



Facultad de Estudios Superiores

**Acatlán**

# **Pronósticos:**

## **Metodología de Box-Jenkins**



Mari Carmen González Videgaray

# **Pronósticos:**

## **Metodología de Box-Jenkins**



## UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**Dr. José Narro Robles**

Rector

**Dr. Eduardo Bárvana García**

Secretario General

## FACULTAD DE ESTUDIOS SUPERIORES ACATLÁN

**Dr. J. Alejandro Salcedo Aquino**

Director

**Dr. Darío Rivera Vargas**

Secretario General

**Mtro. Adalberto López López**

Secretario de Estudios Profesionales

**Mtro. Jorge Luis Suárez Madariaga**

Coordinador de Servicios Académicos

**Mtra. Nora del Consuelo Goris Mayans**

Jefa de la División de Matemáticas e Ingeniería

**D.G. Víctor Hugo Huerta González**

Jefe de la Unidad de Servicios Editoriales



Facultad de Estudios Superiores

**Acatlán**

# **Pronósticos:**

## **Metodología de Box-Jenkins**



Mari Carmen González Videgaray

Título: PRONOSTICOS:  
METODOLOGÍA DE BOX-JENKINS  
Mari Carmen González Videgaray

Primera edición: 2011

Portada: D.G. Juan Humberto Báez Rendón

© UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
Ciudad Universitaria, Delegación Coyoacán,  
C.P. 04510, México, Distrito Federal

© FACULTAD DE ESTUDIOS PROFESIONALES ACATLÁN  
Av. Alcanfores y San Juan Totoltepec, s/n.  
C.P. 53150, Naucalpan de Juárez, Estado de México.  
Unidad de Servicios Editoriales

ISBN: 978-607-02-2223-8

*Prohibida la reproducción total o parcial por  
Cualquier medio sin la autorización escrita del  
Titular de los derechos patrimoniales.*

*Impreso y hecho en México  
Printed and made in Mexico*

# Índice general

Prefacio	9
1 Conceptos básicos	13
1.1 Pronósticos y decisiones . . . . .	13
1.2 Una aproximación al problema . . . . .	14
1.3 Serie de tiempo . . . . .	15
1.3.1 Definición . . . . .	15
1.3.2 Componentes de una serie de tiempo . . . . .	17
1.4 Objetivos del análisis . . . . .	21
1.5 Modelos univariados y multivariados . . . . .	22
1.6 Metodología de Box-Jenkins . . . . .	23
1.7 Funciones de primer y segundo orden . . . . .	25
1.7.1 Función de media . . . . .	26
1.7.2 Función de varianza . . . . .	26
1.7.3 Funciones de relación . . . . .	26
1.8 Periodicidad y su detección . . . . .	27
1.8.1 Propiedades de la función seno . . . . .	27
1.8.2 Función seno generalizada . . . . .	27
1.8.3 Coeficientes de la serie de Fourier . . . . .	32
1.8.4 Espectro lineal . . . . .	36
1.8.5 Periodograma . . . . .	40
1.8.6 Periodograma integrado . . . . .	43
1.9 Funciones muestrales . . . . .	45
1.9.1 Función de media muestral . . . . .	46
1.9.2 Función de varianza muestral . . . . .	46
1.9.3 Funciones muestrales de relación . . . . .	46
1.10 Actividades sugeridas . . . . .	47
2 Modelos estacionarios	49
2.1 Estacionariedad . . . . .	49
2.1.1 Dos conceptos . . . . .	49
2.1.2 Importancia de la estacionariedad . . . . .	53
2.2 Representación matricial . . . . .	55
2.2.1 Matriz de varianzas y autocovarianzas . . . . .	55
2.2.2 Matriz de autocorrelaciones . . . . .	55

## ÍNDICE GENERAL

---

2.3	Ejemplos . . . . .	56
2.3.1	Ejemplo de datos estacionarios . . . . .	56
2.3.2	Ejemplo de datos no estacionarios . . . . .	60
2.4	Modelo lineal general y de forma invertida . . . . .	61
2.5	Función de autocorrelación parcial . . . . .	63
2.6	Modelos autorregresivos . . . . .	65
2.6.1	Modelo AR(1) . . . . .	65
2.6.2	Modelo AR(2) . . . . .	69
2.6.3	Modelo AR( $p$ ) . . . . .	71
2.7	Modelos de medias móviles . . . . .	72
2.7.1	Modelo MA(1) . . . . .	72
2.7.2	Modelo MA(2) . . . . .	74
2.7.3	Modelo MA( $q$ ) . . . . .	76
2.8	Modelos mezclados . . . . .	77
2.8.1	Modelo ARMA(1,1) . . . . .	77
2.8.2	Modelo ARMA( $p,q$ ) . . . . .	78
2.9	Notación abreviada . . . . .	78
2.10	Dualidad de los modelos y resumen . . . . .	79
2.11	Cálculos y gráficos con <i>Statgraphics</i> . . . . .	79
2.12	Actividades sugeridas . . . . .	81
3	Modelos no estacionarios	85
3.1	Introducción . . . . .	85
3.2	Estabilización de la varianza . . . . .	87
3.3	Eliminación de la tendencia . . . . .	88
3.4	Tratamiento de la variación estacional . . . . .	91
3.5	Modelos ARIMA . . . . .	94
3.6	Actividades sugeridas . . . . .	95
4	Pronósticos	97
4.1	Introducción . . . . .	97
4.2	Pronósticos de costo mínimo . . . . .	98
4.3	Cálculo de los pronósticos . . . . .	100
4.4	Límite de los pronósticos . . . . .	101
4.4.1	Pronóstico de un modelo AR(1) . . . . .	102
4.4.2	Pronóstico de un modelo MA(1) . . . . .	103
4.4.3	Pronóstico de un modelo ARMA(1,1) . . . . .	103
4.5	Actualización de los pronósticos . . . . .	104
4.6	Actividades sugeridas . . . . .	104
5	Metodología de Box-Jenkins	105
5.1	Introducción . . . . .	105
5.2	Identificación . . . . .	108
5.2.1	Interpretación de la ACF . . . . .	108
5.2.2	Interpretación de la PACF . . . . .	110
5.2.3	Ejemplos de interpretación de ACF y PACF . . . . .	112

## ÍNDICE GENERAL

---

5.2.4 Identificación con <i>Statgraphics</i> . . . . .	117
5.3 Estimación . . . . .	117
5.3.1 Mínimos cuadrados y máxima verosimilitud . . . . .	119
5.3.2 Procedimiento de estimación . . . . .	120
5.3.3 Inicialización de la serie de tiempo . . . . .	121
5.3.4 Estimación no lineal de los parámetros . . . . .	122
5.3.5 Valor inicial de los parámetros . . . . .	126
5.3.6 Ejemplo: estimación de ARMA(1,1) . . . . .	126
5.4 Diagnóstico . . . . .	128
5.4.1 Análisis de la estacionariedad . . . . .	128
5.4.2 Modelo sobreestimado . . . . .	129
5.4.3 Modelo subestimado . . . . .	132
5.4.4 Datos faltantes y atípicos . . . . .	135
5.4.5 Comparación de modelos . . . . .	136
5.5 Pronóstico . . . . .	137
5.6 Actividades sugeridas . . . . .	139
 6 Modelos con variación estacional	141
6.1 Introducción . . . . .	141
6.2 Modelos estacionales . . . . .	141
6.2.1 Modelos SAR . . . . .	142
6.2.2 Modelos SMA . . . . .	143
6.2.3 Modelos SARIMA . . . . .	146
6.3 Modelos multiplicativos . . . . .	147
6.3.1 Modelo ARIMA(1,0,0)(1,0,0) <sub>s</sub> . . . . .	148
6.3.2 Modelo ARIMA(0,0,1)(0,0,1) <sub>s</sub> . . . . .	149
6.4 Resumen de modelos . . . . .	151
6.5 Actividades sugeridas . . . . .	151
 7 Ejemplo de aplicación	153
7.1 Introducción . . . . .	153
7.2 Las remesas en México . . . . .	153
7.3 Importancia de las remesas . . . . .	154
7.4 Objetivo del estudio . . . . .	155
7.5 Metodología . . . . .	156
7.6 Ajuste del modelo . . . . .	157
7.7 Resultados . . . . .	160
7.8 Pronósticos . . . . .	160
7.9 Discusión . . . . .	161
7.10 Actividades sugeridas . . . . .	163
 Conclusión	165
 Bibliografía	169
 Índice alfabético	175

*PREFACIO*

---

Índice de cuadros	179
Índice de figuras	181

# Prefacio

La vida humana gira en torno a dos polos: los hábitos y las decisiones. Los hábitos representan el mundo del automatismo, la repetición y la rutina, así como los caminos trillados; es la esfera de la inercia, facilidad, seguridad, de las cosas que caen por su propio peso.

Las decisiones son todo lo contrario: el alto en el camino, el lugar de la y griega que hace reconsiderar la ruta, el momento de considerar y ponderar las alternativas, el momento dramático de escoger o desechar perspectivas que también atraen. M. RODRÍGUEZ Y M. MÁRQUEZ

Durante el siglo veinte muchas organizaciones se dieron a la tarea de incorporar las tecnologías de información y comunicación a su trabajo cotidiano. Para ello, se desarrollaron y perfeccionaron gran variedad y número de aplicaciones. Hoy en día, con la aparición y auge tanto de internet como de la *world wide web*, el uso de la computadora y los sistemas automatizados se ha intensificado y ha pasado a formar parte de muchas actividades cotidianas.

Una de las consecuencias más interesantes de este uso es que las empresas e instituciones pueden contar con grandes cantidades de datos, registrados y almacenados con precisión y de manera inmediata. Por ejemplo, el comercio electrónico permite registrar transacciones diversas en tiempo real; los buscadores y servicios de correo electrónico almacenan en forma permanente datos de sus usuarios; los supermercados llevan inventarios que se actualizan con cada compra o recepción de productos; las escuelas cuentan con registros académicos y ambientes virtuales de aprendizaje; en fin, en casi todos los ámbitos es posible recabar, almacenar, organizar y recuperar datos.

Como consecuencia, muchas organizaciones tienen ahora una nueva meta: utilizar esos conjuntos de datos para apoyar y mejorar tanto la investigación como la toma de decisiones. Para ello es factible utilizar modelos matemáticos que reflejen de manera aproximada el comportamiento del evento que se analiza. Un **modelo matemático** es una ecuación o conjunto de ecuaciones que puede usarse con varios fines: describir, explicar, pronosticar o controlar, entre otros.

Una de las actividades más importantes que puede hacerse con un buen conjunto de datos es pronosticar. Algunos de los métodos más usuales para hacer pronósticos, como la regresión lineal o no lineal, requieren de unos

## PREFACIO

---

cuantos datos para generar una curva aproximada a la realidad que, si se extrapolara, puede brindar una predicción más o menos acertada. Sin embargo, al contar con un conjunto más amplio de datos, es posible producir pronósticos que tomen en cuenta otros aspectos importantes como los cambios en las tendencias o variaciones debidas a fenómenos repetitivos como las estaciones del año o las quincenas, de tal modo que los resultados sean aún más precisos y, por lo tanto, más útiles.

Para hacer este tipo de pronósticos resulta indispensable la participación de profesionistas que posean conocimientos de matemáticas y modelos avanzados, de manera que las organizaciones puedan explotar sus datos en la forma óptima.

Los modelos matemáticos pueden ser **deterministas** (cuando los resultados son predecibles con certeza) o **estocásticos** (cuando los resultados no pueden predecirse con certeza, pero sí pueden aproximarse). Por lo general, al tomar decisiones es necesario utilizar modelos estocásticos, ya que se requiere de predicciones sobre un futuro más o menos incierto. En México, donde toda suerte de decisiones políticas (por lo general impredecibles o aún sorpresivas) afectan los procesos económicos y sociales, se hace aún más patente la necesidad de apoyarse en sistemas de pronósticos confiables, precisos y eficientes.

Surge de aquí la oportunidad de conocer los **modelos estocásticos** con fines de predicción o control. Entre ellos destaca la **Metodología de Box-Jenkins** para series de tiempo, es decir, para conjuntos de datos cronológicos. Hace algunos años, esta metodología recibió muchas críticas al ser tachada de “extremadamente difícil y laboriosa”[3, pp. 158-161], por el tipo de cálculos matemáticos que involucra, por lo cual muy pocos tenían acceso a sus beneficios. Hoy en día, se cuenta con excelentes aplicaciones de software, tales como SAS, EViews, Statgraphics, MiniTab, SPSS o R entre otros, que permiten un manejo sencillo, a la vez que generan muy buenos pronósticos. Por otro lado, prácticamente todas las licenciaturas y posgrados del área de matemáticas han incorporado esta metodología a sus planes de estudio.

Por todo lo anterior, este documento pretende ser una guía para el estudiante universitario, al facilitar la comprensión de la teoría y motivar al desarrollo práctico de modelos para fenómenos reales, en un curso introductorio. Se presentan una serie de propuestas didácticas para el proceso enseñanza-aprendizaje de la Metodología de Box-Jenkins que se complementan con los recursos y actividades que se ofrecen en el ambiente virtual de aprendizaje *InteligenciaNet*, ubicado en <http://www.inteligencianet.org>.

El primer Capítulo trata de los conceptos indispensables para la comprensión de los modelos. Contiene algunos aspectos de procesos estocásticos, estadística y series de Fourier. En el Capítulo 2 se explican las propiedades de los modelos para datos estacionarios, es decir, para datos que son relativamente estables en el tiempo; mientras que en el Capítulo 3 se indican las formas de hacer estacionarios los datos cuando no lo son. En el Capítulo 4 se especifica la forma de calcular los pronósticos a partir de un modelo seleccionado. En el Capítulo 5 se detalla la metodología de Box y Jenkins,

para elegir modelos apropiados. En el Capítulo 6 se muestran las propiedades particulares de los modelos para datos con variación estacional. Por último, en el Capítulo 7 se desarrolla una aplicación práctica completa, en la cual se pronostica la tendencia de las remesas enviadas por mexicanos en los Estados Unidos en los últimos años. Cabe señalar también que para el ejemplo de aplicación de este último capítulo se utilizó la estructura conocida como IMRYD: Introducción, Material y métodos, Resultados y Discusión [59].

### Agradecimientos

Muchas gracias por la ayuda de los “*LATEXpertos*”: Mayra Lorena Díaz Sosa, Raúl Uriarte Contreras y Asael Fabian Martínez Martínez, quienes me resolvieron amablemente varias dudas sobre este fascinante software. Gracias también a Jorge Abel Romero Ortiz, quien efectuó con enorme cuidado la primera revisión de estas notas. Un agradecimiento especial para el Mtro. Jorge Luis Suárez Madariaga, el Dr. Carlos Cuevas Covarrubias y la Mtra. María del Carmen Villar Patiño, que amablemente analizaron la primera versión de este documento para su eventual publicación y lo enriquecieron con sus comentarios. Gracias a la Mtra. Nora del Consuelo Goris Mayans, por esmerarse en conservar un ambiente de trabajo propicio para la creación académica.

Muchas gracias al Dr. Jesús H. del Río Martínez, Coordinador de Apoyo a la Investigación de la Universidad Anáhuac México Norte, por su apoyo en mi formación como investigadora y su aportación al último capítulo que escribimos en conjunto, entre muchas otras cosas.

Gracias a Ru y a mis hijos, Tefi y Rubén, por su amor, paciencia y apoyo incondicional.

Gracias, por último, a las más de veinte generaciones de alumnos tanto de Actuaría como de Matemáticas Aplicadas y Computación de la Facultad de Estudios Superiores Acatlán de la Universidad Nacional Autónoma de México, con quienes se fueron construyendo estas ideas a lo largo de muchos años.

Este trabajo se realizó con el patrocinio de la Dirección General de Asuntos del Personal Académico de la UNAM, a través del Proyecto de Investigación PAPIME PE300309, con el título: “InteligenciaNet: Ambiente Virtual y Objetos de Aprendizaje como Apoyo a la Educación Presencial”.

*MariCarmen González Videgaray*

# Capítulo 1

## Conceptos básicos

Es muy difícil hacer predicciones, sobre todo del futuro.  
DECLARACIÓN TOMADA DE UNA CONFERENCIA SOBRE EL MAÑANA.

### 1.1 Pronósticos y decisiones

Una gran cantidad de decisiones, desde las decisiones más sencillas que se toman en la vida diaria, hasta aquéllas de las cuáles depende el desarrollo de una comunidad, involucran un pronóstico. Por ejemplo, para decidir cuál es la hora óptima para salir de casa hacia la escuela o el trabajo, es necesario estimar el tiempo que tomará llegar al lugar correspondiente, así como tener una idea clara con respecto al tráfico en la ruta seleccionada. Asimismo, para determinar si es conveniente aplicar alguna medida que reduzca la inflación o aumente el empleo, es necesario predecir que ocurrirá en diversos escenarios.

Las decisiones tendrán más o menos importancia, en función de sus consecuencias. Desde decisiones personales sencillas, hasta otras que pueden afectar a sociedades o países enteros, pasando por aquellas como elegir una carrera, casarse o no, aceptar o no un trabajo, por ejemplo. Si las consecuencias son de mayor trascendencia, será más importante cuidar la decisión y, sobre todo, acertar en ella.

Para evitar decisiones erróneas o aún catastróficas, deben preverse con la mayor exactitud posible sus consecuencias, esto es, debe hacerse un pronóstico que sea suficientemente parecido al fenómeno real sobre el cual se decide y que permita una visión confiable del proceso bajo estudio. De aquí la importancia de elegir un buen método de pronóstico: *un buen pronóstico prácticamente garantiza una buena decisión.*

Como aquel personaje de Moliére que se sorprendía al percatarse de que siempre hablaba en prosa, tal vez muchas personas se sorprenderían al darse cuenta de la gran cantidad de pronósticos o predicciones que realizan diariamente. Claro está que la mayoría lo hacen de manera intuitiva y sin una aproximación sistemática, sin embargo, resultan bastante satisfactorios. Esto

## 1.2 Una aproximación al problema

---

funciona cuando la información cualitativa es suficiente y los resultados de la decisión no implican grandes costos en caso de error. En cambio, cuando se trata de decisiones que afectarán a muchas personas, cuyos resultados son de gran importancia y si se cuenta con información relativamente amplia y cuantitativa, es más conveniente seleccionar un método también cuantitativo que permita medir, inclusive, la confianza que puede tenerse en el pronóstico. En general, se recomienda combinar un método cuantitativo con uno cualitativo, con objeto de tener una visión panorámica de lo que podría suceder.

El pronóstico cuantitativo no debe verse como una predicción lanzada por un adivino, que debiera ser idéntica al comportamiento futuro del evento analizado. Esto es prácticamente imposible: de hecho, el pronóstico cuantitativo generalmente se da en forma de un intervalo de confianza, es decir, de un rango de valores posibles. Por otro lado, aún si se elige la técnica matemática más sofisticada, es esencial contar con un conocimiento profundo del fenómeno en cuestión, ya que sólo así podrá saberse si el método es adecuado, si el patrón o tendencia es consistente, si el tipo de modelo es apropiado, etcétera. El sentido común, la inteligencia, la experiencia y la información completa deberán anteponerse siempre a los resultados matemáticos.

A lo largo de este texto se considerará como **pronóstico** o **predicción** al valor estimado de un evento que es desconocido por alguna de las siguientes circunstancias:

- No ha transcurrido el tiempo necesario para que el evento suceda, es decir, el evento de interés será conocido en un tiempo futuro.
- No se han dado las circunstancias necesarias para que el evento ocurra, esto es, el evento de interés depende de otros eventos que no han sucedido.

## 1.2 Una aproximación al problema

Existen muchos y muy diversos métodos de pronóstico, tanto cualitativos como cuantitativos. Dentro de los métodos más generales de pronóstico cuantitativo que se basan en datos históricos, existen algunos que pueden manejar casi cualquier tipo de datos numéricos, a través del uso de modelos matemáticos de tipo estocástico. Aunque estos métodos podrían parecer excesivamente laboriosos, existen ya una gran cantidad de paquetes electrónicos para aplicarlos, de manera que resultan sencillos, rápidos y económicos.

Algunas aplicaciones de software útiles para efectuar pronósticos son las siguientes:

- *Statgraphics Centurion* [51].
- *Statgraphics Online* [50], versión libre.
- SAS ETS [45].
- SPSS [49].

- *EViews* [42].
- R (software libre) [44].

Para entender estos métodos puede partirse de la idea de que un conjunto de datos históricos está formado por una serie de **choques aleatorios** (*random shocks*), acumulados y ponderados en el tiempo; de manera tal que al alejarse hacia atrás en el tiempo van perdiendo su influencia en el resultado final. Otra forma de explicar un proceso en el tiempo es considerar que está formado por una ponderación de los valores históricos más algún error o fluctuación aleatoria.

A los modelos anteriores se les llama de manera conjunta **modelos autorregresivos y de medias móviles**<sup>1</sup>.

Así pues, es posible construir modelos de tipo autorregresivo y de medias móviles. Estos modelos se denotan con el nombre abreviado **ARMA** (*Autoregressive Moving Averages*) y se describirán en el Capítulo 2. El proceso por el cual se determina el modelo adecuado no es inmediato y requiere por lo general de varios intentos sucesivos antes de obtener un buen modelo. Esto se tratará en el Capítulo 5.

## 1.3 Serie de tiempo

### 1.3.1 Definición

Uno de los conceptos fundamentales dentro de los pronósticos cuantitativos es el de **serie de tiempo**. Box y Jenkins [12, p. 23] señalan que una serie de tiempo es “un conjunto de observaciones generadas de manera secuencial en el tiempo. Si el conjunto es continuo, se dice que la serie de tiempo es *continua*. Si es discreto, se le llama *discreta*”. Por su parte, tanto Wei [58, p. 1] como Bovas y Ledolter [10, p. 3] las definen como “una secuencia ordenada de observaciones” que también pueden ser continuas o discretas y que, además, pueden tomarse de diversas dimensiones distintas del tiempo, como el espacio por ejemplo. De acuerdo con Brockwell y Davis [15, pp. 1, 35-26], se trata de “un conjunto de observaciones  $x_t$ , cada una registrada en un tiempo específico  $t$ ”, pero también se usa el mismo término para denotar al proceso estocástico o familia de variables aleatorias  $\{X_t, t \in T\}$ . Percival y Walden [39, p. 1] señalan que una serie de tiempo es cualquier secuencia de observaciones asociadas con una variable independiente  $t$ , que puede ser discreta o continua.

Así pues, en este texto se considerará a la **serie de tiempo** como una colección de observaciones cronológicas, es decir, generadas en forma secuencial a través del tiempo. Los datos recabados se ordenan con respecto al tiempo

<sup>1</sup>El nombre fue acuñado en 1976 por Box y Jenkins [12], sin embargo, resulta poco apropiado porque en realidad no se están calculando promedios y, además, existe otra metodología de pronóstico con el mismo nombre de medias móviles y otras bases teóricas. Sin embargo, el nombre ha permanecido así.

### 1.3 Serie de tiempo

---

y las observaciones sucesivas suelen ser dependientes entre sí. Así, la serie de tiempo será un proceso estocástico con espacio de estados  $S$  continuo o discreto y espacio paramétrico  $T$  que también puede ser continuo o discreto. En este trabajo sólo se analizará el segundo caso. De aquí en adelante, la serie de tiempo se denotará como  $\{Y_t\}$ .

Algunos ejemplos de series de tiempo pueden ser:

- La cantidad de personas que usan diariamente el sistema de transporte colectivo metropolitano,
- El número de casos registrados semanalmente de influenza AH1N1,
- El producto nacional bruto (PNB) mensual de México,
- Los ingresos mensuales en México por turismo,
- La temperatura promedio diaria,
- El índice semanal de contaminación por ozono,
- El índice mensual de inflación.

El valor particular  $Y_t$  representa la  $t$ -ésima observación, es decir, el valor que tendrá la serie de tiempo en el momento  $t$ . Dicho valor puede ser conocido para  $t = 1, 2, \dots, N$ , en cuyo caso es una constante; o desconocido, en cuyo caso es una **variable aleatoria**. Cabe destacar que se utiliza la misma notación en ambos casos y se hace la diferencia a través de la posición en el tiempo, es decir, del valor de  $t$ , como se ve en la Figura 1.1, que muestra los diversos valores de la serie.

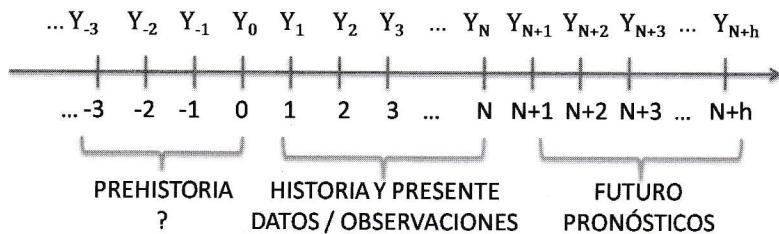


Figura 1.1: La serie de tiempo.

Una serie de tiempo observada, es decir, el conjunto de sus valores o datos conocidos, se acostumbra llamar una **realización** del proceso estocástico. Esto es, la realización es *una* de todo el conjunto de posibles secuencias o resultados del proceso. Se le llama de esta forma porque, si el fenómeno pudiera ocurrir de nuevo (lo cual en la mayoría de los casos reales es imposible), se supone que se obtendrían resultados diferentes, es decir, una realización distinta. Este concepto es fundamental y más adelante se relacionará con el de estacionaridad.

El objeto del **análisis de series de tiempo** es entonces describir el proceso teórico que subyace a la serie de tiempo, en forma de un modelo matemático que tenga propiedades similares al proceso real. Esto permitirá entender mejor dichas propiedades y, sobre todo, elaborar pronósticos con alto grado de precisión y confiabilidad.

Los datos de series de tiempo se obtienen de registros instantáneos; promediados (como temperaturas, índice *Dow Jones*, índice de precios y cotizaciones de la Bolsa Mexicana de Valores); obtenidos de una muestra; o acumulados (producción total de una fábrica en un mes, precipitación pluvial acumulada en un día, accidentes totales de una semana, por ejemplo).

Si bien el parámetro  $t$  se manejará en este trabajo como discreto, el valor de la serie de tiempo  $Y_t$  puede ser discreto o continuo, lo cual dependerá del fenómeno en cuestión. Por ejemplo, las temperaturas o los índices de la Bolsa Mexicana de Valores pueden tomar valores continuos, mientras que el número de accidentes en una carretera será siempre discreto, a menos que se trate de un promedio.

En general, en este texto se considerará que se cuenta con  $N$  datos históricos en una realización, es decir, con las **observaciones**:

$$Y_1, Y_2, \dots, Y_N$$

A los datos  $Y_1, Y_2, \dots, Y_{N-1}$  se les llamará **historia** o datos históricos, mientras que al dato específico  $Y_N$  se le llamará **presente** o dato actual, aún cuando en la mayoría de las ocasiones se refiera a un momento del tiempo que ya transcurrió. Es decir, se trata simplemente del dato más reciente. Si se continúa esta metáfora, los valores anteriores a los datos observados pueden considerarse como la **prehistoria**, mientras que más adelante de  $Y_N$  estará el **futuro**.

Por último, es importante señalar que a veces es necesario corregir o modificar los datos originales, de manera que se facilite su manejo. Por ejemplo:

- Para tener meses de longitud idéntica, pueden redondearse a 30 días. Esto suele hacerse en las áreas de contabilidad o finanzas. Por ejemplo: Días de febrero  $\times 30/28$ .
- Para homogeneizar el valor del dinero, que cambia debido al tiempo y a la inflación, las cantidades pueden convertirse a dólares o deflactarse.

### 1.3.2 Componentes de una serie de tiempo

La serie de tiempo es, por lo general, una mezcla de varios factores o **componentes**. Es importante aprender a distinguir —e inclusive aislar— estos componentes, para identificar su presencia o ausencia en una serie. Esto permitirá comprender mejor sus características y facilitará la selección del modelo adecuado para pronosticar.

Una serie de tiempo estará siempre formada por uno o más de los siguientes componentes:

### 1.3 Serie de tiempo

---

1. **Tendencia.** La tendencia es un cambio sistemático en el patrón de los datos y suele detectarse a simple vista en la gráfica de la serie original. Una serie sin tendencia será semejante o fluctuará alrededor de una recta horizontal, paralela al eje del tiempo. En cambio, una serie con tendencia lineal, será creciente o decreciente en forma aproximada a una línea recta con pendiente positiva o negativa, según el caso. Una serie con patrón cuadrático, en cambio, tendrá semejanza con una parábola. La tendencia se identifica en inglés como *trend*. La Figura 1.2 (Fuente: Elaboración propia con datos de [47]) muestra una serie que tiende a incrementarse. Los valores corresponden a la tasa de desocupación abierta mensual en México. Esto significa que con el paso de los años se observa mayor desempleo en el país. En principio, la tendencia podría ser lineal o cuadrática, sería necesario revisar ambas para identificar cuál se ajusta mejor a los datos, como se verá más adelante.

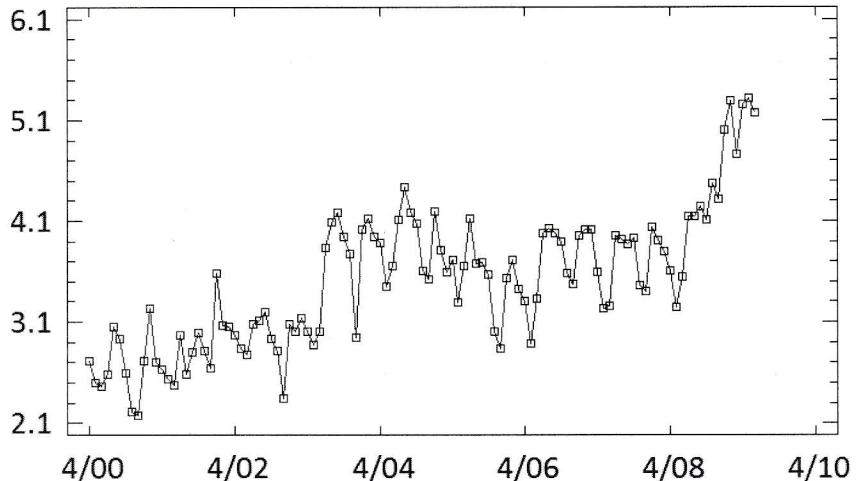


Figura 1.2: Tasa de desocupación abierta mensual en México.

Por supuesto, la tendencia es un componente sustancial que deberá considerarse al pronosticar. Para incluir la tendencia en los modelos de pronóstico existen dos opciones: eliminarla o modelarla. La eliminación puede hacerse —como se verá más adelante— a través de aplicar una transformación (logaritmo, recíproco o raíz cuadrada, por ejemplo), o diferencias sucesivas conocidas como finitas u ordinarias. Algunas aplicaciones de software las identifican como diferencias no estacionales (*nonseasonal differences*). Si bien la tendencia puede modelarse a través de regresión, en los modelos a que se aboca este trabajo se buscará siempre eliminar la tendencia.

**2. Variación estacional.** Es muy común que algunos fenómenos estén vinculados a las estaciones del año, el clima, variaciones económicas repetitivas como las quincenas, entre otras cosas. Cuando se observa un patrón de comportamiento que se repite en períodos menores o iguales a un año, se dice que la serie de tiempo tiene variación estacional (*seasonality* en inglés). Es muy importante considerar esta variación dentro del pronóstico. Para ello puede intentarse su eliminación a través de diferencias estacionales (*seasonal differences*), o puede modelarse como se verá en el Capítulo 6. La Figura 1.3 muestra el turismo receptivo en México (Fuente: Elaboración propia con datos de [48]) es decir, el número de personas que llegan como turistas a algún lugar del país. La tendencia parece creciente hasta el año 2008 y posteriormente parece descender. Sin embargo, es evidente que existe un patrón repetitivo dentro de cada año (marcado con líneas punteadas), es decir, una variación estacional, la cual lógicamente puede atribuirse a las temporadas vacacionales en las cuales se incrementa sistemáticamente el valor de la serie.

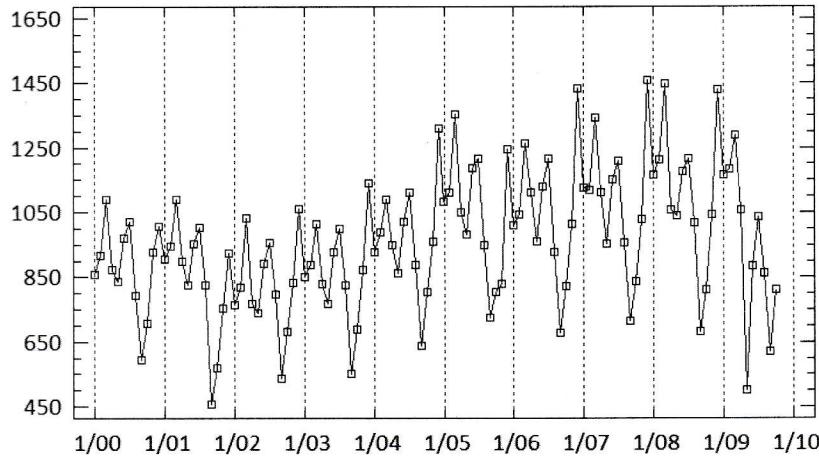


Figura 1.3: Turismo receptivo mensual en México.

**3. Ciclo.** Se trata también de un patrón repetitivo, pero con longitud mayor a un año. Es más difícil de observar puesto que requiere de dar seguimiento a un fenómeno durante mucho tiempo para ser advertido. *E.g.*, para detectar un ciclo que ocurre cada seis años, sería necesario contar al menos con 18 años de observaciones. Un ciclo económico podría deberse, por ejemplo, a los sexenios presidenciales u otro tipo de elecciones dentro de nuestro país; algunas enfermedades tienden a

diseminarse en mayor medida cada cierto número de años, como el sarampión. En las manchas solares se observan ciclos que afectan al clima. En la Figura 1.4 (Fuente: [40]) se observan varios ciclos ocurridos en un lapso de once años, relacionados con las manchas y explosiones solares. El ciclo debe ser considerado en el modelo, si es que se percibe. Por lo general, si no hay datos sistemáticos de periodos largos de tiempo, es imposible notar su ocurrencia. En inglés se conoce como *cycle*.

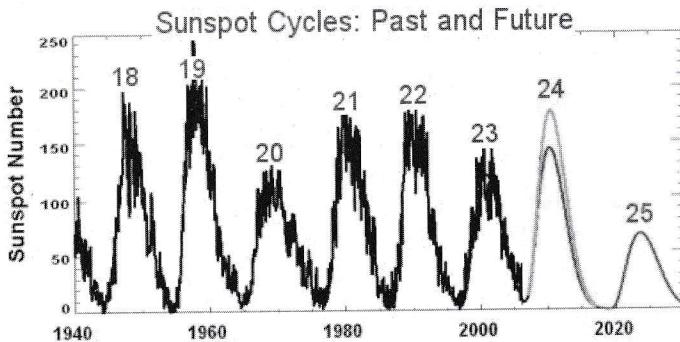


Figura 1.4: Ciclos solares desde 1940.

4. **Fluctuación aleatoria.** Cualquier realización de una serie de tiempo, aún cuando pueda ser parecida a las gráficas que se obtienen a partir de los modelos matemáticos (rectas, paráolas, función logaritmo, etcétera), siempre contendrá un conjunto de variaciones irregulares que representan el resultado de los llamados choques o golpes aleatorios (*random shocks*). Es decir, estas fluctuaciones aleatorias agrupan todo aquello que el modelo matemático no puede explicar. Por ello, para que un modelo sea adecuado, las fluctuaciones aleatorias deben ser prácticamente despreciables, dicho de modo más formal: deben ser estadísticamente insignificantes.

Es decir, si se ha seleccionado un buen modelo, la fluctuación aleatoria ideal debería ser lo que se suele llamar **ruido blanco** (*white noise*). El ruido blanco se define como un conjunto de variables aleatorias independientes, idénticamente distribuidas como Normal con media cero y varianza constante  $\sigma^2$ . En la Figura 1.5 se muestra una gráfica de cien valores de ruido blanco simulados con Excel, con  $\sigma^2 = 1$ .

Esta gráfica no muestra tendencia: los datos están todos alrededor de una media constante cero, la varianza es constante (*i. e.* la serie es homoscedástica), no hay variación estacional ni ciclo. En particular, dado que se trata de una Normal con parámetros  $(0,1)$ , puede verse que la mayoría de los datos están entre la media (cero) más/menos

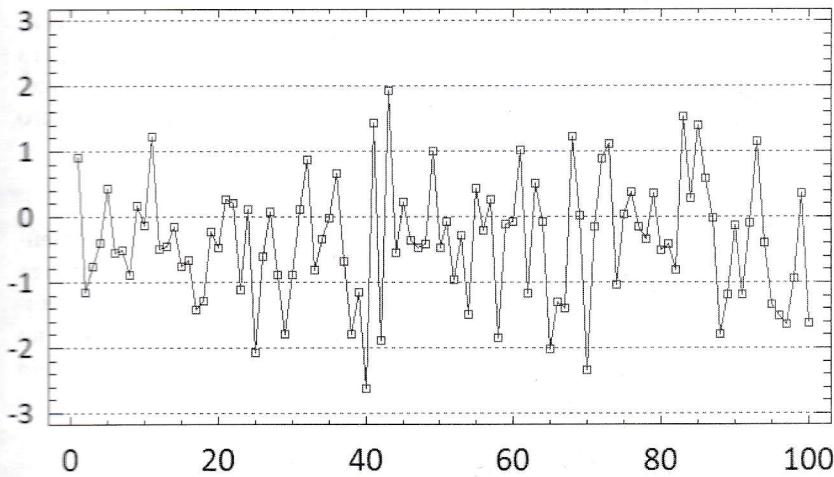


Figura 1.5: Ruido blanco.

dos desviaciones estándar (entre -2 y 2). Éste es el tipo de gráfica que se esperaría encontrar en los residuales que genere un buen modelo de series de tiempo. Además, existen pruebas de hipótesis para verificar que se cumplan las propiedades del ruido blanco.

## 1.4 Objetivos del análisis

El análisis de series de tiempo puede tener uno o más de los siguientes **cuatro** objetivos:

- **Descripción.** Pretende simplemente enunciar el comportamiento de un proceso, permitiendo una visualización clara que generalmente se obtiene de la gráfica de la serie de tiempo. Por ejemplo, se puede observar el comportamiento de las temperaturas diarias para detectar posibles tendencias y/o variaciones estacionales. Responde a la pregunta: ¿Cómo se comporta el fenómeno?
- **Explicación.** Pretende realizar un análisis de tipo causa-efecto al relacionar un fenómeno con el medio circundante o con otras variables. Por ejemplo, puede explicarse la contaminación de acuerdo con el día de la semana, la cantidad de autos en circulación, etcétera. Responde a las preguntas: ¿Por qué se comporta así el fenómeno?, ¿Qué otras variables afectan a la variable de interés?

- **Pronóstico.** Pretende estimar valores futuros de un fenómeno con el fin de tomar decisiones o prever comportamientos. Por ejemplo, es conveniente estimar la cantidad de turismo que llegará a una zona para preparar las condiciones necesarias, sin incurrir en gastos excesivos. Responde a la pregunta: ¿Qué ocurrirá con el fenómeno en un futuro próximo?
- **Control.** Si no sólo se desea observar, sino modificar el comportamiento de un proceso para obtener algún beneficio o para que se mantenga dentro de ciertos límites establecidos, el objetivo es controlar. Por ejemplo, podrían definirse las acciones necesarias para que la inflación no exceda de cierto límite. Responde a la pregunta: ¿Cómo mantener el fenómeno dentro de ciertos límites?

## 1.5 Modelos univariados y multivariados

Los modelos clásicos de regresión relacionan una variable dependiente con una o más variables independientes. La idea general de estos métodos es explicar, pronosticar o ambos, el valor de la variable dependiente a partir de alguna combinación lineal o no lineal de los valores de las variables independientes.

En el análisis de series de tiempo existen tres posibilidades:

1. Utilizar como variables independientes a los valores anteriores o pasados de la misma serie de tiempo que se desea pronosticar. En este caso se hablará de **modelos univariados** de series de tiempo.
2. Utilizar otras series de tiempo independientes como variables explicativas.
3. Mezclar los dos anteriores. En este caso y el anterior se trata de **modelos multivariados** de series de tiempo.

En los dos últimos, pueden manejarse conjuntos de series de tiempo para formar modelos causales, en los cuales deberá encontrarse una **función de transferencia** que relacione las distintas series de tiempo. Este análisis está fuera de los objetivos de este trabajo, pero pueden consultarse las referencias sugeridas en la bibliografía, en particular [58].

Básicamente, el procedimiento que usan los modelos multivariados consiste en formular modelos de tipo regresión lineal o no lineal múltiple, donde, en lugar de colocar una variable como función de otras, se tiene una serie de tiempo en función de otras series de tiempo.

Puesto que en este caso resulta inadecuado hablar de variables dependientes e independientes, ya que la variable dependiente e independiente pueden ser una sola en distintos momentos del tiempo, es mejor utilizar como denominación:

- **Variable exógena:** aquella cuyo valor no depende del resultado del modelo. (Hace la función de variable independiente).
- **Variable endógena:** aquella cuyo valor depende del resultado del modelo. (Hace la función de variable dependiente).

Los modelos que se verán adelante colocan a la variable endógena  $Y_t$  en función de su propia historia, que actuará como un conjunto de variables exógenas.

## 1.6 Metodología de Box-Jenkins

En 1976, George E. P. Box y Gwilym M. Jenkins publicaron su libro *Time Series Analysis. Forecasting and Control* [12], en el cual mencionan cuatro aplicaciones prácticas del pronóstico de series de tiempo: planeación económica y financiera, planeación de la producción, control de inventarios y producción, así como control y optimización de procesos industriales. A partir de la aparición de este libro, muchos autores han seguido los lineamientos de Box y Jenkins, al crear más aplicaciones de estos modelos y facilitar el trabajo por medios electrónicos.

La metodología de Box-Jenkins consiste en extraer los movimientos predecibles de los datos observados y separarlos de la parte no predecible o completamente aleatoria. La serie de tiempo se desagrega en varios componentes, algunas veces llamados **filtros**, precisamente porque la filosofía de este método consiste en detectar los distintos componentes y separarlos usando los “filtros” correspondientes, hasta obtener residuales no predecibles cuyo comportamiento tiene poca influencia en el resultado final y es semejante al ruido blanco. El enfoque de Box-Jenkins hace principalmente uso de tres filtros lineales: el autorregresivo, el de integración y el de medias móviles, como se muestra en la Figura 1.6.

El proceso iterativo de Box-Jenkins para construir modelos lineales de series de tiempo se basa en dos grandes principios:

- **El principio de parsimonia.** Consiste en elegir siempre el modelo más sencillo que sea suficientemente representativo de los datos. Se basa en la filosofía propuesta por William de Occam, conocida como *Occam's razor: All things being equal, the simplest solution tends to be the best one* (De entre cosas iguales, la solución más sencilla tiende a ser la mejor). Dicho de otra manera, no se deben agregar al modelo elementos que no estén plenamente justificados.
- **El principio del mejoramiento iterativo.** Consiste en partir de un modelo sencillo y factible, al cual se harán mejoramientos sucesivos, hasta llegar a un modelo satisfactorio.

De acuerdo con estos dos principios, no es necesario que se identifique un modelo adecuado desde el primer intento, ya que la metodología ofrece

## 1.6 Metodología de Box-Jenkins

---

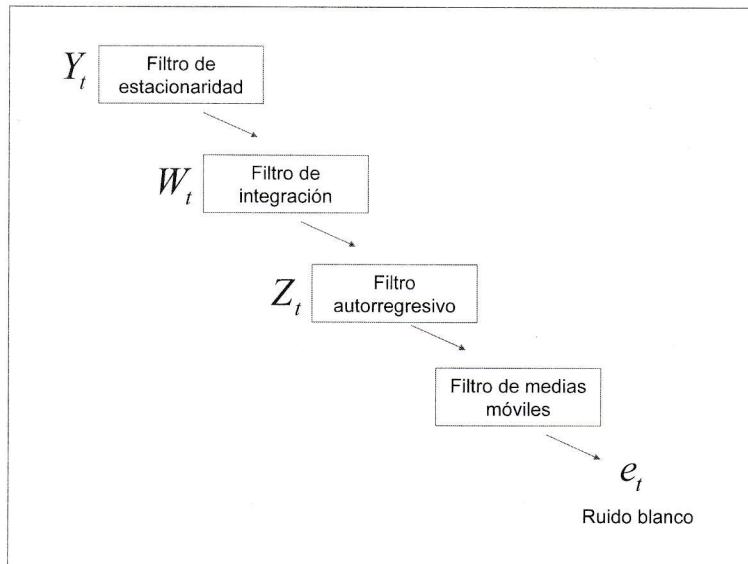


Figura 1.6: Enfoque de Box-Jenkins.

recursos para evaluar y mejorar los diversos modelos tentativos, de manera sucesiva. Además, al tomar en cuenta el principio de parsimonia, deberá iniciarse siempre el proceso con un modelo sencillo, que poco a poco se irá mejorando, a través de las estadísticas que permiten determinar qué es lo que falta —o sobra— en un modelo.

