	+ $-\pi/3$	6
F3= 19,81	0,00000 ***	+ $-\pi/2$	4
F4= 14,10	0,00000 ***	+ $-2\pi/3$	3
F5= 11,86	0,00000 ***	+ $-5\pi/6$	2,40
t2= -4,78	0,00000 ***	π	2
Fs= 17,02	0,00000	All the seasonal cycles	
Ft= 16,37	0,00000	Delta_s (all the seas. + zero freq.)	

7. Feliciano

Canova-Hansen

Degrees of freedom (T-k) = 518, lag order = 7

Statistic	p-value	Ang. Frequency	Period
L_1 = 1,0901	0,00737 ***	+ $-\pi/6$	12
L_2 = 0,8458	0,02837 **	+ $-\pi/3$	6
L_3 = 0,3196	0,42293	+ $-\pi/2$	4
L_4 = 0,8230	0,03207 **	+ $-2\pi/3$	3
L_5 = 0,3627	0,34282	+ $-5\pi/6$	2,40
L_pi= 0,0365	0,96303	π	2
L_f = 2,9514	0,00785 ***	Joint test	

8. San Jaime

Canova-Hansen

Degrees of freedom (T-k) = 522, lag order = 7

Statistic	p-value	Ang. Frequency	Period
L_1 = 0,5396	0,14140	+ $-\pi/6$	12
L_2 = 0,8152	0,03347 **	+ $-\pi/3$	6
L_3 = 0,1601	0,84189	+ $-\pi/2$	4
L_4 = 0,9612	0,01513 **	+ $-2\pi/3$	3
L_5 = 0,1531	0,86135	+ $-5\pi/6$	2,40
L_pi= 0,3139	0,12657	π	2
L_f = 2,3717	0,08439 *	Joint test	

Apéndice XXII: Transformaciones de Doornik-Hansen

La transformación z_1 es como sigue:

$$\beta = \frac{3(n^2 + 27n - 70)(n + 1)(n + 3)}{(n - 2)(n + 5)(n + 7)(n + 9)}$$

$$\omega^2 = -1 + \{2(\beta - 1)\}^{\frac{1}{2}}$$

$$\delta = \frac{1}{\{\log(\sqrt{\omega^2})\}^{\frac{1}{2}}}$$

$$y = \sqrt{b_1} \left\{ \frac{\omega^2 - 1}{2} \cdot \frac{(n + 1)(n + 3)}{6(n - 2)} \right\}^{\frac{1}{2}}$$

$$z_1 = \delta \cdot \log \left\{ y + (y^2 + 1)^{\frac{1}{2}} \right\}$$

Para la transformación z_2 se tiene en cuenta:

$$\delta = (n - 3)(n + 1)(n^2 + 15n - 4)$$

$$a = \frac{(n - 2)(n + 5)(n + 7)(n^2 + 27n - 70)}{6\delta}$$

$$c = \frac{(n - 7)(n + 5)(n + 7)(n^2 + 2n - 5)}{6\delta}$$

$$k = \frac{(n + 5)(n + 7)(n^3 + 37n^2 + 11n - 313)}{12\delta}$$

$$\alpha = a + b_1 \cdot c$$

$$\chi = (b_2 - 1 - b_1)2k$$

$$z_2 = \left\{ \left(\frac{\chi}{2\alpha} \right)^{\frac{1}{3}} - 1 + \frac{1}{9\alpha} \right\} \cdot (9\alpha)^{\frac{1}{2}}$$

Bibliografía

- Aguirre Jaime, A. (1994). *Introducción al tratamiento de series temporales: aplicación a las ciencias de la salud*. Madrid: Díaz de Santos.
- Alonso, J. C., & Semaán, P. (Junio de 2010). Prueba de HEGY en R: Una Guía. (J. C. Alonso, Ed.) *Apuntes de Economía*(23).
- Arnedo, M., Ricarte, J., & Salvador, A. (1999). Un análisis de series temporales interrumpidas: reactividad cardíaca ante un estresor de laboratorio. *Rev. de Psicol. Gral. y Aplic.*, <http://dialnet.unirioja.es/descarga/articulo/2498368.pdf>.
- Beaulieu, J., & Miron, J. (1993). Seasonal Unit Roots in Aggregate U.S. Data. *Journal of Econometrics*, 55, 305-328.
- Bhattacharya, R. N., & Waymire, E. C. (2009). *Stochastic Processes with Applications*. Philadelphia: SIAM (Society for Industrial and Applied Mathematics).
- Bowerman, B. L., O'Connell, R. T., & Koehler, A. B. (2007). *Pronósticos, Series de Tiempo y Regresión. Un enfoque aplicado*. México: Cengage Learning.
- Bowman, A., & Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis*. New York: Clarendon Press. Oxford.
- Box, G., & Pierce, D. A. (1970). Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association*, 65(332), 1509-1526.
- Box, G., Jenkins, G., & Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control*. USA: Prentice-Hall International.
- Brockwell, P., & Davis, R. (1987). *Time Series: Theory and Methods*. New York: Springer.
- Brockwell, P., & Davis, R. (2002). *Introduction to Time Series and Forecasting*. New York: Springer.
- Burden, R., & Faires, J. D. (2009). *Análisis Numérico*. México: Cengage Learning.
- Burkill, J. C. (1951). *The Lebesgue Integral*. USA: Cambridge University Press.
- Canova, F., & Hansen, B. E. (Julio de 1995). Are Seasonal Patterns Constant Over Time? A Test for Seasonal Stability. *Journal of Business & Economic Statistics*, 13(3).
- Chatfield, C. (1995). *The Analysis of Time Series, an Introduction*. Washington, D.C.: Chapman y Hall/CRC.
- Churchill, R. V., & Brown, J. W. (1992). *Variable Compleja y Aplicaciones*. Madrid: McGraw-Hill.
- Cramér, H. (1966). *The Elements of Probability Theory and Some of its Applications*. New York: Science Editions, John Wiley & Sons.