En la Figura 1.7 se presenta un diagrama funcional del método de Box-Jenkins [56, Pág. 63, modificado], que contiene los pasos principales que deberán seguirse para llegar a un modelo adecuado.

Los procedimientos empleados en la identificación son inexactos y requieren de mucho razonamiento y experiencia por parte del analista, así como del conocimiento del fenómeno en cuestión. La metodología parte de elegir un modelo sencillo más o menos adecuado, y su objetivo es tratar de mejorarlo. Por supuesto, no se podrá hablar de un modelo “óptimo”, pero sí de un modelo que genere pronósticos adecuados. Cabe señalar aquí que algunos paquetes estadísticos como SAS y *Statgraphics* ofrecen la posibilidad de generar automáticamente un modelo que pretende ser el mejor, en el sentido de arrojar el menor valor para algún estadístico particular, pero la decisión deberá sujetarse siempre al criterio del analista, ya que puede haber aspectos cualitativos que el software no considere.

La metodología brinda un conjunto de estadísticos y gráficos que serán de gran utilidad para la identificación del modelo y para mejorarlo. Para verificar si un modelo es adecuado, será muy importante también el estudio de los residuales, como se verá más adelante.

Es indispensable hacer notar aquí que en ocasiones hay modelos que se

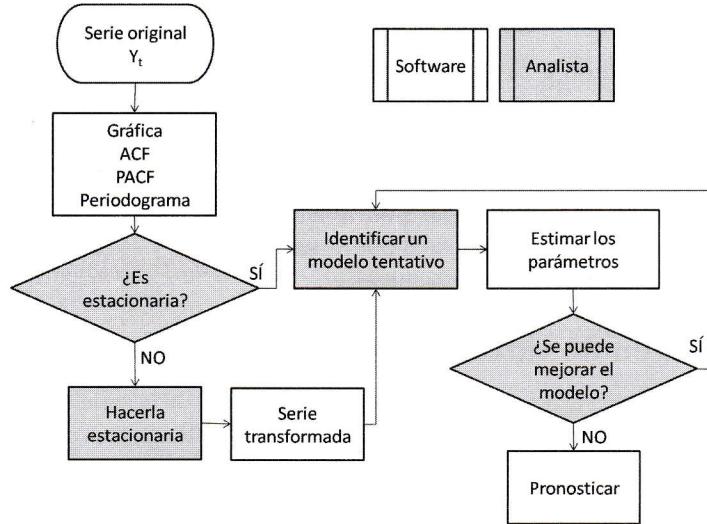


Figura 1.7: Metodología de Box-Jenkins.

ajustan a los datos históricos, pero que generan pronósticos deficientes o inaceptables. En este caso habrá que regresar a la identificación o, inclusive, revisar si el método es realmente adecuado. En general este método requiere de un conjunto amplio de datos, al menos unos sesenta. Por otro lado, de acuerdo con el objetivo final, puede sacrificarse el ajuste por la precisión en los pronósticos, en virtud de que esta última es la meta de interés.

## 1.7 Funciones de primer y segundo orden

De la misma forma en que una variable aleatoria se caracteriza por sus parámetros y entre ellos destacan el de primer orden (la media) y el de segundo orden (la varianza), un proceso estocástico en general y en particular una serie de tiempo, pueden caracterizarse por sus funciones de primer y segundo orden.

La función de primer orden es la función de media. Se habla en este caso de una función porque ya no se trata de un solo punto —como en la variable aleatoria— sino de un conjunto de puntos que irá tomando la media a lo largo del tiempo. Lo mismo ocurre para las funciones de segundo orden: varianza, autocovarianza, autocorrelación y autocorrelación parcial, todas ellas tomarán distintos valores al transcurrir el tiempo. A continuación se definen de manera sucinta.

### 1.7.1 Función de media

La **función de media** es el valor esperado de la serie de tiempo en cada valor del índice tiempo, es decir:

$$\mu_t = E(Y_t) \quad (1.1)$$

Obsérvese que la función anterior podría entonces tomar diversos valores a lo largo del tiempo. En algunos casos podrá ser constante y se dirá entonces que la serie de tiempo no muestra tendencia.

### 1.7.2 Función de varianza

La **función de varianza** mide la dispersión esperada de la serie de tiempo en cada valor del índice tiempo, es decir:

$$\text{Var}(Y_t) = E(Y_t - \mu_t)^2 \quad (1.2)$$

Obsérvese que la función anterior podría tomar diversos valores a lo largo del tiempo. En algunos casos podrá ser constante y se dirá que la serie tiene varianza constante o es **homoscedástica**. Si la varianza cambia a través del tiempo (generalmente es creciente) se dice que la serie es **heteroscedástica**. Una serie así suele tener puntos que se distancian cada vez más de la función de media.

### 1.7.3 Funciones de relación

La función de autocovarianza y la función de autocorrelación miden la forma en que, al variar la serie en un punto, se afecta su valor en otro punto del tiempo.

Así, la **función de autocovarianza** se calcula como:

$$\text{Autocov}(Y_t, Y_{t-k}) = E[(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})] \quad (1.3)$$

Por otra parte, la **función de autocorrelación** se calcula como:

$$\text{Autocorr}(Y_t, Y_{t-k}) = \frac{\text{Autocov}(Y_t, Y_{t-k})}{\sqrt{\text{Var}(Y_t)\text{Var}(Y_{t-k})}} \quad (1.4)$$

Como puede verse, ambas funciones dependen tanto de la posición en el tiempo  $t$ , como de la magnitud del intervalo que separa a cada pareja de variables, es decir,  $k$ . Como la función de autocorrelación se utilizará muy a menudo a lo largo de este texto, su nombre será sustituido por las siglas **ACF**, provenientes de su equivalente en inglés: *AutoCorrelation Function*.

Por su parte, la **función de autocorrelación parcial** mide la contribución al modelo de cada nueva variable endógena  $Y_{t-k}$  (variable histórica o autorregresiva) que se agrega al modelo de pronóstico. Esta función se abreviará como **PACF** (por su equivalencia en inglés: *Partial AutoCorrelation Function*). En el Capítulo 2 se explicará la forma de obtenerla.

## 1.8 Periodicidad y su detección

Como se ha dicho, una serie de tiempo puede contener un componente periódico, ya sea ciclo o variación estacional. Existen varias formas de detectar este componente. Puede ser visible desde la gráfica original de los datos, como en la Figura 1.3. También se puede observar, como se verá más adelante, en la ACF, PACF o ambas. Sin embargo, también es posible construir un estadístico para verificar su relevancia dentro de los datos. Para ello se utilizará la serie de Fourier, de la forma que se explica en seguida.

### 1.8.1 Propiedades de la función seno

En la gráfica de la función seno más sencilla,  $f(x) = \sin(x)$ , para  $x \in [0, 2\pi]$  radianes, es posible detectar con facilidad cuatro propiedades esenciales:

- **Amplitud o fuerza:** es la distancia máxima de la curva al eje horizontal. También se le llama fuerza de la función. Se denota por  $A$ .
- **Longitud de periodo:** distancia de cresta a cresta o de valle a valle. Es el lapso que tarda la curva en repetirse. Se denota por  $L$ .
- **Frecuencia:** recíproco de la longitud. Mide la proporción del periodo que se recorre en una unidad. Por supuesto, a mayor longitud menor frecuencia y viceversa. Se denota por  $f$ .
- **Fase:** es el desplazamiento horizontal de la curva, con respecto al eje vertical. Se denota por  $\phi$ .

En la Figura 1.8 se han marcado la amplitud, longitud y fase. Es sencillo verificar en la gráfica que en este caso la amplitud es 1, la longitud es  $2\pi$  (aproximadamente 6.2832), la frecuencia es  $1/2\pi$  y la fase es 0.

### 1.8.2 Función seno generalizada

¿Cómo variar las propiedades de la función seno para hacerla más general? Por ejemplo, si se desea obtener una función cuya amplitud tenga un valor de  $A$ , la ecuación correspondiente sería:

$$f(x) = A \sin(x)$$

Así, en la Figura 1.9 se observa la función seno anterior junto con una nueva función seno cuya amplitud es de  $A = 2$ . Como puede verse, la amplitud se incrementó.

Modificar la fase es muy sencillo, basta con sumar (o restar) una constante al argumento de la función seno, de la siguiente forma:

$$f(x) = A \sin(x + \phi)$$

### 1.8 Periodicidad y su detección

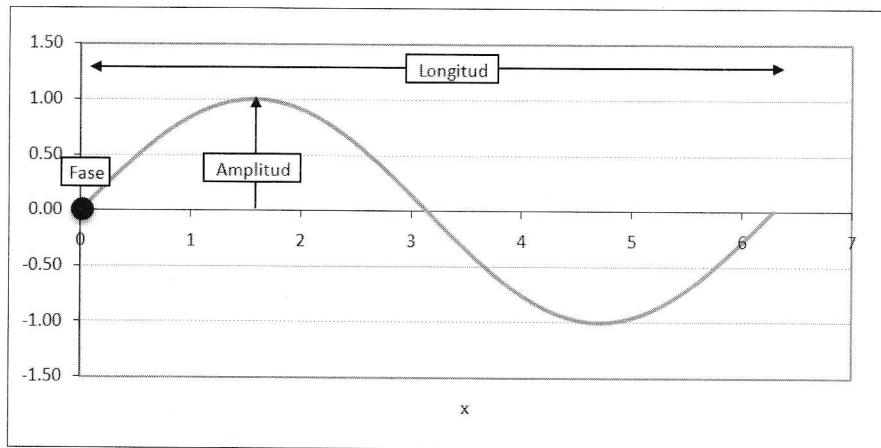


Figura 1.8: Función seno.

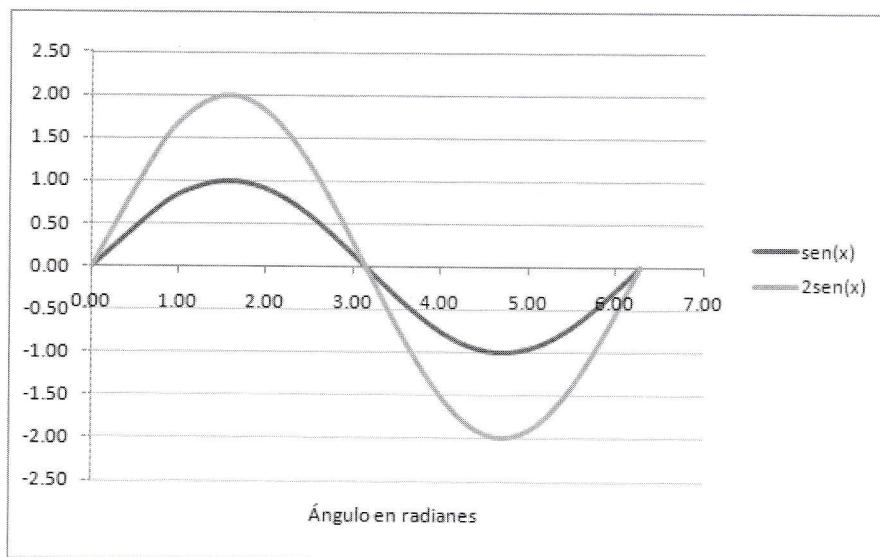


Figura 1.9: Funciones seno con diferentes amplitudes.

Por ejemplo, en la Figura 1.10 se observa la función seno original junto con la nueva función  $f(x) = \operatorname{sen}(x - 1)$ ; al modificar la fase, la función seno sufre un desplazamiento horizontal.

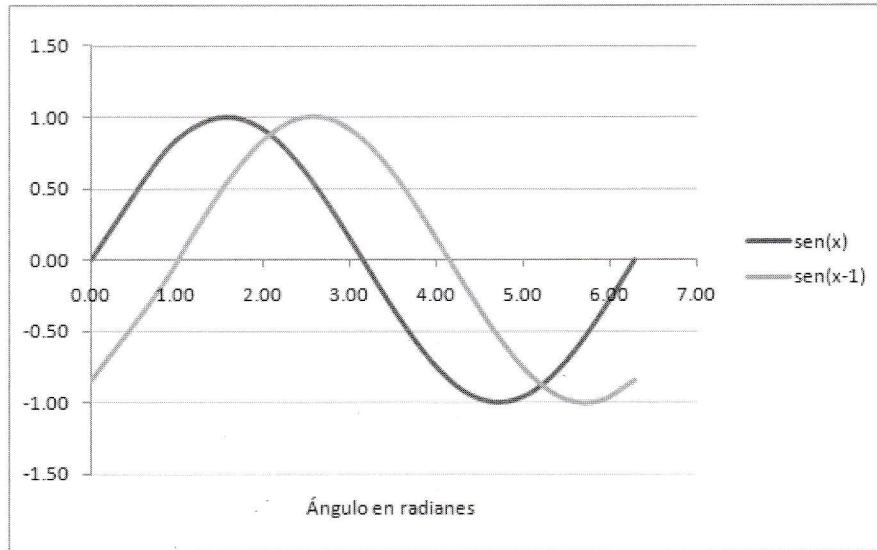


Figura 1.10: Funciones seno con diferentes fases.

Ahora, para manipular la longitud —y por lo tanto la frecuencia— de la función, puede considerarse que la función seno pasa por un conjunto de  $N$  puntos discretos equidistantes, de tal manera que, si la longitud es  $2\pi$  como en el caso inicial ( $f(x) = \operatorname{sen}(x)$ ) y corresponde a los  $N$  puntos, la frecuencia en cada uno de los puntos sería de  $2\pi/N$ .

Así, la función seno con longitud  $L = N$  puntos, frecuencia  $f = 1/N$ , amplitud  $A$  y fase  $\phi$ , se expresaría como:

$$f(t) = A \operatorname{sen}\left(\frac{2\pi t}{N} + \phi\right)$$

Para  $t = 1, 2, \dots, N$ .

En la Figura 1.11 se muestran dos funciones seno para longitud  $L = 32$  puntos, frecuencia  $f = 1/32$  y amplitud  $A = 3$ . La primera con fase  $\phi = 0$  y la segunda con  $\phi = 2$ .

En cambio, en la Figura 1.12 se muestran otras dos funciones seno para longitud  $L = 16$  puntos, frecuencia  $f = 2/32$  y amplitud  $A = 3$ . La primera con fase  $\phi = 0$  y la segunda con  $\phi = 2$ .

Por último, también es posible hacer una translación de la función seno con respecto al eje horizontal si se agrega una constante  $A_0$ . Esta constante occasionará un desplazamiento rígido de la función hacia arriba si  $A_0 > 0$  o hacia abajo si  $A_0 < 0$ .

## 1.8 Periodicidad y su detección

---

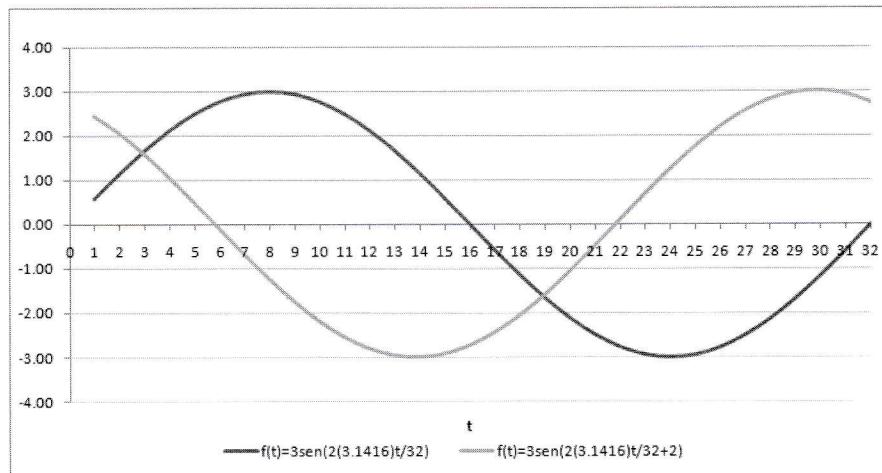


Figura 1.11: Funciones seno con longitud de 32 puntos.

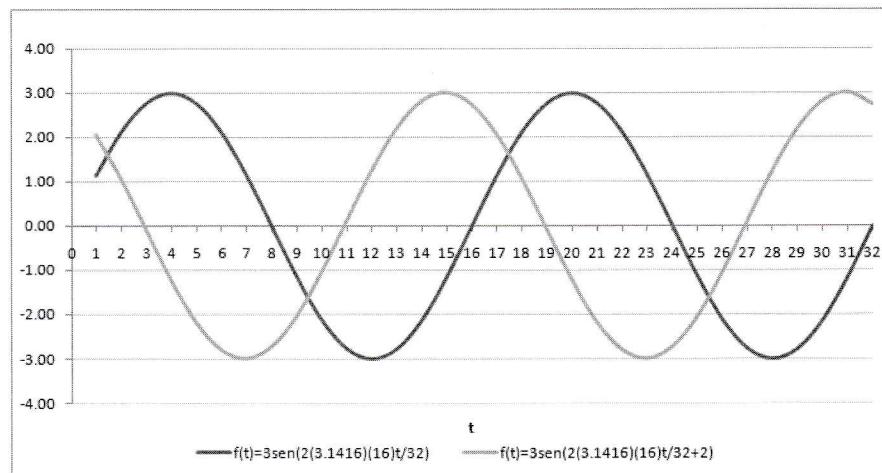


Figura 1.12: Funciones seno con longitud de 16 puntos.

Ahora, supóngase que se tiene un conjunto de  $N$  puntos que corresponden a las observaciones de una serie de tiempo  $\{Y_t\}$ :  $Y_1, Y_2, \dots, Y_N$ . Se desea generar una combinación lineal de funciones seno que pasen por todos y cada uno de los  $N$  puntos. Para ello podrían ajustarse las funciones seno con las siguientes frecuencias:

$$\begin{aligned} f_1 &= 1/N \\ f_2 &= 2/N \\ f_3 &= 3/N \\ &\vdots \\ f_{[N/2]} &= 1/2 \end{aligned}$$

Donde:

$$[N/2] = \begin{cases} N/2 & \text{si } N \text{ es par} \\ (N-1)/2 & \text{si } N \text{ es impar} \end{cases}$$

Las frecuencias anteriores suelen llamarse **armónicas de Fourier**, por haber sido creadas precisamente por este matemático. Corresponden a los recíprocos de las longitudes respectivas, es decir:

$$\begin{aligned} L_1 &= N \\ L_2 &= N/2 \\ L_3 &= N/3 \\ &\vdots \\ L_{[N/2]} &= 2 \end{aligned}$$

Obsérvese que la máxima longitud posible consiste en una función seno que toma todos los puntos para concluir un solo periodo, mientras que la mínima longitud consiste en que se concluya todo un periodo cada dos puntos.

Así, la combinación lineal de  $[N/2]$  funciones seno que pasan por los  $N$  puntos de la serie de tiempo puede expresarse de acuerdo con la ecuación 1.5 de la **función seno generalizada**:

$$Y_t = \sum_{k=0}^{[N/2]} A_k \operatorname{sen} \left( \frac{2\pi kt}{N} + \phi_k \right) \quad (1.5)$$

Con objeto de separar los dos términos que conforman el argumento de la función seno, de manera que sea posible estimarlos con modelos lineales, puede usarse la identidad trigonométrica:

$$\operatorname{sen}(u+v) = \operatorname{sen} u \cos v + \cos u \operatorname{sen} v$$

Así, la ecuación 1.5 quedará como:

$$Y_t = \sum_{k=0}^{[N/2]} \left[ A_k \operatorname{sen} \left( \frac{2\pi kt}{N} \right) \cos \phi_k + A_k \cos \left( \frac{2\pi kt}{N} \right) \operatorname{sen} \phi_k \right]$$

Si se renombran las constantes como:

$$\alpha_k = A_k \sin \phi_k \quad (1.6)$$

$$\beta_k = A_k \cos \phi_k \quad (1.7)$$

Y se sustituyen en la ecuación anterior, se obtiene lo siguiente:

$$Y_t = \sum_{k=0}^{[N/2]} \left[ \alpha_k \cos \left( \frac{2\pi k t}{N} \right) + \beta_k \sin \left( \frac{2\pi k t}{N} \right) \right] \quad (1.8)$$

Que no es otra cosa más que la conocida **serie de Fourier**. Es posible demostrar que esta serie es una combinación lineal de  $N$  funciones ortogonales (seno y coseno), de tal manera que forma una base del espacio vectorial de dimensión  $N$  y, por lo tanto, cualquier conjunto de observaciones de una serie de tiempo  $Y_t$  puede modelarse con esta función, si se encuentran los coeficientes correspondientes. En el caso que nos ocupa, estos coeficientes permitirán, a su vez, reconstruir la función seno generalizada (1.5) con sus respectivas amplitudes, longitudes, frecuencias y fases.

### 1.8.3 Coeficientes de la serie de Fourier

Puede demostrarse que las ecuaciones para calcular los **coeficientes de la serie de Fourier** [58] son las siguientes:

$$\begin{aligned} a_k &= \frac{1}{N} \sum_{t=1}^N Y_t \cos \left( \frac{2\pi k t}{N} \right) \quad \text{si } k = 0, \text{ y } k = N/2 \text{ cuando } N \text{ par} \\ a_k &= \frac{2}{N} \sum_{t=1}^N Y_t \cos \left( \frac{2\pi k t}{N} \right) \quad \text{si } k = 1, 2, \dots, \left[ \frac{N-1}{2} \right] \\ b_k &= \frac{2}{N} \sum_{t=1}^N Y_t \sin \left( \frac{2\pi k t}{N} \right) \quad \text{si } k = 1, 2, \dots, \left[ \frac{N}{2} \right] \end{aligned} \quad (1.9)$$

Donde  $a_k$  y  $b_k$  son los valores numéricos calculados para  $\alpha_k$  y  $\beta_k$ , respectivamente. Ahora, sean  $\omega_k = 2\pi k / N$  para  $k = 0, 1, \dots, [N/2]$ . Estos valores se llaman **frecuencias de Fourier o armónicas**.

**Ejemplo 1: Serie de tiempo con longitud 16.** Supongamos que se tiene una serie de tiempo sumamente sencilla, con 16 valores cuya gráfica se presenta en la Figura 1.13 y que se obtienen evaluando la función  $Y_t = \sin \left( \frac{2\pi k t}{N} \right)$  para  $k = 1, 2, \dots, 16$ .

Para estos datos es posible calcular los coeficientes de Fourier, tal como se muestra en el Cuadro 1.1.

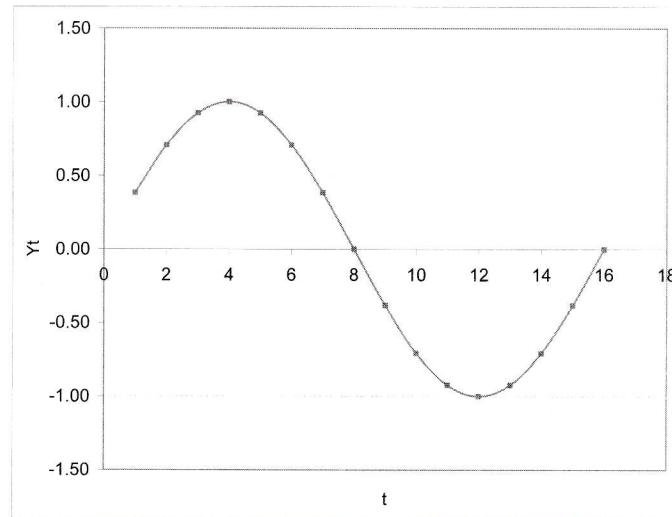


Figura 1.13: Ejemplo 1 - Función seno con  $A = 1$ ,  $L = 16$ ,  $f = 1/16$ ,  $\phi = 0$ .

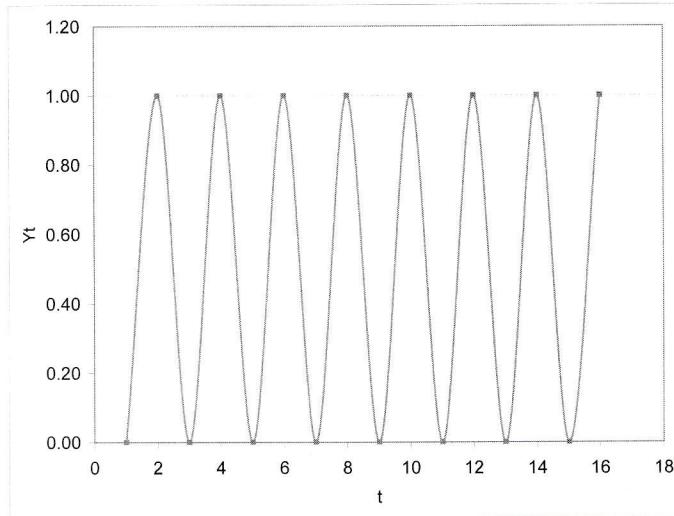
Cuadro 1.1: Coeficientes de Fourier del Ejemplo 1.

$k$	$a_k$	$b_k$
0	0.0	-
1	0.0	1.0
2	0.0	0.0
3	0.0	0.0
4	0.0	0.0
5	0.0	0.0
6	0.0	0.0
7	0.0	0.0
8	0.0	0.0

**Ejemplo 2: Serie de tiempo con longitud 2.** Se tiene una serie de tiempo sumamente sencilla, con 16 valores que son 0 y 1 en forma alternada (datos del Cuadro 1.2, los cuales corresponden a la gráfica que se presenta en la Figura 1.14. Para estos datos es posible calcular los coeficientes de Fourier que se muestran en el Cuadro 1.3:

Cuadro 1.2: Datos del Ejemplo 2.

$t$	$Y_t$	$t$	$Y_t$
1	0	9	0
2	1	10	1
3	0	11	0
4	1	12	1
5	0	13	0
6	1	14	1
7	0	15	0
8	1	16	1

Figura 1.14: Ejemplo 2 - Función seno con  $A = 1$ ,  $L = 2$ ,  $f = 8/16$ ,  $\phi = 0$ .

**Ejemplo 3: Serie de tiempo aleatoria.** Se tiene una serie de tiempo con 16 valores (Cuadro 1.4) que se obtuvieron utilizando la instrucción `=aleatorio()` en la hoja de cálculo *Microsoft Excel*. Estos datos produjeron la gráfica que se presenta en la Figura 1.15.

Cuadro 1.3: Coeficientes de Fourier del Ejemplo 2.

$k$	$a_k$	$b_k$
0	0.5	-
1	0.0	0.0
2	0.0	0.0
3	0.0	0.0
4	0.0	0.0
5	0.0	0.0
6	0.0	0.0
7	0.0	0.0
8	0.5	0.0

Cuadro 1.4: Datos del Ejemplo 3.

$t$	$Y_t$	$t$	$Y_t$
1	0.76	9	0.63
2	0.67	10	0.13
3	0.99	11	0.57
4	0.81	12	0.37
5	0.21	13	0.53
6	0.46	14	0.08
7	0.68	15	0.44
8	0.10	16	0.75

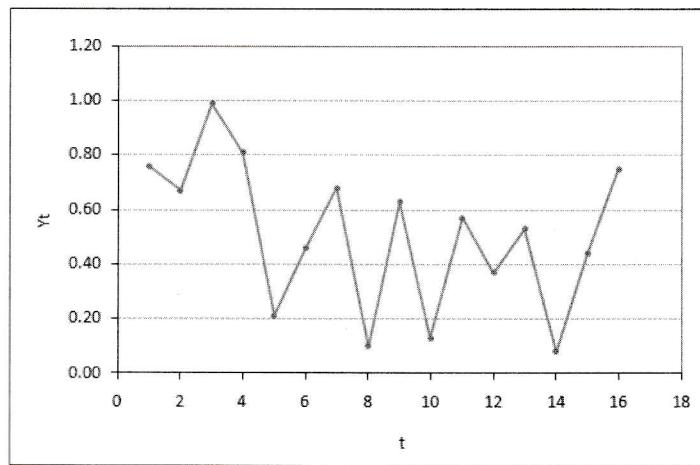


Figura 1.15: Ejemplo 3 - Función con valores aleatorios.

Para estos datos es posible calcular los coeficientes de Fourier, que se muestran en el Cuadro 1.5.

Cuadro 1.5: Coeficientes de Fourier del Ejemplo 3.

$k$	$a_k$	$b_k$
0	0.51	-
1	0.12	0.17
2	-0.02	0.13
3	-0.02	0.06
4	0.09	-0.07
5	0.16	0.01
6	-0.06	0.06
7	0.07	-0.11
8	-0.09	0.00

#### 1.8.4 Espectro lineal

Los coeficientes de Fourier pueden obtenerse de forma manual o, preferentemente, con un programa u hoja de cálculo como *Excel*. Una vez que se cuente con ellos para cada una de las frecuencias armónicas, será posible obtener los valores correspondientes de la función seno generalizada, a través de las ecuaciones 1.6 y 1.7, como se explica a continuación.

La constante, término independiente o desplazamiento vertical, se calcula como:

$$A_0 = a_0 \quad (1.10)$$

Las amplitudes correspondientes a cada frecuencia serán:

$$A_k = \sqrt{a_k^2 + b_k^2} \quad \text{para } k = 1, 2, \dots, [N/2] \quad (1.11)$$

Y las fases se obtienen como:

$$\phi_k = \arctan\left(\frac{a_k}{b_k}\right) \quad \text{para } k = 1, 2, \dots, [N/2] \quad (1.12)$$

Observarás que utilizamos el arco tangente (y no el arco seno o arco coseno) para encontrar rápidamente el ángulo que satisface simultáneamente ambas propiedades, es decir, que está en el cuadrante correspondiente del círculo trigonométrico (Figura 1.16). Sin embargo, si utilizas una calculadora, hoja de cálculo o lenguaje de programación, es probable que la operación arco tangente devuelva ángulos negativos. En este caso, será necesario convertirlos en positivos, identificando correctamente el cuadrante que corresponde. Por ejemplo, si tanto  $a_k$  como  $b_k$  son positivos, el ángulo está en el primer cuadrante y no es necesario modificarlo. Si  $a_k > 0$  y  $b_k < 0$ , significa que el ángulo está en el segundo cuadrante. Si se obtiene un ángulo negativo, podrá

hacerse positivo sumando el valor de  $\pi$ . De cualquier forma, en adelante no se utiliza la fase.

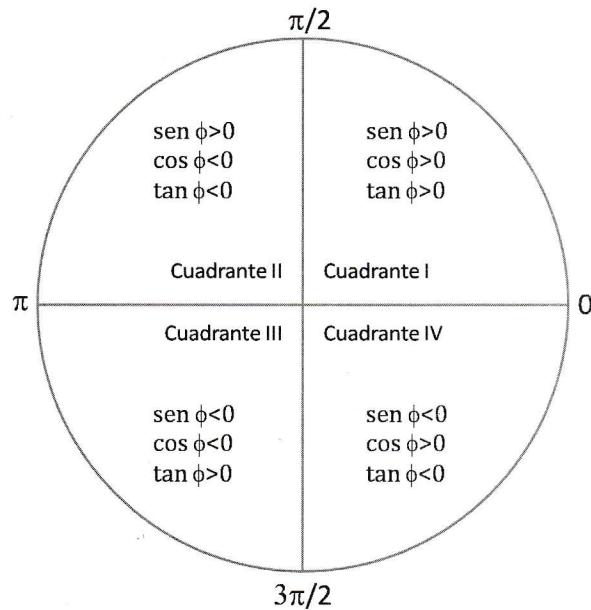


Figura 1.16: Cuadrantes del círculo trigonométrico.

¿Cuál es la utilidad de estos resultados dentro del análisis de series de tiempo? Lo que se ha hecho es descomponer a la serie en un conjunto de funciones seno, cada una con su respectiva longitud, frecuencia, amplitud y fase. Si la serie real contiene algún tipo de periodicidad, la función seno cuya frecuencia corresponda con ella deberá mostrar una amplitud más destacada. Por supuesto, si una serie tiene periodicidad de orden  $k$ , entonces también tendrá periodicidad en los órdenes correspondientes a los múltiplos:  $2k, 3k, 4k, \dots$ . Así, si se elabora una gráfica de las amplitudes (eje vertical) contra las frecuencias (eje horizontal), se espera que las amplitudes correspondientes a la frecuencia correspondiente al periodo subyacente, así como sus múltiplos, destaque con respecto a las demás. Esta gráfica recibe el nombre de **espectro lineal**.

**Ejemplo 1: Serie de tiempo con longitud 16.** Con los datos ya expuestos y los coeficientes de Fourier, es posible obtener los valores de las amplitudes y fases que se muestran en el Cuadro 1.6. A su vez, estos valores permiten graficar el espectro lineal, como se muestra en la Figura 1.17. En ella es posible apreciar que la única frecuencia cuya amplitud es distinta de cero es la armónica  $f_1 = 1/16$ , la cual corresponde a una longitud de 16, la misma señalada al describir el ejemplo por vez primera.

Cuadro 1.6: Amplitudes y fases del Ejemplo 1.

$k$	$A_k$	$\phi_k$
0	0.0	-
1	1.0	1.57
2	0.0	-
3	0.0	-
4	0.0	-
5	0.0	-
6	0.0	-
7	0.0	-
8	0.0	-

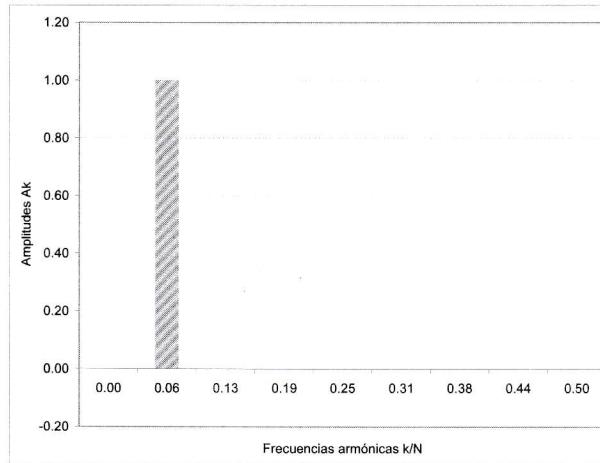


Figura 1.17: Espectro lineal del Ejemplo 1.

**Ejemplo 2: Serie de tiempo con longitud 2.** Con los datos ya expuestos y los coeficientes de Fourier, es posible obtener los valores de las amplitudes y fases del Cuadro 1.7:

Cuadro 1.7: Amplitudes y fases del Ejemplo 2.

$k$	$A_k$	$\phi_k$
0	0.5	-
1	0.0	-
2	0.0	-
3	0.0	-
4	0.0	-
5	0.0	-
6	0.0	-
7	0.0	-
8	0.5	0.0

Con las amplitudes y fases es posible entonces graficar el espectro lineal correspondiente, como se muestra en la Figura 1.18. En ella es posible apreciar que las dos frecuencias cuya amplitud es distinta de cero son las correspondientes al término constante y la armónica  $f_8 = 8/16$ , la cual corresponde a una longitud de 2. De esta forma comprobamos que, a través de los coeficientes de Fourier, se recuperan las características de la función seno original.

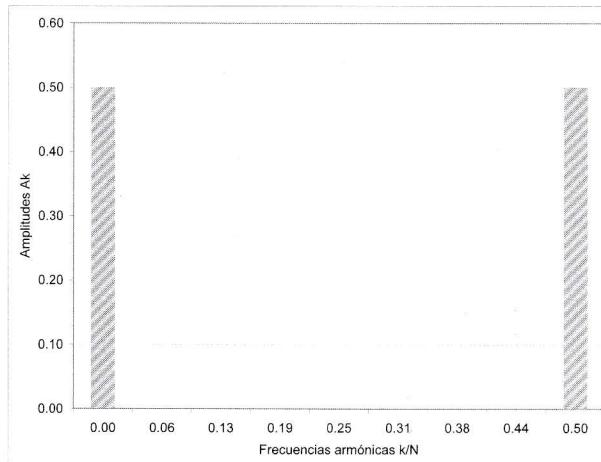


Figura 1.18: Espectro lineal del Ejemplo 2.

Cuadro 1.8: Amplitudes y fases del Ejemplo 3.

$k$	$A_k$	$\phi_k$
0	0.51	-
1	0.20	0.62
2	0.13	-0.17
3	0.07	-0.35
4	0.11	-0.90
5	0.16	1.50
6	0.09	-0.75
7	0.13	-0.60
8	0.09	-1.57

**Ejemplo 3: Serie de tiempo aleatoria.** Con los datos ya expuestos y los coeficientes de Fourier, es posible obtener los valores de las amplitudes y fases que se muestran en el Cuadro 1.8:

Estos valores generan el espectro lineal que se muestra en la Figura 1.19. En ella es posible apreciar que sobresalen prácticamente todas las amplitudes y ninguna de ellas domina. Esto significa que la serie carece de periodicidad, lo cual es lógico considerando que procede de datos completamente aleatorios.

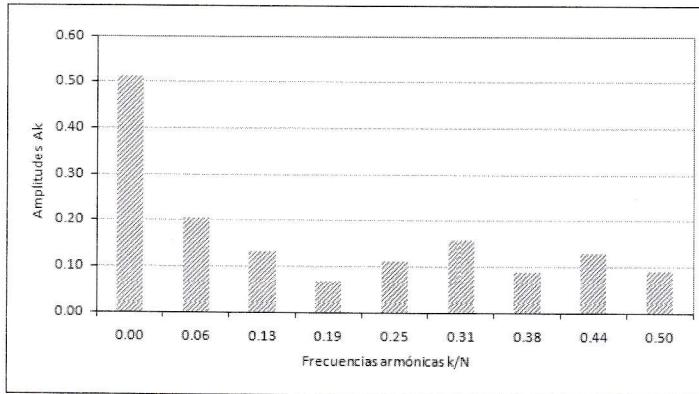


Figura 1.19: Espectro lineal del Ejemplo 3.

### 1.8.5 Periodograma

Si bien el espectro lineal es útil para detectar posibles componentes periódicos en una serie de tiempo, en el caso particular en que exista periodicidad pero con amplitudes relativamente bajas, es posible que no sea fácil visualizarla en la gráfica. Entonces, una forma de hacerla más visible es “exagerar”

la amplitud. Con este objeto Schuster (1898) propuso utilizar la gráfica del **periodograma**, el cual se calcula como se indica en la ecuación 1.13:

$$I(\omega_k) = \begin{cases} Na_0^2 & k = 0, \\ \frac{N}{2} (a_k^2 + b_k^2) & k = 1, 2, \dots, [\frac{N-1}{2}], \\ Na_{N/2}^2 & k = \frac{N}{2} \text{ cuando } N \text{ es par.} \end{cases} \quad (1.13)$$

El periodograma se define de esta forma para que la suma de todos sus valores sea igual a  $\sum_{t=1}^N Y_t^2$ , es decir, a la suma de cuadrados de la serie de tiempo. Al tomar los cuadrados de las amplitudes y multiplicarlos por el valor de  $N$  los valores grandes se harán aún mayores, mientras que los pequeños disminuyen. Así se incrementa el contraste entre unos y otros. La mayor parte de los programas estadísticos, como SPSS, *Statgraphics* y SAS, sólo colocan en la gráfica los valores del periodograma para las frecuencias mayores que cero. Como ejemplo puede verse la Figura 1.20, en la cual se observa el periodograma del Ejemplo 3, elaborado con SPSS. La mayoría de los paquetes estadísticos no incluyen el valor correspondiente a la frecuencia cero, ya que no brinda información acerca de la periodicidad.

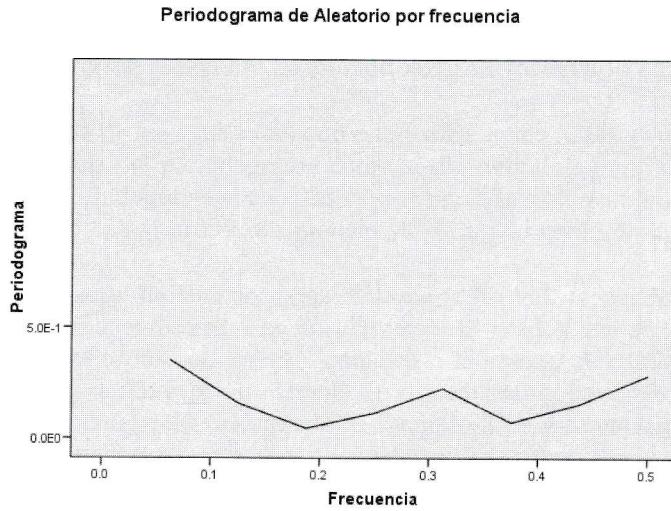


Figura 1.20: Periodograma de valores aleatorios en SPSS.

Las Figuras 1.21 a 1.23 muestran los periodogramas de los Ejemplos 1 a 3, respectivamente.

## 1.8 Periodicidad y su detección

---

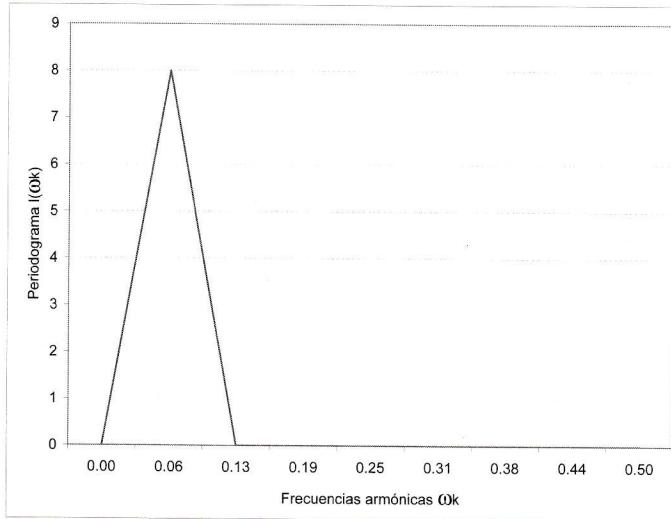


Figura 1.21: Periodograma del Ejemplo 1.

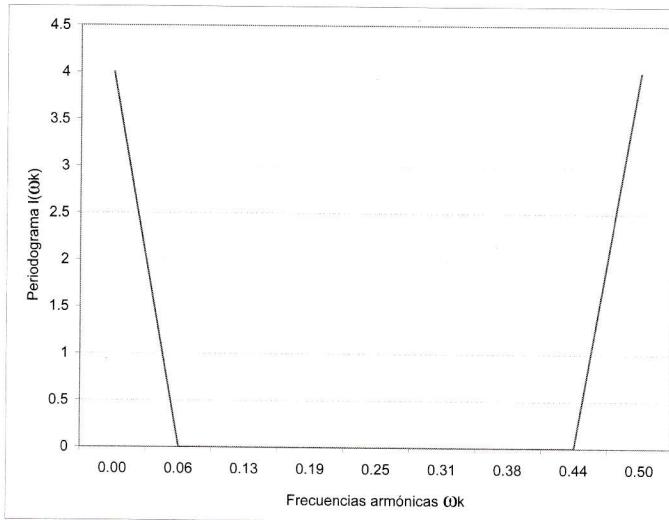


Figura 1.22: Periodograma del Ejemplo 2.

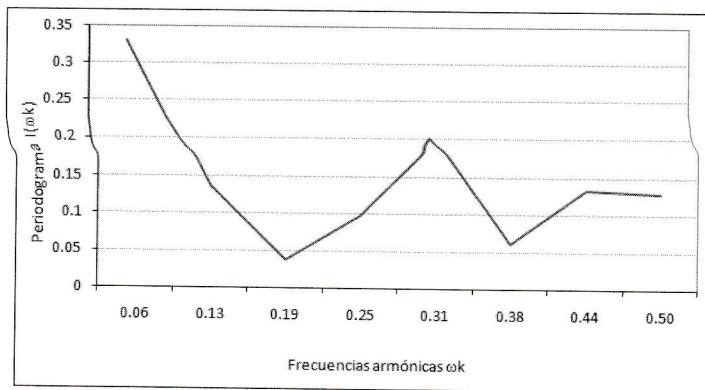


Figura 1.23: Periodograma del Ejemplo 3.

### 1.8.6 Periodograma integrado

Otra forma de analizar con rapidez el periodograma es hacer la suma normalizada de sus valores y representarlo a través de una prueba de bondad de ajuste no paramétrica de Kolmogorov-Smirnov. La suma normalizada se calcula como:

$$IP(\omega_k) = \sum_{j=0}^k \frac{I(\omega_j)}{\sum_{t=1}^N Y_t^2} \quad (1.14)$$

La gráfica de la suma normalizada contra los valores de las frecuencias armónicas  $f_k = k/N$  recibe el nombre de **periodograma integrado** (*integrated periodogram*).

La mayor parte de las aplicaciones de estadística, como *Statgraphics*, ofrecen tanto la gráfica del periodograma como la del periodograma integrado. La ventaja del periodograma integrado es que permite hacer un análisis visual rápido tanto del ajuste de un modelo ya estimado como de la posible periodicidad.

Cuando una serie de tiempo carezca de periodicidad, el periodograma integrado tenderá a parecerse a la línea recta a 45 grados que sería la suma de amplitudes todas con valores mínimos e idénticos. A continuación, en la Figura 1.24 se muestra un ejemplo de periodograma integrado sin señales de periodicidad, que corresponde al Ejemplo 3 de este capítulo. (Los periodogramas integrados que se muestran aquí se obtuvieron con el software *Statgraphics*).

En cambio, cuando una serie de tiempo contenga periodicidad, el periodograma presentaría valores altos tanto en la frecuencia correspondiente como en sus múltiplos. Entonces, en el periodograma integrado se observarán una especie de “escalones” indicando ya sea variación estacional o ciclo. Las Figuras 1.25 y 1.26 muestran los periodogramas integrados para los Ejemplos 2 y 3, respectivamente.

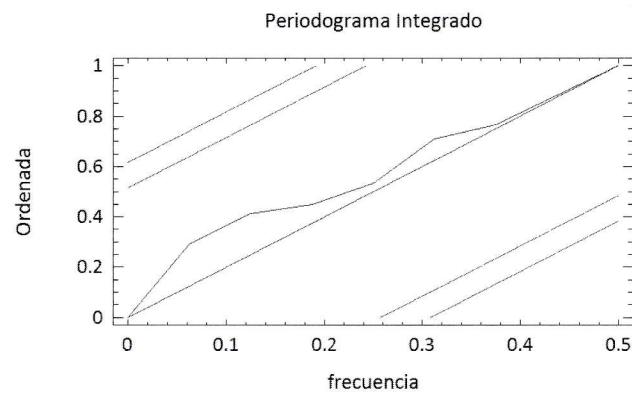


Figura 1.24: Periodograma Integrado Ejemplo 3.

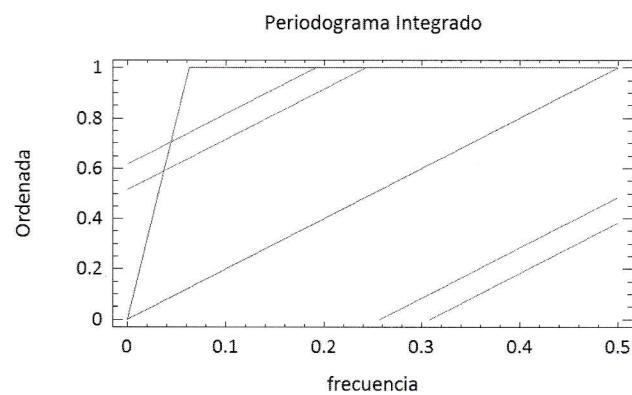


Figura 1.25: Periodograma Integrado Ejemplo 1.

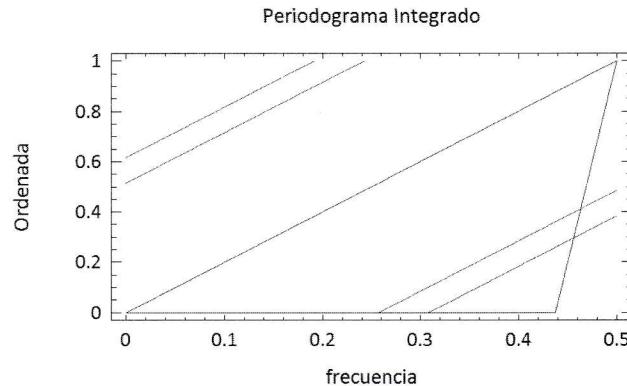


Figura 1.26: Periodograma Integrado Ejemplo 2.

Finalmente, es importante observar qué significa si el periodograma integrado muestra un valor muy alto al inicio, es decir, en frecuencias cercanas a cero: esto implica que la serie tiene tendencia.

## 1.9 Funciones muestrales

De acuerdo con los principios de la estadística inferencial, es posible obtener conclusiones acerca de las propiedades y características de una población, a través de una muestra representativa, es decir, de un subconjunto de elementos tomados de la población. En el caso de las series de tiempo, la única muestra con la que se cuenta es el conjunto de observaciones históricas:  $Y_1, Y_2, \dots, Y_N$ . Puesto que, en general, la serie de tiempo ocurre por única vez y es imposible replicarla, se acostumbra llamar a estas observaciones una **realización** del proceso estocástico, de entre todas las realizaciones posibles.

Como se verá en el Capítulo 2, esta realización sólo podrá considerarse como una muestra representativa de tamaño  $N$  bajo ciertas condiciones. Cuando esto ocurra, tendrá sentido calcular a través de estos datos los estimadores de la media, varianza, autocovarianzas y autocorrelaciones. Así, en esta sección se presentan los estimadores bajo los siguientes tres supuestos:

1. La función de media es constante en el tiempo, no hay tendencia. Para verificar este supuesto se observa la gráfica de la serie de tiempo, si los puntos se mantienen alrededor de una línea horizontal, puede considerarse que no hay tendencia.
2. La función de varianza es constante en el tiempo (la serie es homoscedástica). Para verificar este supuesto se observa la gráfica de la serie de

tiempo, si los puntos se mantienen dentro de una franja o banda constante alrededor de la media, se considera que la varianza es constante.

3. La función de autocorrelación (o autocovarianza) no depende del tiempo  $t$ , sino únicamente del intervalo entre dos variables,  $k$ . Esta suposición no puede hacerse de manera visual, más adelante se revisarán las herramientas para tratar este aspecto.

### 1.9.1 Función de media muestral

Si la función de media es constante, puede estimarse a través del promedio aritmético de los datos, de la siguiente forma:

$$\hat{\mu}_Y = \bar{Y} = \frac{1}{N} \sum_{t=1}^N Y_t \quad (1.15)$$

### 1.9.2 Función de varianza muestral

Si la función de varianza es constante a través del tiempo (la serie de tiempo es homoscedástica), podrá ser estimada con:

$$\hat{\text{Var}}(Y_t) = \hat{\gamma}_0 = S_Y^2 = \frac{\sum_{t=1}^N (Y_t - \bar{Y})^2}{N - 1} \quad (1.16)$$

### 1.9.3 Funciones muestrales de relación

Se considera, además, que la función de autocovarianza permanece constante en el tiempo, para variables que están separadas por un intervalo de la misma longitud  $k$ . Si la media es cero y la varianza es constante, la función de autocovarianza muestral se calcula como:

$$\begin{aligned} \gamma_k &= \text{Cov}(Y_t, Y_{t+k}) \\ &= E[(Y_t - \mu_Y)(Y_{t+k} - \mu_Y)] \end{aligned} \quad (1.17)$$

Y el estimador de la autocovarianza es:

$$\hat{\gamma}_k = \sum_{t=1}^N \frac{(Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{N - 1} \quad (1.18)$$

Mientras que la función de autocorrelación puede expresarse como:

$$\begin{aligned} \rho_k &= \text{Corr}(Y_t, Y_{t+k}) \\ &= \frac{\text{Cov}(Y_t, Y_{t+k})}{\sqrt{\text{Var}(Y_t)\text{Var}(Y_{t+k})}} \\ &= \frac{\gamma_k}{\gamma_0} \end{aligned} \quad (1.19)$$

Donde el valor de  $\rho_k$  debe ser tal que  $-1 \leq \rho_k \leq 1$ , porque tanto la matriz de autocorrelación como la de autocovarianzas son definidas positivas.

Finalmente, el estimador de la función de autocorrelación en este caso será:

$$r_k = \hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (1.20)$$

## 1.10 Actividades sugeridas

1. Esta actividad se sugiere para un foro virtual. Localiza la gráfica de una serie de tiempo real. Indica:
  - a) De qué son los datos.
  - b) La fuente en que obtuviste los datos, de manera que cualquier otra persona pueda consultarlos también.
  - c) Si se observa o no cada uno de los cuatro componentes (tendencia, variación estacional, ciclo, fluctuación aleatoria) en tu serie y por qué lo crees así.
  - d) Adjunta la gráfica con copia-pegado. Revisa que se vea correctamente.
  - e) Antes de colocar tu respuesta, verifica que tu serie sea distinta de las que ya están colocadas, a menos que quieras enfatizar alguna diferencia de opinión con una de las que ya aparecen.
2. Con algún programa u hoja de cálculo, calcula 60 puntos ( $t=1, 2, \dots, 60$ ) de las siguientes funciones. Grafica y une los puntos. Señala en cada gráfica tanto el valor numérico como la interpretación visual de amplitud, longitud, frecuencia y fase:
  - a)  $f(t) = \sin\left(\frac{2\pi t}{60}\right)$
  - b)  $f(t) = 5 \sin\left(\frac{2\pi t}{60}\right)$
  - c)  $f(t) = 5 \sin\left(\frac{2\pi 10t}{60}\right)$
  - d)  $f(t) = 8 \sin\left(\frac{2\pi 10t}{60} + 15\right)$
  - e)  $f(t) = 20 + 8 \sin\left(\frac{2\pi 10t}{60} + 15\right)$

### 1.10 Actividades sugeridas

$$f) \quad f(t) = 20 + 8 \operatorname{sen}\left(\frac{2\pi 10t}{60} + 15\right) + 2 \operatorname{sen}\left(\frac{2\pi 5t}{60} + 5\right)$$

3. Para las siguientes series de 16 datos, calcula los coeficientes de Fourier, grafica el espectro lineal y el periodograma e indica, finalmente, si existe periodicidad y en su caso a qué frecuencias corresponde:
- 1,0,1,-1,0,1,-1,0,1,-1,0,1,-1,0,1,-1.
  - 1,2,3,4,5,6,7,8,1,2,3,4,5,6,7,8.
  - 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16.
  - 16 datos generados con la función `=aleatorio()` de Excel.
- e)  $f(t) = 10 + 8 \operatorname{sen}\left(\frac{2\pi 2t}{16} + 5\right) + 2 \operatorname{sen}\left(\frac{2\pi 4t}{16}\right).$
4. Para los siguientes datos, elabora con *Statgraphics* la gráfica de la serie original. Indica si la media es constante o no. Indica si la varianza es constante o no. Agrega las gráficas de periodograma y periodograma integrado; indica si existe o no periodicidad y, en su caso, cuál es su longitud.
- 48 valores simulados de ruido blanco con desviación estándar  $\sigma = 1$ . Deben ser valores entre -3 y 3.
  - 48 datos calculados con  $f(t) = 5 + \operatorname{sen}\left(\frac{2\pi 3t}{48}\right).$
  - 48 datos calculados sumando los puntos a y b.
  - Al menos 48 datos de una serie de tiempo real, en la cual observes variación estacional. Indica la referencia de dónde obtuviste la serie, usando algún administrador de referencias como: *Zotero*, *My EndNote Web*, *BibTex*, *JabRef*, entre otros.
5. Experimenta y juega con las series de Fourier en el sitio: [http://phet.colorado.edu/simulations/sims.php?sim=Fourier\\_Making\\_Waves](http://phet.colorado.edu/simulations/sims.php?sim=Fourier_Making_Waves). También puedes descargar el programa.

# Capítulo 2

## Modelos estacionarios

Todos los modelos son falsos, pero algunos modelos son útiles.  
GEORGE E. P. BOX

### 2.1 Estacionaridad

En este capítulo se revisarán los conceptos de estacionaridad y su importancia para el modelo matemático. Se explicarán los diversos modelos estacionarios y sus características.

#### 2.1.1 Dos conceptos

La estacionaridad será crucial para identificar el modelo más adecuado para un conjunto de los datos. El concepto de **estacionaridad** (*stationarity*) puede definirse de dos formas:

**Concepto de estacionaridad en sentido fuerte o estricto.** Se dice que una serie de tiempo  $Y_t$  es estacionaria en sentido estricto o fuerte, si su función de distribución conjunta de probabilidad permanece idéntica en el tiempo. Es decir, es estacionaria en sentido estricto si y sólo si:

$$f(Y_t, Y_{t-1}, \dots, Y_{t-k}) = f(Y_{t+\tau}, Y_{t-1+\tau}, \dots, Y_{t-k+\tau}), \quad \forall \tau \in \mathbb{R} \quad (2.1)$$

Por lo general esta propiedad no se puede verificar en datos reales, pero puede explicarse con un ejemplo sencillo. Si la serie de tiempo  $Y_t$  representa el resultado del  $t$ -ésimo lanzamiento de una moneda cuyos dos resultados posibles son águila ( $Y_t = 0$ ) o sol ( $Y_t = 1$ )<sup>1</sup>, puede considerarse que se trata

<sup>1</sup>El origen de la frase “águila o sol”, que en otros países es “cara o cruz”, es una moneda de veinte centavos mexicanos que circuló entre los años cincuenta y setenta. Sus caras tenían efectivamente un águila en el anverso y, en el reverso, la Pirámide del Sol de Teotihuacán bajo un gorro frigio resplandeciente, que es un símbolo de la libertad. La Figura 2.1 muestra la imagen de esta moneda (Fuente: Modificado de [8])

## 2.1 Estacionaridad

---

de una serie de tiempo estacionaria en sentido estricto, ya que la función de distribución de probabilidad conjunta no depende del tiempo: los lanzamientos son independientes entre sí y la probabilidad de obtener águila o sol es idéntica a lo largo del tiempo. En este caso particular se trata de un proceso estocástico Bernoulli, el cual es estacionario en sentido estricto (Figura 2.2).



Figura 2.1: Moneda antigua de veinte centavos mexicanos.

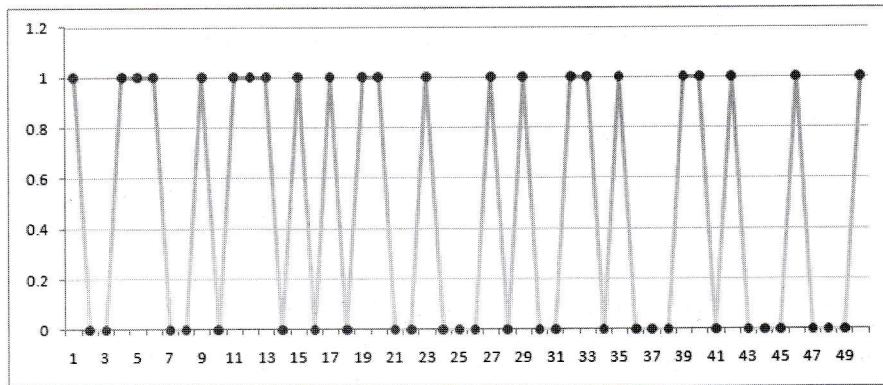


Figura 2.2: Ejemplo de realización de un proceso Bernoulli.

En cambio, si  $Y_t$  representa ahora la suma de las águilas obtenidas en el  $t$ -ésimo lanzamiento de una moneda, se tendrá un proceso binomial. Es decir,  $Y_t = \sum_{n=0}^{n=t} X_n$ , donde  $X_n$  es un proceso Bernoulli. Para el tiempo  $t = 1$  se tendrá una probabilidad binomial con  $n = 1$  y  $p = 0.5$ ; para el tiempo  $t = 2$  se tendrá una probabilidad binomial con  $n = 2$  y  $p = 0.5$ , y así sucesivamente. Es decir, la función de probabilidad cambia con el número de lanzamiento. En general, para el lanzamiento  $t$ , la distribución de probabilidad será una binomial( $t, 0.5$ ), cuyo valor dependerá entonces de  $t$ . De tal manera que el valor esperado  $E(Y_t) = 0.5t$ , cuya representación geométrica es una recta

creciente; mientras que la varianza será  $\text{Var}(Y_t) = 0.25t$ , lo cual significa una dispersión también creciente. Por lo tanto, el proceso binomial no es estacionario en sentido estricto (Figura 2.3).

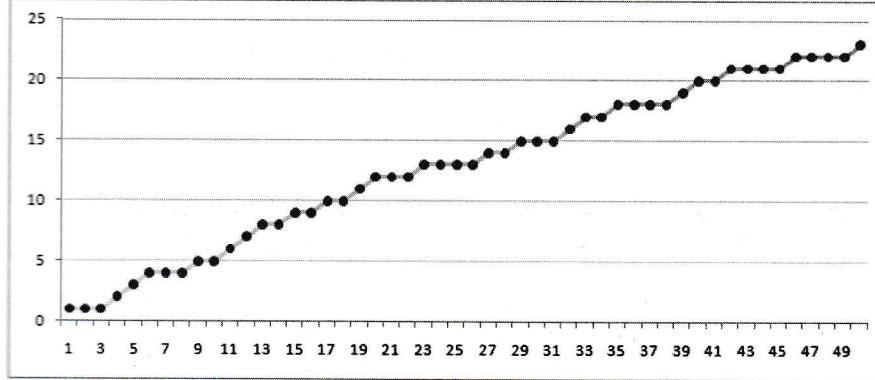


Figura 2.3: Ejemplo de realización de un proceso binomial.

**Concepto de estacionaridad en sentido débil o amplio.** En virtud de que la estacionaridad en sentido estricto es difícil de probar, puede utilizarse un concepto alterno más sencillo. Se dice que una serie de tiempo es estacionaria en sentido amplio o débil, si y sólo si satisface tres condiciones:

1. *Su media es constante en el tiempo.* Es decir,

$$E(Y_t) = E(Y_{t+\tau}) = \mu, \quad \forall \tau \in \mathbb{R}. \quad (2.2)$$

Aunque aparentemente parece un problema trivial, si la serie muestra tendencia (generalmente creciente en muchos casos reales), el problema se complica. Si se ve la serie como una realización de todas las series de tiempo que podrían ser generadas por un mismo mecanismo o función, se cuenta tan solo con una muestra de tamaño uno. Con esta muestra deberá estimarse la media, que en algunos casos será constante y en otros será una función del tiempo. Si no hay tendencia, puede asumirse que la media es constante y que el valor observado para cada periodo puede representarse por una constante que será la media muestral.

En este caso la **media** de la serie podrá estimarse a través del promedio aritmético, como se muestra en la ecuación 2.3.

$$\hat{\mu} = \bar{Y} = \frac{1}{N} \sum_{t=1}^N Y_t \quad (2.3)$$

## 2.1 Estacionaridad

---

2. *La varianza es constante en el tiempo.* Una segunda condición es que la dispersión del proceso sea también invariante en el tiempo, *i.e.:*

$$\text{Var}(Y_t) = \text{Var}(Y_{t+\tau}) = \gamma_0, \quad \forall \tau \in \mathbb{R}. \quad (2.4)$$

Si la **varianza** es constante, puede estimarse a través de la ecuación 2.5

$$\hat{\gamma}_0^2 = S^2 = \frac{\sum(Y_t - \bar{Y})^2}{N - 1} \quad (2.5)$$

Que es el estimador insesgado<sup>2</sup> para la varianza. Recuérdese que si se divide entre  $N$  el estimador sería asintóticamente insesgado.

3. *La función de autocorrelación es independiente del tiempo.* Finalmente, debe imponerse una condición más sobre la correlación entre los datos. La autocorrelación<sup>3</sup> mide la posible dependencia entre un valor observado  $Y_t$  y otro  $Y_{t-k}$  que está separado por un intervalo de longitud  $k$ :

$$\text{Autocorr}(Y_t, Y_{t-k}) = \text{Autocorr}(Y_{t+\tau}, Y_{t+\tau-k}) = \rho_k, \quad \forall \tau \in \mathbb{R}. \quad (2.6)$$

Si la media y la varianza son constantes, entonces la **autocorrelación** puede expresarse como se ve en la ecuación 2.9.

$$\rho_k = \text{Corr}(Y_t, Y_{t-k}) \quad (2.7)$$

$$= \frac{\text{Cov}(Y_t, Y_{t-k})}{\sqrt{\text{Var}(Y_t)\text{Var}(Y_{t-k})}} \quad (2.8)$$

$$= \frac{\gamma_k}{\gamma_0} \quad (2.9)$$

Así, para una serie de tiempo estacionaria, este valor sólo depende de  $k$ , esto es, la relación entre dos variables sólo depende del intervalo que las separa y no de su posición en el tiempo,  $t$ . Por lo tanto, la función de autocorrelación se estima con la ecuación 2.10

$$r_k = \hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (2.10)$$

Obsérvese que, si la posición en el tiempo  $t$  no es relevante, las autocorrelaciones serán simétricas con respecto al origen, puesto que las

<sup>2</sup>Un estimador es insesgado si su valor esperado es igual al parámetro correspondiente.

<sup>3</sup>Recibe el nombre de *autocorrelación* porque se trata de la correlación entre valores de la misma variable observada en diferentes momentos del tiempo.

variables  $Y_t$  y  $Y_{t-k}$  están separadas por un intervalo de longitud  $k$ , de la misma forma que lo están  $Y_t$  y  $Y_{t+k}$ . Por lo tanto:

$$\rho_k = \rho_{-k} \quad (2.11)$$

Por lo anterior, únicamente se considerarán las autocorrelaciones positivas. La gráfica de las autocorrelaciones estimadas  $r_k$  contra los valores de los intervalos  $k$  suele llamarse **correlograma**. Esta gráfica permite detectar aquellas autocorrelaciones que destacan sobre las otras, esto es, que sobresalen del intervalo de confianza. El correlograma también actuará como un refuerzo para identificar posible periodicidad (variación estacional o ciclo), puesto que en ese caso destacarán las autocorrelaciones para los períodos iguales a la longitud. Por ejemplo, si destacan las autocorrelaciones en  $k = 12$  y los datos son mensuales, se puede concluir que existe variación estacional. Esto se complementará con la información que arrojen el periodograma y el periodograma integrado.

Por último, debe señalarse que si bien todo proceso estocástico en sentido estricto lo será también en sentido amplio, el inverso no es necesariamente cierto. Sin embargo, en el análisis de los datos reales, por lo general sólo será posible determinar la estacionaridad en sentido débil o amplio.

### 2.1.2 Importancia de la estacionaridad

¿Por qué resulta indispensable analizar si una serie de tiempo es o no estacionaria? Si el fenómeno real bajo estudio es estacionario, entonces los valores de la media, la varianza y la función de autocorrelación pueden estimarse considerando a la serie de tiempo con  $N$  datos (una realización particular), como una muestra de tamaño  $N$ , cuyos parámetros son constantes a través del tiempo, donde  $N$  es el número de observaciones disponibles que provienen todas de una misma población. Esto se ilustra en la Figura 2.4. En este caso, además, pueden usarse las ecuaciones correspondientes para estimar los parámetros que no dependen del tiempo.

En cambio, si la serie de tiempo no es estacionaria, el problema es que los  $N$  datos representan en realidad  $N$  muestras de tamaño 1. Una muestra de tamaño uno *no* es representativa, por lo cual es prácticamente imposible inferir las propiedades de la población a partir de los datos, o estimar sus parámetros. Esto se ilustra en la Figura 2.5.

Ahora bien, ¿qué hacer si el fenómeno no es estacionario? En ese caso deberá utilizarse alguna transformación matemática temporal para hacerlo estacionario. Una vez hecho esto, se seguirán los pasos ya descritos en la Figura 1.7 del Capítulo 1. Posteriormente, cuando se obtengan los pronósticos, se realizará la transformación inversa para regresar a los datos originales. Este procedimiento se revisará en el Capítulo 3.

## 2.1 Estacionaridad

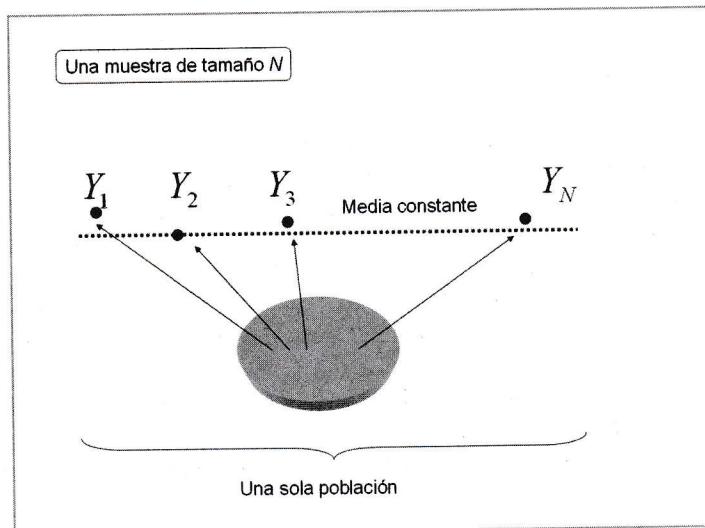


Figura 2.4: Proceso estocástico estacionario.

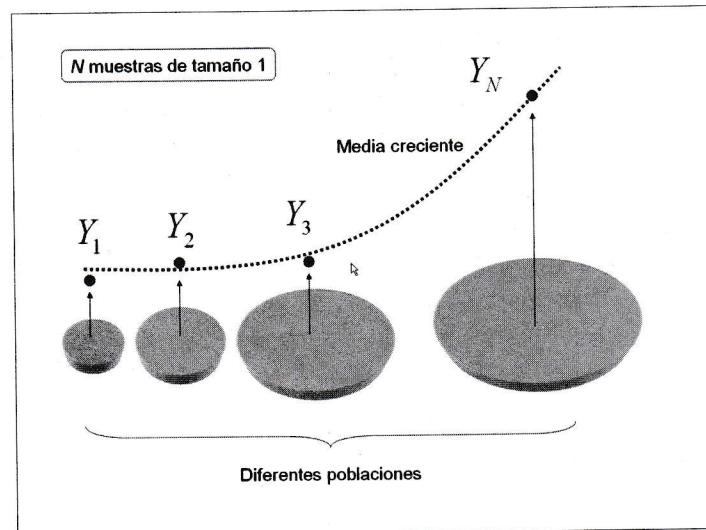


Figura 2.5: Proceso estocástico no estacionario.

## 2.2 Representación matricial

### 2.2.1 Matriz de varianzas y autocovarianzas

Si se cuenta con  $N$  datos provenientes de un procesos estocástico estacionario (ya sea en sentido amplio o estricto), las varianzas y autocovarianzas pueden ordenarse en una matriz simétrica de dimensión  $N \times N$ , en la cual el elemento  $(i, j) \forall i, j \in \{1, 2, \dots, N\}$  representa la autocovarianza entre  $Y_i$  y  $Y_j$  para  $i \neq j$ , y la varianza para  $i = j$ .

Así, si se denota por  $\Gamma$  (Gamma) a la **matriz de varianzas y autocovarianzas**, ésta quedará conformada como se indica en la ecuación 2.12.

$$\Gamma = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{N-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \dots \\ \gamma_{N-1} & \gamma_{N-2} & \gamma_{N-3} & \dots & \gamma_0 \end{bmatrix} \quad (2.12)$$

Cabe señalar que esta matriz es definida positiva, puesto que el producto matricial para cualquier vector renglón  $\mathbf{x}$  de dimensión  $1 \times N$ :

$$\mathbf{x}\Gamma\mathbf{x}^T = \text{Var}(\mathbf{x}) > 0$$

### 2.2.2 Matriz de autocorrelaciones

De la misma forma se construye la **matriz de autocorrelaciones** (denotada por la letra griega Rho), al dividir la matriz de varianzas y autocovarianzas entre la varianza de la serie de tiempo, es decir:

$$\mathbf{P} = \frac{1}{\gamma_0} \Gamma$$

Entonces:

$$\mathbf{P} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{N-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \dots \\ \rho_{N-1} & \rho_{N-2} & \rho_{N-3} & \dots & 1 \end{bmatrix} \quad (2.13)$$

Si  $\Gamma$  es definida positiva, entonces  $\mathbf{P}$  también lo es. Esto implica que los menores principales de la matriz serán mayores que cero, por lo tanto:

$$\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix} > 0$$

$$\begin{vmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{vmatrix} > 0$$

Y así sucesivamente. En general:

$$\begin{vmatrix} 1 & \rho_k \\ \rho_k & 1 \end{vmatrix} > 0$$

## 2.3 Ejemplos

---

Por lo tanto:

$$\begin{aligned}1 - \rho_k^2 &> 0 \\-1 < \rho_k &< 1\end{aligned}\tag{2.14}$$

Este resultado es muy útil. En general, si la autocorrelación  $\rho_k$  se acerca a 1, se dice que existe una relación (o proporción) lineal directa entre las dos variables; si se acerca a -1, la relación es inversa (al crecer una variable, la otra decrece y viceversa). Finalmente, si se acerca a cero, se considera que las variables son independientes.

## 2.3 Ejemplos

A continuación se presentan dos casos reales de series de tiempo mexicanas, que permiten establecer la forma de diferenciar un conjunto de datos estacionarios de otro que no lo es.

### 2.3.1 Ejemplo de datos estacionarios

El primer paso en el análisis de series de tiempo es hacer la gráfica las observaciones disponibles contra el tiempo. Esto sirve para visualizar aspectos tales como:

- Posible tendencia.
- Posible heteroscedasticidad.
- Posible variación estacional o ciclo.
- Posibles discontinuidades, datos faltantes o perdidos (*missing data*).
- Posibles datos atípicos (*outliers*), es decir, que salen del comportamiento general de la mayoría de los datos.

En la Figura 2.6 se presenta la gráfica de la inflación mensual en México<sup>4</sup>, es decir, los niveles mensuales del Índice Nacional de Precios al Consumidor (INPC).

Se observan media y varianza constantes, es decir, el fenómeno parece estacionario en sentido amplio, ya que los datos permanecen alrededor de una recta horizontal, dentro de una franja o banda homogénea. La media aritmética es de 0.4134 y la desviación estándar estimada es de 0.3242. La Figura 2.7 muestra gráficamente esta idea: prácticamente todos los datos se encuentran ubicados entre la media más/menos dos desviaciones estándar.

<sup>4</sup>Banco de México. *Índices Mensuales de Precios al Consumidor*. <http://www.banxico.gob.mx/polmoneinflacion/estadisticas/indicesPrecios/indicesPreciosConsumidor.html>. Periodo de enero 2000 a febrero 2007. Consultado en marzo 2007. México.

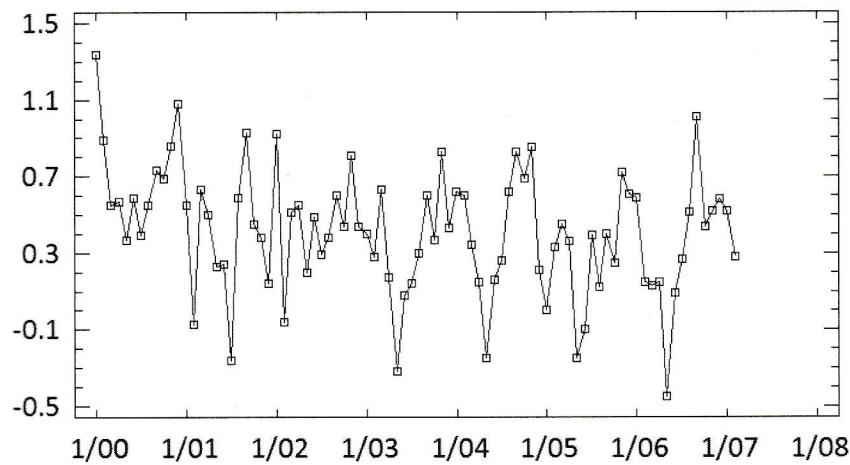


Figura 2.6: Tasa de inflación mensual en México.

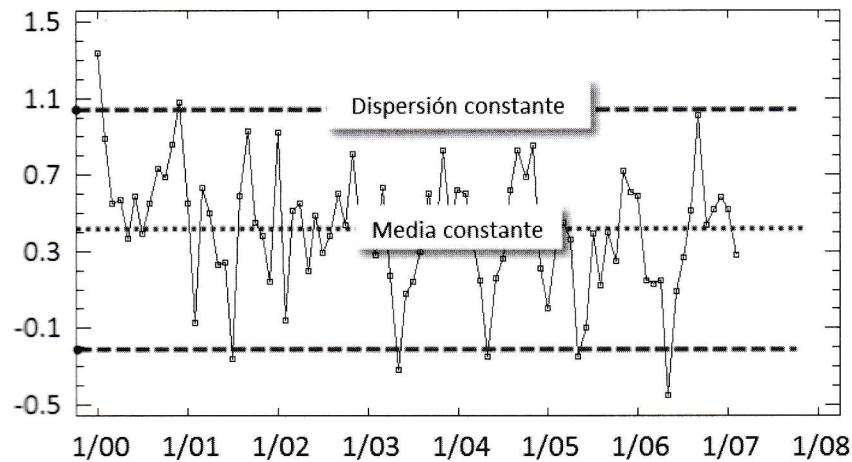


Figura 2.7: La tasa de inflación es estacionaria.

### 2.3 Ejemplos

---

Por otro lado, tanto la función de autocorrelación (ACF) mostrada en la Figura 2.8 como el periodograma y el periodograma integrado que se muestran en las Figuras 2.9 y 2.10, respectivamente, indican la existencia de una posible variación estacional cada doce meses. La función de autocorrelación muestra un valor relativamente alto en 12 que coincide con el pico y el escalón del periodograma. En la tabla del periodograma que calcula *Statgraphics* (Figura 2.11) puede notarse que el valor más alto de la ordenada corresponde a una longitud aproximada de 12 meses, lo cual coincide sin problemas con la interpretación: el comportamiento del INPC se repite en forma semejante cada año. Además, destaca el hecho de que la ACF decrece en forma bastante rápida.

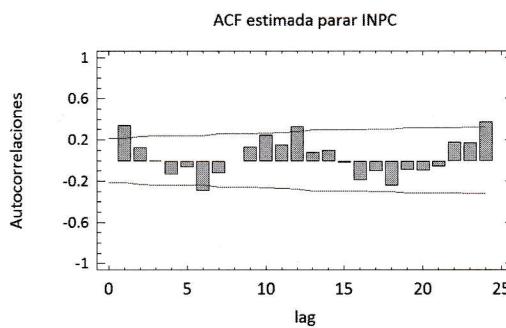


Figura 2.8: ACF de la tasa de inflación mensual en México.

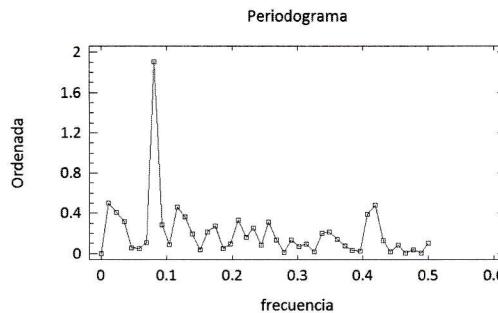


Figura 2.9: Periodograma de la tasa de inflación mensual en México.

Así pues, con estos datos será posible trabajar sin mayores problemas, ya que parecen cumplir las tres condiciones de la estacionaridad. Sin embargo, muestran variación estacional. Como se verá más adelante, existen dos formas de tratar con este tipo de periodicidad:

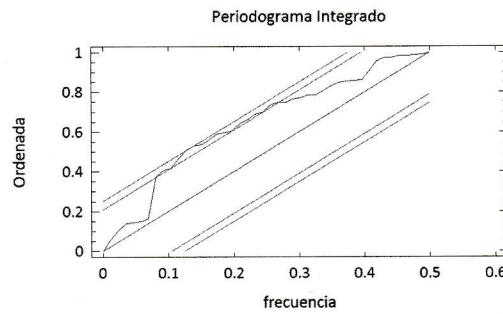


Figura 2.10: Periodograma integrado de la tasa de inflación mensual en México.

<i>i</i>	Frequency	Period	Ordinate	Cumulative	Integrated
<i>i</i>				Sum	Periodogram
0	0.0		6.76292E-28	6.76292E-28	7.57039E-29
1	0.0116279	86.0	0.502723	0.502723	0.0562746
2	0.0232558	43.0	0.411313	0.914035	0.102317
3	0.0348837	28.6667	0.317243	1.23128	0.137829
4	0.0465116	21.5	0.0565571	1.28784	0.14416
5	0.0581395	17.2	0.0493069	1.33714	0.149679
6	0.0697674	14.3333	0.109341	1.44648	0.161919
7	0.0813953	12.2857	1.90178	3.34826	0.374803
8	0.0930233	10.75	0.281534	3.6298	0.406318
9	0.104651	9.55556	0.0914279	3.72122	0.416553
10	0.116279	8.6	0.461512	4.18274	0.468214
11	0.127907	7.81818	0.362652	4.54539	0.508809

Figura 2.11: Tabla de periodograma para INPC (fragmento).

### 2.3 Ejemplos

1. Intentar eliminarla a través de transformaciones matemáticas sencillas (diferencias estacionales).
2. Incluirla en el modelo a través de elementos apropiados para ello.

Por lo general será necesario probar ambos caminos antes de decidirse por uno de ellos.

#### 2.3.2 Ejemplo de datos no estacionarios

Ahora se revisará un ejemplo con características distintas. En el Cuadro 2.1 se muestran los datos la tasa mensual de desocupación abierta en México:<sup>5</sup>

Cuadro 2.1: Tasa de desocupación abierta mensual en México.

Mes	Tasa de Desocupación							
	2002	2003	2004	2005	2006	2007	2008	2009
Enero	3.04	2.72	3.81	3.75	3.53	3.96	4.04	5.00
Febrero	2.64	2.74	3.92	3.93	3.72	4.02	3.91	5.30
Marzo	2.77	2.82	3.86	3.95	3.43	4.01	3.80	4.76
Abril	2.75	2.95	3.58	4.20	3.31	3.60	3.61	5.25
Mayo	2.68	2.74	3.48	4.29	2.88	3.23	3.24	5.31
Junio	2.32	3.27	3.78	3.54	3.33	3.26	3.55	5.17
Julio	2.93	3.52	3.75	3.50	3.98	3.95	4.15	6.12
Agosto	2.77	3.96	4.35	4.05	4.03	3.92	4.15	6.28
Septiembre	3.09	3.87	4.01	3.69	3.98	3.87	4.25	6.41
Octubre	2.66	3.63	3.60	3.57	3.89	3.93	4.11	5.94
Noviembre	2.64	3.80	3.78	3.01	3.58	3.46	4.47	
Diciembre	2.12	2.96	3.04	2.83	3.47	3.40	4.32	

La gráfica de la serie de tiempo que producen estos datos se muestra en la Figura 2.12. Si se observa cuidadosamente esta figura y se revisan las condiciones que se mencionaron como necesarias para que el proceso sea estacionario en sentido amplio, puede destacarse lo siguiente:

- La varianza parece ser constante (la serie es homoscedástica).
- Existe una tendencia creciente bastante clara<sup>6</sup>.
- A simple vista no parece existir variación estacional, aunque puede surgir a lo largo del análisis. Es común que al separar alguno de los componentes, resalte otro que no había sido advertido.
- Existen fluctuaciones aleatorias.

<sup>5</sup>Instituto Nacional de Estadística y Geografía. *Tasa de desocupación abierta mensual en áreas urbanas*. Periodo de enero 2002 a octubre 2009. México. Fuente [29]

<sup>6</sup>Por tratarse del desempleo, esta tendencia es un hecho negativo para el país, que se verá reflejado en el ejemplo de aplicación que se introduce en el Capítulo 7, en el cual se analiza el comportamiento de las remesas que ingresan de Estados Unidos a México.

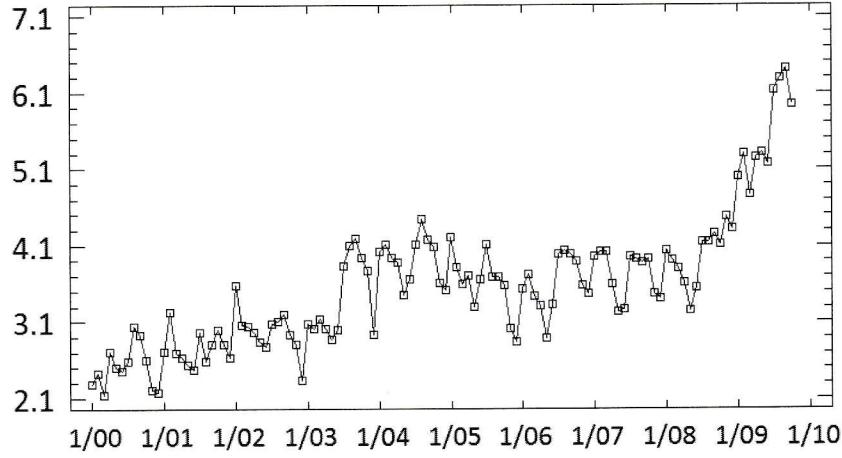


Figura 2.12: Tasa de desocupación abierta mensual.

- Por lo tanto, se concluye que estos datos corresponden a un proceso que no es estacionario en sentido amplio y que, por lo tanto, requerirán modificaciones para modelarse.

Para poder trabajar con este tipo de series es necesario, antes de identificar un posible modelo tentativo, lograr que sean estacionarias. Esto se hará a través de la estabilización de la varianza, la eliminación de la tendencia y la eliminación de la variación estacional, si es que esto último es posible, como se verá más adelante. Por ahora se revisarán únicamente los modelos para series estacionarias.

## 2.4 Modelo lineal general y de forma invertida

Si se considera que una serie de tiempo  $Y_t$  está conformada ya sea por una combinación lineal o por una suma ponderada de choques aleatorios sucesivos:

$e_t, e_{t-1}, e_{t-2}, \dots$ , puede construirse un modelo con las ponderaciones de dichos errores. El **modelo lineal general**, entonces, estará formado por los errores aleatorios ponderados, de la siguiente forma:

$$Y_t = e_t + \Pi_1 e_{t-1} + \Pi_2 e_{t-2} + \Pi_3 e_{t-3} + \dots \quad (2.15)$$

A continuación se demostrará que, bajo ciertas condiciones, este modelo lineal general equivale a lo que se llama **modelo de forma invertida**, en donde la serie de tiempo es una suma ponderada de sus valores anteriores, más el error aleatorio del tiempo  $t$ :

$$Y_t = \Psi_1 Y_{t-1} + \Psi_2 Y_{t-2} + \Psi_3 Y_{t-3} + \dots + e_t \quad (2.16)$$

## 2.4 Modelo lineal general y de forma invertida

---

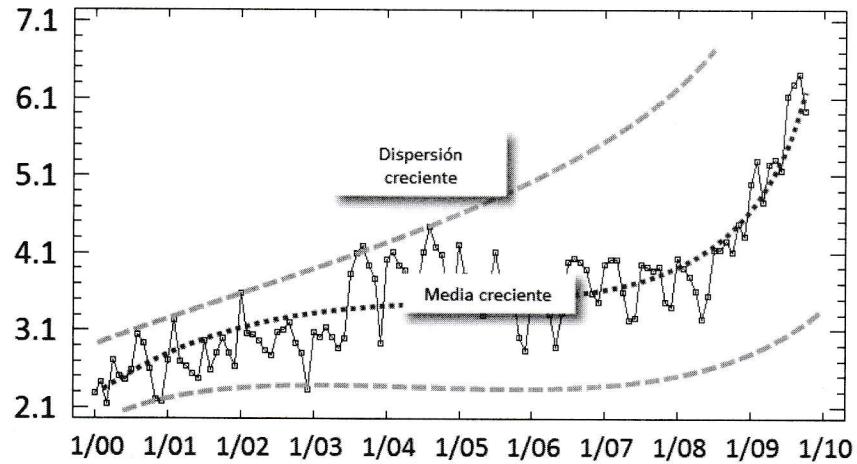


Figura 2.13: Características de la tasa de desocupación.