- DeGroot, M. H. (1988). *Probabilidad y Estadística*. Wilmington, Delaware, U.S.A.: Addison-Wesley Iberoamericana.
- Díaz-Emparanza, I. (2011). Numerical Distribution Functions for Seasonal Unit Root Tests. (D. d. Vasco, Ed.) *Documentos de Trabajo Biltoki*. Recuperado el 24 de Enero de 2015, de <https://addi.ehu.es/bitstream/10810/5568/1/hegyg-pt.pdf>
- Díaz-Emparanza, I., & Moral, M. P. (2013). Seasonal Stability Tests in gretl. An Application to International Tourism Data. (D. d. Vasco, Ed.) *Documentos de Trabajo Biltoki*. Recuperado el 24 de Enero de 2015, de <https://addi.ehu.es/bitstream/10810/10577/1/2013.3.pdf>
- Dickey, D., & Fuller, W. (1979). Distributions of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, 74, 427 - 431. Recuperado el 02 de Marzo de 2014, de <http://www.deu.edu.tr/userweb/onder.hanedar/dosyalar/1979.pdf>
- Donner, R., & Barbosa, S. (2008). *Nonlinear Time Series Analysis in the Geosciences. Applications in Climatology, Geodynamics and Solar-Terrestrial Physics*. Germany: Springer.
- Doornik, J., & Hansen, H. (1994). An Omnibus Test for Univariate and Multivariate Normality. *Working Paper, Nuffield College, Oxord*.
- Elliott, G., Rothenberg, T., & Stock, J. (Julio de 1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica*, 64(4), 813 - 836. Recuperado el 22 de Julio de 2014, de http://scholar.harvard.edu/files/stock/files/efficient_tests_for_an_autoregressive_unit_root.pdf
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of the U.K. Inflation. *Econométrica*, 50(4), 987-1008.
- Evans, M. J., & Rosenthal, J. S. (2003). *Probabilidad y Estadística. La Ciencia de la Incertidumbre*. Barcelona: Reverté.
- Franses, P. H., & Hobijn, B. (1997). Critical values for unit root tests in seasonal time series. *Journal of Applied Statistics*, 24(1), 25 - 47. Recuperado el 24 de Agosto de 2014, de http://www.tandfonline.com/doi/abs/10.1080/02664769723864#.U_nW7_I5P3M
- Freund, J., & Walpole, R. (1990). *Estadística Matemática con Aplicaciones*. México: Prentice-Hall Hispanoamericana.
- Fuller, W. A. (1996). *Introduction to Statistical Time Series*. U.S.A.: Wiley.
- Gómez, V., & Taguas, D. (1995). Detección y Corrección Automática de Outliers con TRAMO: una aplicación al IPC de bienes industriales no energéticos. *Universidad de Zaragoza - Departamento de Análisis Económico*.
- Grossman, S. (1996). *Álgebra Lineal*. México: McGraw-Hill.