Por ejemplo, supóngase el modelo finito de forma lineal general:

$$Y_t = e_t + \Pi_1 e_{t-1} \quad (2.17)$$

Donde  $\Pi_2 = \Pi_3 = \dots = 0$ . Si se despeja de 2.17 el valor de  $e_t$  se tendría:

$$e_t = Y_t - \Pi_1 e_{t-1}$$

Si se aplica esta misma ecuación para calcular  $e_{t-1}$ , se verá que:

$$e_{t-1} = Y_{t-1} - \Pi_1 e_{t-2} \quad (2.18)$$

Ahora, si se sustituye la ecuación 2.18 en 2.17, se obtiene:

$$\begin{aligned} Y_t &= e_t + \Pi_1 (Y_{t-1} - \Pi_1 e_{t-2}) \\ Y_t &= \Pi_1 Y_{t-1} + e_t - \Pi_1^2 e_{t-2} \end{aligned}$$

Si se continúa sustituyendo de este modo, se llegará a:

$$Y_t = \Pi_1 Y_{t-1} - \Pi_1^2 Y_{t-2} + \Pi_1^3 Y_{t-3} + \dots + e_t \quad (2.19)$$

Esta última ecuación corresponde al modelo de forma invertida donde:  $\Psi_1 = \Pi_1$ ,  $\Psi_2 = -\Pi_1^2$ ,  $\Psi_3 = \Pi_1^3$ , etcétera. Aquí se observa entonces que, si se toma como punto de partida el modelo lineal general *finito* de la ecuación 2.17, éste será equivalente a la expresión del modelo de forma invertida pero con un *infinito* número de términos. Por supuesto, para que esta equivalencia tenga sentido, se esperaría que  $|\Pi_1| < 1$ , de manera que los términos más lejanos en el tiempo tengan menor peso en el modelo.

De la misma forma es posible demostrar que un modelo finito de forma invertida con la ecuación 2.20 puede transformarse en un modelo lineal general infinito, tan solo al sustituir los valores sucesivos de  $Y_{t-1}, Y_{t-2}, \dots$  y dejar la ecuación en función de los errores aleatorios (véase la Actividad 1 de este capítulo).

$$Y_t = \Psi_1 Y_{t-1} + e_t \quad (2.20)$$

A partir de las demostraciones anteriores y de las ecuaciones 2.15 y 2.16 se concluye que:

1. Los modelos lineal general y de forma invertida no cumplen con el principio de parsimonia ya que tienen un número infinito de términos, por lo cual resultarían poco prácticos.
2. Es lógico pensar que, a medida que las variables se alejan en el tiempo sus coeficientes, es decir, sus ponderaciones, vayan disminuyendo en su valor absoluto y acercándose cada vez más a cero. Es decir, que los valores más lejanos en el tiempo tengan menos influencia en el modelo matemático.
3. Puede demostrarse que, bajo ciertas condiciones, el modelo lineal general puede cambiarse a la forma invertida y viceversa.
4. Puede notarse que un modelo finito de la forma lineal general se transforma en un modelo infinito de forma invertida y viceversa. De acuerdo con el principio de parsimonia, se preferirá siempre aquel modelo que tenga menos variables.

Todo lo anterior lleva a pensar que es factible plantear modelos finitos, cuyo uso será sencillo y práctico, que reflejen de manera apropiada el comportamiento de algunas series de tiempo y permitan obtener pronósticos con bastante precisión. En las secciones 2.6, 2.7 y 2.8 se describen estos modelos.

## 2.5 Función de autocorrelación parcial

Antes de continuar, es preciso definir un instrumento matemático adicional que será necesario para identificar los diversos modelos que pueden ajustarse a un conjunto de datos. Se ha dicho que un modelo lineal general puede transformarse en un modelo de forma invertida y viceversa. Si se toma ahora como punto de partida el modelo de forma invertida, es posible definir la **autocorrelación parcial** (*partial autocorrelation*) de orden  $k$  como la contribución marginal o el peso de cada nueva variable autorregresiva  $Y_{t-k}$  que se agrega al modelo. La autocorrelación parcial de orden  $k$  se suele denotar por  $\rho_{kk}$ .

De acuerdo con la definición anterior,  $\rho_{11}$  será el coeficiente  $\Psi_1$  en el modelo  $Y_t = \Psi_1 Y_{t-1} + e_t$ ;  $\rho_{22}$  será el coeficiente  $\Psi_2$  en el modelo  $Y_t = \Psi_1 Y_{t-1} + \Psi_2 Y_{t-2} + e_t$ ;  $\rho_{33}$  será el coeficiente  $\Psi_3$  en el modelo  $Y_t = \Psi_1 Y_{t-1} + \Psi_2 Y_{t-2} + \Psi_3 Y_{t-3} + e_t$ , y así sucesivamente.

La función de autocorrelación parcial o PACF (*Partial AutoCorrelation Function*) será la gráfica de los valores de  $\rho_{kk}$  contra los valores de los intervalos  $k$ , para  $k = 1, 2, \dots$

El cálculo de la PACF se hace a través de las ecuaciones resultantes para la ACF. Para ello, se parte del modelo en forma invertida (ecuación 2.16), pero con sólo los primeros  $p$  coeficientes distintos de cero, esto es:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t$$

La ecuación anterior se multiplica por  $Y_{t-k}$ , para después tomar el valor esperado de ambos lados. Si se supone que la media es constante e igual a cero (lo cual podría obtenerse con facilidad si se resta  $\bar{Y}$  a cada observación) se tendrá, para  $k \neq t$ :

$$\begin{aligned} \gamma_k &= \text{Autocov}(Y_t, Y_{t-k}) \\ &= E(Y_t Y_{t-k}) \\ &= E((\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p}) Y_{t-k}) \\ &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p \end{aligned} \quad (2.21)$$

Si la ecuación 2.21 se divide entre  $\gamma_0$  se obtiene:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad (2.22)$$

Ahora, la ecuación 2.22 puede plantearse para  $p = 1, 2, \dots$ . De la ecuación para  $p = 1$  se despeja el valor de  $\phi_1$  que será igual a  $\rho_{11}$ ; de la ecuación para  $p = 2$  se despeja el valor de  $\phi_2$  que será igual a  $\rho_{22}$ , etcétera.

Así, para obtener  $\rho_{11}$  se despeja  $\phi_1$  de la ecuación 2.22 con  $p = 1$  y  $k = 1$ :

$$\begin{aligned} \rho_1 &= \phi_1 \rho_0 \\ \rho_1 &= \rho_{11}(1) \\ \Rightarrow \rho_{11} &= \rho_1 \end{aligned} \quad (2.23)$$

Entonces, la **autocorrelación parcial estimada o muestral** de orden uno será la ecuación 2.24.

$$r_{11} = \hat{\rho}_{11} = r_1 \quad (2.24)$$

De la misma forma, para obtener  $\rho_{22}$  se despeja  $\phi_2$  de las ecuaciones simultáneas 2.22 con  $p = 2$  y  $k = 1, 2$ :

$$\begin{aligned} \rho_1 &= \phi_1 \rho_0 + \phi_2 \rho_1 \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \rho_0 \end{aligned} \quad (2.25)$$

Entonces:

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \end{aligned} \quad (2.26)$$

De donde:

$$\rho_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \quad (2.27)$$

Con la ecuación 2.27 puede plantearse el estimador correspondiente:

$$r_{22} = \hat{\rho}_{22} = \frac{r_2 - r_1^2}{1 - r_1^2} \quad (2.28)$$

Y así sucesivamente. Por último, puede verse que la PACF se estimará como  $\hat{\rho}_{kk} = r_{kk}$  a través de los estimadores  $r_k$ .

## 2.6 Modelos autorregresivos

Si se toma como punto de partida el modelo de forma invertida, se puede construir un modelo más económico que únicamente contenga algunas de las variables históricas ponderadas, más un error aleatorio. El modelo de la ecuación 2.29 se llama **modelo autorregresivo (AutoRegressive Model)** y se denota por **AR( $p$ )** puesto que contiene  $p$  variables históricas autorregresivas.

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \dots + \phi_p Y_{t-p} + e_t \quad (2.29)$$

Donde las  $e_t$  son **ruido blanco**, es decir, variables aleatorias independientes que siguen una distribución Normal con media cero y varianza constante  $\sigma^2$ .

Así por ejemplo el modelo AR(1) será:

$$Y_t = \phi_1 Y_{t-1} + e_t \quad (2.30)$$

Mientras que el modelo AR(2) es:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t \quad (2.31)$$

Ahora cabe preguntarse: ¿Estos modelos son estacionarios en sentido amplio? A continuación se revisarán los modelos AR(1) y AR(2) para después generalizar esta respuesta.

### 2.6.1 Modelo AR(1)

Para verificar si el modelo AR(1) es estacionario en sentido amplio, se revisan las tres condiciones establecidas en la Sección 2.1.1:

Sí la serie de tiempo  $Y_t$  es estacionaria y su media es constante, debería cumplirse que  $E(Y_t) = E(Y_{t-1}) = \mu$  y por lo tanto:

$$\begin{aligned}\mu &= \phi_1\mu \\ (1 - \phi_1)\mu &= 0 \\ \mu &= \frac{0}{1 - \phi_1} \\ \mu &= 0\end{aligned}\tag{2.32}$$

Así, la media será cero siempre y cuando  $\phi_1 \neq 1$ . Puede notarse que también es posible agregar una constante al modelo y en ese caso, la media sería igual a dicha constante dividida entre  $(1 - \phi_1)$ .

2. La varianza debe ser constante:

$$\begin{aligned}\text{Var}(Y_t) &= \text{Var}(\phi_1 Y_{t-1} + e_t) \\ &= \text{Var}(\phi_1 Y_{t-1}) + \text{Var}(e_t) \\ &= \phi_1^2 \text{Var}(Y_{t-1}) + \sigma^2\end{aligned}$$

Si la serie de tiempo es estacionaria y la varianza es constante, debería cumplirse que  $\text{Var}(Y_t) = \text{Var}(Y_{t-1}) = \gamma_0$  y por lo tanto:

$$\begin{aligned}\gamma_0 &= \phi_1^2 \gamma_0 + \sigma^2 \\ (1 - \phi_1^2) \gamma_0 &= \sigma^2 \\ \gamma_0 &= \frac{\sigma^2}{1 - \phi_1^2}\end{aligned}\tag{2.33}$$

Puesto que la varianza debe ser positiva ( $\gamma_0 > 0$ ), será constante siempre y cuando se cumpla que:

$$\begin{aligned}1 - \phi_1^2 &> 0 \\ \phi_1^2 &< 1 \\ |\phi_1| &< 1\end{aligned}\tag{2.34}$$

Así, de acuerdo con la ecuación 2.34 el modelo AR(1) tiene una **condición de estacionaridad** que incluye la condición que se había planteado anteriormente para la media constante.

1. La media debe ser constante:

$$\begin{aligned} E(Y_t) &= E(\phi_1 Y_{t-1} + e_t) \\ &= E(\phi_1 Y_{t-1}) + E(e_t) \\ &= \phi_1 E(Y_{t-1}) \end{aligned}$$

Si la serie de tiempo  $Y_t$  es estacionaria y la media es constante, debería cumplirse que  $E(Y_t) = E(Y_{t-1}) = \mu$  y por lo tanto:

$$\begin{aligned} \mu &= \phi_1 \mu \\ (1 - \phi_1) \mu &= 0 \\ \mu &= \frac{0}{1 - \phi_1} \\ \mu &= 0 \end{aligned} \tag{2.32}$$

Así, la media será cero siempre y cuando  $\phi_1 \neq 1$ . Puede notarse que también es posible agregar una constante al modelo y en ese caso, la media sería igual a dicha constante dividida entre  $(1 - \phi_1)$ .

2. La varianza debe ser constante:

$$\begin{aligned} \text{Var}(Y_t) &= \text{Var}(\phi_1 Y_{t-1} + e_t) \\ &= \text{Var}(\phi_1 Y_{t-1}) + \text{Var}(e_t) \\ &= \phi_1^2 \text{Var}(Y_{t-1}) + \sigma^2 \end{aligned}$$

Si la serie de tiempo es estacionaria y la varianza es constante, debería cumplirse que  $\text{Var}(Y_t) = \text{Var}(Y_{t-1}) = \gamma_0$  y por lo tanto:

$$\begin{aligned} \gamma_0 &= \phi_1^2 \gamma_0 + \sigma^2 \\ (1 - \phi_1^2) \gamma_0 &= \sigma^2 \\ \gamma_0 &= \frac{\sigma^2}{1 - \phi_1^2} \end{aligned} \tag{2.33}$$

Puesto que la varianza debe ser positiva ( $\gamma_0 > 0$ ), será constante siempre y cuando se cumpla que:

$$\begin{aligned} 1 - \phi_1^2 &> 0 \\ \phi_1^2 &< 1 \\ |\phi_1| &< 1 \end{aligned} \tag{2.34}$$

Así, de acuerdo con la ecuación 2.34 el modelo AR(1) tiene una **condición de estacionaridad** que incluye la condición que se había planteado anteriormente para la media constante.

3. La función de autocorrelación ACF debe depender únicamente del intervalo  $k$  entre dos variables y no de su posición en el tiempo  $t$ . Si se revisa cada una de las autocovarianzas considerando que la media es cero, se tendrá para  $k = 1$ :

$$\begin{aligned}\gamma_1 &= \text{Autocov}(Y_t, Y_{t-1}) \\ &= E[(\phi_1 Y_{t-1} + e_t) Y_{t-1}] \\ &= \phi_1 E[Y_{t-1}^2] + E[e_t Y_{t-1}] \\ &= \phi_1 \gamma_0\end{aligned}\tag{2.35}$$

Para  $k = 2$ :

$$\begin{aligned}\gamma_2 &= \text{Autocov}(Y_t, Y_{t-2}) \\ &= E[(\phi_1 Y_{t-1} + e_t) Y_{t-2}] \\ &= \phi_1 E[Y_{t-1} Y_{t-2}] + E[e_t Y_{t-2}] \\ &= \phi_1 \gamma_1 \\ &= \phi_1^2 \gamma_0\end{aligned}\tag{2.36}$$

Para  $k = 3$ :

$$\begin{aligned}\gamma_3 &= \text{Autocov}(Y_t, Y_{t-3}) \\ &= E[(\phi_1 Y_{t-1} + e_t) Y_{t-3}] \\ &= \phi_1 E[Y_{t-1} Y_{t-3}] + E[e_t Y_{t-3}] \\ &= \phi_1 \gamma_2 \\ &= \phi_1^3 \gamma_0\end{aligned}\tag{2.37}$$

Y así sucesivamente, de manera que para  $k = 1, 2, 3, \dots$ :

$$\begin{aligned}\gamma_k &= \text{Autocov}(Y_t, Y_{t-k}) \\ &= E[(\phi_1 Y_{t-1} + e_t) Y_{t-k}] \\ &= \phi_1 E[Y_{t-1} Y_{t-k}] + E[e_t Y_{t-k}] \\ &= \phi_1 \gamma_{k-1} \\ &= \phi_1^k \gamma_0\end{aligned}\tag{2.38}$$

A partir de los resultados anteriores, se concluye que la función de autocorrelación o ACF de un modelo AR(1) estacionario será:

$$\begin{aligned}\rho_k &= \frac{\gamma_k}{\gamma_0} \\ &= \phi_1^k\end{aligned}\tag{2.39}$$

Puesto que la condición de estacionaridad implica que  $|\phi_1| < 1$ , la gráfica de la ACF tendrá dos formas posibles, mismas que se ejemplifican en las Figuras 2.14 y 2.15. Como puede verse, la ACF de un modelo AR(1) será siempre decreciente infinita.

## 2.6 Modelos autorregresivos

---

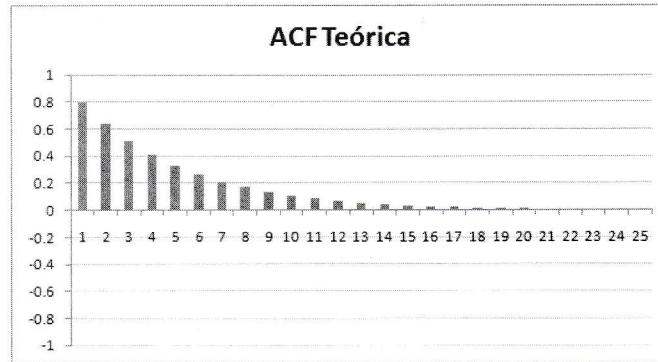


Figura 2.14: ACF de un modelo AR(1) con coeficiente positivo.

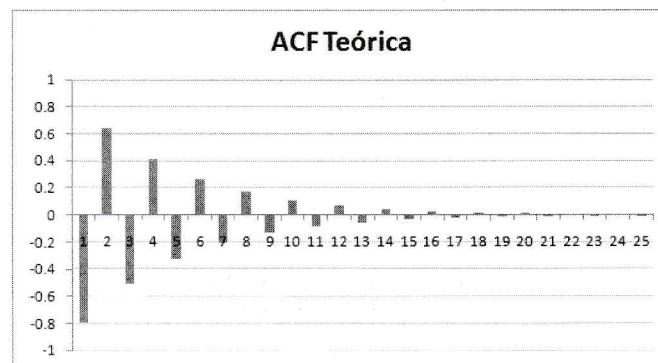


Figura 2.15: ACF de un modelo AR(1) con coeficiente negativo.

Por otro lado, es posible analizar la función de autocorrelación parcial o PACF de este modelo. Puesto que la única variable autorregresiva que tiene contribución en este caso es  $Y_{t-1}$  y dicha contribución tiene el valor del coeficiente  $\phi_1$ , se tiene que la PACF se trunca en el valor de  $k = 1$  y es cero para el resto de los casos. Así, las Figuras 2.16 y 2.17 muestran ejemplos de los dos casos posibles.

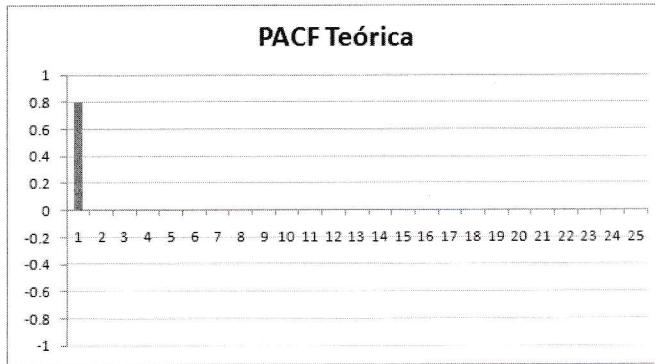


Figura 2.16: PACF del modelo AR(1) con coeficiente positivo.

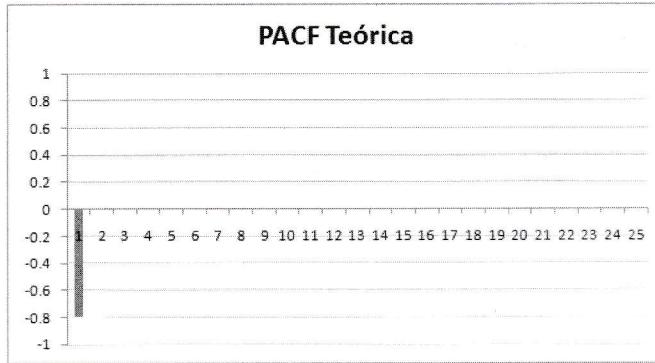


Figura 2.17: PACF del modelo AR(1) con coeficiente negativo.

### 2.6.2 Modelo AR(2)

Al igual que en la sección anterior, puede verificarse que el modelo AR(2):

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t \quad (2.40)$$

## 2.6 Modelos autorregresivos

---

Debe cumplir las siguientes condiciones en sus parámetros para ser estacionario [12, 58]:

$$\begin{aligned}\phi_1 + \phi_2 &< 1 \\ \phi_2 - \phi_1 &< 1 \\ |\phi_2| &< 1\end{aligned}\tag{2.41}$$

Si se cumplen estas condiciones, la media del proceso AR(2) será:

$$E(Y_t) = \mu = 0\tag{2.42}$$

Su varianza se calcula como:

$$\begin{aligned}\gamma_0 &= \text{Var}(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t) \\ &= \text{Var}(\phi_1 Y_{t-1}) + \text{Var}(\phi_2 Y_{t-2}) + \text{Var}(e_t) \\ &\quad + 2\text{Autocov}(\phi_1 Y_{t-1}, \phi_2 Y_{t-2}) \\ &= \phi_1^2 \gamma_0 + \phi_2^2 \gamma_0 + 2\phi_1 \phi_2 \gamma_1 + \sigma^2 \\ (1 - \phi_1^2 - \phi_2^2)\gamma_0 &= 2\phi_1 \phi_2 \gamma_1 + \sigma^2 \\ \Rightarrow \gamma_0 &= \frac{2\phi_1 \phi_2 \gamma_1 + \sigma^2}{1 - \phi_1^2 - \phi_2^2}\end{aligned}\tag{2.43}$$

Las autocovarianzas pueden calcularse tomando el modelo 2.31, multiplicándolo por  $Y_{t-k}$  y obteniendo la esperanza. De ahí se obtiene la ACF al dividir entre  $\gamma_0$ , con lo cual se tendrá la ecuación recursiva:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}\tag{2.44}$$

De aquí se concluye que para  $k = 1$ :

$$\begin{aligned}\rho_1 &= \phi_1 \rho_0 + \phi_2 \rho_1 \\ &= \phi_1 + \phi_2 \rho_1 \\ (1 - \phi_2)\rho_1 &= \phi_1 \\ \rho_1 &= \frac{\phi_1}{1 - \phi_2}\end{aligned}\tag{2.45}$$

Para  $k = 2$ :

$$\begin{aligned}\rho_2 &= \phi_1 \rho_1 + \phi_2 \rho_0 \\ &= \phi_1 \rho_1 + \phi_2 \\ &= \frac{\phi_1^2}{1 - \phi_2} + \phi_2\end{aligned}\tag{2.46}$$

De aquí en adelante se usa la ecuación recursiva para  $k = 3, 4, 5, \dots$ :

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}\tag{2.47}$$

Así, dadas las condiciones de estacionariedad, la gráfica de la ACF tendrá algunas variantes, pero siempre será decreciente infinita.

Por otro lado, la PACF de este tipo de modelos se trunca en  $k = 2$  puesto que únicamente contribuyen al modelo  $Y_{t-1}$  y  $Y_{t-2}$  con sus respectivos coeficientes.

### 2.6.3 Modelo AR( $p$ )

La ecuación general del modelo AR( $p$ ) es:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t \quad (2.48)$$

La función de media de este modelo será:

$$\begin{aligned} E(Y_t) &= E(\phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t) \\ \mu &= \phi_1 \mu + \dots + \phi_p \mu \\ &= \frac{0}{1 - \phi_1 - \dots - \phi_p} \\ &= 0 \text{ si } \phi_1 + \dots + \phi_p \neq 1 \end{aligned} \quad (2.49)$$

Ahora, para obtener la varianza y las covarianzas, se procede como sigue:

$$\begin{aligned} E(Y_t Y_{t-k}) &= E(\phi_1 Y_{t-1} Y_{t-k} + \dots + \phi_p Y_{t-p} Y_{t-k} + e_t Y_{t-k}) \\ \gamma_k &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} \dots + \phi_p \gamma_{k-p} + E(e_t Y_{t-k}) \end{aligned} \quad (2.50)$$

Entonces, si  $k = 0$ , se tendrá la ecuación para la varianza:

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 \dots + \phi_p \gamma_p + \sigma^2 \quad (2.51)$$

En otro caso, si  $k \neq 0$ , se trata de la ecuación para las autocovarianzas:

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} \dots + \phi_p \gamma_{k-p} \quad (2.52)$$

Ahora, si la ecuación 2.52 se divide entre  $\gamma_0$ , se obtiene:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \dots + \phi_p \rho_{k-p} \quad (2.53)$$

Con ello puede plantearse un conjunto de  $p$  ecuaciones simultáneas:

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \dots + \phi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} \dots + \phi_p \end{aligned} \quad (2.54)$$

Las ecuaciones 2.54 se conocen como **Ecuaciones de Yule-Walker** y pueden usarse también para la estimación, si se sustituye  $\rho_k$  por  $r_k$ .

De acuerdo con lo que se ha visto hasta aquí, para el modelo AR( $p$ ) puede generalizarse que:

- Tendrá condiciones de estacionaridad para sus parámetros.
- La ACF es decreciente infinita, con variantes que dependen de los signos y magnitudes de los coeficientes.
- La PACF se trunca en  $k = p$ .

## 2.7 Modelos de medias móviles

Si se toma como punto de partida el modelo lineal general, se puede construir un modelo más económico que únicamente contenga algunos de los errores aleatorios ponderados. El modelo de la ecuación 2.55 es el **modelo de medias móviles** y se denota por  $\text{MA}(q)$  puesto que contiene  $q$  variables históricas autorregresivas (*Moving Averages Model*). Este modelo se acostumbra formular con signos negativos en los coeficientes, con objeto de facilitar el uso de la notación abreviada que se revisará en la Sección 2.9. Únicamente deberá recordarse que si el coeficiente aparece en el modelo con valor negativo, esto implica que su valor es positivo y viceversa.

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3} + \dots - \theta_q e_{t-q} \quad (2.55)$$

Donde las  $e_t$  son ruido blanco, es decir, variables aleatorias independientes que siguen una distribución Normal con media cero y varianza constante  $\sigma^2$ .

Así por ejemplo el modelo MA(1) será:

$$Y_t = e_t - \theta_1 e_{t-1} \quad (2.56)$$

Mientras que el modelo MA(2) es:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \quad (2.57)$$

Ahora cabe preguntarse: ¿estos modelos son estacionarios en sentido amplio? A continuación se revisarán los modelos MA(1) y MA(2) para después generalizar esta respuesta.

### 2.7.1 Modelo MA(1)

Para verificar si el modelo MA(1) es estacionario en sentido amplio, se revisan las tres condiciones establecidas en la Sección 2.1.1:

1. La media debe ser constante:

$$\begin{aligned} E(Y_t) &= E(e_t - \theta_1 e_{t-1}) \\ &= E(e_t) - E(\theta_1 e_{t-1}) \\ &= 0 \end{aligned} \quad (2.58)$$

Por lo cual la media es constante independientemente del valor del coeficiente  $\theta_1$ .

2. La varianza debe ser constante:

$$\begin{aligned} \text{Var}(Y_t) &= \text{Var}(e_t - \theta_1 e_{t-1}) \\ &= \text{Var}(e_t) + \text{Var}(-\theta_1 e_{t-1}) \\ &= \sigma^2 + \theta_1^2 \sigma^2 \\ &= (1 + \theta_1^2) \sigma^2 \end{aligned} \quad (2.59)$$

Por lo tanto, la varianza es constante (y positiva) independientemente del valor de  $\theta_1$ . Esto significa que el modelo MA(1) **no tiene condiciones de estacionaridad**.

3. La función de autocorrelación ACF debe depender únicamente del intervalo  $k$  entre dos variables y no de su posición en el tiempo  $t$ . Si se revisa cada una de las autocovarianzas considerando que la media es cero, se tendrá para  $k = 1$ :

$$\begin{aligned}\gamma_1 &= \text{Autocov}(Y_t, Y_{t-1}) \\ &= E[(e_t - \theta_1 e_{t-1})(e_{t-1} - \theta_1 e_{t-2})] \\ &= -\theta_1 E[e_{t-1}^2] \\ &= -\theta_1 \sigma^2\end{aligned}\tag{2.60}$$

Para  $k = 2$ :

$$\begin{aligned}\gamma_2 &= \text{Autocov}(Y_t, Y_{t-2}) \\ &= E[(e_t - \theta_1 e_{t-1})(e_{t-2} - \theta_1 e_{t-3})] \\ &= 0\end{aligned}\tag{2.61}$$

Para  $k = 3, 4, 5, \dots$ :

$$\gamma_k = 0\tag{2.62}$$

A partir de los resultados anteriores, se concluye que la función de autocorrelación o ACF de un modelo MA(1) será:

$$\begin{aligned}\rho_k &= \frac{\gamma_k}{\gamma_0} \\ &= \begin{cases} \frac{-\theta_1}{1 + \theta_1^2} & \text{para } k = 1 \\ 0 & \text{para } k = 2, 3, 4, \dots \end{cases}\end{aligned}\tag{2.63}$$

Si bien se ha concluido que el modelo MA(1) no tiene condiciones de estacionaridad, recuérdese que este modelo equivale a un modelo MA( $\infty$ ):

$$\begin{aligned}Y_t &= e_t - \theta_1 e_{t-1} \\ e_t &= Y_t + \theta_1 e_{t-1} \\ e_{t-1} &= Y_{t-1} + \theta_1 e_{t-2}\end{aligned}\tag{2.64}$$

$$\begin{aligned}\Rightarrow Y_t &= e_t - \theta_1(Y_{t-1} + \theta_1 e_{t-2}) \\ &= e_t - \theta_1 Y_{t-1} - \theta_1^2 e_{t-2} \\ &= e_t - \theta_1 Y_{t-1} - \theta_1^2 Y_{t-2} - \theta_1^3 Y_{t-3} + \dots\end{aligned}\tag{2.65}$$

## 2.7 Modelos de medias móviles

---

Por lo tanto, para que la ecuación anterior tenga sentido, se requiere imponer ahora una **condición de invertibilidad**, la cual implica que  $|\theta_1| < 1$ . Asimismo, de la ecuación 2.65 se desprenden los valores de la PACF que representan las contribuciones parciales de cada  $Y_{t-k}$ . Así pues tanto la gráfica de la ACF como la de la PACF tendrán dos formas posibles, mismas que se ejemplifican en la Figuras 2.18, 2.19, 2.20 y 2.21, respectivamente. Como puede verse, la ACF de un modelo MA(1) se trunca en  $k = 1$  mientras que su PACF es decreciente infinita.

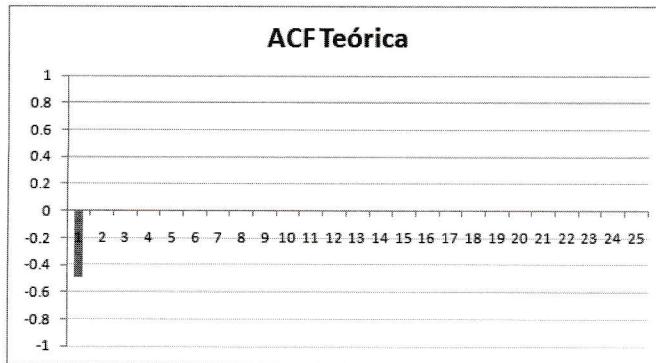


Figura 2.18: ACF de un modelo MA(1) con coeficiente positivo.

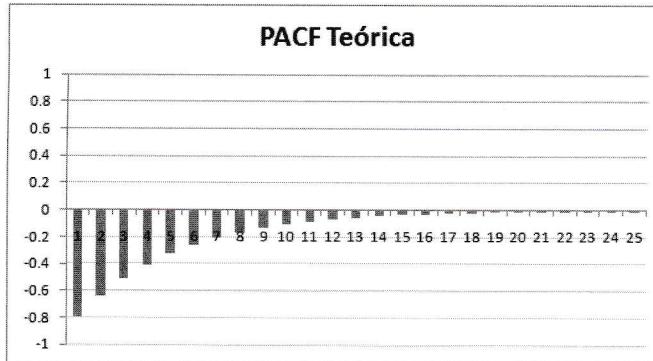


Figura 2.19: PACF de un modelo MA(1) con coeficiente positivo.

### 2.7.2 Modelo MA(2)

El análisis anterior para verificar la estacionaridad en sentido amplio puede hacerse también para el modelo MA(2), de la siguiente forma:

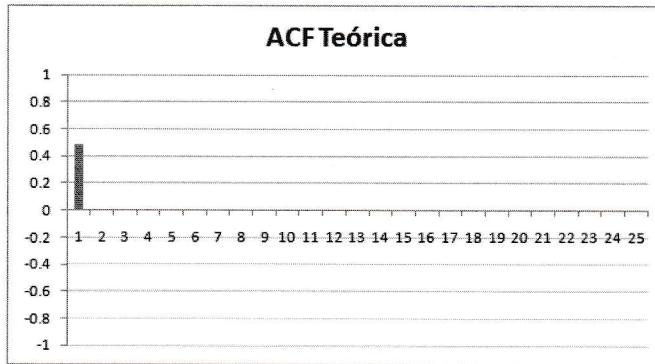


Figura 2.20: ACF de un modelo MA(1) con coeficiente negativo.

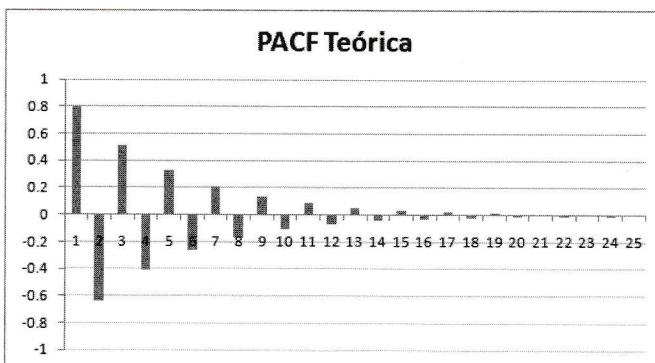


Figura 2.21: PACF de un modelo MA(1) con coeficiente negativo.

1. La función de media debe ser constante:

$$\begin{aligned} E(Y_t) &= E(e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}) \\ &= 0 \end{aligned} \quad (2.66)$$

2. La función de varianza debe ser constante:

$$\begin{aligned} \text{Var}(Y_t) &= \text{Var}(e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}) \\ \gamma_0 &= (1 + \theta_1^2 + \theta_2^2) \end{aligned} \quad (2.67)$$

3. La función de autocorrelación debe depender de  $k$  y no de  $t$ :

$$\begin{aligned} \rho_k &= \frac{\gamma_k}{\gamma_0} \\ &= \begin{cases} \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} & \text{para } k=1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{para } k=2 \\ 0 & k>2 \end{cases} \end{aligned} \quad (2.68)$$

Como puede verse, el modelo MA(2) carece de condiciones de estacionaridad; sin embargo, por ser equivalente a un modelo AR( $\infty$ ), tiene un conjunto de condiciones de invertibilidad [12, 58]:

$$\begin{aligned} \theta_1 + \theta_2 &< 1 \\ \theta_2 - \theta_1 &< 1 \\ |\theta_2| &< 1 \end{aligned} \quad (2.69)$$

### 2.7.3 Modelo MA( $q$ )

De acuerdo con lo que se ha visto hasta aquí, puede generalizarse para el modelo MA( $q$ ) lo siguiente:

- Carece de condiciones de estacionaridad para sus parámetros, es decir, *siempre* es estacionario.
- Tiene condiciones de invertibilidad para sus parámetros.
- La ACF se trunca en  $k = q$ .
- La PACF es decreciente infinita, con variantes que dependen de los signos y magnitudes de los coeficientes.

## 2.8 Modelos mezclados

Por supuesto, es posible mezclar los modelos AR( $p$ ) con los modelos MA( $q$ ), con lo cual se obtiene una clase más amplia, los modelos ARMA( $p, q$ ) que contienen  $p + q$  parámetros, como puede verse en la ecuación 2.70. Estos modelos tendrán tanto condiciones de estacionaridad como condiciones de invertibilidad.

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (2.70)$$

Además, es sencillo verificar que tanto su ACF como su PACF serán decrecientes infinitas.

### 2.8.1 Modelo ARMA(1,1)

El modelo mezclado con menos parámetros es el ARMA(1,1) que se expresa con la ecuación 2.71.

$$Y_t = \phi_1 Y_{t-1} + e_t - \theta_1 e_{t-1} \quad (2.71)$$

Sus características son las siguientes:

1. Función de media constante:

$$\begin{aligned} E(Y_t) &= E(\phi_1 Y_{t-1} + e_t - \theta_1 e_{t-1}) \\ &= 0 \end{aligned}$$

2. Función de varianza constante:

$$\begin{aligned} \text{Var}(Y_t) &= \text{Var}(\phi_1 Y_{t-1} + e_t - \theta_1 e_{t-1}) \\ \gamma_0 &= \frac{(1 + \theta_1^2 - 2\phi_1\theta_1)\sigma^2}{1 - \phi_1^2} \end{aligned}$$

3. Función de autocorrelación que depende de  $k$  y no de  $t$ :

$$\rho_k = \begin{cases} 1 & \text{para } k = 0 \\ \frac{(\phi_1 - \theta_1)(1 - \phi_1\theta_1)}{1 + \theta_1^2 - 2\phi_1\theta_1} & \text{para } k = 1 \\ \phi_1\rho_{k-1} & \text{para } k \geq 2 \end{cases}$$

Puede verificarse con facilidad que el modelo ARMA(1,1) equivale a un modelo AR( $\infty$ ), por lo cual su PACF será decreciente infinita. A su vez, puede mostrarse también que equivale a un modelo MA( $\infty$ ). Por ello, este modelo tiene una condición de estacionaridad ( $|\phi_1| < 1$ ) y una condición de invertibilidad ( $|\theta_1| < 1$ ).

### 2.8.2 Modelo ARMA(p,q)

El modelo **ARMA**( $p, q$ ) se forma al mezclar el modelo AR( $p$ ) con el modelo MA( $q$ ). Contiene  $p + q$  parámetros, correspondientes a  $p$  variables autorregresivas y  $q + 1$  errores aleatorios:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$$

Estos modelos, como puede suponerse tendrán tanto condiciones de estacionaridad como de invertibilidad, puesto que poseen parte AR y parte MA. De la misma forma, su ACF será decreciente infinita, con un efecto de la parte MA que termina en  $k = q$  y su PACF será también decreciente infinita, con un efecto de la parte AR que desaparece después de  $k = p$ .

## 2.9 Notación abreviada

Puesto que los modelos de tipo ARMA requieren de una expresión relativamente extensa, es posible desarrollar para ellos una **notación abreviada**. Con este fin se define en primer lugar un **operador de salto hacia atrás** (*backward shift operator*)  $B$  de la siguiente forma:

$$\begin{aligned} BY_t &= Y_{t-1} \\ B^2Y_t &= Y_{t-2} \\ Be_t &= e_{t-1} \\ B^2e_t &= e_{t-2} \end{aligned} \tag{2.72}$$

Por ejemplo, si se utiliza este operador, el modelo AR(1) podría expresarse como:

$$Y_t = \phi_1 BY_t + e_t$$

Lo cual permite reorganizar la ecuación como:

$$(1 - \phi_1 B)Y_t = e_t$$

Así, el modelo general ARMA( $p, q$ ) se expresaría como:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)Y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)e_t$$

Si se definen ahora dos nuevas variables de notación:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \tag{2.73}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \tag{2.74}$$

Entonces los modelos generales ARMA( $p, q$ ) pueden escribirse de forma abreviada y más elegante<sup>7</sup> como:

$$\phi(B)Y_t = \theta(B)e_t \tag{2.75}$$

<sup>7</sup>Esta notación, sin embargo, es menos intuitiva para su comprensión.

Cuadro 2.2: Dualidad y resumen de modelos ARMA					
Modelo	Equivale a	CE	CI	ACF	PACF
AR( $p$ )	MA( $\infty$ )	Sí	No	Decreciente infinita	Se trunca en $k = p$
MA( $q$ )	AR( $\infty$ )	No	Sí	Se trunca en $k = q$	Decreciente infinita
ARMA( $p, q$ )	AR( $\infty$ ) MA( $\infty$ )	Sí	Sí	Decreciente infinita	Decreciente infinita

## 2.10 Dualidad de los modelos y resumen

Como se ha visto en las secciones anteriores, existe una especie de **dualidad** entre los procesos AR y los procesos MA. En el Cuadro 2.2 se resume la información de todos los modelos explicados hasta el momento (CE=Condiciones de estacionaridad; CI=Condiciones de invertibilidad):

## 2.11 Cálculos y gráficos con *Statgraphics*

Como se ha visto, existen una serie de cálculos y gráficos que resultarán de gran utilidad para analizar una serie de tiempo, identificar sus componentes y plantear algún modelo que permita pronosticarla de manera adecuada. Sin embargo, también resulta evidente que este trabajo será laborioso, sobre todo en la medida en que se cuente con datos abundantes. Por ello, se han desarrollado aplicaciones de software para esta metodología, como las que ofrecen *Statgraphics*, *Statgraphics Online*, SPSS, SAS o R, entre otros.

A continuación se muestra la forma de efectuar estos procedimientos en *Statgraphics*. Cabe señalar que en todas las aplicaciones esto se hace de manera similar. Lo importante es saber elegir e interpretar cada herramienta dentro del paquete estadístico.

Por ejemplo, para obtener la gráfica original de una serie de tiempo, pueden introducirse los datos directamente en *Statgraphics*, en una hoja con renglones y columnas, a través de la cuadrícula que presenta para este fin (Figura 2.22). Sin embargo, también es posible utilizar datos ya capturados en otros formatos, tales como *Excel* o el portapapeles. Para ello se puede utilizar la función copia-y-pegar o bien hacer la lectura indicando cuál es el formato apropiado.

Una vez que se ha capturado una serie de tiempo, es posible obtener su gráfica, así como los valores y gráficos de ACF, PACF, periodograma y periodograma integrado. Una forma de hacer esto es utilizar la opción **Describe / Time Series / Descriptive Methods**, como se ve en la Figura 2.23.

## 2.11 Cálculos y gráficos con Statgraphics

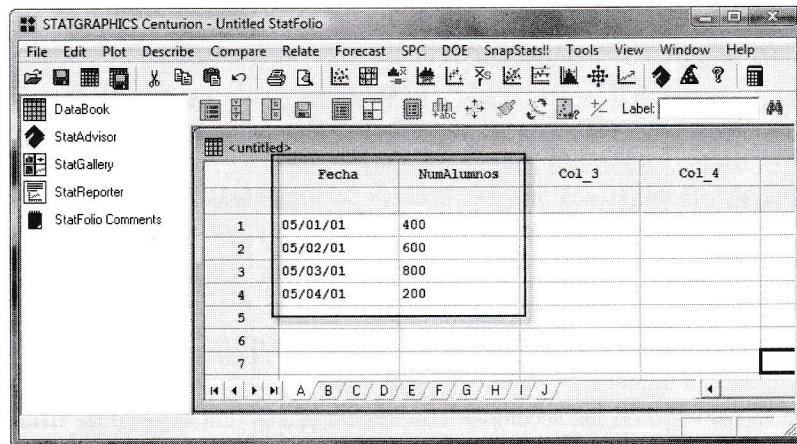


Figura 2.22: Pantalla de captura en *Statgraphics*.

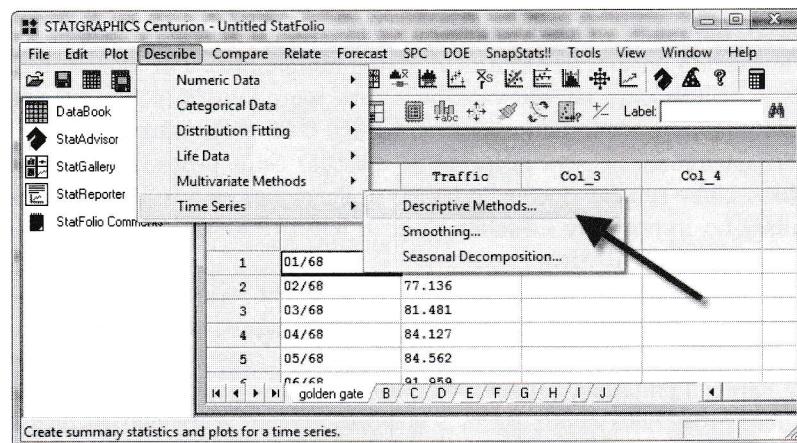


Figura 2.23: Elegir descripción de la serie de tiempo.

Hecho esto, debe seleccionarse la columna en la cual se encuentran ubicados los datos, de los cuales se obtendrán las diversas gráficas y cálculos. Para ello se marca el nombre de la serie y después el botón con una flecha. Es importante indicar cuáles son las unidades de tiempo (días, meses, años, etcétera), así como el punto inicial del primer dato. La Figura 2.24 muestra un ejemplo.

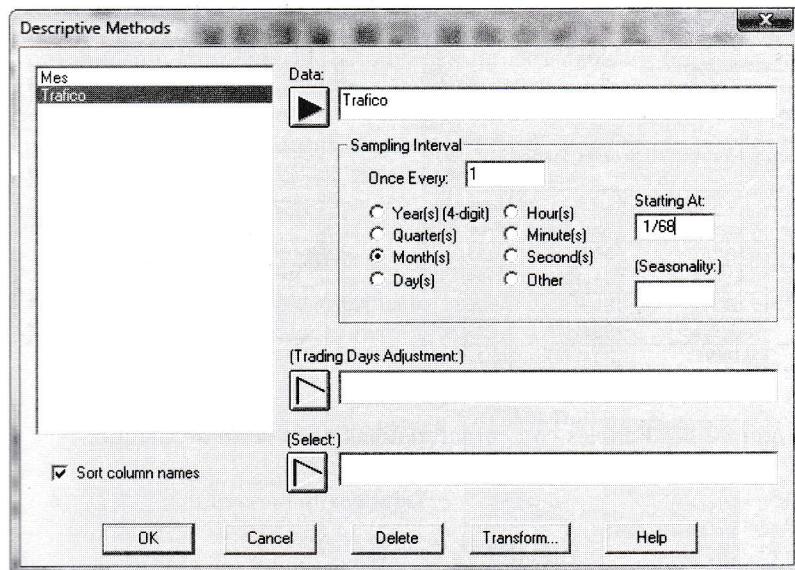


Figura 2.24: Selección de la serie de tiempo.

Al aceptar, el programa arrojará una pantalla similar a la de la Figura 2.25. Se observan cuatro paneles con la siguiente información: Descripción general de la serie; la gráfica de la serie de tiempo; las autocorrelaciones estimadas y la gráfica de la ACF. Sin embargo, es muy probable que se desee ver algunas otras gráficas y tal vez algunos otros datos. Para ello existen dos botones en la parte superior de los paneles.

El primero:

El segundo:

el segundo:

## 2.12 Actividades sugeridas

1. Demuestra que el caso particular finito de la forma invertida, dado por el modelo:  $Y_t = \Psi_1 Y_{t-1} + e_t$ , puede transformarse en el modelo lineal general, con un número infinito de términos.

## 2.12 Actividades sugeridas

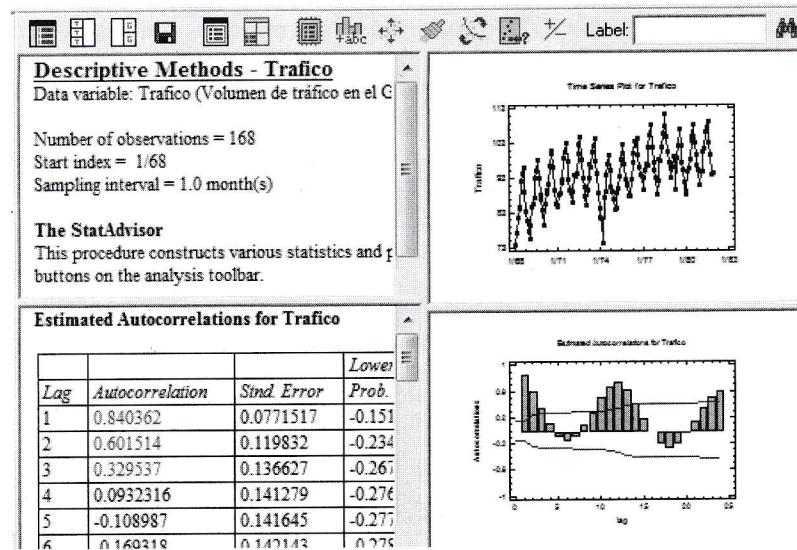


Figura 2.25: Paneles con características de la serie de tiempo.

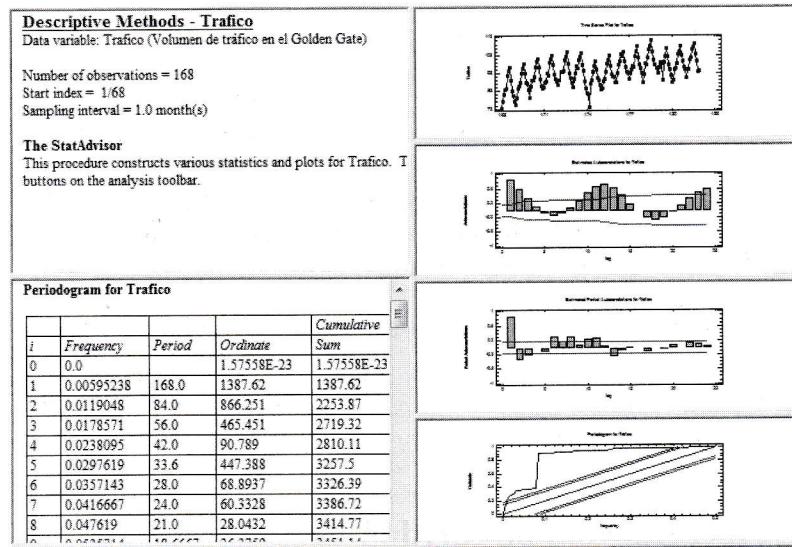


Figura 2.26: Paneles con otras características de la serie de tiempo.

2. Señala cómo debería de ser el coeficiente  $\Psi_1$  del ejercicio anterior para que el modelo tenga sentido, es decir, para que los valores de los errores más lejanos en el tiempo tengan menor influencia en el modelo.
3. Utilizando la página *2008 Forecasting Software Survey*, ubicada en <http://www.lionhrtpub.com/orms/surveys/FSS/fssmain.html> que se muestra en esta sección, indica cuáles de los paquetes de software incluyen el pronóstico a través de la metodología de Box-Jenkins. Señala su nombre, quién los produce, su sitio web principal y su precio.
4. Demuestra que el proceso ARMA(1,1):  $Y_t = \phi_1 Y_{t-1} + e_t - \theta_1 e_{t-1}$  tiene tanto condiciones de estacionaridad como de invertibilidad.
5. Responde las siguientes preguntas, a través de localizar los datos e internet. No olvides colocar las referencias de manera apropiada:
  - a) ¿Dónde y cuándo nació George Edward Pelham Box?
  - b) ¿En qué universidad trabaja?
  - c) ¿Qué honores o premios importantes ha recibido?
  - d) Box se casó (en segundo matrimonio) con la hija de un importante estadístico. ¿De qué estadístico hablamos y cómo se llama la esposa de Box?
  - e) ¿Con quién hizo la transformación de Box-Cox?
  - f) Incluye una fotografía de G. Box
  - g) ¿Cuándo y dónde nació Gwilym Meirion Jenkins?
  - h) ¿En qué año murió?
  - i) ¿De qué murió?
6. ¿Qué es IDAMS y cómo se obtiene?

## 2.12 Actividades sugeridas

7. Utiliza *Excel* para graficar la ACF y PACF teóricas de los siguientes modelos, e indica de qué modelo se trata en cada caso:

a)  $Y_t = 0.5Y_{t-1} + e_t$

b)  $Y_t = -0.5Y_{t-1} + e_t$

c)  $Y_t = Y_{t-1} + e_t$ . Ojo: Es un modelo no estacionario.

d)  $Y_t = 2Y_{t-1} + e_t$ . Ojo: Es un modelo no estacionario.

e)  $Y_t = e_t - 0.5e_{t-1}$

f)  $Y_t = e_t + 0.5e_{t-1}$

g)  $Y_t = e_t + e_{t-1}$ . Ojo: Es un modelo no invertible.

h)  $Y_t = 0.4Y_{t-1} - 0.3Y_{t-2} + e_t$

i)  $Y_t = -0.4Y_{t-1} + 0.3Y_{t-2} + e_t$

j)  $Y_t = -0.4Y_{t-1} + e_t + 0.3e_{t-1}$

k)  $Y_t = Y_{t-1} + e_t - e_{t-1}$ . Ojo: Es un modelo no estacionario y no invertible.

8. Demuestra que el modelo ARMA(1,1) equivale a un modelo AR( $\infty$ ).

9. Demuestra que el modelo ARMA(1,1) equivale a un modelo MA( $\infty$ ).

## Capítulo 3

# Modelos no estacionarios

El matemático es una especie de diseñador de modas  
totalmente inconsciente de las criaturas para quienes diseña.  
TOBÍAS DANTZIG.

### 3.1 Introducción

Hasta el momento se han revisado únicamente los modelos ARMA( $p,q$ ) que pueden utilizarse para aquellos datos que cumplen con las condiciones de estacionariedad en sentido amplio. Sin embargo, en muchos de los casos reales se encuentran series de tiempo cuya varianza es creciente, con tendencia, con variación estacional, con combinaciones de estas características o con todas ellas. ¿Cómo se procede entonces?

En este tipo de casos será necesario modificar temporalmente los datos, a través de operaciones matemáticas sencillas, de manera que se ajusten a las condiciones de estacionariedad. Hecho esto, será posible identificar alguno de los modelos ya enunciados en el capítulo anterior. Con el modelo se generan los pronósticos y por último, se “deshace” la transformación matemática (o las transformaciones si es que fueron varias), al aplicar las funciones inversas, para regresar los pronósticos a las unidades de los datos originales. Esto se hará tanto para el pronóstico puntual como para el pronóstico por intervalo.

Para lograr este propósito, es conveniente fijar un orden en el procedimiento con que se tratarán los **modelos no estacionarios**, como se propone a continuación:

1. **Estabilización de la varianza.** El primer paso es estabilizar la varianza<sup>1</sup>, es decir, lograr que los datos sean homoscedásticos y su apariencia sea la de una banda o franja del mismo ancho, alrededor de la media, sin importar si la media es o no constante. Esto se hará a través de

<sup>1</sup>La varianza sólo puede estabilizarse, no eliminarse. Esto último significaría convertir datos estocásticos en datos determinísticos.

### *3.1 Introducción*

---

transformaciones matemáticas tales como raíz cuadrada, logaritmo o recíproco. Una vez que los datos sean homoscedásticos, se procede al paso siguiente. Por supuesto, si los datos ya son homoscedásticos, no es necesario efectuar transformación<sup>2</sup>.

2. **Eliminación de la tendencia.** Si los datos ya tienen varianza constante, se procede a revisar si existe algún tipo de tendencia. La tendencia se eliminará a través de las llamadas diferencias ordinarias o finitas. Una vez que los datos tengan media constante, es decir, que se vean como fluctuaciones alrededor de una recta horizontal, todas dentro de una franja del mismo ancho, conviene revisar si existe variación estacional, como se explica en el siguiente punto.
3. **Tratamiento de la variación estacional.** En algunos datos ocurre que la variación estacional es visible desde la gráfica de los datos originales o al estabilizar la varianza. En otros datos es posible que se advierta hasta después de eliminar la tendencia. Por supuesto, también habrá datos que carecen de variación estacional. Si existe variación estacional, en ocasiones es posible eliminarla con diferencias estacionales. También puede ser que la variación estacional sea tan poderosa dentro de la estructura de los datos, que sea imposible eliminarla. En este último caso convendrá modelarla, como se verá en el Capítulo 6. Así pues, existen dos caminos para tratar con la variación estacional:
  - a) Eliminarla a través de diferencias estacionales.
  - b) Representarla a través de modelos multiplicativos que incluyan el componente estacional.

Puesto que no será posible saber cuál de estos dos caminos funciona mejor hasta probarlos y llegar a la generación de los pronósticos, será indispensable recorrer ambos y, una vez que se cuente con todos los resultados, elegir el mejor.

Es fundamental comenzar siempre el análisis de la estacionaridad por la estabilización de la varianza, ya que las transformaciones matemáticas que se requieren para ello pudieran resultar imposibles para datos negativos o cercanos a cero y las diferencias pueden generar resultados de este tipo. En cambio, se verá que el orden en cuanto a la eliminación de la tendencia y el tratamiento de la variación estacional puede modificarse sin problema. La única recomendación esencial es que, atendiendo al principio de parsimonia, las transformaciones se hagan de una en una, analizando siempre el resultado

---

<sup>2</sup>Conviene recordar que cada transformación produce errores en los datos y, además, una transformación innecesaria será contraria al principio de parsimonia.

obtenido, para evitar un modelo sobreestimado o con parámetros excesivos que no se justifiquen.

En seguida se explica cada uno de los procedimientos para convertir los datos en estacionarios.

## 3.2 Estabilización de la varianza

Existen varias **transformaciones** posibles para inducir una varianza constante. La idea básica es transformar los datos de tal modo que la dispersión sea siempre semejante alrededor de la media, independientemente de que la media sea o no constante. Las transformaciones más utilizadas con este fin son aquellas que ocasionan que los datos que son muy diferentes entre sí, resulten más parecidos. Entre las más utilizadas están la raíz cuadrada, el logaritmo (natural o decimal) y el recíproco.

En general, estas transformaciones son útiles cuando:

- La varianza es proporcional al crecimiento de la serie. Esto es muy común en las series financieras o económicas, cuyos datos se manejan en unidades monetarias, ya que al aumentar los precios, los incrementos tienden a ser mayores conforme transcurre el tiempo, aunque los porcentajes de aumento sean similares. Como gran parte de las variables económicas en México están asociadas al fenómeno inflacionario, este comportamiento se encontrará en prácticamente cualquier serie cuyas unidades sean pesos. Para eliminar el problema de la inflación se recomienda utilizar el procedimiento conocido como **deflactar** los datos o usar como unidad una moneda con mayor estabilidad como podría ser el dólar.
- El paso del tiempo introduce mayor variabilidad (o volatilidad) en el fenómeno. Por ejemplo, en la tasa de desocupación urbana en México (véase Figura 2.12), además de observarse una tendencia creciente (lo cual por supuesto preocupa), puede verse que cada vez hay datos más extremos en sus valores.

Entre las transformaciones más utilizadas para estabilizar la varianza destacan las siguientes:

- La raíz cuadrada:  $Z_t = \sqrt{Y_t}$ .
- El logaritmo (natural o decimal):  $Z_t = \ln Y_t$  o  $Z_t = \log Y_t$ .
- El recíproco o inverso:  $Z_t = 1/Y_t$ .
- El recíproco de la raíz cuadrada:  $Z_t = 1/\sqrt{Y_t}$ .
- El recíproco del logaritmo:  $Z_t = 1/\ln Y_t$ .

### 3.3 Eliminación de la tendencia

Estas operaciones pueden ampliarse y generalizarse en la llamada **transformación de Box-Cox** que suele expresarse como:

$$Z_t = \begin{cases} \frac{(Y_t + \lambda_2)^{\lambda_1} - 1}{\lambda_1} & \text{si } \lambda_1 \neq 0 \\ \ln(Y_t + \lambda_2) & \text{si } \lambda_1 = 0 \end{cases} \quad (3.1)$$

Dentro de *Statgraphics* es posible hacer estas transformaciones de la siguiente forma. Una vez que se tiene la serie de tiempo como archivo de datos (*data file*) y que se ha seleccionado la opción **Describe / Time Series / Descriptive Methods**, con el clic derecho del ratón se obtiene el menú desplegable que se muestra en la Figura 3.1.

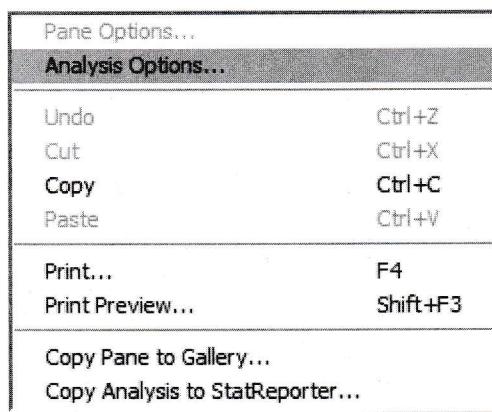


Figura 3.1: Menú desplegable en *Statgraphics*.

Al elegir el rubro **Analysis Options** se obtendrá la ventana de diálogo ilustrada en la Figura 3.2, en la cual es posible efectuar las transformaciones necesarias a los datos.

Para activar las opciones relacionadas con la variación estacional (*Seasonal* y *Seasonal Order*), es necesario haber definido desde el inicio, en la selección de los datos de la serie de tiempo, cuál podría ser el orden de la posible variación estacional, como se indica en la Figura 3.3.

## 3.3 Eliminación de la tendencia

Una vez que los datos son homoscedásticos, debe revisarse la posible tendencia de la serie. Se ha definido a la tendencia como un cambio sistemático en el nivel de una serie de tiempo. Es importante notar que lo que parece un cambio de nivel en una serie pequeña puede no serlo al tener más datos, convirtiéndose en un movimiento periódico. De nuevo, será indispensable el conocimiento del fenómeno para establecer el modelo adecuado.

### 3.3 Eliminación de la tendencia

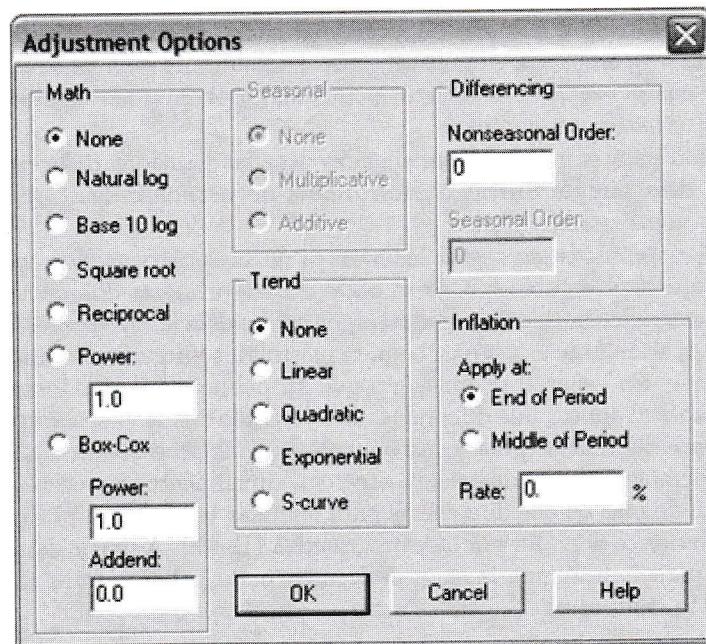


Figura 3.2: Ventana para hacer transformaciones de datos.

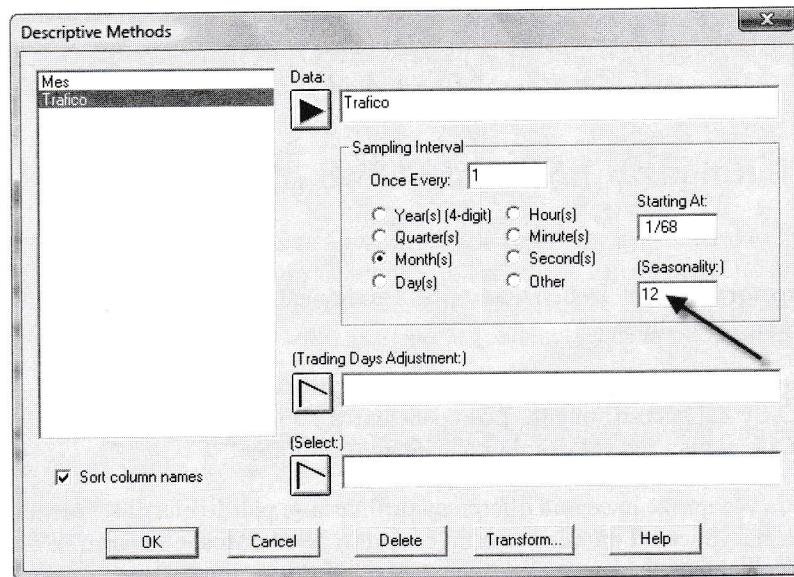


Figura 3.3: Orden de la posible variación estacional.

### 3.3 Eliminación de la tendencia

---

Existen varios métodos para modelar la tendencia, uno de ellos es usar un modelo de regresión. Esto será adecuado cuando la tendencia sea fija y determinística. Sin embargo, en muchas de las series de tiempo reales, la tendencia tiene componentes estocásticos, por lo cual es conveniente encontrar una forma alternativa de eliminarla.

Una opción para eliminar la tendencia es usar las llamadas **diferencias ordinarias, diferencias finitas o diferencias no estacionales** (*ordinary differences o nonseasonal differences*).

Este método consiste en restar los valores de las observaciones uno de otro, en un orden preestablecido. Si se obtienen las primeras diferencias de una serie con tendencia lineal, por ejemplo, esta tendencia desaparecerá. Si la tendencia es semejante a una parábola, requerirá dos diferencias ordinarias para eliminarse. En general, se considera que un polinomio de grado uno se vuelve constante al aplicar una diferencia; uno de grado dos se vuelve constante al aplicar dos diferencias y así sucesivamente [46]. En procesos económicos, financieros, demográficos, industriales, etcétera, difícilmente aparecen series que requieran de más de dos diferencias para eliminar la tendencia. Las diferencias ordinarias se denotan por la letra griega  $\Delta$  (Delta). A partir de ahora se usará  $Z_t$  para indicar que la serie de tiempo pudo haber sufrido alguna transformación. Entonces:

$$\Delta Z_t = Z_t - Z_{t-1} \quad (3.2)$$

$$\begin{aligned} \Delta^2 Z_t &= \Delta Z_t - \Delta Z_{t-1} \\ &= (Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2}) \\ &= Z_t - 2Z_{t-1} + Z_{t-2} \end{aligned} \quad (3.3)$$

$$\Delta^3 Z_t = Z_t - 3Z_{t-1} + 3Z_{t-2} - Z_{t-4} \quad (3.4)$$

⋮

**Notas importantes.** Deben observarse los siguientes aspectos con relación a las diferencias ordinarias:

- Las segundas diferencias ordinarias no equivalen a la diferencia  $Z_t - Z_{t-2}$ , y así sucesivamente. Los coeficientes de los términos de las diferencias siguen los valores de un triángulo de Pascal.
- Cada vez que se hace una diferencia ordinaria se pierde una observación. Así, si se hacen  $d$  diferencias a un conjunto de  $N$  datos, quedarán  $N - d$  observaciones.
- Puede hacerse una equivalencia entre las diferencias ordinarias y el operador de salto hacia atrás:  $\Delta = 1 - B$ .

- Una vez eliminada la tendencia, seguir haciendo diferencias producirá también series sin tendencia; sin embargo, además de que se pierden observaciones, el modelo se complicará de manera innecesaria, contraviniendo el principio de parsimonia. Véase el siguiente ejemplo:

$$\begin{array}{ll} Z_t &= e_t && \text{Ruido blanco} \\ \Delta Z_t &= Z_t - Z_{t-1} = e_t - e_{t-1} && \text{ARMA}(0,1) \end{array}$$

La idea básica de eliminar la tendencia no es olvidar que existe, sino obtener una nueva serie que pueda ser analizada más fácilmente, y después volver a introducir la tendencia. Por ejemplo, si se obtuvieron las primeras diferencias

$$Z_t = Y_t - Y_{t-1}$$

Puede reconstruirse la serie original  $Y_t$ , en forma recursiva, a través de la ecuación:

$$\begin{aligned} Y_t &= Z_t + Y_{t-1} \\ &= Z_t + Z_{t-1} + Y_{t-2} \\ &\vdots \\ &= Z_t + Z_{t-1} + Z_{t-2} + \dots \end{aligned} \tag{3.5}$$

A este proceso de sumas sucesivas se le suele llamar **integración** de la serie.

¿Cómo efectuar las diferencias ordinarias con *Statgraphics*? Para ello se utiliza la ventana de diálogo **Adjustment Options** de la Figura 3.2, indicando en el espacio marcado con *Nonseasonal Order* el orden de las diferencias no estacionales, como se ve en la Figura 3.4.

### 3.4 Tratamiento de la variación estacional

Una vez que la varianza se ha estabilizado (es decir, que la serie es homoscedástica) y se ha eliminado la tendencia (la media es constante), resultará muy evidente la variación estacional, si es que existe. En algunos casos, la variación estacional es evidente aún con la tendencia y/o la heteroscedasticidad.

En algunos fenómenos reales, la variación estacional es tan importante dentro de la serie de tiempo, que será imposible eliminarla. Sin embargo, es conveniente tratar de hacerlo a través de **diferencias estacionales** (*seasonal differences*).

Sea  $s$  la longitud del periodo de la variación estacional. Es decir, si un comportamiento se repite de manera semejante cada 12 meses (un año con unidades mensuales),  $s = 12$ ; si se repite cada 4 trimestres (un año con unidades trimestrales),  $s = 4$ ; cada 7 días (una semana con unidades diarias),  $s = 7$ ; etcétera. Las diferencias estacionales se denotarán por  $\Delta_s$  (Delta

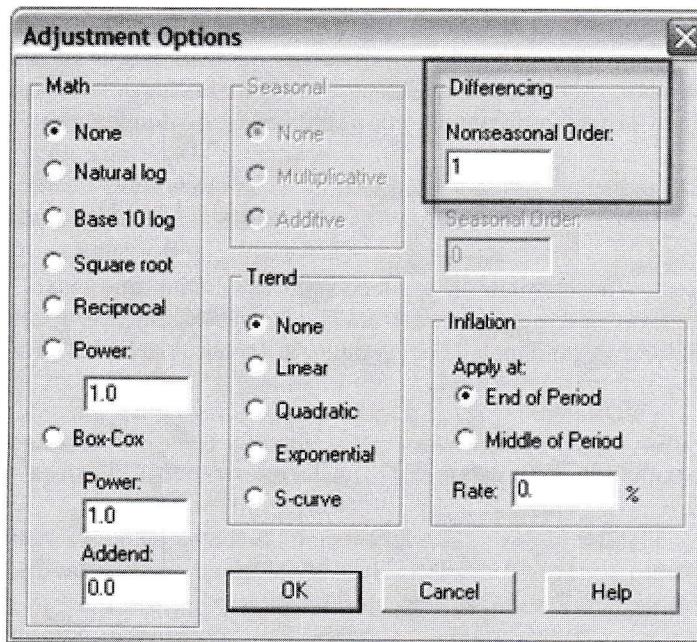


Figura 3.4: Orden de las diferencias no estacionales.

subíndice  $s$ ) y se calculan de la siguiente manera:

$$\Delta_s Z_t = Z_t - Z_{t-s} \quad (3.6)$$

$$\begin{aligned} \Delta_s^2 Z_t &= \Delta_s Z_t - \Delta_s Z_{t-s} \\ &= (Z_t - Z_{t-s}) - (Z_{t-s} - Z_{t-2s}) \\ &= Z_t - 2Z_{t-s} + Z_{t-2s} \end{aligned} \quad (3.7)$$

$$\Delta_s^3 Z_t = Z_t - 3Z_{t-s} + 3Z_{t-2s} - Z_{t-3s} \quad (3.8)$$

⋮

Por lo general, sólo se requiere efectuar una o dos diferencias estacionales. Ya sea que logre eliminarse o no la variación estacional, es conveniente probar uno o más modelos estacionales, con objeto de seleccionar la opción que ofrezca los mejores resultados.

**Notas importantes.** Deben observarse los siguientes aspectos con relación a las diferencias estacionales:

- Las segundas diferencias estacionales no equivalen a la diferencia estacional  $Z_t - Z_{t-s}$ , y así sucesivamente. Los coeficientes de los términos de

las diferencias estacionales siguen los valores de un triángulo de Pascal.

- Cada vez que se hace una diferencia ordinaria se pierden  $s$  observaciones. Así, si se hacen  $D$  diferencias estacionales a un conjunto de  $N$  datos, quedarán  $N - Ds$  observaciones.
- Puede hacerse una equivalencia entre las diferencias estacionales y el operador de salto hacia atrás:  $\Delta_s = 1 - B^s$ .
- Efectuar diferencias estacionales innecesarias ocasionará que los datos adquieran una variación estacional inexistente, contraviniendo el principio de parsimonia y el sentido lógico. Véase el siguiente ejemplo:

$$\begin{array}{ll} Z_t &= e_t && \text{Ruido blanco} \\ \Delta_s Z_t &= Z_t - Z_{t-s} = e_t - e_{t-s} && \text{Modelo con variación estacional} \end{array}$$

- Es posible que un conjunto de datos requiera tanto diferencias ordinarias como diferencias estacionales. En este caso puede demostrarse que es indistinto el orden en que éstas se efectúen. Por ejemplo, supóngase que se aplican primero las diferencias ordinarias y posteriormente las estacionales, entonces:

$$\begin{aligned} \Delta_s \Delta Z_t &= \Delta_s(Y_t - Y_{t-1}) \\ &= Y_t - Y_{t-1} - Y_{t-s} + Y_{t-1-s} \end{aligned} \quad (3.9)$$

Si, por el contrario, se efectúa primero la diferencia estacional y luego la ordinaria:

$$\begin{aligned} \Delta \Delta_s Z_t &= \Delta(Y_t - Y_{t-s}) \\ &= Y_t - Y_{t-s} - Y_{t-1} + Y_{t-s-1} \end{aligned} \quad (3.10)$$

Puede verificarse con facilidad que las ecuaciones 3.9 y 3.10 son algebraicamente equivalentes.

- No debe olvidarse que habrá que probar también, como se verá más adelante, el modelado de la variación estacional (ver Capítulo 6).

Dentro de *Statgraphics*, es necesario indicar que una serie de tiempo podría tener variación estacional para que se activen las opciones correspondientes a ella. Esto se hace al seleccionar los datos, en la ventana de diálogo **Time**, como se ve en la Figura 3.5.

Si has seleccionado los datos y no señalaste la variación estacional, no te preocupes, puedes hacerlo en cualquier momento a través del botón para identificar los datos: .

### 3.5 Modelos ARIMA

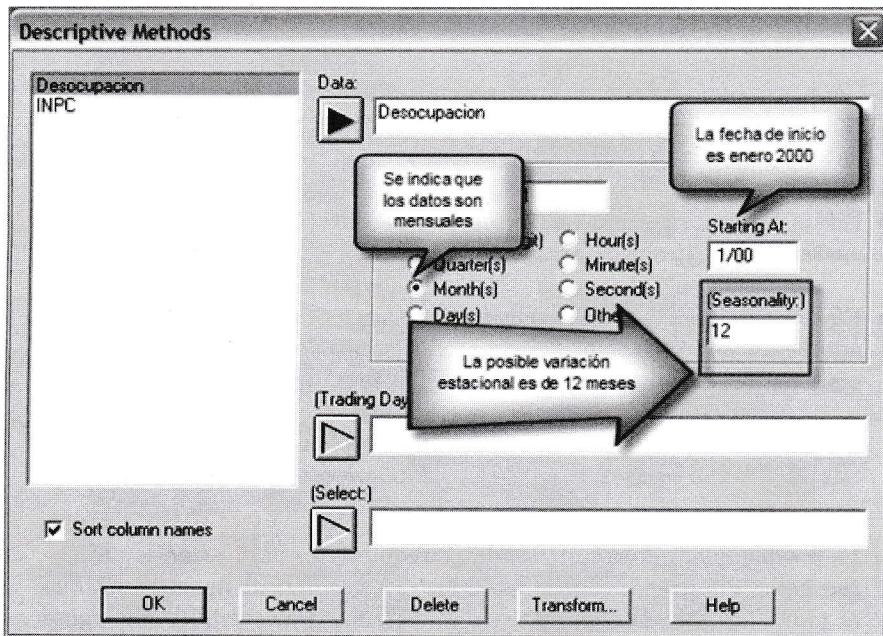


Figura 3.5: Longitud de la variación estacional.

## 3.5 Modelos ARIMA

Por último, es posible señalar tanto las transformaciones como las diferencias (ordinarias, estacionales o ambas) dentro de los modelos. Como se ha dicho, al utilizar  $Z_t$  en lugar de  $Y_t$  (observaciones originales), se está indicando que pudo haber ya alguna transformación, misma que habrá de deshacerse (a través de la transformación inversa), una vez seleccionado el modelo y generados los pronósticos.

Por otro lado, se ha dicho que, si se efectúan diferencias ordinarias, los datos originales pueden recuperarse a través de sumas sucesivas, por un proceso llamado de "integración". Por esta razón los modelos más generales se llamarán ahora **modelos integrados autorregresivos y de medias móviles** o simplemente **modelos ARIMA**( $p,d,q$ ), de su equivalente en inglés: *Autorregresive Integrated Moving Averages*.

Hasta ahora se han considerado modelos cuya media es cero, con el objeto de facilitar los cálculos. Para crear modelos aún más generales y versátiles, puede agregarse una constante o término independiente, denotada por  $\delta_0$  (delta subíndice cero). Así, el modelo ARIMA( $p,d,q$ ) se expresará como:

$$W_t = \delta_0 + \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (3.11)$$

O en su forma abreviada como:

$$\phi(B)W_t = \delta_0 + \theta(B)e_t \quad (3.12)$$

### 3.6 Actividades sugeridas

Donde:  $W_t = \Delta^d Z_t$ ;  $Z_t$  es alguna transformación de  $Y_t$ ;  $\delta_0 \in \mathbb{R}$ ; y  $B$  es el operador de salto hacia atrás.

Obsérvese que, tanto en la ecuación 3.11 como en la 3.12, el término constante  $\delta_0$  no es la media, excepto en el caso de que el modelo seleccionado fuera únicamente de medias móviles, de otro modo:

$$\begin{aligned} E(W_t) &= \delta_0 + \phi_1 E(W_{t-1}) + \dots + \phi_p E(W_{t-p}) + 0 \\ (1 - \phi_1 - \dots - \phi_p)E(W_t) &= \delta_0 \\ E(W_t) &= \frac{\delta_0}{1 - \phi_1 - \dots - \phi_p} \end{aligned} \quad (3.13)$$

Para concluir, si se hubieran efectuado diferencias estacionales de orden  $D$ , el modelo sería el mismo pero con  $W_t = \Delta_s^D \Delta^d Z_t$ , en cuyo caso cabría preguntarse... ¿cuántas observaciones se habrán perdido?

## 3.6 Actividades sugeridas

1. Simula 100 valores de la serie  $Y_t = e_t$  (ruido blanco), por el método que gustes. Coloca en tu tarea la gráfica de estos datos e indica el método utilizado para simularlos. Además:
  - a) Elabora con *Statgraphics* las gráficas: ACF, PACF, Periodograma y Periodograma integrado. Escribe brevemente tu interpretación de las gráficas.
  - b) Ahora obtén las primeras diferencias de la serie simulada, es decir:  $Z_t = \Delta e_t = e_t - e_{t-1}$ , o cada valor menos el anterior. Te quedarán 99 valores. Agrega la gráfica de estos datos.
  - c) Elabora con *Statgraphics* las gráficas: ACF, PACF, Periodograma y Periodograma integrado de  $Z_t$ . Interpreta brevemente las gráficas resultantes.
2. Simula 100 valores de la serie  $Y_t = Y_{t-1} + e_t$  (puedes usar el ruido blanco del ejercicio anterior), por el método que gustes. Coloca en tu tarea la gráfica de estos datos e indica el método utilizado para simularlos. Además:
  - a) Elabora con *Statgraphics* las gráficas: ACF, PACF, Periodograma y Periodograma integrado. Escribe brevemente tu interpretación de las gráficas.
  - b) Ahora obtén las primeras diferencias de la serie simulada, es decir:  $Z_t = \Delta Y_t$ , o cada valor menos el anterior. Te quedarán 99 valores. Agrega la gráfica de estos datos.

### *3.6 Actividades sugeridas*

---

- c) Elabora con *Statgraphics* las gráficas: ACF, PACF, Periodograma y Periodograma integrado de  $Z_t$ . Interpreta brevemente las gráficas resultantes.
3. Localiza una serie de tiempo real que presente con claridad algún tipo de tendencia lineal. Coloca en tu tarea la gráfica de estos datos e indica de manera correcta la fuente. Además:
  - a) Elabora con *Statgraphics* las gráficas: ACF, PACF, Periodograma y Periodograma integrado. Escribe brevemente tu interpretación de las gráficas.
  - b) Ahora obtén las primeras diferencias de la serie, es decir:  $Z_t = \Delta Y_t$ , o cada valor menos el anterior. Agrega la gráfica de estos datos.
  - c) Elabora con *Statgraphics* las gráficas: ACF, PACF, Periodograma y Periodograma integrado de  $Z_t$ . Interpreta brevemente las gráficas resultantes.
4. Demuestra si el siguiente modelo es o no estacionario en sentido amplio:

$$Y_t = e_t + e_{t-1} + e_{t-2} + \dots + e_2 + e_1$$

5. Demuestra si el siguiente modelo es o no estacionario en sentido amplio:

$$Y_t = 0.8Y_{t-1} + 0.8Y_{t-2} + e_t$$

## Capítulo 4

# Pronósticos

El azar es el seudónimo que usa Dios cuando no quiere firmar.

ANATOLE FRANCE

### 4.1 Introducción

Por supuesto, al seleccionar un modelo adecuado para un conjunto de datos reales, además de describirlos y comprenderlos mejor, el fin último es pronosticar sus valores para apoyar la toma de decisiones correcta. En ese sentido, tanto los modelos ARMA como los más generales ARIMA, permiten generar pronósticos con facilidad, especialmente si se utiliza algún software estadístico adecuado como *Statgraphics*, *Statgraphics Online*, SAS, EViews, SPSS, R, etcétera.

Para obtener los pronósticos, como se ha dicho, habrá que elegir un modelo específico. Sin embargo y aunque parezca paradójico, también será conveniente utilizar el propio pronóstico como elemento de juicio para valorar y comparar los modelos. Esto significa que la elaboración de pronósticos formará parte del proceso de identificación de modelos. En realidad, en este caso no suele interesar tanto el ajuste del modelo a los datos reales (interpolación), sino su funcionalidad en cuanto a la generación de valores futuros, fuera de los datos (extrapolación). Así, será conveniente en algunas ocasiones seleccionar algún modelo que, aún cuando no brinde el mejor ajuste, sí provea los pronósticos más precisos y confiables.

Un **pronóstico** es un estimador del valor desconocido que tendrá la variable aleatoria en el futuro. Como todo estimador, puede obtenerse de dos maneras:

- **Puntual.** Se trata de un número específico que pretende ser aproximado al valor que tendrá, en su momento, el fenómeno real. Se calcula a través de algún método como mínimos cuadrados, máxima verosimilitud o momentos.

- **Por intervalo.** Se trata de formular los límites inferior y superior de un rango que, con cierto nivel de confianza, deberá contener al valor desconocido futuro de la variable aleatoria. Para calcularlo se usa el estimador puntual como centro. Además, se supone que el estimador sigue –aproximadamente– una distribución de probabilidad teórica y que, por lo tanto, puede estimarse que el valor real estará dentro del intervalo con un nivel de confianza del 90 o 95 %, si se considera un nivel de significancia<sup>1</sup> denotado por  $\alpha$  (alfa) de 0.10 ó 0.05, respectivamente.

En general, es conveniente contar con los dos tipos de pronóstico, puesto que cada uno de ellos tiene ventajas particulares: el pronóstico puntual brinda una idea aproximada del valor real, pero es casi seguro que será diferente de éste. El intervalo de confianza, por su parte, es menos preciso, pero existe una probabilidad alta ( $1 - \alpha$ ), de que contenga al valor real.

## 4.2 Pronósticos de costo mínimo

Como todo estimador, el pronóstico es una aproximación al valor real desconocido que tendrá la variable en un punto futuro del tiempo. Por ello, todo pronóstico está sujeto a error, de manera inevitable. Este error, a su vez, ocasionará un costo, tanto si se trata de un error por exceso, como si es por defecto.

Por ejemplo, supóngase que una empresa comercializadora de productos perecederos (*e.g.* fruta, lácteos, carnes) desea pronosticar las cantidades de cada producto que debe tener en su almacén para ofrecerlos en venta. Si el pronóstico queda por debajo del valor real de la demanda de los clientes, no será posible atenderlos adecuadamente y se tendrá un costo de pérdida de oportunidad o, incluso, de clientes que dejen de asistir al negocio por no haber encontrado el producto. En cambio, si el pronóstico excede al valor real, habrá una cantidad de producto en almacén que no se venderá. Este producto deberá continuar almacenado (si es posible), desecharse o procesarse en otra forma (jugos, nieves, helados, paletas), con el consiguiente costo para la organización.

¿Cómo encontrar entonces un pronóstico cuyo costo esperado sea mínimo?

Para comenzar, después de la revisión de la estacionalidad, se cuenta con  $N - d - sD$  datos observados<sup>2</sup> de la serie de tiempo:

$$Z_t: Z_{d+sD+1}, Z_{d+sD+2}, \dots, Z_N.$$

<sup>1</sup>El **nivel de significancia** es la probabilidad de rechazar una hipótesis nula dado que dicha hipótesis es verdadera.

<sup>2</sup>A partir de ahora se utilizará la notación  $Z_t$  para denotar a la serie de tiempo que ya es estacionaria, es decir, que puede incluir alguna transformación, alguna diferencia ordinaria y/o alguna diferencia estacional. En ese sentido, es posible que ya no se cuente con los  $N$  datos originales, sino únicamente con  $N - d - sD$  datos. Por supuesto,  $d$  y  $D$  pueden ser ceros.

Así pues, estos datos pueden dividirse en tres grandes categorías:

1. Los valores  $Z_{d+sD+1}, Z_{d+sD+2}, \dots, Z_{N-1}$  son los **datos históricos**.
2. Al valor  $Z_N$  se le considera el **dato presente o actual**. Es probable que no corresponda en realidad el presente como tal, sino que sea únicamente el dato más reciente con el que se cuenta.
3. Por último, se tienen los valores antiguos desconocidos:  $Z_0, Z_{-1}, Z_{-2}, \dots$ , a los cuales se puede llamar **datos prehistóricos**, ya que no existe registro de ellos (ver Figura 1.1).

El objetivo ahora es construir un modelo matemático, a partir de los datos presente e históricos, que permita pronosticar los valores de la serie para un **horizonte  $h$**  en particular. Es decir, se desea estimar los valores futuros:  $Z_{N+1}, Z_{N+2}, \dots, Z_{N+h}$ . Para ello es conveniente definir una notación que represente al estimador o **pronóstico**. Esta notación se observa en la ecuación 4.1.

$$\hat{Z}_{N+h} = Z_N(h) \quad (4.1)$$

Así pues, el **error** que se cometerá al pronosticar la variable para un horizonte de longitud  $h$ , será la diferencia entre el valor real (desconocido) y el valor estimado o pronóstico, lo cual se expresa como:

$$e_N(h) = Z_{N+h} - Z_N(h) \quad (4.2)$$

Como se ha dicho, este error implica un costo, tanto si 4.2 es positivo como si es negativo. Esto sugiere que el **costo del error** podría estimarse a través de una función del valor absoluto del mismo:

$$C(e_N(h)) = c |Z_{N+h} - Z_N(h)| \quad (4.3)$$

Donde  $c \in \mathbb{R}^+$ . Sin embargo, al tratar de minimizar el **costo del pronóstico** se encontraría que la función 4.3 no es derivable (tiene pico, por lo cual la derivada por la izquierda es distinta de la derivada por la derecha). Por lo tanto, se puede considerar que el costo es aproximado a una función cuadrática del error, como se ve en la Figura 4.1.

Esto lleva a plantear el costo del error de acuerdo con la ecuación 4.4.

$$C(e_N(h)) = c (Z_{N+h} - Z_N(h))^2 \quad (4.4)$$

Así, a medida que el error crezca en valor absoluto, ya sea positivo o negativo, el costo se incrementará. Puesto que se cometerá un error de manera inevitable, lo ideal será encontrar el pronóstico que minimice el costo esperado de dicho error.

Es relativamente sencillo verificar que el mejor pronóstico será el *valor esperado* de la variable aleatoria desconocida, puesto que minimiza el costo esperado del pronóstico; es decir, conviene elegir:

$$Z_N(h) = E(Z_{N+h}) = m_h \quad (4.5)$$

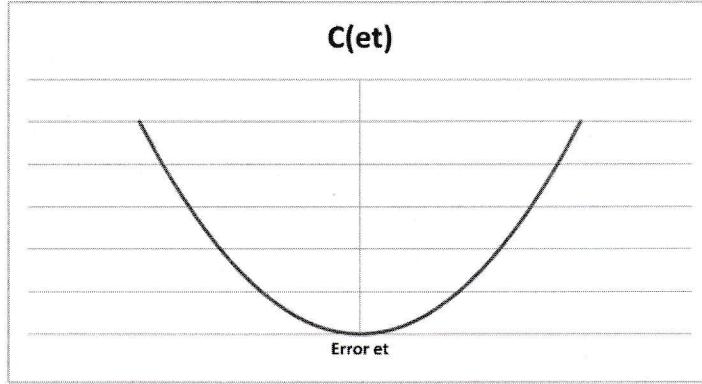


Figura 4.1: Costo del error de pronóstico.