- Gujarati, D., & Porter, D. (2010). *Econometría*. México: McGraw-Hill.
- Gutiérrez, R. B., & Cintas, P. G. (2004). *55 Respuestas a Dudas Típicas de Estadística*. Madrid: Díaz de Santos.
- Harvey, A. C. (1993). *Time Series Models*. Cambridge, Massachusetts: The MIT Press.
- Hendry, D., & Mizon, G. (1978). Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England. *The Economic Journal*, 88(351), 549-563.
- Hylleberg, S. (1995). Test for Seasonal Unit Roots: General to Specific or Specific to General. *Journal of Econometrics*, 69, 5-25.
- Hylleberg, S., Engle, R., Granger, C., & Yoo, B. (1990). Seasonal Integration and Cointegration. *Journal of Econometrics*, 44, 215 - 238. Recuperado el 17 de Agosto de 2014, de <http://www.deu.edu.tr/userweb/onder.hanedar/dosyalar/HEGY.pdf>
- Juarez Cerrillo, S. F. (s.f.). *Análisis de la variabilidad climática*. Recuperado el 18 de Febrero de 2013, de http://www.peccuv.mx/descargas/pdf/reportes_investigacion/Capítulo%203%201%20Análisis%20Estadístico%20Historico%20de%20Variabilidad%20Climática.pdf
- Kwiatkowski, D., Phillips, P., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54, 159 - 178. Recuperado el 20 de Julio de 2014, de <http://www.ccee.edu.uy/ensenian/catmetec/material/KPSS.pdf>
- Landro, A., & González, M. (2009). *Elementos de econometría de los fenómenos dinámicos. Parte I. Los procesos estocásticos lineales unidimensionales*. Buenos Aires: Ediciones Cooperativas.
- Lehmann, E. L., & Casella, G. (1998). *Theory of Point Estimation*. New York: Springer.
- Lelys Guenni, E. D. (2008). Análisis de la Tendencia y la Estacionalidad de la Precipitación Mensual en Venezuela. *Revista Colombiana de Estadística*, volumen 31, no. 1, pp. 41 a 65.
- Ljung, G. M., & Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297-303.
- Marsden, J., & Tromba, A. (2004). *Cálculo Vectorial*. Madrid: Pearson - Addison Wesley.
- Mauricio, J. A. (2007). *Introducción al Análisis de Series Temporales*. Recuperado el 20 de Febrero de 2013, de <http://www.ucm.es/info/ecocuan/jam/JAM-IASST-Libro.pdf>
- Mudelesee, M. (2010). *Climate Time Series Analysis, Classical Statistical and Bootstrap Methods. Atmospheric and Oceanographic Sciences Library*. New York: Springer.

- Pérez Ramírez, F. O., & Correa Restrepo, O. (2007). *Introducción a las Series de Tiempo. Métodos Paramétricos*. Medellín, Colombia: Colección 5 Sello Editorial Universidad de Medellín.
- Pérez, C. (2001). *Técnicas estadísticas con SPSS*. Madrid: Pearson Prentice Hall.
- Petrov, V., & Mordecki, E. (2008). *Teoría de la Probabilidad*. Montevideo: DIRAC.
- Phillips, P., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335 - 346. Recuperado el 22 de Julio de 2014, de <http://www.ssc.wisc.edu/~bhansen/718/PhillipsPerron1988.pdf>
- Quarteroni, A., & Saleri, F. (2006). *Cálculo Científico con Matlab y Octave*. Milano, Italia: Springer.
- Rey Pastor, J., Pi Calleja, P., & Trejo, C. A. (1959). *Análisis Matemático. Análisis Funcional y Aplicaciones* (Vol. III). Buenos Aires: Kapelusz.
- Ruiz-Ramírez, J., Hernández-Rodríguez, G., & Zulueta-Rodríguez, R. (2011). Análisis de series del tiempo en el pronóstico de la producción de caña de azúcar. *Terra Latinoamericana*, Vol. 29, num. 1, pp-103-109; Recuperado el 1 de marzo de 2013 de <http://redalyc.uaemex.mx/redalyc/pdf/573/57319955011.pdf>.
- Saez, M., Pérez-Hoyos, S., Tobías, A., Saurina, C., Barceló, M. A., & Ballester, F. (1999). Métodos de series temporales en los estudios epidemiológicos sobre contaminación atmosférica. *Revista Española de Salud Pública*, 1; Recuperado el 1 de marzo de <http://www.bvsde.paho.org/bvsea/e/fulltext/metodo/metodo.pdf>.
- Schwarz, G. (1978). Estimating the Dimension of a model. *Annals of Statistics*, 6, 461-464.
- Schwert, W. (1989). Tests for Unit Root: A Monte Carlo Investigation. *Journal of Business & Economic Statistics*, 7(2), 147 - 159. Recuperado el 22 de Julio de 2014, de <http://schwert.ssb.rochester.edu/jbes89.pdf>
- Shumway, R. H., & Stoffer, D. S. (2011). *Time Series Analysis and Its Applications*. New York: Springer.
- Spivak, M. (2005). *Calculus. Cálculo Infinitesimal*. Barcelona: Reverté.
- Stirzaker, D. (2005). *Stochastic Processes and Models*. New York: OXFORD University Press.
- Tiao, G. C., & Box, G. E. (Diciembre de 1981). Modeling Multiple Times Series with Applications. *Journal of the American Statistical Association*, 76(376), 802-816.
- Triola, M. (2004). *Estadística*. México: Pearson Addison Wesley.
- Wasserman, L. (2006). *All of Nonparametric Statistics*. New York: Springer.
- Wei, W. W. (2006). *Time Series Analysis. Univariate and multivariate methods*. U.S.A.: Pearson Addison Wesley.