Para demostrarlo, supóngase que se toma como pronóstico cualquier valor distinto de 4.5, tal como  $Z_N(h) = E(Z_{N+h}) + d = m$  con  $d \in \mathbb{R}$ . Entonces el **costo esperado del error** sería:

$$\begin{aligned} E[C(e_N(h))] &= E[c(Z_{N+h} - (m_h + d))^2] \\ &= c[E(Z_{N+h} - m_h)^2 - 2dE(Z_{N+h} - m_h) + d^2] \\ &= c[E(Z_{N+h} - m_h)^2 + d^2] \end{aligned}$$

Se observa que, para minimizar 4.6, conviene elegir  $d = 0$ , con lo cual se demuestra que el pronóstico óptimo para  $Z_{N+h}$  se obtiene con  $m = m_h = E(Z_{N+h})$ .

### 4.3 Cálculo de los pronósticos

Se ha establecido ya que el mejor pronóstico será el valor esperado de la variable real desconocida. Supóngase entonces que el modelo elegido para el pronóstico es un ARMA( $p,q$ ) general que cumple las condiciones de estacionaridad e invertibilidad. Para obtener el pronóstico que corresponde al horizonte  $h$ , se siguen estos pasos:

1. Se obtienen los estimadores (como se verá en la Sección 5.3) de los  $p+q+1$  parámetros desconocidos, es decir, de:  $\delta_0, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ .
2. Se reemplaza el error presente  $e_N$  y los errores históricos  $e_{N-1}, e_{N-2}, e_{N-3}, \dots, e_{N-d-SD+1}$ , con los **residuales** respectivos. Esto es, se reemplazan por la diferencia entre el valor real  $Z_t$  y el estimado por el modelo elegido, para cada valor de  $t$ .

3. Se reemplaza cada error futuro  $e_{N+j}$ ,  $0 < j \leq h$ , por su valor esperado, es decir, por cero.
4. Se reemplazan los valores históricos y presente  $Z_{N+j}$ ,  $j \leq 0$ , por los datos observados.
5. Se reemplaza cada valor futuro desconocido,  $Z_{N+j}$ ,  $1 \leq j < h$ , de manera recursiva, por su pronóstico respectivo,  $Z_N(j)$ .

**Ejemplo: Pronóstico de un modelo ARMA(1,1)** El primer pronóstico de un modelo ARMA(1,1) sería:

$$Z_N(1) = \hat{\delta}_0 + \hat{\phi}_1 Z_N + 0 - \hat{\theta}_1 \hat{e}_N$$

Donde:

$\hat{\delta}_0$  es el estimador de la constante del modelo.

$\hat{\phi}_1$  es el estimador del primer coeficiente autorregresivo.

$Z_N$  es el valor observado para el momento  $N$ .

0 es el estimador del error futuro  $e_{N+1}$ , es decir, su valor esperado.

$\hat{\theta}_1$  es el estimador del primer coeficiente de medias móviles.

$\hat{e}_N$  es el residual del intervalo  $N$ , es decir, la diferencia del valor observado  $Z_N$  menos el estimado por el modelo:  $\hat{Z}_N$ .

El segundo pronóstico sería:

$$Z_N(2) = \hat{\delta}_0 + \hat{\phi}_1 Z_N(1) + 0 - \hat{\theta}_1(0)$$

Donde:

$Z_N(1)$  es el pronóstico ya calculado.

El primer 0 es el valor esperado del error  $e_{N+2}$ .

El segundo 0 es el valor esperado del error  $e_{N+1}$ .

## 4.4 Límite de los pronósticos

Como se verá ahora, la funcionalidad de los resultados anteriores es relativamente limitada. En general, los pronósticos serán apropiados para un horizonte relativamente corto. A partir de ahí, los pronósticos se convierten en la media de la serie. Esto no debe desanimar al analista, puesto que las

#### 4.4 Límite de los pronósticos

---

predicciones obtenidas serán de gran calidad y podrán actualizarse conforme se agreguen nuevos datos, con lo cual se generarán los pronósticos sucesivos.

##### 4.4.1 Pronóstico de un modelo AR(1)

Supóngase que para un conjunto de datos estacionarios en sentido amplio se identificó un modelo AR(1), en el cual el valor esperado o media teórica es  $\mu$ , es decir:

$$E(Z_t) = \mu$$

Entonces, si se resta dicha media a cada observación, el modelo AR(1) identificado tendrá una constante  $\delta_0 = 0$  y se expresará como:

$$\begin{aligned} Z_t - \mu &= \phi_1(Z_{t-1} - \mu) + e_t \\ Z_t &= (1 - \phi_1)\mu + \phi_1 Z_{t-1} + e_t \end{aligned}$$

El pronóstico para una unidad de tiempo adelante sería:

$$Z_N(1) = E(Z_{N+1}) = (1 - \phi_1)\mu + \phi_1 Z_N \quad (4.6)$$

Y el pronóstico para un horizonte  $h$  sería:

$$Z_N(h) = E(Z_{N+h}) = (1 - \phi_1)\mu + \phi_1 Z_N(h-1)$$

Como se ve, esto genera una serie de ecuaciones recursivas, donde el pronóstico para dos unidades de tiempo adelante es:

$$Z_N(2) = E(Z_{N+2}) = (1 - \phi_1)\mu + \phi_1 Z_N(1) \quad (4.7)$$

Si se sustituye 4.6 en la ecuación anterior, se obtiene:

$$\begin{aligned} Z_N(2) &= (1 - \phi_1)\mu + \phi_1((1 - \phi_1)\mu + \phi_1 Z_N) \\ &= (1 - \phi_1)(1 + \phi_1)\mu + \phi_1^2 Z_N \end{aligned}$$

Si se continúa con este proceso, puede generalizarse la ecuación recursiva para  $Z_N(h)$ :

$$Z_N(h) = (1 - \phi_1)(1 + \phi_1 + \phi_1^2 + \dots + \phi_1^{h-1})\mu + \phi_1^h Z_N \quad (4.8)$$

Para obtener el pronóstico a largo plazo, basta con calcular el límite de la ecuación anterior cuando  $h \rightarrow \infty$ , bajo el supuesto de que se cumple la condición de estacionaridad, por la cual  $|\phi_1| < 1$ . Como la suma del segundo factor es una progresión geométrica decreciente que tiende a:

$$\frac{1}{1 - \phi_1}$$

Entonces se tendrá que:

$$\lim_{h \rightarrow \infty} Z_N(h) = \mu \quad (4.9)$$

Esto significa que, a largo plazo, el pronóstico de un modelo de tipo AR(1) tiende a la media de la serie. Lo mismo ocurrirá en general para todos los modelos de tipo AR.

#### 4.4.2 Pronóstico de un modelo MA(1)

Supóngase ahora que para un conjunto de datos se identificó un modelo MA(1), en el cual el valor esperado o media es  $\mu$ , es decir:

$$E(Z_t) = \mu$$

Entonces, si nuevamente se resta la media a todas las observaciones, con lo cual el modelo carecerá de término constante, el modelo MA(1) resultante se expresaría como:

$$\begin{aligned} Z_t - \mu &= e_t + \theta_1 e_{t-1} \\ Z_t &= \mu + e_t + \theta_1 e_{t-1} \end{aligned}$$

El pronóstico para una unidad de tiempo adelante sería:

$$Z_N(1) = E(Z_{N+1}) = \mu + \theta_1 \hat{e}_N \quad (4.10)$$

Donde  $\hat{e}_N$  es el residual que se obtiene como la diferencia entre el valor real y el estimado por el modelo:

$$\hat{e}_N = Z_N - \mu + \theta_1 \hat{e}_{N-1}$$

A partir de ahí, los siguientes pronósticos serán:

$$\begin{aligned} Z_N(2) &= \mu \\ Z_N(3) &= \mu \\ &\vdots \\ Z_N(h) &= \mu \end{aligned}$$

Es decir, inmediatamente después de la primera unidad de tiempo pronosticada, los siguientes valores serán iguales a la media de la serie. De la misma forma puede demostrarse que para un MA(2) se obtienen dos pronósticos y luego la media de la serie; para un MA(3) se obtienen tres pronósticos y luego la media de la serie, etcétera.

#### 4.4.3 Pronóstico de un modelo ARMA(1,1)

Puede demostrarse de manera sencilla que, si se ha elegido como modelo un ARMA(1,1), el pronóstico a largo plazo de este modelo también tiende a la media de la serie (ver Sección 4.6). Lo mismo ocurre con los modelos generales ARMA( $p,q$ ). Por ello, debe tenerse precaución al pronosticar y confiar poco en los pronósticos muy lejanos en el tiempo. Esta precaución es razonable puesto que, mientras más adelante esté el futuro, menos se sabe acerca de cómo serán sus circunstancias.

## 4.5 Actualización de los pronósticos

Como se ha demostrado, los modelos de tipo ARMA( $p,q$ ) generan pronósticos que podrán ser muy precisos y confiables, a corto y tal vez a mediano plazo, pero no a largo plazo. Mientras más estable sea un fenómeno, será más razonable confiar en los pronósticos lejanos en el tiempo, y viceversa.

Por otro lado, conforme se recaben más datos históricos, será conveniente incorporarlos al cálculo –y eventualmente la identificación– del modelo, para actualizar los pronósticos siguientes. Esto puede hacerse de dos formas:

1. **Actualización secuencial.** El nuevo dato recabado se incorpora a los valores observados y con ello se estiman de nuevo los parámetros, pero se sigue utilizando el mismo modelo ya seleccionado para pronosticar. Se recomienda usar esta forma cuando el fenómeno analizado es muy estable y los pronósticos anteriores han resultado adecuados.
2. **Actualización adaptativa.** El nuevo dato recabado se incorpora a los valores observados y se revisa el modelo para verificar si continúa siendo el mismo o debe cambiarse. Se recomienda cuando el fenómeno analizado es poco estable o los pronósticos han sido deficientes.

## 4.6 Actividades sugeridas

1. Demuestra que el pronóstico del modelo ARMA(1,1) tiende a la media  $\mu$  conforme el horizonte  $h$  tiende a infinito.
2. Supongamos que los datos reales de un fenómeno se ajustan a un modelo AR(1):  $Z_t = 10 + 0.8Z_{t-1} + e_t$ , donde  $Z_t = \Delta Y_t = Y_t - Y_{t-1}$ . ¿A qué valor tienden los pronósticos de  $Y_t$ ?
3. Supongamos que los datos reales de un fenómeno se ajustan a un modelo MA(1):  $Z_t = 10 + e_t - 0.8e_{t-1}$ , donde  $Z_t = \Delta Y_t = Y_t - Y_{t-1}$ . ¿A qué valor tienden los pronósticos de  $Y_t$ ?
4. Selecciona los datos de alguna serie de tiempo real estacionaria, y colócalos en un software estadístico como *Statgraphics*, SPSS o SAS. Elige algún modelo ARMA( $p,q$ ) y obtén los pronósticos a corto, mediano y largo plazo. Anota tus observaciones acerca de las gráficas resultantes.

## Capítulo 5

# Metodología de Box-Jenkins

Si Dios hubiera esperado a tener todo perfecto, no estaríamos aquí.

J. A. CRONIN

### 5.1 Introducción

Hasta el momento, se han revisado las características de los modelos teóricos de tipo ARIMA( $p,d,q$ ) que representan un conjunto amplio de posibilidades para el analista. Con esta revisión se tiene un panorama de los modelos y sus características esenciales. El siguiente paso consiste en confrontar los datos reales con estos modelos, de una manera sistemática y ordenada que permita elegir una opción que satisfaga las expectativas en cuanto a los pronósticos.

Conviene recordar aquí que la metodología de Box-Jenkins se basa en dos principios para el modelado: el de parsimonia y el de aproximaciones sucesivas. Por ello, el primer objetivo será identificar un **modelo tentativo o preliminar** para las observaciones. Posteriormente será posible mejorarlo y construir otros modelos para que, de todos ellos, se elija el que más convenga.

Con esta finalidad, la metodología ofrece el siguiente procedimiento:

1. En primera instancia debe verificarse que la serie sea estacionaria e intentar eliminar la variación estacional, si ésta existe. Para ello se sugiere:
  - a) Revisar la posible existencia de heteroscedasticidad y, en su caso, efectuar la transformación apropiada.
  - b) Observar si se presenta algún tipo de tendencia. En este caso, hacer las diferencias ordinarias que se requieran.

## *5.1 Introducción*

---

- c) Determinar si el fenómeno contiene alguna variación estacional que, en su caso, debe corresponder a una interpretación lógica. Si esto ocurre, debe probarse la aplicación de diferencias estacionales.
2. Una vez que se cuenta con una serie estacionaria, deberá identificarse un modelo tentativo, de acuerdo con las características de la serie. Cabe señalar que, si existieron varias formas alternativas de conseguir la estacionariedad, será necesario continuar con el procedimiento para cada una de ellas, de tal manera que al final, al obtener los pronósticos, pueda elegirse el mejor modelo. Para efectuar la identificación de un modelo, conviene:
  - a) Observar la gráfica de la ACF muestral y determinar si es decreciente infinita o si se trunca y, en este último caso, en qué valor ocurre esto. Para ello se contará también con pruebas de hipótesis para cada autocorrelación y para un conjunto de ellas.
  - b) Revisar la gráfica de la PACF muestral y determinar si es decreciente infinita o si se trunca y, en este último caso, en qué valor ocurre esto. Para ello se contará también con pruebas de hipótesis para cada autocorrelación parcial.
  - c) Analizar tanto el periodograma como periodograma integrado, de manera conjunta con ACF y PACF, para evaluar si existe algún tipo de variación estacional no eliminada por las diferencias estacionales. A través de los valores numéricos del periodograma podrá constatarse el orden de esta variación, que debe coincidir con las expectativas del analista. En caso de obtener algún valor sorprendente, habría que analizar si se ha descubierto algo o si simplemente existe algún error en el procedimiento.
  - d) Con los resultados de los análisis anteriores, el analista podrá formular uno o varios modelos tentativos. Cada modelo tentativo señala un camino que puede dar resultados interesantes, por lo cual habrá que seguirlo hasta el último paso, es decir, hasta calcular los pronósticos correspondientes.
3. Una vez identificado, para cada modelo tentativo habrán de estimarse los parámetros, es decir, el término constante y los coeficientes autorregresivos y de medias móviles que correspondan al orden de cada parte. Esto se hace, por lo general, a través del **método de máxima verosimilitud** que, en este caso, resulta equivalente al **método de mínimos cuadrados**.

**cuadrados.** Este paso se realiza con el apoyo de una aplicación de software y requiere de las siguientes acciones:

- a) Puesto que el método de estimación será de tipo numérico e involucra la utilización de los datos que hemos llamado prehistóricos, será necesario asignar valores a estas variables. Esto se llama **inicialización de la serie de tiempo**.
  - b) Por otro lado, el modelo generará ecuaciones donde el valor que se desea minimizar, es decir, el error, presenta una forma no lineal. Con objeto de linealizarlo, como se verá más adelante, será necesario contar con valores iniciales de los parámetros que deben estimarse.
4. Una vez que se ha elegido el modelo tentativo y se han estimado sus parámetros, habrá que diagnosticar el modelo. Para ello se recomienda efectuar el análisis siguiente:
- a) Aunque pareciera repetitivo, verificar nuevamente que los datos cumplan con las condiciones de estacionariedad. Si esto no se satisface, el modelo no resultará funcional.
  - b) Revisar si el modelo tentativo pudiera estar sobre especificado, es decir, si le sobran variables que no aportan un mejor ajuste. Esto se hará a través de pruebas de hipótesis para cada uno de los parámetros estimados.
  - c) Revisar si el modelo tentativo pudiera estar subespecificado, esto es, si le hacen falta variables para mejorar su ajuste a los datos. Esto se hará a través de evaluar el comportamiento de los residuales, que debieran aproximarse al ruido blanco. En caso de no ser así, sus propiedades indicarán con bastante precisión qué es lo que debe agregarse al modelo.
  - d) Observar en la gráfica original si existen **datos atípicos** (*outliers*). Se trata de valores cuyo comportamiento es muy distinto de los demás. Como se verá adelante, pueden ser atípicos influyentes o discrepantes. Aquí jugará un papel importante el conocimiento del analista sobre el fenómeno que se estudia.
  - e) Una vez que se haya agotado este proceso, elaborar un cuadro comparativo con todos los modelos factibles. Este cuadro permitirá juzgar la bondad de cada uno de ellos y elegir el que tenga las mejores características.

5. Cuando se cuente ya con un modelo apropiado, será posible cumplir el objetivo: pronosticar los valores de la serie de tiempo para un horizonte determinado. En este momento corresponderá al analista ofrecer los valores puntuales y por intervalo, así como su interpretación, recomendaciones, discusión y sugerencias.

En las siguientes secciones se describen con detalle los procedimientos de identificación, estimación, diagnóstico y pronóstico, con el auxilio del programa *Statgraphics* y algunos ejemplos seleccionados.

## 5.2 Identificación

El procedimiento para identificar el modelo adecuado para una serie de tiempo observada, consiste en contar con una especie de “catálogo” de los diversos modelos ARMA y sus características esenciales, para compararlas con las de los datos reales y seleccionar las más parecidas.

Los elementos básicos que permiten la identificación de un modelo son:

- La función de autocorrelación (ACF)
- La función de autocorrelación parcial (PACF)
- El periodograma
- El periodograma integrado

Puesto que ya se ha revisado la interpretación tanto del periodograma como del periodograma integrado, se analizará ahora la forma de utilizar las funciones de autocorrelación y autocorrelación parcial.

### 5.2.1 Interpretación de la ACF

En general, se desea determinar si debe considerarse que la ACF es decreciente infinita o si se trunca en un cierto valor de  $k$ . Para ello, se utiliza en primera instancia una inspección visual de la gráfica (correlograma), auxiliada por un intervalo de confianza que permitirá probar la siguiente hipótesis para cada una de las autocorrelaciones teóricas  $\rho_k$ :

$$\begin{aligned} H_0: \rho_k &= 0 \\ &\text{versus} \\ H_a: \rho_k &\neq 0 \end{aligned}$$

La primera se llama **hipótesis nula**, ya que representa la igualdad o nula diferencia entre el valor del parámetro y el cero; mientras que la segunda se denomina **hipótesis alternativa**, ya que representa la circunstancia que se cumplirá en caso de no considerar aceptable la primera hipótesis.

De acuerdo con lo anterior, si se acepta la hipótesis nula, se considera que la autocorrelación correspondiente al intervalo  $k$  es estadísticamente insignificante, es decir, no hay autocorrelación entre  $Y_t$  y  $Y_{t-k}$ . En el caso contrario, si se rechaza  $H_0$ , la autocorrelación  $\rho_k$  es **significativa** o relevante.

Para realizar la **prueba de hipótesis** es necesario construir un estadístico (o estadística) de prueba que relacione el valor del estimador con el valor supuesto del parámetro, y que siga una distribución de probabilidad conocida. En este caso el estadístico se obtiene de la siguiente forma:

$$t_{r_k} = \frac{r_k - 0}{s_{r_k}} \quad (5.1)$$

Donde la varianza del estimador  $r_k$  [9] está dada por:

$$\text{Var}(r_k) \approx \frac{1}{N} [1 + 2(\rho_1^2 + \rho_2^2 + \dots + \rho_q^2)] \quad \forall k > q \quad (5.2)$$

Por lo tanto, el estimador de dicha varianza será:

$$\hat{\text{Var}}(r_k) = s_{r_k}^2 \approx \frac{1}{N} [1 + 2(r_1^2 + r_2^2 + \dots + r_q^2)] \quad \forall k > q \quad (5.3)$$

Ahora, el estadístico formulado en la ecuación 5.1 sigue aproximadamente una distribución de probabilidad t-de-Student, con tantos grados de libertad como datos disponibles, para un nivel de confianza de  $(1 - \alpha)100\%$ . Puesto que la metodología de Box-Jenkins requiere de contar por lo menos con ochenta datos, para que sea posible detectar tendencia y variación estacional, así como hacer diferencias de cualquier tipo sin reducir demasiado el tamaño de muestra, puede considerarse que el valor de la t-de-Student, para  $\alpha = 0.05$  es 1.96 o, aproximadamente, 2.

Por ello, puede establecerse como regla de decisión la siguiente:

- Aceptar  $H_0$  si  $|t_{r_k}| < 2$ , es decir, considerar que  $\rho_k$  es estadísticamente insignificante.
- Rechazar  $H_0$  en caso contrario, es decir, considerar que  $\rho_k$  es significativo.

Resulta conveniente contar con una herramienta visual para efectuar este análisis de manera rápida. Puede utilizarse la prueba de hipótesis anterior para construir un intervalo de confianza que se agrega a la gráfica de la ACF, con el objeto de identificar rápidamente cuáles valores sobresalen de dicho intervalo, que se construye como:

$$P[\rho_k - 2s_{r_k} < r_k < \rho_k + 2s_{r_k}] = 0.95 \quad (5.4)$$

Así, pueden agregarse a la gráfica de la ACF los límites inferior y superior del intervalo de confianza, para la hipótesis, de modo que será evidente de

## 5.2 Identificación

---

manera inmediata cuáles son las  $r_k$  cuyo valor queda por fuera de  $[-2s_{r_k}, 2s_{r_k}]$  y que son, por lo tanto, significativas con un 95 % de confianza.

Por otro lado, también puede resultar útil probar si un conjunto de  $K$  autocorrelaciones son o no todas insignificantes, es decir, probar la hipótesis:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_K = 0 \\ versus \\ H_a: \text{Al menos una } \rho_k \neq 0, \forall k \leq K.$$

Para ello se utiliza el **estadístico de Box-Pierce**[13], con el cual se efectúa la llamada prueba *Portmanteau*<sup>1</sup>. Dicho estadístico es:

$$Q(K) = N(N+2) \sum_{k=1}^K \frac{r_k^2}{N-k} \quad (5.5)$$

Que se distribuye aproximadamente como una  $\chi^2$  (ji-cuadrada) con  $K - p - q - 1$  grados de libertad, para un nivel de confianza del  $(1 - \alpha)100\%$ , donde  $p$  es el orden de la parte autorregresiva y  $q$  el orden de la parte de medias móviles. Se resta además uno por la estimación de la constante. Por lo general los paquetes estadísticos utilizan  $K = 12$  o  $K = 24$ , aunque el analista puede modificar estos valores.

La regla de decisión se formula como:

- Aceptar  $H_0$  si  $Q(K) < \chi_{N,1-\alpha}^2$ . En este caso, se considera que las  $K$  primeras autocorrelaciones son estadísticamente insignificantes, con lo cual puede considerarse aceptable que los datos analizados sean ruido blanco, a reserva de verificar esto con otras pruebas adicionales.
- En caso contrario, rechazar  $H_0$ , por lo cual se considera que al menos una de las autocorrelaciones es significativa. Para identificar cuál o cuáles, se utiliza la prueba descrita anteriormente.

Existe una regla de decisión alternativa que utiliza la probabilidad de que la hipótesis nula sea verdadera, llamada **P-value**: Se acepta  $H_0$  si este valor es mayor que  $\alpha$  y se rechaza en caso contrario.

### 5.2.2 Interpretación de la PACF

En general, se desea determinar si debe considerarse que la PACF es decreciente infinita o si se trunca en un cierto valor de  $k$ . Para ello, se utiliza una inspección visual de la gráfica correspondiente, auxiliada por un intervalo de confianza que permitirá probar la siguiente hipótesis para cada una de las autocorrelaciones parciales  $\rho_{kk}$ :

<sup>1</sup>Significa “maleta” en francés, ya que contiene a un conjunto de autocorrelaciones.

$$\begin{aligned} H_0: \rho_{kk} &= 0 \\ &\text{versus} \\ H_a: \rho_{kk} &\neq 0 \end{aligned}$$

Si se acepta la hipótesis nula, se considera que la autocorrelación parcial correspondiente al intervalo  $k$  es estadísticamente insignificante, es decir, la variable autorregresiva  $Y_{t-k}$  no tiene aportación valiosa para el modelo AR( $k$ ). En el caso contrario, la autocorrelación parcial correspondiente es significativa.

Para realizar la prueba de hipótesis es necesario construir un estadístico que relacione el valor del estimador con el valor supuesto del parámetro y que siga una distribución de probabilidad conocida. El estadístico se obtiene de la siguiente forma:

$$t_{r_{kk}} = \frac{r_{kk} - 0}{s_{r_{kk}}} \quad (5.6)$$

Donde la varianza del estimador  $r_k$  [43] está dada por:

$$\text{Var}(r_{kk}) \approx \frac{1}{N} \quad \forall k > p \quad (5.7)$$

Por lo tanto, el estimador de dicha varianza será igual al valor anterior:

$$\hat{\text{Var}}(r_{kk}) = s_{r_{kk}}^2 \approx \frac{1}{N} \quad \forall k > p \quad (5.8)$$

Ahora, el estadístico formulado en la ecuación 5.1 sigue aproximadamente una distribución t-de-Student, con tantos grados de libertad como datos disponibles, para un nivel de confianza de  $(1-\alpha)100\%$ . Puesto que la metodología de Box-Jenkins requiere de contar por lo menos con 80 datos, para que sea posible detectar tendencia y variación estacional, así como hacer diferencias sin reducir demasiado el tamaño de muestra, puede considerarse que el valor en tablas de la t-de-Student, para  $\alpha = 0.05$  es 1.96 o, aproximadamente, 2.

Por ello, puede establecerse como regla de decisión la siguiente:

1. Aceptar  $H_0$  si  $|t_{r_{kk}}| < 2$ , es decir, considerar que  $\rho_{kk}$  es estadísticamente insignificante.
2. Rechazar  $H_0$  en caso contrario, es decir, considerar que  $\rho_{kk}$  es significativo.

Resulta conveniente contar con una herramienta visual para efectuar este análisis de manera rápida. Puede utilizarse la prueba de hipótesis anterior para construir un intervalo de confianza que se agrega a la gráfica de la PACF, con el objeto de identificar rápidamente cuáles valores sobresalen de dicho intervalo que se construye como:

$$P[\rho_{kk} - 2s_{r_{kk}} < r_{kk} < \rho_{kk} + 2s_{r_{kk}}] = 0.95 \quad (5.9)$$

Así, puede agregarse este intervalo a la gráfica de la ACF, de modo que será evidente de manera inmediata cuáles son las  $r_{kk}$  cuyo valor queda por fuera de  $[-2s_{r_{kk}}, 2s_{r_{kk}}]$  y que son, por lo tanto, significativas con un 95 % de confianza.

### 5.2.3 Ejemplos de interpretación de ACF y PACF

**Ejemplo 1.** Supongamos que una serie de tiempo con 80 observaciones presenta la gráfica original de la Figura 5.1; la ACF de la Figura 5.2; la PACF de la Figura 5.3; y el periodograma integrado de la Figura 5.4. ¿Cómo se interpretan estas gráficas y cuál sería el modelo tentativo para estos datos?

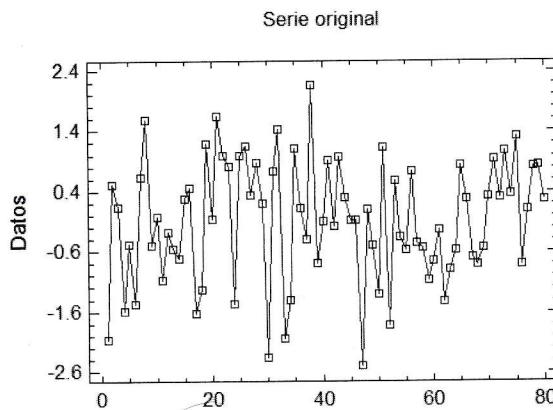


Figura 5.1: Gráfica original del Ejemplo 1.

Además, el software arroja el siguiente resultado para la prueba *Portman-teau* o de Box-Pierce:

```
Box-Pierce Test
Test based on first 24 autocorrelations
Large sample test statistic = 21.5952
P-value = 0.603417
```

**Interpretación** En la gráfica original no se aprecia tendencia o algún tipo de periodicidad, lo cual se refuerza con la imagen del periodograma integrado. Tanto en la ACF como en la PACF todos los valores están dentro de los respectivos intervalos de confianza. Además, el estadístico de Box-Pierce indica que puede aceptarse la hipótesis nula de que las primeras 24 autocorrelaciones son todas insignificantes, ya que la probabilidad de que la hipótesis nula sea verdadera (*P-value*) es mayor que  $\alpha = 0.05$ .

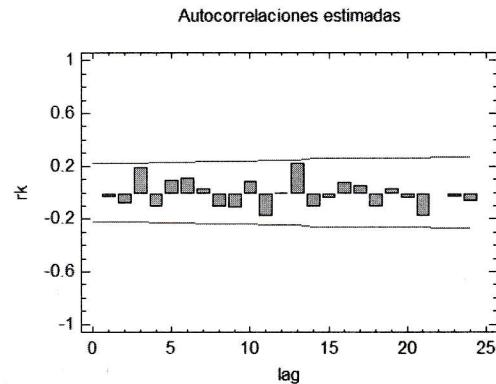


Figura 5.2: ACF del Ejemplo 1.

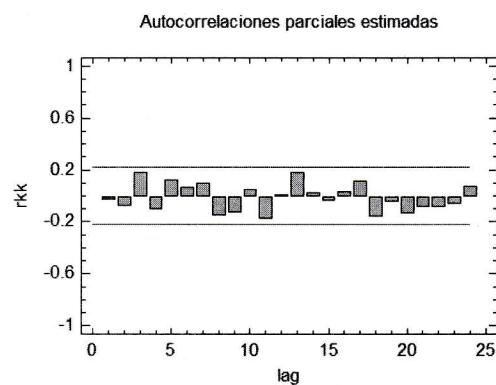


Figura 5.3: PACF del Ejemplo 1.

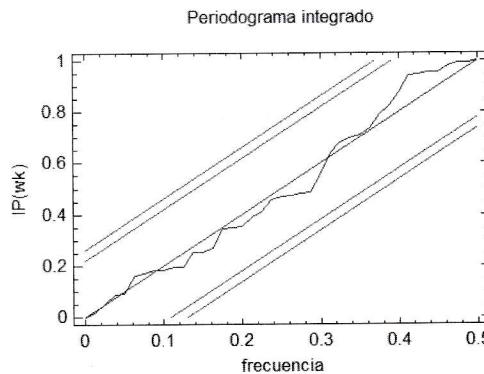


Figura 5.4: Periodograma integrado del Ejemplo 1.

Todo esto sugiere que la serie de tiempo es sencillamente ruido blanco, es decir, variables aleatorias independientes, con media cero y varianza constante. Por ello, el único modelo factible a considerar sería un ARMA(0,0), es decir,  $Y_t = e_t$  o ruido blanco.

Cabe destacar que éste es el tipo de resultados que se deseará encontrar en los residuales de un modelo, ya que así se mostrará un buen ajuste de dicho modelo a los datos.

**Ejemplo 2.** Supongamos que una serie de tiempo con 80 observaciones presenta la gráfica original de la Figura 5.5; la ACF de la Figura 5.6; la PACF de la Figura 5.7; y el periodograma integrado de la Figura 5.8. ¿Cómo se interpretan estas gráficas y cuál sería el modelo tentativo para estos datos?

Además, el software arroja el siguiente resultado para la prueba *Portmanteau* o de Box-Pierce:

```
Box-Pierce Test
Test based on first 24 autocorrelations
Large sample test statistic = 49.6996
P-value = 0.00154595
```

**Interpretación** En la gráfica original no se aprecia tendencia o algún tipo de periodicidad, lo cual se refuerza con la imagen del periodograma integrado. En la ACF sobresalen del intervalo de confianza el valor correspondiente a  $k = 1$ . Además, la prueba de Box-Pierce indica que no es aceptable considerar que las primeras 24 autocorrelaciones son ceros, ya que la probabilidad de que la hipótesis nula sea verdadera (*P-value*) es menor que  $\alpha = 0.05$ . En la PACF destacan los valores correspondientes a  $k = 1$  y  $k = 2$ .

Así pues, las imágenes sugieren que un primer modelo tentativo a probar

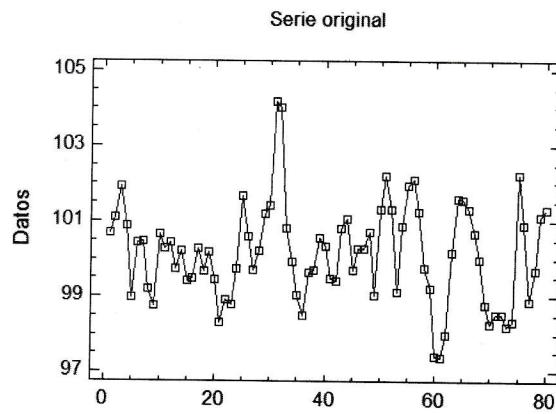


Figura 5.5: Gráfica original del Ejemplo 2.

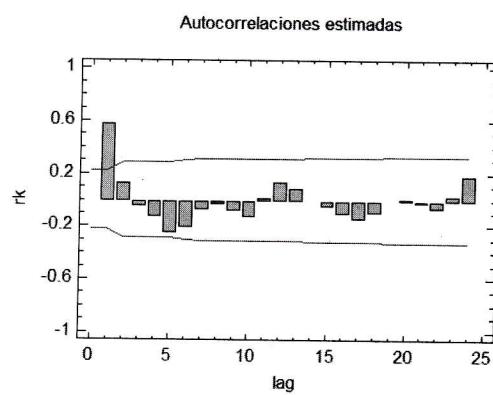


Figura 5.6: ACF del Ejemplo 2.

## 5.2 Identificación

---

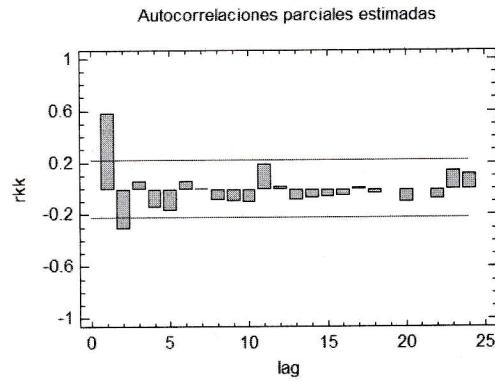


Figura 5.7: PACF del Ejemplo 2.

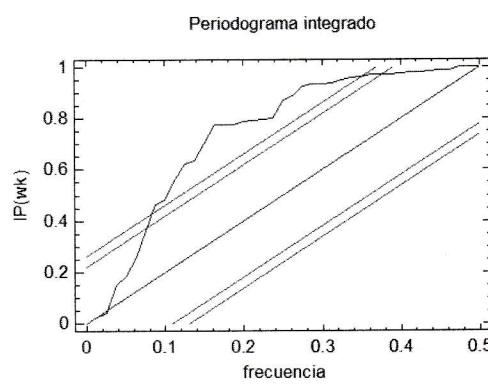


Figura 5.8: Periodograma integrado del Ejemplo 2.

sería el MA(1), si se considera que la ACF se trunca en  $k = 1$  y la PACF es decreciente infinita.

Por supuesto, existen otros modelos factibles como un ARMA(1,1) o un AR(2), sin embargo, es deseable comenzar con el más sencillo para analizar después si está sub estimado. En este caso el modelo más económico es el MA(1). Cabe señalar también que el periodograma integrado muestra únicamente una falta de ajuste, ya que está fuera del intervalo de confianza.

#### 5.2.4 Identificación con *Statgraphics*

Como se ha visto, el proceso de identificación puede apoyarse en algún software estadístico, como es el caso de *Statgraphics*. Para ello, la aplicación ofrece el análisis descriptivo que se muestra en el menú **Describe / Time Series / Descriptive Methods**. Una vez realizada la identificación, será necesario indicar al software el modelo elegido. Para ello se utilizará ahora la sección **Forecast** de la barra principal de opciones y, dentro de ella, se elige **User-Specified Model...**, como se muestra en la Figura 5.9.

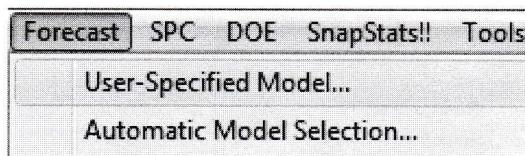


Figura 5.9: Pronósticos en *Statgraphics*.

Al hacer esta elección, el software presentará nuevamente la ventana para elegir los datos de la serie de tiempo. Después de seleccionar la serie, se ofrecerá un modelo por omisión, no indicado por el usuario. Por lo general se verá el modelo de **caminata aleatoria** (*random walk*), que tiene como pronóstico el dato más reciente de la serie.

Para especificar el modelo deseado es necesario dar clic derecho sobre cualquier parte de los resultados y luego marcar **Analysis Options**, con lo cual se dispondrá de la pantalla llamada **Model Specification Options**, que se muestra en la Figura 5.10, y en la que se ha seleccionado ya la opción **ARIMA Model**. Al hacer esto quedarán disponibles las ventanas **ARy MA**, en las cuales se colocará el orden de la parte autorregresiva y de medias móviles, respectivamente. En estas ventanas se escriben los valores de  $p$  y  $q$  del modelo a probar. Puede observarse que las opciones relacionadas con la variación estacional sólo se activarán si se indica el orden de dicha variación en la pantalla de selección de datos.

### 5.3 Estimación

En esta sección se discute el proceso por medio del cual se estiman los parámetros de un modelo  $\text{ARIMA}(p,d,q)$ . En general, para la estimación

### 5.3 Estimación

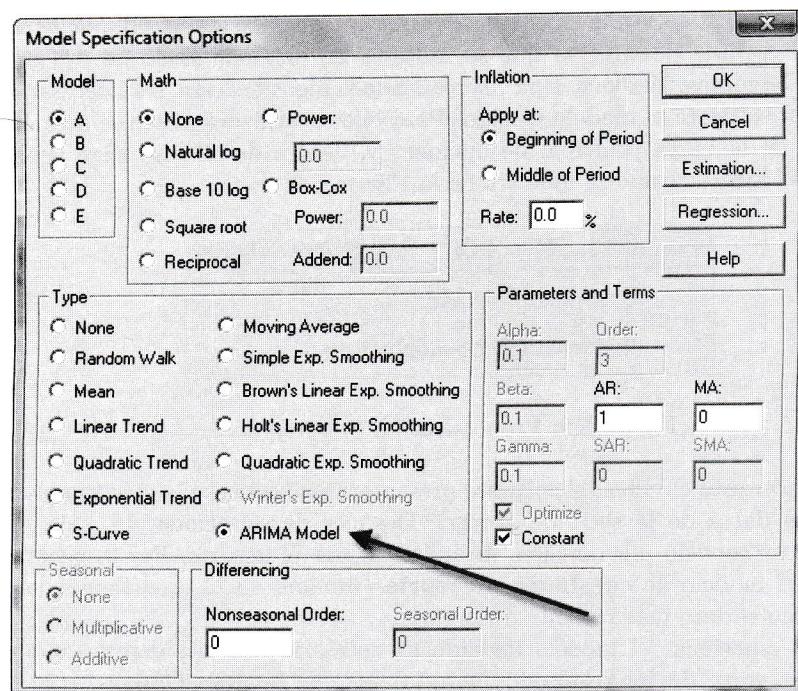


Figura 5.10: Especificación de modelos.

de parámetros puede usarse algún método como: momentos, máxima verosimilitud o mínimos cuadrados. De acuerdo con William S. Wei [58], lo más recomendable es utilizar máxima verosimilitud que resultará equivalente a mínimos cuadrados, como se verá más adelante. Sin embargo, como este procedimiento generará un sistema de ecuaciones no lineales, será necesario echar mano de otras herramientas matemáticas para linealizar estas expresiones y llegar a un resultado satisfactorio.

A continuación se verifica la equivalencia entre el método de máxima verosimilitud y el de mínimos cuadrados, y se detallan los pasos que comprende la estimación.

### 5.3.1 Mínimos cuadrados y máxima verosimilitud

Una vez que se ha realizado la especificación del modelo tentativo, es decir, una vez elegidos los valores del orden autorregresivo:  $p$ ; las diferencias ordinarias:  $d$ ; el orden de medias móviles:  $q$ , y las diferencias estacionales:  $D$ , se tendrá el modelo general que se muestra en la ecuación 5.10.

$$\phi(B)Z_t = \delta_0 + \theta(B)e_t \quad (5.10)$$

Ahora deben obtenerse los **estimadores** para los  $p + q + 1$  **parámetros** desconocidos:  $\delta_0, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ .

Si se aplica el método numérico de **mínimos cuadrados**, deberán encontrarse aquellos estimadores que minimicen la suma de cuadrados de las diferencias entre la serie real  $Z_t$  y la serie ajustada por el modelo  $\hat{Z}_t$ . Esto es, si se considera cada error igual al valor real de la observación menos el valor estimado por el modelo para este mismo punto, esta diferencia puede obtenerse despejando el error  $e_t$  de la ecuación 5.10:

$$e_t = \theta^{-1}(B)(\phi(B)Z_t - \delta_0) \quad (5.11)$$

Donde la ecuación 5.11 representará a los **residuales**, una vez que se sustituyan los parámetros por sus valores estimados respectivos.

El objetivo de la estimación será entonces encontrar un vector de parámetros autorregresivos:  $\phi = [\phi_1, \dots, \phi_p]$ ; un vector de parámetros de medias móviles:  $\theta = [\theta_1, \dots, \theta_q]$ ; y una constante  $\delta_0$ , tales que minimicen la suma de los errores al cuadrado, denotada por:

$$S(\delta_0, \phi, \theta) = \sum_{t=d+Ds+1}^N e_t^2 \quad (5.12)$$

Así, los vectores que minimizan la ecuación 5.12 pueden escribirse como:

$$\hat{\phi} = [\hat{\phi}_1, \dots, \hat{\phi}_p]$$

$$\hat{\theta} = [\hat{\theta}_1, \dots, \hat{\theta}_p]$$

### 5.3 Estimación

---

Y los residuales asociados con dichos estimadores serán:

$$\hat{e}_t = \hat{\theta}^{-1}(B) (\hat{\phi}(B) Z_t - \hat{\delta}_0) \quad (5.13)$$

Por lo tanto, la **suma de los cuadrados de los residuales** se expresará por:

$$S(\hat{\delta}_0, \hat{\phi}, \hat{\theta}) = \sum_{t=d+Ds+1}^N \hat{e}_t^2 \quad (5.14)$$

Puede verse ahora que el proceso de estimación de los parámetros presentará dos problemas:

1. El primer error de la serie de tiempo,  $e_1$ , depende de valores pasados no observados (lo que hemos llamado prehistoria), es decir, depende de:  $Z_0, Z_{-1}, \dots, Z_{-p+1}$  y de  $e_0, e_{-1}, \dots, e_{-q+1}$ . Esto hará necesario el uso de algún método que permita inicializar la serie, antes de aplicar el proceso de estimación no lineal.
2. Como se verá en la siguiente sección, la estimación de 5.14 se complica si están presentes términos de medias móviles, ya que la ecuación 5.11 es no lineal con respecto a los parámetros  $\theta(B)$ . Por esta razón será necesario usar un método iterativo de estimación no lineal.

Una vez estimados los parámetros del modelo, deberá verificarse si los residuales correspondientes se comportan como ruido blanco o si es necesario hacer algún ajuste. Esto se hará revisando los residuales y sus características.

#### 5.3.2 Procedimiento de estimación

El problema ahora es estimar los vectores de parámetros  $\phi$  y  $\theta$ , más la constante  $\delta_0$ . Para ello se supone que los errores  $e_t$  son variables aleatorias independientes idénticamente distribuidas como Normal con media cero y varianza constante, es decir, son ruido blanco. Esto ocurrirá precisamente cuando el modelo seleccionado sea apropiado para los datos reales.

Entonces la **función de verosimilitud condicional** asociada con los parámetros  $(\delta_0, \phi, \theta, \sigma)$  estará dada por:

$$\begin{aligned} F(\delta_0, \phi, \theta, \sigma) &= \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^{N-d-Ds} \exp \frac{-S(\hat{\delta}_0, \hat{\phi}, \hat{\theta})}{2\sigma^2} \\ &= (2\pi\sigma^2)^{-(N-d-Ds)/2} \exp \frac{-S(\hat{\delta}_0, \hat{\phi}, \hat{\theta})}{2\sigma^2} \end{aligned}$$

Para simplificar los cálculos y dado que maximizar el logaritmo de una función equivale a maximizar la función original, conviene tomar el logaritmo

natural, con lo cual se obtiene la **función de verosimilitud logarítmica condicional**:

$$L(\delta_0, \phi, \theta, \sigma) = \frac{-(N - d - Ds) \ln(2\pi\sigma^2)}{2} - \frac{S(\hat{\delta}_0, \hat{\phi}, \hat{\theta})}{2\sigma^2} \quad (5.15)$$

Como puede verse, el valor de la ecuación 5.15 está condicionado tanto a los estimadores como a los datos no observados o prehistóricos que se requerirán para calcular los residuales. Ahora el objetivo es encontrar los estimadores que maximicen esta función de verosimilitud. Es sencillo observar que los mejores estimadores serán aquellos que minimicen  $S(\hat{\delta}_0, \hat{\phi}, \hat{\theta})$ , puesto que es una suma de cuadrados que tiene signo negativo, y el resto de la ecuación son valores constantes. Esto lleva a concluir que los estimadores de máxima verosimilitud serán los mismos que los obtenidos por mínimos cuadrados. En la siguiente sección se revisa la forma de inicializar los valores prehistóricos.

### 5.3.3 Inicialización de la serie de tiempo

Ya que la función suma de mínimos cuadrados  $S(\hat{\delta}_0, \hat{\phi}, \hat{\theta})$  requiere de valores pasados no observables, tanto de  $Z_t$  como de  $e_t$ , es necesario estimarlos de alguna manera. Para ello pueden considerarse dos caminos:

1. **Valores esperados incondicionales.** Una primera solución es sustituir los valores desconocidos  $Z_0, Z_{-1}, \dots, Z_{-p+1}$  y  $e_0, e_{-1}, \dots, e_{q+1}$  por sus valores esperados incondicionales, es decir,  $\mu$  y cero, respectivamente. Si la serie es estacionaria y se cuenta con una muestra relativamente grande, esta alternativa brindará resultados razonablemente adecuados.
2. **Valores esperados condicionales.** La segunda solución consiste en tomar como punto de partida los valores esperados incondicionales y, a partir de ellos, estimar los parámetros del modelo. Hecho esto, se utiliza el modelo obtenido para efectuar un **pronóstico hacia atrás** (*backcasting* o *backforecasting*) y obtener valores de los datos desconocidos. Para ello se define el **operador de salto hacia adelante**  $F$  como:

$$FZ_t = Z_{t+1} \quad (5.16)$$

Con lo cual el pronóstico hacia atrás se obtiene a través del modelo:

$$\phi(F)Z_t = \delta_0 + \theta(F)e_t \quad (5.17)$$

Este procedimiento puede repetirse hasta que ocurra la primera de tres alternativas:

### 5.3 Estimación

---

- a) Los valores estimados de los parámetros prácticamente no cambien, es decir, la diferencia entre una iteración y la siguiente sea, en valor absoluto, menor que un error específico  $\epsilon_1$ .
- b) La suma de cuadrados de los residuales prácticamente no cambie, es decir, la diferencia entre una iteración y la siguiente sea, en valor absoluto, menor que un error específico  $\epsilon_2$ .
- c) Se concluya un número máximo de iteraciones, en caso de que ninguna de las opciones anteriores sea convergente.

Los valores esperados condicionales son adecuados cuando la serie de tiempo se acerca a la no estacionaridad, la muestra es relativamente pequeña o ambas cosas.

En general no es posible saber de antemano cuál de las dos alternativas anteriores (valores esperados incondicionales o condicionales) generará los mejores resultados para el modelo. Por ello, se sugiere probar la estimación por ambas vías y finalmente, a través de los resultados generados, decidir cuál de las dos se prefiere. Los paquetes estadísticos que ya se han mencionado ofrecen estas dos posibilidades.

#### 5.3.4 Estimación no lineal de los parámetros

Como se ha explicado, se desea encontrar valores de los parámetros  $\delta_0$ ,  $\phi$  y  $\theta$  que minimicen la suma de cuadrados de los errores, condicionada a dichos parámetros, así como a la serie de tiempo observada y no observada, es decir, que minimicen la función:

$$S(\hat{\delta}_0, \hat{\phi}, \hat{\theta}) = \sum_{t=d+Ds+1}^N [e_t / \delta_0, \phi, \theta, Z]^2 \quad (5.18)$$

Donde  $Z$  es el vector de los datos observados —y no observados—  $Z_t$ .

Si se supone un modelo completamente autorregresivo, es decir, de la forma:

$$\phi(B)Z_t = \delta_0 + e_t \quad (5.19)$$

Que también puede expresarse como:

$$Z_t = \delta_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + e_t \quad (5.20)$$

Puede observarse que la ecuación 5.20 sigue exactamente la estructura de un modelo de **regresión lineal múltiple**, es decir, es de la forma:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_p X_{pt} + e_t \quad (5.21)$$

Por lo tanto, los parámetros de la ecuación 5.20 pueden estimarse utilizando la solución matricial correspondiente a un modelo de regresión lineal múltiple, por lo cual:

$$\hat{\phi} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (5.22)$$

Donde la matriz  $\mathbf{X}$  contiene a las **variables independientes** o **exógenas**:

$$\mathbf{X} = \begin{bmatrix} Z_p & Z_{p-1} & \dots & Z_1 \\ Z_{p+1} & Z_p & \dots & Z_2 \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N-d-sD-1} & Z_{N-d-sD-2} & \vdots & Z_{N-d-sD-p} \end{bmatrix} \quad (5.23)$$

Mientras que el vector  $\mathbf{Y}$  contiene a las **variables dependientes** o **endógenas**:

$$\mathbf{Y} = \begin{bmatrix} Z_{p+1} \\ Z_{p+2} \\ \vdots \\ Z_{N-d-sD} \end{bmatrix} \quad (5.24)$$

Como puede verse, para un modelo que es únicamente autorregresivo, el proceso de estimación de los parámetros es esencialmente una regresión lineal múltiple. Sin embargo, si el modelo contiene medias móviles, no será posible aplicar este tipo de estimación. Obsérvese que si se obtiene el error  $e_t$  a partir del modelo general ARIMA, los parámetros  $\theta(B)$  resultan no lineales:

$$e_t = \theta^{-1}(B)\phi(B)Z_t + \delta_0 \quad (5.25)$$

A pesar de ello, la ecuación 5.25 puede linearse a través de una rutina iterativa, utilizando los dos primeros términos de una **serie de Taylor** [16]. La **linealización** se hará alrededor de algún valor inicial de cada uno de los elementos de los vectores  $\phi$  y  $\theta$ . Hecho esto, se efectúa la regresión con la ecuación linealizada, con lo cual se obtendrán estimadores de mínimos cuadrados para  $\phi$  y  $\theta$ , mismos que se utilizarán como punto de partida para una nueva linealización, y así sucesivamente, hasta que ocurra la convergencia o se concluya un número de iteraciones preestablecido.

Para observar este proceso iterativo con más detalle, conviene generar un nuevo vector  $\beta$  que contendrá los  $p + q + 1$  parámetros que van a estimarse, es decir:  $[\delta_0, \phi, \theta]$ . Así pues, el objetivo será elegir los valores del vector  $\beta$  que minimicen:

$$S(\beta) = \sum_{t=1}^{N-d-sD} [e_t / \mathbf{Z}, \beta]^2 = \sum_{t=1}^{N-d-sD} [e_t]^2 \quad (5.26)$$

Ahora puede hacerse la expansión de  $[e_t]$  alrededor de un **valor inicial**

### 5.3 Estimación

---

del vector, es decir, de  $\beta_0$ , de la siguiente forma:

$$\begin{aligned}[e_t] &= [e_t / \mathbf{Z}, \beta_0] + \sum_{i=1}^{p+q+1} (\beta_i - \beta_{i,0}) \frac{\partial[e_t]}{\partial\beta_i} \Big|_{\beta=\beta_0} + \\ &+ \frac{1}{2} \sum_{i=1}^{p+q+1} (\beta_i - \beta_{i,0})^2 \frac{\partial^2[e_t]}{\partial\beta_i^2} \Big|_{\beta=\beta_0} + \dots \end{aligned}\quad (5.27)$$

Donde  $\beta_{i,0}$  es el valor inicial para el parámetro desconocido  $\beta_i$ , es decir, es el  $i$ -ésimo componente del vector de valores iniciales de los parámetros  $\beta_0$ . La serie de Taylor de la ecuación anterior se trunca en el segundo término para obtener una ecuación lineal.

Sean ahora:

$$x_{i,t} = - \frac{\partial[e_t]}{\partial\beta_i} \Big|_{\beta=\beta_0} \quad (5.28)$$

$$[e_{t,0}] = [e_t / \mathbf{Z}, \beta_0] \quad (5.29)$$

Se sigue entonces que la ecuación 5.27 puede escribirse como:

$$[e_t] \approx [e_{t,0}] - \sum_{i=1}^{p+q+1} (\beta_i - \beta_{i,0}) x_{i,t} \quad (5.30)$$

Lo cual a su vez puede reescribirse como un modelo de regresión lineal:

$$[e_{t,0}] + \sum_{i=1}^{p+q+1} \beta_{i,0} x_{i,t} \approx \sum_{i=1}^{p+q+1} \beta_i x_{i,t} + [e_t] \quad (5.31)$$

El lado izquierdo de la ecuación anterior puede verse como una variable dependiente compuesta, que tendrá diferentes valores para  $t = 1, 2, \dots, N - d - sD$ . Obsérvese que  $[e_{t,0}]$  es el valor del residual correspondiente al tiempo  $t$ , dado el vector inicial  $\beta_0$ . Del lado derecho de la ecuación hay  $p + q + 1$  variables independientes, multiplicadas por el mismo número de parámetros desconocidos  $\beta_i$ , así como un error aditivo.

Puesto que se tiene ya un modelo de regresión lineal múltiple, los parámetros  $\beta_i$  pueden estimarse con la ecuación 5.22, en la cual:

$$\mathbf{Y} = \begin{bmatrix} [e_{N-d-sD+1,0}] + \sum_{i=1}^{p+q+1} \beta_{i,0} x_{i,N-d-sD+1} \\ [e_{N-d-sD+2,0}] + \sum_{i=1}^{p+q+1} \beta_{i,0} x_{i,N-d-sD+2} \\ \vdots \\ [e_{N,0}] + \sum_{i=1}^{p+q+1} \beta_{i,0} x_{i,N} \end{bmatrix} \quad (5.32)$$

$$\mathbf{X} = \begin{bmatrix} x_{1,N-d-sD+1} & x_{2,N-d-sD+1} & \dots & x_{p+q+1,N-d-sD+1} \\ x_{1,N-d-sD+2} & x_{2,N-d-sD+2} & \dots & x_{p+q+1,N-d-sD+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,N} & x_{2,N} & \dots & x_{p+q+1,N} \end{bmatrix} \quad (5.33)$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{p+q+1} \end{bmatrix} \quad (5.34)$$

Se realiza entonces la regresión ordinaria por mínimos cuadrados para producir un nuevo estimador de  $\boldsymbol{\beta}$ , al cual se llamará  $\boldsymbol{\beta}_1$ . En seguida, utilizando una nueva expansión de Taylor para  $e_t$  alrededor de  $\boldsymbol{\beta}_1$ , se obtiene una nueva versión de la ecuación de regresión, cuyos parámetros se estiman nuevamente por mínimos cuadrados ordinarios, generando el nuevo vector  $\boldsymbol{\beta}_2$ .

Este proceso se repite hasta que se satisfaga la primera de tres condiciones:

1. **Criterio 1:** En inglés *stopping criterion 1*. El vector de parámetros estimados prácticamente no varía de una estimación a la siguiente, es decir, para la  $k$ -ésima iteración se cumple que:

$$|\boldsymbol{\beta}_k - \boldsymbol{\beta}_{k-1}| \leq \epsilon_1 \quad (5.35)$$

2. **Criterio 2:** En inglés *stopping criterion 2*. La suma de cuadrados de los residuales prácticamente no varía de una estimación a la siguiente, es decir, para la  $k$ -ésima iteración se cumple que:

$$|S(\boldsymbol{\beta}_k) - S(\boldsymbol{\beta}_{k-1})| \leq \epsilon_2 \quad (5.36)$$

3. El **número máximo de iteraciones** (*maximum iterations*) llega a su fin.

Se llama entonces  $k$  al número de convergencia, esto es, al número de iteraciones necesarias para que se cumpla alguna de las dos primeras condiciones. Por lo tanto,  $\boldsymbol{\beta}_k$  será el estimador final de los parámetros.

Los errores estándar y los estadísticos de los estimadores se calculan a partir de la última linealización, de la misma forma que en cualquier procedimiento de estimación lineal. A partir de ellos se realizará —como se verá más adelante— el diagnóstico del modelo.

Es importante señalar que no hay garantía de que el proceso de estimación descrito sea siempre convergente a un estimador final de los parámetros. Es posible que exista divergencia o un conjunto de soluciones múltiples. En este caso, habrá que revisar la especificación del modelo y el cumplimiento de las condiciones de estacionaridad. Algunos paquetes estadísticos como SPSS permiten la modificación de los valores iniciales de los parámetros, pero la mayoría no.

Por último, la Figura 5.11 muestra las opciones que brinda Statgraphics para el proceso de estimación de los parámetros del modelo.

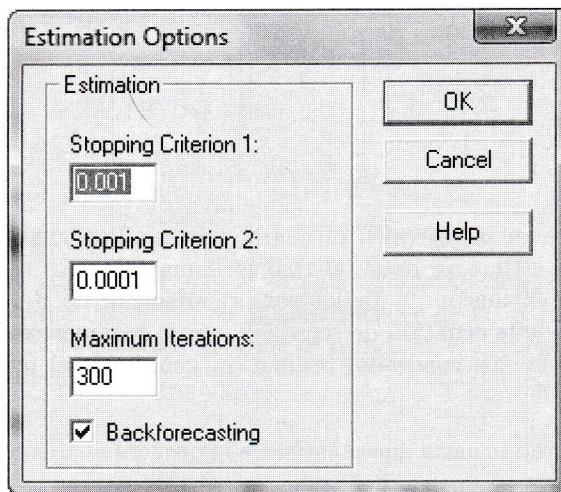


Figura 5.11: Opciones de estimación.

### 5.3.5 Valor inicial de los parámetros

Como se ha explicado, para realizar la estimación no lineal, es necesario inicializar los valores de los parámetros contenidos en el vector  $\beta_0 = (\delta_0, \phi_0, \theta_0)$ . Por supuesto, la convergencia en el proceso de estimación será más rápida si los valores iniciales son buenos, es decir, cercanos a los valores finales.

Una forma de obtener valores iniciales de los parámetros es despejar sus valores de las ecuaciones de Yule-Walker, sustituyendo las autocorrelaciones muestrales en lugar de las teóricas. Otra opción es usar tablas como las que ofrece el libro de Box y Jenkins [12]. Bowermann y O'Connell [11] sugieren usar 0.1 como primera aproximación, mientras que Pindyck y Rubinfeld [41] han tomado como valor inicial cero, con buenos resultados.

### 5.3.6 Ejemplo: estimación de ARMA(1,1)

Para observar de manera más clara el proceso de estimación no lineal, puede revisarse el caso del modelo mezclado ARMA(1,1), en el cual la constante  $\delta_0$  es cero y no se efectuaron diferencias ( $d = 0, D = 0$ ):

$$\begin{aligned} (1 - \phi_1 B) Z_t &= (1 - \theta_1 B) e_t \\ e_t &= \frac{1 - \phi_1 B}{1 - \theta_1 B} Z_t \end{aligned} \quad (5.37)$$

La ecuación anterior es no lineal con respecto a  $\theta_1$ , pero puede aproximarse a través de la expansión en serie de Taylor, utilizando los dos primeros términos. Esto implica calcular las primeras derivadas de  $e_t$  con respecto a  $\phi_1$  y  $\theta_1$ , evaluadas en los valores iniciales correspondientes  $\phi_{1,0}$  y  $\theta_{1,0}$ :

$$\begin{aligned} x_{1,t} &= -\frac{\partial e_t}{\partial \phi_1} \Big|_{\theta_{1,0}} \phi_{1,0} \\ &= \frac{B}{1 - \theta_{1,0}B} Z_t \end{aligned} \quad (5.38)$$

$$\begin{aligned} x_{2,t} &= -\frac{\partial e_t}{\partial \theta_1} \Big|_{\theta_{1,0}} \phi_{1,0} \\ &= \frac{B - \phi_{1,0}B^2}{(1 - \theta_{1,0}B)^2} Z_t \end{aligned} \quad (5.39)$$

Pueden entonces calcularse series de tiempo numéricas para  $x_{1,t}$  y  $x_{2,t}$ , para los valores de  $t = 1, 2, \dots, N$ . Por ejemplo:

$$\begin{aligned} (1 - \theta_{1,0}B)x_{1,t} &= BZ_t \\ x_{1,t} - \theta_{1,0}x_{1,t-1} &= Z_{t-1} \\ x_{1,t} &= \theta_{1,0}x_{1,t-1} + Z_{t-1} \end{aligned} \quad (5.40)$$

Si se inicializa con  $Z_0 = x_{1,0} = 0$ , puede resolverse la ecuación 5.40 en forma sucesiva para generar la serie completa  $x_{1,t}$ :

$$\begin{aligned} x_{1,1} &= 0 \\ x_{1,2} &= Z_1 \\ x_{1,3} &= \theta_{1,0}Z_1 + Z_2 \\ &\vdots \end{aligned}$$

Y se efectúa el mismo procedimiento para  $x_{2,t}$ :

$$x_{2,t} = 2\theta_{1,0}x_{2,t-1} - \theta_{1,0}^2x_{2,t-2} - Z_{t-1} + \phi_{1,0}Z_{t-2} \quad (5.41)$$

Si esta serie se inicializa con  $x_{2,1} = Z_0 = Z_{-1} = 0$  se obtienen las variables:

$$\begin{aligned} x_{2,1} &= 0 \\ x_{2,2} &= Z_1 \\ x_{2,3} &= 2\theta_{1,0}x_{2,2} - Z_2 + \phi_{1,0}Z_1 \\ &\vdots \end{aligned}$$

Con los valores iniciales  $\phi_{1,0}$  y  $\theta_{1,0}$  se calculan los residuales  $e_{t,0}$ , con lo cual puede formularse ya el modelo de regresión lineal como:

$$e_{t,0} + \phi_{1,0}x_{1,t} + \theta_{1,0}x_{2,t} \approx \phi_1 x_{1,t} + \theta_1 x_{2,t} + e_t \quad (5.42)$$

## 5.4 Diagnóstico

Una vez que se ha identificado un modelo preliminar y se han estimado sus parámetros, es necesario verificar si este modelo puede mejorarse y de qué forma. La metodología de Box-Jenkins no produce modelos “óptimos”, sin embargo, si se aplica correctamente y se prueban todos los modelos factibles, el analista tendrá la seguridad de haber elegido el mejor modelo entre ellos. Para ello, se parte de un modelo inicial tentativo que se va mejorando de acuerdo con los principios de parsimonia y aproximaciones sucesivas.

Las herramientas básicas para el diagnóstico de modelos son:

1. Los residuales, es decir, las diferencias entre cada dato y el modelo.
2. Los pronósticos de los últimos datos, que pueden retirarse temporalmente del modelo para ser pronosticados por él y compararse con los datos reales. Este número de datos se identifica generalmente con la opción de retenerlos para **validar el modelo** (*withhold for validation*).

Al diagnosticar un modelo conviene entonces considerar los siguientes pasos:

1. Analizar nuevamente la estacionaridad de los datos.
2. Verificar si el modelo está sobreestimado, en cuyo caso habrá que eliminar parámetros.
3. Verificar si el modelo está subestimado, en cuyo caso habrá que agregar parámetros.
4. Revisar la posible ocurrencia de datos atípicos o faltantes, para darles el tratamiento adecuado.

Por último, se toman todos los modelos que han resultado de interés para el analista y se comparan de manera global, ya sea a través de un cuadro o utilizando las opciones que ofrece el software para ello (*Model Comparisons* en *Statgraphics*). A continuación se explican y ejemplifican las etapas mencionadas, cada una de las cuales cuenta con pruebas de hipótesis, estadísticas y gráficas auxiliares.

### 5.4.1 Análisis de la estacionaridad

Como se ha dicho, antes de identificar un modelo tentativo debe verificarse que la serie de tiempo sea estacionaria en sentido amplio. De cualquier forma, es conveniente constatar esto en los residuales de cada modelo. Si la serie es estacionaria, deberán serlo también sus residuales.

En caso de que los residuales muestren tendencia, heteroscedasticidad o ambas, es indispensable corregir esto desde la serie original, de la forma ya señalada en el Capítulo 3. Esto es, si existe varianza creciente, efectuar la transformación adecuada; si existe tendencia, efectuar las diferencias ordinarias que se requieran. En caso de que los residuales muestren variación

estacional, debe probarse tanto la posibilidad de eliminarla con diferencias estacionales, como de modelarla. Será hasta el cuadro comparativo final de los modelos que se determinará cuál de las dos opciones es más conveniente.

#### 5.4.2 Modelo sobrestimado

Al estimar un modelo, los paquetes estadísticos ofrecen un conjunto de herramientas que permiten evaluar si cada uno de los parámetros del modelo se puede considerar significativo o no. Si un parámetro es significativo, deberá permanecer en el modelo; en caso contrario, deberá omitirse. Cuando uno o más parámetros del modelo son insignificantes, se dice que el modelo está sobrestimado.

Las herramientas estadísticas para este efecto son:

- La desviación estándar estimada del estimador del parámetro:  $S_{\hat{\beta}_j}$ .
- El estadístico  $t$  del estimador del parámetro:  $t_{\hat{\beta}_j}$ .
- El valor-P o *P-value* del estimador del parámetro.

Todos ellos se utilizan para probar la siguiente hipótesis para cada parámetro:

$$\begin{aligned} H_0: \beta_j &= 0 \\ &\text{versus} \\ H_a: \beta_j &\neq 0 \end{aligned}$$

Donde  $\beta_j$  es el  $j$ -ésimo parámetro del modelo. Para probar esta hipótesis se construye el estadístico:

$$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j - 0}{S_{\hat{\beta}_j}} \quad (5.43)$$

El cual se distribuye aproximadamente como una t-de-Student con  $N - d - sD - p - q - 1$  grados de libertad (el número de variables independientes menos el número de parámetros estimados), para un nivel de confianza de  $(1 - \alpha)\%$ . De aquí puede formularse una regla de decisión sencilla:

- Aceptar  $H_0$  si  $|t_{\hat{\beta}_j}| < 2$ . En este caso, el parámetro  $\beta_j$  es estadísticamente insignificante y debe retirarse del modelo.
- En caso contrario, rechazar  $H_0$  y dejar el parámetro en el modelo.

La regla de decisión puede construirse también utilizando el *P-value*, de la siguiente forma:

- Aceptar  $H_0$  si  $P-value > \alpha$ . En este caso, el parámetro  $\beta_j$  es estadísticamente insignificante y debe retirarse del modelo.

## 5.4 Diagnóstico

---

- En caso contrario, rechazar  $H_0$  y dejar el parámetro en el modelo.

Es importante aquí hacer una consideración. Si el parámetro que resulta estadísticamente insignificante es el de mayor orden, es decir, se trata de  $\mu$ ,  $\phi_p$  o  $\theta_q$ , para eliminarlo bastará con quitar la constante o reducir el orden del modelo a  $p - 1$  o  $q - 1$ , según corresponda. Sin embargo, si el parámetro a eliminar no es el de mayor orden, habrá de revisarse la posibilidad de agregar un modelo estacional de tipo multiplicativo, como se verá más adelante.

**Ejemplo: Modelo sobreestimado.** En el análisis de la serie de tiempo de la inflación en México se efectuó una diferencia estacional ( $s = 12$ ) y se identificó un modelo MA(1) para esta serie diferenciada. El software *Statgraphics* arroja los resultados de la Figura 5.12.

ARIMA Model Summary				
Parameter	Estimate	Stnd. Error	t	P-value
MA(1)	-0.409307	0.121206	-3.37695	0.001185
Mean	-0.0832393	0.0576141	-1.44477	0.152859
Constant	-0.0832393			

Backforecasting: yes  
Estimated white noise variance = 0.105975 with 72 degrees of freedom  
Estimated white noise standard deviation = 0.325538  
Number of iterations: 3

Figura 5.12: Ejemplo de modelo sobreestimado.

¿Qué se desprende de la tabla anterior? El parámetro MA(1), es decir,  $\theta_1$ , se puede considerar como significativo puesto que el valor absoluto de su estadístico t es mayor que dos (-3.38) y su P-value es menor que 0.05 (0.0012). En cambio, la media tiene un estadístico t de -1.44 y un P-value de 0.1529, que indican que es estadísticamente insignificante, por lo cual debe eliminarse la constante del modelo para estimar de nuevo los parámetros. En la parte inferior de la tabla se muestran tanto la varianza como el error estándar estimados de los residuales, cuyos 72 grados de libertad (*degrees of freedom*) corresponden al total de datos originales (86), menos los correspondientes a una diferencia estacional (12), menos 2 parámetros estimados ( $\delta_0$  y  $\theta_1$ ). La pantalla indica también que se ha elegido un modelo ARIMA(0,0,1) con una diferencia estacional.

Para modificar el modelo en *Statgraphics* se usa el clic derecho del ratón sobre la ventana de resultados y se elige de nuevo *Analysis Options*. En este caso debe eliminarse la marca de la ventana *Constant*, como se ve en la Figura 5.13.

Al estimar nuevamente el modelo sin la constante, se produce el resultado que se muestra en la Figura 5.14.

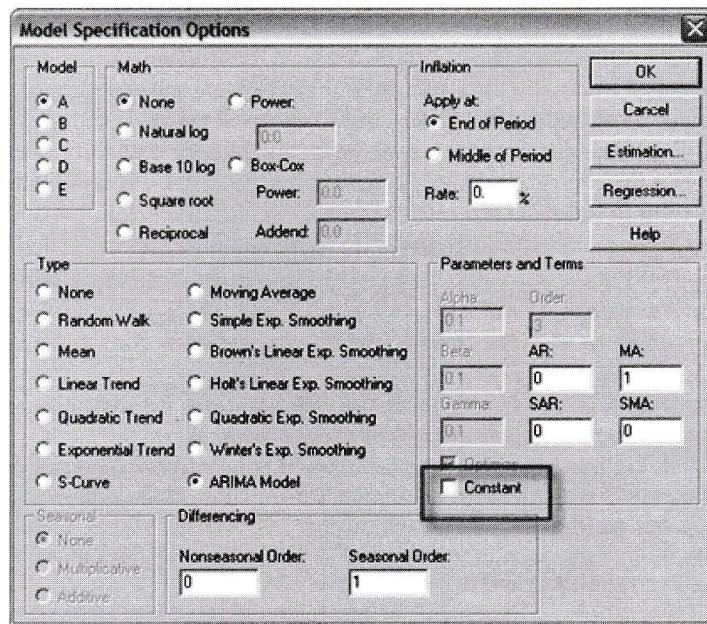


Figura 5.13: *Model Specification Options*.

```

ARIMA Model Summary
Parameter      Estimate     Stnd. Error    t      P-value
-----
MA(1)        -0.429835    0.119493   -3.59716    0.000582

Backforecasting: yes
Estimated white noise variance = 0.107774 with 73 degrees of freedom
Estimated white noise standard deviation = 0.328289
Number of iterations: 3

```

Figura 5.14: Modelo sin el parámetro sobrante.

### 5.4.3 Modelo subestimado

Se ha visto ya el caso en el cual debe eliminarse algún parámetro del modelo. Por supuesto, existe el caso contrario, en el cual debe agregarse algún parámetro. En este caso se dice que el modelo está subestimado, es decir, le faltan parámetros.

Si un modelo es adecuado para una serie de tiempo, se espera que sus residuales se comporten como ruido blanco. Por lo tanto, si esto no ocurre, es una señal de que el modelo requiere de ajustes que pueden ser transformaciones, diferencias (ordinarias o estacionales) o nuevos parámetros. Para determinar qué es lo que falta al modelo se repite el proceso de identificación, pero ahora tomando como punto de partida la serie de tiempo conformada por los residuales. Si los residuales son ruido blanco, el modelo es factible (aunque no necesariamente sea el mejor modelo); por el contrario, si los residuales siguen algún comportamiento distinto, habrá que incorporar este comportamiento al modelo que se está probando.

Para hacer este análisis de los residuales, los paquetes estadísticos suelen ofrecer las siguientes herramientas:

- La gráfica de los residuales (*Residual plot*). Esta gráfica permite observar si los residuales son estacionarios y si existen datos atípicos o faltantes.
- La ACF de los residuales o RACF (*Residual AutoCorrelation Function*). Esta gráfica se interpreta de la misma forma que se indicó en la Sección 5.2 y en seguida se verá cómo se incorpora el resultado al modelo.
- La PACF de los residuales o RPACF (*Residual Partial AutoCorrelation Function*). Esta gráfica se interpreta de la misma forma que se indicó en la Sección 5.2 y en seguida se verá cómo se incorpora el resultado al modelo.
- El periodograma de los residuales, que se utiliza para detectar tendencia, falta de ajuste o variación estacional, de la misma forma que antes. Como se ha dicho, esta interpretación debe ser congruente con el fenómeno observado y con las gráficas RACF y RPACF.
- Las estadísticas clásicas para medir el ajuste de los datos, que se usan también para comparar los diferentes modelos, tanto en su ajuste como en su capacidad de pronóstico. Dichas estadísticas incluyen:
  - La raíz cuadrada del error cuadrático medio (*Root Mean Squared Error*: RMSE).
  - El error absoluto medio (*Mean Absolute Error*: MAE).

- El error absoluto medio porcentual (*Mean Absolute Percentage Error*: MAPE).
  - El error medio (*Mean Error*: ME).
  - El error medio porcentual (*Mean Percentage Error*: MPE).
- Las pruebas de aleatoriedad de los residuales, entre las cuales destacan:
    - La prueba de corridas arriba y abajo de la mediana (*Runs above and below median*).
    - La prueba de corridas arriba y abajo (*Runs up and down*).
    - La prueba de Box-Pierce o prueba *Portmanteau* de las autocorrelaciones de los residuales.
    - La prueba de diferencias entre las medias de la primera y segunda mitad de los residuales para verificar si existe alguna tendencia.
    - La prueba de diferencias entre las varianzas de la primera y segunda mitad de los residuales, para verificar posible heteroscedasticidad.

A través de los instrumentos anteriores puede verificarse si los residuales se comportan como ruido blanco o si todavía reflejan la existencia de algún patrón o modelo que deba incorporarse al que ya se ha formulado. La forma de hacer esto es bastante sencilla y puede explicarse con un par de ejemplos.

**Ejemplo de modelo MA(1) con residuales MA(1).** Supóngase que a un conjunto de datos se ajusta un modelo MA(1) en el cual los residuales no son ruido blanco sino que, a su vez, siguen también un modelo MA(1). Es decir, al ajustar el modelo:

$$Z_t = (1 - \theta_1 B)a_t \quad (5.44)$$

Se encuentra que  $a_t$  sigue el modelo:

$$a_t = (1 - \lambda B)e_t \quad (5.45)$$

Donde  $e_t$  sí es ruido blanco. ¿Cómo se habría detectado esto? Al ajustar el primer modelo MA(1) se encontraría que la ACF de los residuales se trunca en  $k = 1$  y la PACF correspondiente es decreciente infinita. Ahora, si se sustituye la ecuación 5.45 en 5.44, se obtiene:

$$\begin{aligned} Z_t &= (1 - \theta_1 B)(1 - \lambda B)e_t \\ Z_t &= (1 - (\theta_1 + \lambda)B + \theta_1 \lambda B^2)e_t \end{aligned} \quad (5.46)$$

Como puede verse, la ecuación 5.46 es sencillamente un modelo MA(2). Entonces se concluye que, si se parte de un modelo MA(1) y sus residuales son también MA(1), debe probarse un modelo MA(2).

**Ejemplo de modelo AR(1) con residuales AR(1).** Supóngase que a un conjunto de datos se ajusta un modelo AR(1) en el cual los residuales no son ruido blanco sino que, a su vez, siguen también un modelo AR(1). Es decir, al ajustar el modelo:

$$(1 - \phi_1 B)Z_t = a_t \quad (5.47)$$

Se encuentra que  $a_t$  sigue el modelo:

$$\begin{aligned} (1 - \lambda B)a_t &= e_t \\ a_t &= (1 - \lambda B)^{-1}e_t \end{aligned} \quad (5.48)$$

Donde  $e_t$  sí es ruido blanco. ¿Cómo se habría detectado esto? Al ajustar el primer modelo AR(1) se encontraría que la ACF de los residuales es decreciente infinita y la PACF correspondiente se trunca en  $k = 1$ . Ahora, si se sustituye la ecuación 5.48 en 5.47, se obtiene:

$$\begin{aligned} (1 - \phi_1 B)Z_t &= (1 - \lambda B)^{-1}e_t \\ (1 - \lambda B)(1 - \phi_1 B)Z_t &= e_t \\ (1 - (\lambda + \phi_1)B + \lambda\phi_1 B)Z_t &= e_t \end{aligned} \quad (5.49)$$

Como puede verse, la ecuación 5.49 es sencillamente un modelo AR(2). Entonces se concluye que, si se parte de un modelo AR(1) y sus residuales son también AR(1), debe probarse un modelo AR(2).

**Generalización.** A partir de los ejemplos anteriores y de pruebas semejantes, se concluye que en general, la decisión será:

- Si los residuales siguen un comportamiento de tipo AR, incrementar en 1 el orden de la parte AR del modelo original.
- Si los residuales siguen un comportamiento de tipo MA, incrementar en 1 el orden de la parte MA del modelo original.

Es importante señalar aquí que, de acuerdo con el principio de parsimonia, los parámetros deben agregarse de uno en uno y, cada vez que se añadan, deberá probarse si el resultado no fue un modelo sobreestimado.

Para explicar los riesgos de agregar más de un parámetro a la vez, supóngase que se ha identificado un modelo ARMA( $p, q$ ), es decir, el modelo:

$$(1 - \phi_1 B - \dots - \phi_p B^p)Z_t = (1 - \theta_1 B - \dots - \theta_q B^q)e_t \quad (5.50)$$

Si la ecuación anterior se multiplica a ambos lados por el factor  $(1 - B)$ , aunque algebraicamente no se altera, se obtendrá un modelo que se ha complicado de manera artificial:

$$(1 - B)(1 - \phi_1 B - \dots - \phi_p B^p)Z_t = (1 - B)(1 - \theta_1 B - \dots - \theta_q B^q)e_t \quad (5.51)$$

La ecuación 5.51 seguirá un modelo ARMA( $p + 1, q + 1$ ) que será funcional y cuyos parámetros pudieran ser significativos, pero que no aporta nuevas explicaciones a los datos.

#### 5.4.4 Datos faltantes y atípicos

Una parte importante del diagnóstico consiste en revisar si se carece de uno o más datos (**datos faltantes** o *missing values*) o si existen valores que son excesivamente distintos de la mayoría de los puntos de la serie (**datos atípicos** o *outliers*). Cada uno de ellos debe recibir un tratamiento especial, como se explica a continuación:

- **Datos faltantes:** Si se carece de algún dato, ya sea porque no se registró, porque corresponde a un día inhábil o cualquier otra razón, es fundamental dejar su espacio correspondiente vacío, para que la serie mantenga sus características originales en cuanto al tiempo. Esto es, si por ejemplo se tienen datos mensuales y se elimina un registro, se perderá la posición de cada mes. Los paquetes estadísticos contienen instrucciones especiales para tratar estos datos faltantes. Generalmente se deja como un espacio vacío o el programa lo sustituye por una interpolación de los datos adyacentes.
- **Datos atípicos:** Los datos atípicos suelen destacar en la gráfica original de la serie y en los residuales, por ser muy diferentes de la mayoría. Los datos atípicos se clasifican en dos categorías:
  - **Datos discrepantes.** Se trata de datos muy distintos que obedecen a alguna causa —inclusive a un error de captura— que no afecta a los datos subsiguientes. Esto puede observarse en la gráfica de la serie. Es importante determinar la causa para constatar que se trata de datos discrepantes. La acción conveniente en este caso es dejarlos como datos faltantes, para que el software se encargue de sustituirlos por aproximaciones que no sobresalgan del resto de la serie.
  - **Datos influyentes.** Se trata de datos que no sólo son muy diferentes del resto, sino que además tienen un efecto que modifica a los datos subsiguientes. Generalmente se deben a causas que pueden identificarse con claridad. Por ejemplo, en el caso del valor de dólar o la inflación, el llamado “error de diciembre” de 1995 fue un dato influyente. Lo mismo ocurrió con el terremoto de 1985, que afectó por ejemplo al turismo de las fechas inmediatas posteriores. Estos datos no deben omitirse, pues su efecto es sustancial. En muchas ocasiones conviene, inclusive, tomar la serie a partir de ellos y eliminar los datos anteriores que carecían de este efecto. Cabe señalar que existe toda una rama de las series de tiempo, el **análisis de intervención** [14], que se dedica a verificar estadísticamente la existencia de datos influyentes.

#### 5.4.5 Comparación de modelos

Una vez que se han formulado todos los modelos que parezcan factibles para el conjunto de datos, es conveniente compararlos en un cuadro que permita tener una mirada global a las características de todos ellos. El software *Statgraphics* ofrece esta opción a través de identificar cada modelo con letras (de la “A” a la “E”), es decir, pueden compararse hasta cinco modelos de manera simultánea. En caso necesario, puede enviarse esta comparación al *StatReporter* para luego agregar otras.

Para efectuar la comparación, es conveniente dividir los datos en tres segmentos:

- Los datos que se utilizarán para la estimación.
- Los datos que se utilizarán para la validación.
- Los pronósticos.

Una vez que se defina cuál es el mejor modelo para los propósitos del analista, puede eliminarse el segmento de validación, cuyo uso es únicamente temporal.

Es muy recomendable utilizar la comparación que genera el paquete estadístico y trasladarla a un procesador de textos, por ejemplo, para enriquecer el cuadro con aspectos tales como:

- El modelo particular seleccionado.
- Las transformaciones y diferencias aplicadas a los datos.
- El número de parámetros y cuáles son significativos. Recuérdese que en el modelo final todos los parámetros deberán ser significativos, en la medida de lo posible.
- Las gráficas RACF, RPACF y el periodograma integrado de los residuales.
- Los estadísticos y pruebas de hipótesis con respecto al periodo de estimación.
- Los estadísticos y pruebas de hipótesis con respecto al periodo de validación.
- Un análisis de la bondad de los pronósticos. Recuérdese que éste es el objetivo esencial, así que en un momento dado puede sacrificarse el tener un modelo sobreestimado o subestimado, si este modelo genera buenos pronósticos.

A partir de este cuadro más completo será sencillo visualizar todos los aspectos y tomar la mejor decisión en la selección del modelo final.

## 5.5 Pronóstico

Una vez que se haya elegido el mejor modelo, podrán obtenerse los pronósticos de manera muy sencilla. El paquete *Statgraphics*, por ejemplo, genera estos valores con la opción **Forecast Table** y las gráficas correspondientes con **Time Series Plot** y **Forecast Plot**. Todas ellas pueden guardarse o copiarse y pegarse en otro software. Los pronósticos incluyen tanto el valor puntual como el intervalo de confianza para el valor de  $\alpha$  que defina el analista.

Si bien el software se encarga de deshacer las diferencias ordinarias y estacionales, habrá que revisar si ocurre lo mismo con las transformaciones. *Statgraphics* y SAS, por ejemplo, realizan la transformación inversa de forma automática.

**Ejemplo: Tasa de desocupación.** Si se toman los datos de la tasa de desocupación abierta urbana en México, pueden obtenerse los cinco modelos factibles que se enlistan en la Figura 5.15. Por su parte, la Figura 5.16 muestra la comparación de los cinco modelos anteriores durante el periodo de estimación, mientras que la Figura 5.17 contiene las estadísticas correspondientes al periodo de validación.

```

Model Comparison
-----
Data variable: Desocupacion
Number of observations = 68
Start index = 1/00
Sampling interval = 1.0 month(s)
Length of seasonality = 12
Number of periods withheld for validation: 12

Models
-----
(A) ARIMA(0,1,1) with constant
(B) ARIMA(1,1,0) with constant
(C) ARIMA(0,1,1) with constant
    Math adjustment: Square root
(D) ARIMA(1,1,0) with constant
    Math adjustment: Square root
(E) ARIMA(1,1,1) with constant

```

Figura 5.15: Comparación de modelos.

En este caso, aunque el Modelo (E) tiene la menor suma de cuadrados de los residuales (RMSE), puede verse en el cuadro de estimación (Figura 5.18) que el parámetro AR(1) sobra y debe, por lo tanto, eliminarse.

Por ello, probablemente convenga conservar como modelo el ARIMA(0,1,1), cuyos pronósticos para un año se muestran en la Figura 5.19 y cuyos residuales parecen ser ya ruido blanco.

## 5.5 Pronóstico

---

Estimation Period					
Model	RMSE	MAE	MAPE	ME	MPE
(A)	0.316263	0.246241	9.03234	-0.000703072	-1.40547
(B)	0.357373	0.278425	10.1256	0.00000196509	-1.03876
(C)	0.316318	0.245013	8.93463	0.00963748	-1.02447
(D)	0.358117	0.278635	10.1205	0.00291828	-0.911503
(E)	0.319166	0.246148	9.02591	0.00067289	-1.34302

Model	RMSE	RUNS	RUNM	AUTO	MEAN	VAR
(A)	0.316263	OK	OK	OK	OK	OK
(B)	0.357373	OK	OK	OK	OK	OK
(C)	0.316318	OK	OK	OK	OK	OK
(D)	0.358117	OK	OK	OK	OK	OK
(E)	0.319166	OK	OK	OK	OK	OK

Figura 5.16: Comparación del periodo de estimación.

Validation Period					
Model	RMSE	MAE	MAPE	ME	MPE
(A)	0.153417	0.312331	8.75877	-0.111149	-3.82486
(B)	0.197858	0.351667	9.71301	-0.0605519	-2.2864
(C)	0.154581	0.308013	8.66385	-0.122727	-4.13619
(D)	0.199275	0.352451	9.73776	-0.0632186	-2.35637
(E)	0.153413	0.311615	8.74192	-0.112598	-3.86361

Figura 5.17: Comparación del periodo de validación.

```

ARIMA Model Summary
Parameter      Estimate     Stnd. Error      t      P-value
-----      -----
AR(1)          0.00000121702  0.00000741531  0.164122  0.870152
MA(1)          0.673319    0.106061      6.34844   0.000000
Mean           0.0320153   0.0152149     2.1042   0.039295
Constant       0.0320152
-----
Backforecasting: yes
Estimated white noise variance = 0.101926 with 64 degrees of freedom
Estimated white noise standard deviation = 0.319259
Number of iterations: 10

```

Figura 5.18: Estimación del modelo ARMA(1,1,1).

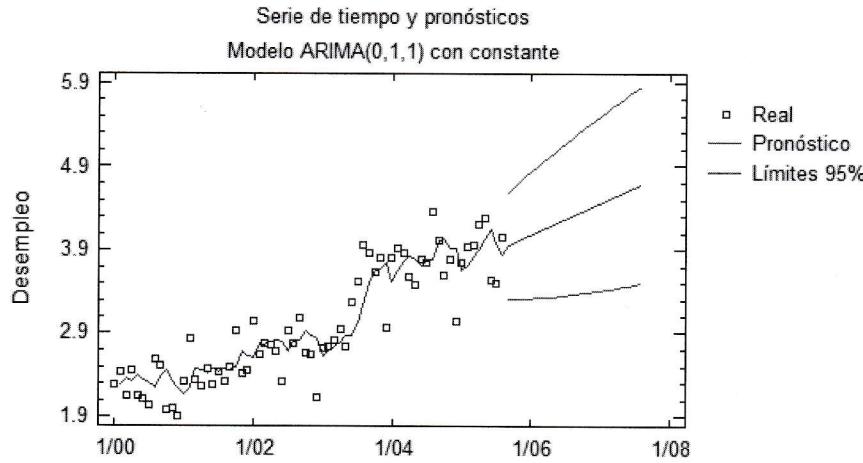


Figura 5.19: Pronósticos modelo ARIMA(0,1,1).

## 5.6 Actividades sugeridas

1. Cada estudiante puede simular 100 datos de algún modelo ARIMA y presentar su correspondiente gráfica, ACF, PACF y periodograma integrado, sin indicar el modelo correspondiente. El resto del grupo deberá señalar cuál es el modelo simulado. Ejemplos de modelos a simular son:

- a)  $Y_t = 10 + 0.8Y_{t-1} + e_t$
- b)  $Y_t = 10 - 0.8Y_{t-1} + e_t$
- c)  $Y_t = 10 + e_t - 0.8e_{t-1}$
- d)  $Y_t = 10 + e_t + 0.8e_{t-1}$
- e)  $Y_t = 10 + 0.3Y_{t-1} + e_t - 0.8e_{t-1}$
- f)  $Y_t = 10 - 0.3Y_{t-1} + e_t - 0.8e_{t-1}$
- g)  $Y_t = 10 + 0.3Y_{t-1} + e_t + 0.8e_{t-1}$
- h)  $Y_t = 10 - 0.3Y_{t-1} + e_t - 0.8e_{t-1}$
- i)  $Y_t = Y_{t-1} + e_t$
- j)  $Y_t = 0.6 + 0.8Y_{t-4} + e_t$

### *5.6 Actividades sugeridas*

---

2. Selecciona alguna serie de tiempo real con 80 o más datos, e identifica un modelo tentativo. Verifica sus propiedades en un software como SPSS, SAS o *Statgraphics*.
3. Demuestra que el modelo AR(1) puede estimarse de manera lineal, a través de la serie de Taylor.
4. Transforma al modelo de regresión lineal el modelo MA(1).
5. Supóngase que se prueba un modelo ARMA(1,1). Se encuentra que los residuales, en lugar de ser ruido blanco, siguen el comportamiento de un modelo AR(1). Demuestra cuál es el siguiente modelo que debe probarse.
6. Supóngase que se prueba un modelo AR(2). La ACF de sus residuales parece truncarse en  $k = 2$ , mientras que la RPACF tiene un aspecto decreciente infinito. Demuestra cuál es el siguiente modelo que debe probarse.
7. Supóngase que se trata de ajustar un modelo para la venta mensual de juguetes. El analista prueba un modelo AR(12) que parece dar buenos pronósticos. Sin embargo, los parámetros  $\phi_3, \phi_4, \dots, \phi_{10}$  resultan no significativos de acuerdo con su estadístico  $t$ . ¿Qué deberá hacer el analista en este caso?
8. Utiliza los datos del turismo nacional en México para identificar un modelo tentativo y efectuar el diagnóstico correspondiente.
9. Utiliza los datos del número de personas que han fallecido en México a causa de la influenza AH1N1 y colócalos en un software como *Statgraphics* o SAS. Identifica un modelo preliminar para estos datos.
10. Entra al sitio <http://www.statgraphicsonline.com> y regístrate para tener una cuenta personal. Hecho esto, utiliza este software en internet para elaborar pronósticos con la metodología de Box-Jenkins.

## Capítulo 6

# Modelos con variación estacional

Los callados estadísticos han cambiado nuestro mundo,  
no sólo al descubrir nuevos hechos o desarrollos técnicos, sino al modificar la  
forma en que razonamos, experimentamos y nos formamos opiniones.

IAN HACKING

### 6.1 Introducción

Como ya se ha visto, es común que muchas series de tiempo como aquellas relacionadas con aspectos turísticos, de salud o económicos, tengan algún tipo de variación estacional. Se ha dicho que esta variación puede intentar eliminarse a través de diferencias estacionales. Sin embargo, también es posible modelarla a través de la inclusión de variables apropiadas en los modelos. Para este fin, existen dos grandes tipos de modelos:

- **Modelos estacionales**, es decir, aquellos que únicamente reflejan un comportamiento periódico.
- **Modelos multiplicativos**, esto es, son modelos que involucran tanto la parte ordinaria que ya se ha revisado, como la parte estacional.

En las siguientes secciones se explican y ejemplifican ambos casos.

### 6.2 Modelos estacionales

Es posible que una serie de tiempo tenga un comportamiento que puede explicarse a partir de las relaciones que existen entre las variables y sus errores en intervalos de longitud  $s$ . Se suele llamar a  $s$  el orden de la parte estacional. En este caso, será conveniente utilizar **modelos exclusivamente**

**estacionales** (*seasonal models*). Estos modelos son semejantes a los modelos ordinarios ARIMA excepto que, en lugar de avanzar en periodos unitarios, avanzan –o retroceden– en periodos de longitud  $s$ .

Se denotan por una de estas dos formas: SARMA( $P,D,Q$ ) o la expresión ARMA( $P,D,Q_s$ ), donde<sup>1</sup>:

- $P$  es el orden de la parte autorregresiva estacional SAR.
- $D$  es el orden de las diferencias estacionales de longitud  $s$ .
- $Q$  es el orden de la parte de medias móviles estacionales SMA.

A continuación se describen en forma breve estos modelos.

### 6.2.1 Modelos SAR

El **modelo general autorregresivo estacional** se denota por SAR( $P$ ) y se escribe de la siguiente forma:

$$Z_t = \delta_0 + \Phi_1 Z_{t-s} + \Phi_2 Z_{t-2s} + \dots + \Phi_P Z_{t-Ps} + e_t \quad (6.1)$$

Obsérvese que ahora se utilizan tanto las letras griegas como la letra “P” en mayúscula, para diferenciar estos parámetros de los utilizados en los modelos ordinarios.

Resulta sencillo demostrar que este modelo tiene condiciones de estacionaridad pero no de invertibilidad. Por ejemplo, puede probarse que para el modelo SAR(1) debe cumplirse que  $|\Phi_1| < 1$ , como condición de estacionaridad.

Asimismo, la ACF de los modelos SAR( $P$ ) es decreciente infinita, pero sólo con valores significativos en múltiplos de  $s$ . En particular, la ACF del modelo SAR(1) sería:

$$\begin{aligned} \rho_1 &= 0 \\ \rho_2 &= \\ \rho_s &= \Phi_1 \\ \rho_{s+1} &= 0 \\ &\vdots \\ \rho_{2s} &= \Phi_1^2 \end{aligned} \quad (6.2)$$

$$\rho_{2s+1} = 0 \quad (6.3)$$

$$\begin{aligned} &\vdots \\ \rho_{3s} &= \Phi_1^3 \\ \rho_{2s+1} &= 0 \end{aligned} \quad (6.4)$$

⋮

---

<sup>1</sup>Obsérvese que para estos modelos se utilizan letras mayúsculas, tanto latinas como griegas.

Por otro lado, la PACF de los modelos SAR( $P$ ) se trunca en  $k = P$ , puesto que las únicas variables autorregresivas que aportan valor al modelo son  $Z_{t-s}, Z_{t-2s}, \dots, Z_{t-Ps}$ , y por la misma razón sólo tiene valores significativos en múltiplos de  $s$ . Por ejemplo, para el modelo SAR(1) la PACF sería:

$$\begin{aligned}
 \rho_{11} &= 0 \\
 \rho_{22} &= 0 \\
 &\vdots \\
 \rho_{ss} &= \Phi_1 & (6.5) \\
 \rho_{s+1,s+1} &= 0 \\
 &\vdots \\
 \rho_{2s,2s} &= 0 \\
 \rho_{2s+1,2s+1} &= 0 \\
 &\vdots
 \end{aligned}$$

En la Figura 6.1 se muestran la ACF y PACF de un modelo SAR(1) simulado con  $s = 4$ . Puede observarse que en la ACF destacan claramente los valores múltiplos de cuatro; mientras que en la PACF la gráfica parece truncarse en  $k = 4$ . Por otro lado, el periodograma integrado (Figura 6.2) contiene escalones que refuerzan la necesidad de utilizar un modelo estacional.

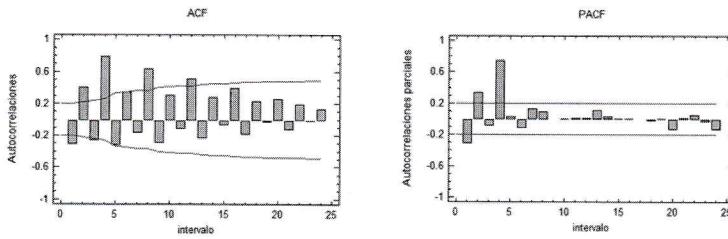


Figura 6.1: ACF y PACF del modelo SAR(1).

### 6.2.2 Modelos SMA

El **modelo general de medias móviles estacional** se denota con la expresión SMA( $Q$ ) y se escribe de la siguiente forma:

$$Z_t = \delta_0 + e_t - \Theta_1 e_{t-s} - \Theta_2 e_{t-2s} - \dots - \Theta_Q e_{t-Qs} \quad (6.6)$$

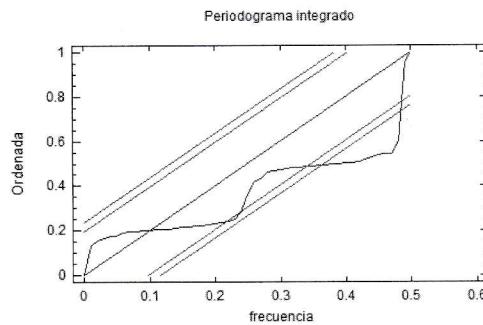


Figura 6.2: Periodograma integrado del modelo SAR(1).

Obsérvese que ahora se utilizan tanto las letras griegas como la letra “Q” en mayúscula, para diferenciar los parámetros de los modelos ordinarios.

Resulta sencillo demostrar que este modelo tiene condiciones de invertibilidad pero no de estacionaridad. Por ejemplo, puede probarse que para el modelo SMA(1) debe cumplirse que  $|\Theta_1| < 1$ .

Asimismo, la ACF de los modelos SMA( $Q$ ) se trunca en  $k = Q$ , pero sólo contiene valores significativos en múltiplos de  $s$ . En particular, la ACF del modelo SMA(1) será:

$$\begin{aligned} \rho_1 &= 0 \\ \rho_2 &= \\ \rho_s &= -\frac{\Theta_1}{1 + \Theta_1^2} \\ \rho_{s+1} &= 0 \\ &\vdots \end{aligned} \tag{6.7}$$

$$\begin{aligned} \rho_{2s} &= 0 \\ \rho_{2s+1} &= 0 \\ &\vdots \end{aligned} \tag{6.8}$$

Por otro lado, la PACF de los modelos SMA( $Q$ ) es decreciente infinita, puesto que equivale a un modelo AR( $\infty$ ) que sólo tiene valores significativos

en múltiplos de  $s$ . Por ejemplo, para el modelo SMA(1) la PACF será:

$$\rho_{11} = 0 \quad (6.9)$$

$$\rho_{22} = 0$$

$$\vdots$$

$$\rho_{ss} = -\Theta_1 \quad (6.9)$$

$$\rho_{s+1,s+1} = 0$$

$$\vdots$$

$$\rho_{2s,2s} = -\Theta_1^2 \quad (6.10)$$

$$\rho_{2s+1,2s+1} = 0$$

$$\vdots$$

$$\rho_{3s,3s} = -\Theta_1^3 \quad (6.11)$$

$$\rho_{2s+1,2s+1} = 0$$

$$\vdots$$

En la Figura 6.3 se muestran la ACF y PACF de un modelo SMA(1) simulado con  $s = 4$ . Puede observarse que la ACF resulta decreciente infinita con valores únicamente en múltiplos de cuatro, en tanto que la PACF se trunca en  $k = 4$ . Por su parte, el periodograma integrado (Figura 6.4) tiene claramente los escalones que denotan la variación estacional.

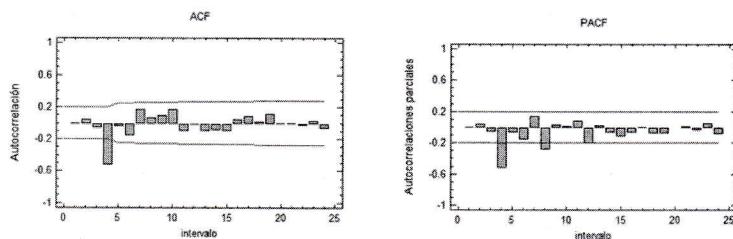


Figura 6.3: ACF y PACF del modelo SMA(1).

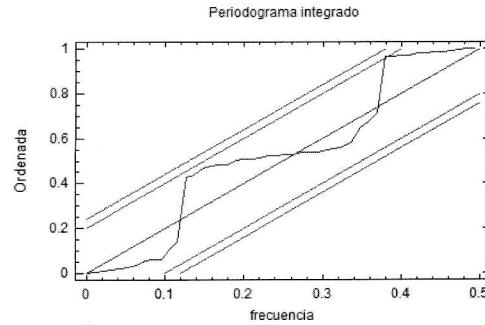


Figura 6.4: Periodograma integrado del modelo SMA(1).

### 6.2.3 Modelos SARIMA

. De acuerdo con lo anterior, el **modelo general estacional** estacional de la forma SARIMA( $P,D,Q$ ) se representaría como:

$$\begin{aligned} Z_t &= \delta_0 + \Phi_1 Z_{t-s} + \dots + \Phi_P Z_{t-Ps} + e_t \\ &\quad - \Theta_1 e_{t-s} - \dots - \Theta_Q e_{t-Qs} \end{aligned} \quad (6.12)$$

Donde:

$$Z_t = \Delta_s^D W_t$$

Aquí  $W_t$  puede ser la serie de tiempo original  $Y_t$  o alguna transformación de ella, si es que fue necesario por existir heteroscedasticidad. Si se utiliza el operador de salto hacia atrás  $B$ , la ecuación 6.12 puede expresarse con la notación abreviada:

$$\Phi(B^s)Z_t = \delta_0 + \Theta(B^s)e_t \quad (6.13)$$

Donde:

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps} \quad (6.14)$$

$$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs} \quad (6.15)$$

¿Cómo serán las ACF y PACF de los modelos SARIMA? Como puede esperarse, ambas serán decrecientes infinitas pero únicamente tendrán valores significativos en múltiplos de  $s$ . Por otro lado, sus parámetros tendrán tanto condiciones de estacionariedad como de invertibilidad.

En la Figura 6.5 se muestran la ACF y PACF de un modelo SARIMA(1,1) simulado con  $s = 4$ . Se observa que ambas funciones tienen un aspecto decreciente infinito, en el cual sobresalen los valores correspondientes a múltiplos de cuatro.

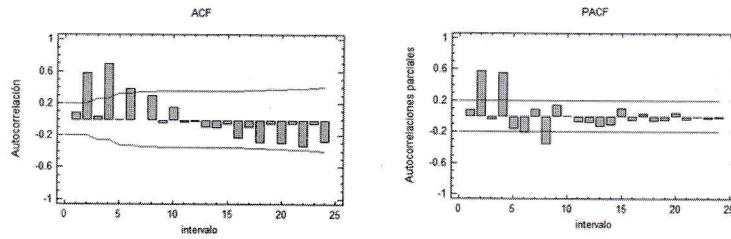


Figura 6.5: ACF y PACF del modelo SARMA(1,1).

### 6.3 Modelos multiplicativos

Por último, resultará de gran interés construir modelos que agrupen tanto la parte ordinaria como la parte estacional, en una sola clase general. Con estos modelos se logra gran versatilidad y posibilidad de ajuste, que generará buenos pronósticos en la mayoría de los casos. Se llamarán **modelos multiplicativos** porque ésta es la operación que conjunta ambas partes.

El **modelo general multiplicativo** se denota por:

$$\text{ARIMA } (p,d,q) \times (P,D,Q)_s$$

O simplemente:

$$\text{ARIMA } (p,d,q) (P,D,Q)_s$$

Donde:

- El valor  $p$  es el orden de la parte AR.
- El valor  $d$  es el orden de las diferencias ordinarias o no estacionales.
- El valor  $q$  es el orden de la parte MA.
- El valor  $P$  es el orden de la parte SMA.
- El valor  $D$  es el orden de las diferencias estacionales.
- El valor  $Q$  es el orden de la parte SMA.

Este modelo general multiplicativo suele expresarse en forma abreviada como:

$$\Phi(B^s)\phi(B)Z_t = \delta_0 + \Theta(B^s)\theta(B)e_t \quad (6.16)$$

Por supuesto, de manera indistinta puede colocarse primero la parte ordinaria y luego la estacional o viceversa, ya que el producto es conmutativo. Recuérdese que en el modelo anterior  $Z_t = \Delta_s^D \Delta^d W_t$  y  $W_t$  puede ser la serie de tiempo original  $Y_t$  o alguna transformación de ella.

Para comprender mejor el funcionamiento de estos modelos, a continuación se muestran algunos casos particulares.

### 6.3.1 Modelo ARIMA(1,0,0)(1,0,0)<sub>s</sub>

Supóngase, por ejemplo, que se observa de manera evidente que un conjunto de datos trimestrales puede modelarse como un SAR(1) con  $s = 4$ , sin constante. En este caso se habría encontrado una ACF que decrece de manera infinita pero que tiene valores que destacan en  $k = 4, 8, 12, 16, \dots$  y una PACF con un valor sobresaliente en  $k = 4$  y que parece truncarse después de él. Sin embargo, los residuales del modelo SAR(1) no resultan ruido blanco sino que son, a su vez, un modelo ordinario AR(1). Esto significa que la RACF es decreciente infinita y la RPACF se trunca en  $k = 1$ .

Así, el modelo original SAR(1) que se probó puede escribirse como:

$$\begin{aligned} Z_t &= \Phi_1 Z_{t-4} + a_t \\ (1 - \Phi_1 B^4)Z_t &= a_t \end{aligned} \quad (6.17)$$

Donde  $a_t$  se comporta como un AR(1), entonces:

$$\begin{aligned} a_t &= \phi_1 a_{t-1} + e_t \\ (1 - \phi_1 B)a_t &= e_t \\ a_t &= (1 - \phi_1 B)^{-1} e_t \end{aligned} \quad (6.18)$$

En la cual  $e_t$  es ruido blanco. Si se sustituye la ecuación 6.18 en 6.17, se obtiene un modelo ARIMA(1,0,0)(1,0,0)<sub>4</sub>:

$$\begin{aligned} (1 - \Phi_1 B^4)Z_t &= (1 - \phi_1 B)^{-1} e_t \\ (1 - \Phi_1 B^4)(1 - \phi_1 B)Z_t &= e_t \end{aligned} \quad (6.19)$$

Es posible analizar un poco más este último modelo. Si se efectúan los productos, se obtiene lo siguiente:

$$\begin{aligned} (1 - \Phi_1 B^4 - \phi_1 B + \Phi_1 \phi_1 B^5)Z_t &= e_t \\ Z_t &= \phi_1 Z_{t-1} + \Phi_1 Z_{t-4} \\ &\quad - \Phi_1 \phi_1 Z_{t-5} + e_t \end{aligned} \quad (6.20)$$

Obsérvese ahora que la ecuación 6.20 no es otra cosa más que un modelo AR(5) en el cual:

$$\begin{aligned} \phi_1 &= \phi_1 \\ \phi_2 &= 0 \\ \phi_3 &= 0 \\ \phi_4 &= \Phi_1 \\ \phi_5 &= \Phi_1 \phi_1 \end{aligned}$$

Así pues, el modelo ARMA(1,0,0)(1,0,0)<sub>4</sub> o simplemente AR(1) × SAR(1), equivale a un modelo AR(5). Sin embargo, es preferible el primer modelo de acuerdo con el principio de parsimonia, puesto que sólo tiene dos parámetros, mientras que el segundo tiene cinco.

De aquí se concluye también que la ACF de este modelo será decreciente infinita, mientras que su PACF se truncará en  $k = 5$ , con valores que destacan en 1, 4 y 5. También puede comprenderse ahora por qué en algunos casos existen parámetros intermedios no significativos, como se comentó en la sección del diagnóstico de modelos.

En la Figura 6.6 se muestran la ACF y PACF de un modelo AR(1)  $\times$  SAR(1) simulado con  $s = 4$ .

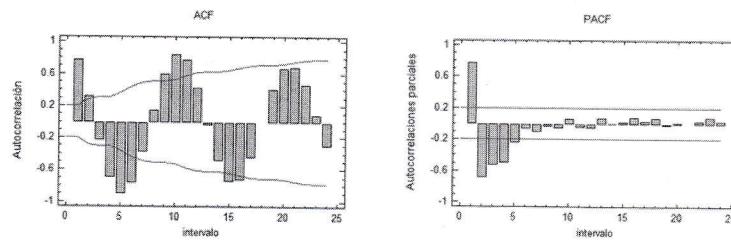


Figura 6.6: ACF y PACF del modelo AR(1)SAR(1).

En general, puede entonces esperarse que en un modelo ARIMA( $p,0,0$ )  $\times$  ( $P,0,0$ ) $_s$  la ACF será decreciente infinita con valores que destacan en múltiplos de  $s$ , mientras que la PACF se truncará en  $k = p + Ps$ .

### 6.3.2 Modelo ARIMA(0,0,1)(0,0,1) $_s$

Supóngase ahora que se observa de manera evidente que un conjunto de datos trimestrales puede modelarse como un SMA(1) con  $s = 4$  sin constante. En este caso se habría encontrado una ACF con un valor sobresaliente en  $k = 4$  y una PACF decreciente infinita con valores que destacan en  $k = 4, 8, 12, 16, \dots$ . Sin embargo, los residuales del modelo SMA(1) no resultan ruido blanco sino que son, a su vez, un modelo ordinario MA(1). Esto significa que la RACF se trunca en  $k = 1$  y la RPACF es decreciente infinita.

Así, el modelo original SMA(1) que se probó puede escribirse como:

$$\begin{aligned} Z_t &= a_t - \Theta_1 a_{t-4} \\ Z_t &= (1 - \Theta_1 B^4) a_t \end{aligned} \quad (6.21)$$

Donde  $a_t$  se comporta como un MA(1), entonces:

$$\begin{aligned} a_t &= e_t - \theta_1 e_{t-1} \\ a_t &= (1 - \theta_1 B) e_t \end{aligned} \quad (6.22)$$

Donde  $e_t$  es ruido blanco. Si se sustituye la ecuación 6.22 en 6.21, se obtiene un modelo ARIMA(0,0,1)(0,0,1) $_4$ :

$$Z_t = (1 - \Theta_1 B^4)(1 - \theta_1 B) e_t \quad (6.23)$$

### 6.3 Modelos multiplicativos

Es posible analizar un poco más este último modelo. Si se efectúan los productos, se obtiene lo siguiente:

$$\begin{aligned} Z_t &= (1 - \Theta_1 B^4 - \theta_1 B + \Theta_1 \theta_1 B^5) e_t \\ Z_t &= e_t - \theta_1 e_{t-1} - \Theta_1 e_{t-4} + \Theta_1 \theta_1 e_{t-5} \end{aligned} \quad (6.24)$$

Obsérvese ahora que la ecuación 6.24 no es otra cosa más que un modelo MA(5) en el cual:

$$\begin{aligned} \theta_1 &= \theta_1 \\ \theta_2 &= 0 \\ \theta_3 &= 0 \\ \theta_4 &= \Theta_1 \\ \theta_5 &= \Theta_1 \theta_1 \end{aligned}$$

Así pues, el modelo ARMA(0,0,1) (0,0,1)<sub>4</sub>, que también puede expresarse de forma más sencilla como MA(1)  $\times$  SMA(1), equivale a un modelo MA(5). Sin embargo, es preferible el primer modelo de acuerdo con el principio de parsimonia, puesto que sólo tiene dos parámetros, mientras que el segundo tiene 5.

De aquí se concluye también que la ACF de este modelo se truncará en  $k = 5$ , con valores que destacan en 1, 4 y 5, mientras que su PACF será decreciente infinita. También puede comprenderse ahora por qué en algunos casos existen parámetros intermedios no significativos, como se comentó en la sección del diagnóstico de modelos.

En la Figura 6.7 se muestran la ACF y PACF de un modelo MA(1)  $\times$  SMA(1) simulado con  $s = 4$ .

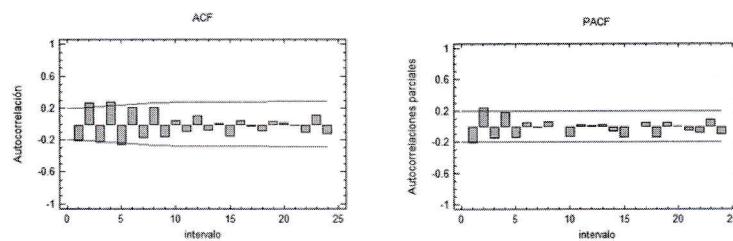


Figura 6.7: ACF y PACF del modelo MA(1)SMA(1).

En general, puede entonces esperarse que en un modelo ARIMA(0,0, $q$ )  $\times$  (0,0, $Q$ ) <sub>$s$</sub>  la ACF se truncará en  $k = q + Qs$  mientras que la PACF será decreciente infinita con valores que destacan en múltiplos de  $s$ .

## 6.4 Resumen de modelos

En el Cuadro 6.1 se resume la información de todos los modelos ARIMA (CE=Condiciones de estacionaridad; CI=Condiciones de invertibilidad):

## 6.5 Actividades sugeridas

1. Simula un modelo de tipo SAR(1) con  $s = 12$  y obtén su ACF, PACF y periodograma integrado. Interpreta los resultados.
2. Simula un modelo de tipo SMA(1) con  $s = 12$  y obtén su ACF, PACF y periodograma integrado. Interpreta los resultados.
3. Simula un modelo de tipo SARMA(1) con  $s = 12$  y obtén su ACF, PACF y periodograma integrado. Interpreta los resultados.
4. Simula un modelo multiplicativo de tipo SMA(1)×AR(1) con  $s = 12$  y obtén su ACF, PACF y periodograma integrado. Interpreta los resultados.
5. Simula un modelo multiplicativo de tipo SAR(1)×MA(1) con  $s = 12$  y obtén su ACF, PACF y periodograma integrado. Interpreta los resultados.
6. Supóngase que se prueba un modelo SAR(1) con  $s = 12$  y se encuentra que sus residuales no son ruido blanco, sino que siguen un modelo de tipo MA(1). ¿Cuál será el siguiente modelo que debe probarse?
7. Supóngase que se prueba un modelo SMA(1) con  $s = 12$  y se encuentra que sus residuales no son ruido blanco, sino que siguen un modelo de tipo MA(1). ¿Cuál será el siguiente modelo que debe probarse?
8. Supóngase que se prueba un modelo SAR(1) con  $s = 7$  y se encuentra que sus residuales no son ruido blanco, sino que siguen un modelo de tipo SAR(1). ¿Cuál será el siguiente modelo que debe probarse?
9. Elige una serie de tiempo real que contenga variación estacional. Determina el mejor modelo para dicha serie.
10. Determina el mejor modelo para pronosticar los datos del turismo receptivo nacional en México. Utiliza los datos que brinda el Instituto Nacional de Estadística y Geografía (INEGI).

Cuadro 6.1: Resumen de modelos ARIMA

Modelo	CE	CI	ACF	PACF
AR( $p$ )	Sí	No	Decreciente infinita	Se trunca en $p$
MA( $q$ )	No	Sí	Se trunca en $q$	Decreciente infinita
ARMA( $p, q$ )	Sí	Sí	Decreciente infinita	Decreciente infinita
SAR( $P$ )	Sí	No	Decreciente infinita en múltiplos de $s$	Se trunca en $P$
SMA( $Q$ )	No	Sí	Se trunca en $Qs$	Decreciente infinita en múltiplos de $s$
SARMA( $P, Q$ )	Sí	Sí	Decreciente infinita en múltiplos de $s$	Decreciente infinita en múltiplos de $s$
AR( $p$ )SAR( $P$ )	Sí	No	Decreciente infinita, valores mayores en múltiplos de $s$	Se trunca en $p + sP$
MA( $q$ )SMA( $Q$ )	No	Sí	Se trunca en $q + Qs$	Decreciente infinita, valores mayores en múltiplos de $s$
Otras mezclas	Sí	Sí	Decreciente infinita, valores mayores en múltiplos de $s$	Decreciente infinita, valores mayores en múltiplos de $s$

## Capítulo 7

# Ejemplo de aplicación

La información disponible en línea, en las bibliotecas o en archivos,  
pueden ser útiles a condición de tener algo que decir,  
algo que preguntarse, o algo que investigar.

GREGORIO HERNÁNDEZ ZAMORA

### 7.1 Introducción

En esta sección se presenta un ejemplo en el cual se aplica la metodología de Box-Jenkins para analizar el comportamiento presente y futuro de las remesas en México. Se expone en forma de artículo de investigación original, con los correspondientes apartados de: contexto, antecedentes, justificación, objetivo, material y métodos, resultados y discusión.

La idea de esta aplicación es que pudiera servir como modelo a los estudiantes, tanto para utilizar los modelos ARIMA en problemas reales, como para tener una visión acerca de la forma de elaborar artículos científicos.

### 7.2 Las remesas en México

Las Naciones Unidas reconocieron desde mediados del siglo XX el derecho de moverse con libertad dentro y fuera de las fronteras de los países, al firmar la Declaración Universal de los Derechos Humanos, que en su artículo 13 establece que “toda persona tiene derecho a circular libremente y a elegir su residencia en el territorio de un estado” y “a salir de cualquier país, incluso del propio, y a regresar a su país” [4, 36]. Un caso particular de estos movimientos lo constituye la migración hacia los países con mejores condiciones económicas.

En los países que les acogen, los migrantes transnacionales generan diversas respuestas de carácter legal y étnico cuyo gradiente va desde la aceptación resignada hasta la deportación, el rechazo, el resurgimiento del racismo, la intolerancia o la violación de sus derechos fundamentales. En gran medida,

### **7.3 Importancia de las remesas**

---

la aceptación se debe a que los migrantes perciben salarios bajos y sin prestaciones laborales, cubren necesidades de mano de obra y realizan trabajos que, por su naturaleza sucia, desagradable o peligrosa, los nacionales no están dispuestos a realizar [1].

Sin embargo, estos trabajos permiten que los migrantes del mundo envíen dinero a sus países de origen que, a su vez, suelen caracterizarse por tener tasas de inflación altas y una oferta escasa o poco atractiva de empleos [1]. Estos envíos de dinero se conocen como remesas. En particular, el Banco de México define de manera formal las remesas familiares como la:

Cantidad en moneda nacional o extranjera proveniente del exterior, transferida a través de empresas, originada por un remitente (persona física residente en el exterior que transfiere recursos económicos a sus familiares en México) para ser entregada en territorio nacional a un beneficiario (persona física residente en México que recibe los recursos que transfiere el remitente) [5].

Por su parte, los familiares de los migrantes, que permanecen en sus países de origen, no sólo reciben las remesas, sino que deben adaptarse a las circunstancias que implica la ausencia temporal o definitiva de quien emigró. El costo familiar de esta ausencia es alto, pues el cuidado y la educación informal de los hijos suelen quedar a cargo de uno solo de los padres o de algún otro pariente o amigo, con las consecuentes repercusiones psicológicas y en el desempeño escolar [1] de niños y adolescentes. A pesar de ello, los grupos sociales consideran que el beneficio económico recibido amerita este sacrificio.

## **7.3 Importancia de las remesas**

Como se ha dicho, las remesas están conformadas por el dinero que envían los migrantes mexicanos de los Estados Unidos a sus familias en México. Las remesas familiares totales incluyen transferencias electrónicas, *money orders*, cheques personales, efectivo y bienes en especie.

México es el segundo receptor de remesas en el mundo, después de la India [20], y el mayor en América Latina [57]. De acuerdo con la clasificación del Banco Mundial, México puede considerarse como un país de ingreso medio, pero con una gran polaridad [61], es decir, con una distribución inequitativa de la riqueza. Esta inequidad promueve que los sectores menos favorecidos de la población tiendan a elegir la opción de conseguir trabajo en los Estados Unidos, país vecino y desarrollado, como forma de mejorar su nivel de vida y el de sus familias.

En principio, la mayoría de los migrantes mexicanos planean regresar: buscan trabajar en Estados Unidos por períodos cortos para generar una fuente alternativa de ingresos o acumular ahorros [23], que les permitan vivir mejor en México. La forma más común de usar el dinero devengado es enviarlo a sus familiares a través de las remesas, ya sea para gasto o ahorro. En este último caso, algunas familias tratan de conservar las remesas en dólares

para evitar el problema de las devaluaciones [57], de tal modo que el monto acumulado eventualmente garantice un futuro más atractivo para todos sus miembros.

A decir de Morán Quiroz [22], la mayor parte de los mexicanos que trabajan en los Estados Unidos consideran esta situación como una etapa provisoria en sus vidas:

Este destino en el extranjero suele ser percibido como temporal y meramente laboral, mientras la localidad de origen es vista como el lugar al que volverán de manera definitiva sea con ánimos empresariales (con el dinero ahorrado en el extranjero), sea con ánimos de retiro del trabajo (con la pensión devengada en Estados Unidos).

Más de once millones de personas nacidas en México [19] han pagado mucho dinero y han arriesgado sus vidas para ir a los Estados Unidos. Las razones que aducen para ello son diversas: buscar el “sueño americano”, reunir a sus familias, trabajar y enviar remesas a sus familias en México. Muchos de ellos explican con sencillez: “sólo quiero juntar un dinerito y regresar a México” [28]. Por su parte, “los familiares cultivan celosamente la idea del retorno del migrante, y así se lo hacen saber; a su vez, éste retroalimenta estas expectativas con sus contactos de mayor o menor frecuencia, pero también dejando signos visibles de su intención” [33].

Sin embargo, en la actualidad, el endurecimiento en las medidas de vigilancia en la frontera hace que los trabajadores indocumentados tiendan a permanecer durante más tiempo en los Estados Unidos, ya que el reingreso a dicho país se ha tornado más difícil. Además, las redes sociales han propiciado la conformación de comunidades mexicanas que conviven y se apoyan mutuamente. Algunos migrantes, poco a poco, se han llevado a sus familias completas, inclusive a niños y adultos mayores.

Estos escenarios actuales y futuros seguramente redundarán en el monto y frecuencia de las remesas. En su artículo acerca de las migrantes mexicanas, Marroni [33] prevé que:

Se consolidará la tendencia al espaciamiento del retorno a México frente al aumento de los riesgos y de los costos del reingreso a Estados Unidos. En este sentido, se observará una propensión al establecimiento de los migrantes en la Unión Americana; situación que converge con la del gran número de jóvenes solteros —hombres y mujeres— que constituyen familias en aquel país. Estos nuevos núcleos matrimoniales disminuirán, posiblemente, sus contactos con la región.

## **7.4 Objetivo del estudio**

Como ha podido apreciarse, la magnitud de los impactos económicos, sociales, culturales y políticos de la migración y del envío subsecuente de remesas es muy importante para México. En este sentido, es fundamental tratar de responder a preguntas como las que plantean Aragonés *et al.* [2] en

## 7.5 Metodología

---

cuanto a si es válido y eficiente depositar el peso del desarrollo y la superación de la pobreza en los migrantes y sus familias.

El escenario actual de recesión internacional no ofrece un panorama alentador para los migrantes, lo que lleva a preguntarse cuál será el comportamiento de la migración y del envío de remesas para los próximos años. La crisis económica actual ha ocasionado una recesión en los Estados Unidos y es de esperarse que se privilegien las fuentes de empleo para los ciudadanos de dicho país, y como consecuencia los inmigrantes vean reducidas sus posibilidades de contratarse y, aún eventualmente puedan ser expulsados. Además, la globalización favorece que trabajadores de otros países compitan por las fuentes de trabajo con los inmigrantes mexicanos, dada la mayor facilidad de comunicación y el abatimiento en los costos de transporte.

Por otro lado, el flujo de remesas puede disminuir, dado que muchos migrantes han adquirido la nacionalidad norteamericana y se han llevado consigo a sus familiares, quienes eran los receptores de las remesas. Además, los migrantes de segundas y terceras generaciones no sienten el mismo compromiso con sus familiares mexicanos para apoyarles económicamente.

Todo lo anterior contribuye a indicar que el flujo de remesas, en cantidad, frecuencia y monto, puede ir a la baja. Para confirmar esta suposición existen métodos objetivos de tipo cuantitativo, tales como los modelos matemáticos.

Los modelos matemáticos ofrecen instrumentos apropiados para describir el comportamiento actual y para proyectar la tendencia futura. Dentro de los modelos más utilizados para obtener pronósticos destacan los llamados autoregresivos integrados y de medias móviles (*autorregresive integrated moving averages*), también conocidos como modelos ARIMA. Estos modelos brindan pronósticos de alta precisión a corto y mediano plazo y permiten incluir aspectos como tendencia, heteroscedasticidad y variación estacional.

El propósito de este estudio es complementar los análisis reportados en la revisión de la literatura, con un análisis cuantitativo del comportamiento histórico de las remesas familiares de migrantes mexicanos, sus tendencias y perspectivas, a través de modelos ARIMA, con base en la metodología de Box-Jenkins. A partir del pronóstico obtenido se revisará nuevamente la conveniencia de que el modelo económico descance de manera importante en las remesas.

## 7.5 Metodología

Para efectuar este ejercicio se contó con los datos correspondientes a las remesas familiares totales por mes, en millones de dólares, procedentes del Banco de México [7]. Se obtuvieron las cifras de enero de 1995 a diciembre 2008, mismas que se observan en la Figura 7.1.

Los datos recabados se procesaron con el software estadístico *Statgraphics Centurion XV* versión 15.01.2 [51], con la opción Forecasting, dentro de la cual se utilizaron Descriptive Time Series Methodsy los modelos de tipo autoregresivo integrado y de medias móviles (ARIMA) en User Specified Model.

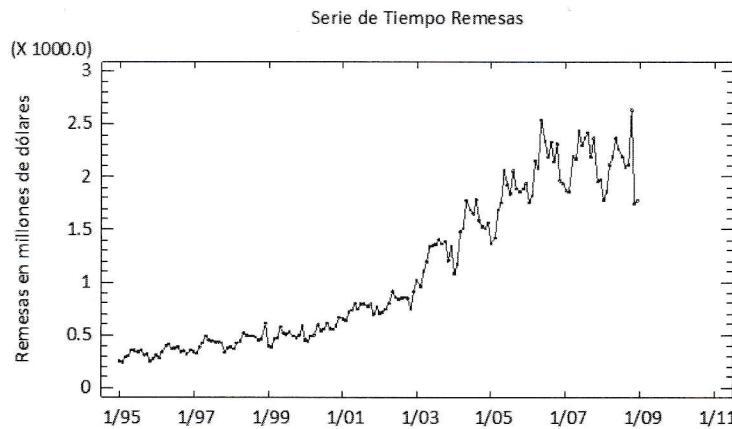


Figura 7.1: Remesas familiares totales.

Se eligieron estos modelos debido a que generan los pronósticos de mayor precisión, si se toma como base el Criterio de Información de Akaike (AIC), que es un instrumento para comparar modelos estadísticos [53]. El mejor modelo será aquel cuyo AIC sea menor para el periodo de validación, es decir, para “pronosticar” los datos que se han retirado de manera temporal para efectuar estas pruebas.

## 7.6 Ajuste del modelo

De acuerdo con la Figura 7.1, los datos son heteroscedásticos, es decir, muestran varianza creciente. Por ello, la varianza se estabilizó con una transformación logaritmo natural cuya gráfica se ve en la Figura 7.2.

Una vez homogeneizada la varianza, es necesario eliminar la tendencia, lo cual se hizo a través de una diferencia ordinaria cuyos resultados se muestran en la Figura 7.3.

La serie modificada ya es estacionaria aunque todavía presenta variación estacional, como puede verse en la Figura 7.4, correspondiente a la ACF; en la Figura 7.5 que muestra la PACF; y en la Figura 7.6 que incluye el periodograma integrado.

Se probaron varios modelos para esta nueva serie, con doce datos para validación y un pronóstico a tres años (36 datos). El mejor modelo resultante de este proceso de comparación fue ARIMA(2,1,0)×(2,0,0)<sub>12</sub>, para el logaritmo natural de los datos originales. Esto significa que el modelo refleja tendencia (una diferencia ordinaria) y variación estacional. Dentro de la estimación de los parámetros se utilizó el pronóstico hacia atrás (*backforecasting*).

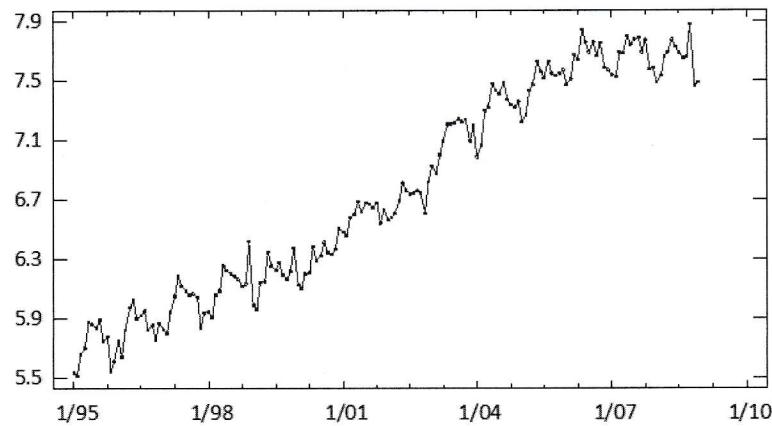


Figura 7.2: Logaritmo natural de las remesas.

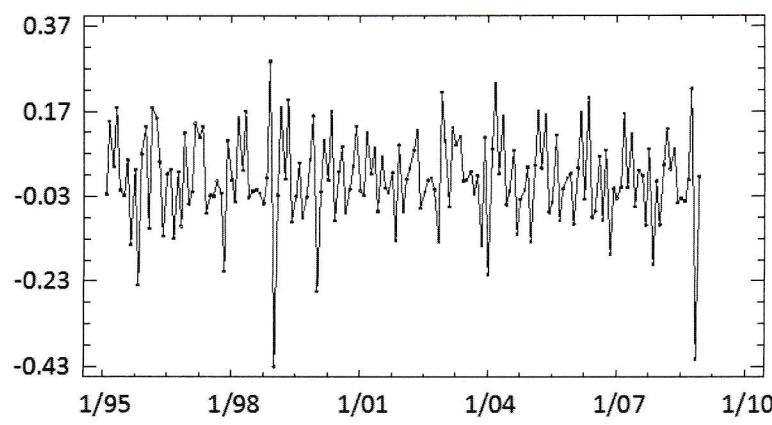


Figura 7.3: Diferencia del logaritmo natural de las remesas.

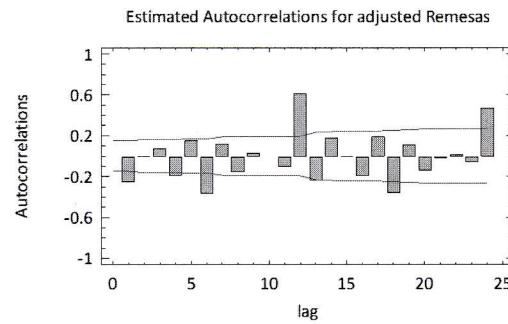


Figura 7.4: ACF de la serie modificada.

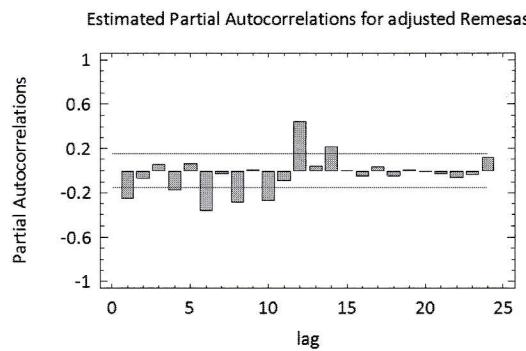


Figura 7.5: PACF de la serie modificada.

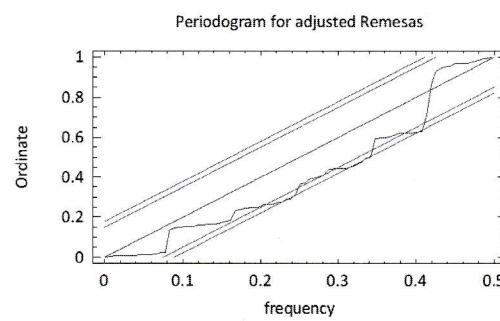


Figura 7.6: Periodograma integrado de la serie modificada.

## 7.7 Resultados

### 7.7 Resultados

Los valores resultantes de la estimación de parámetros del modelo seleccionado, es decir, del modelo ARIMA(2,1,0)  $\times$  (2,0,0)<sub>12</sub>, se muestran en el Cuadro 7.1. En este mismo cuadro se verifica que todos los parámetros resultan significativos, ya que sus estadísticos  $t$  son mayores a dos y sus valores-P son menores que  $\alpha = 0.05$ . La constante se eliminó por no ser significativa.

Cuadro 7.1: Parámetros del modelo ARIMA.				
Parámetro	Estimador	Error estándar	$t$	Valor-P
AR(1)	-0.311880	0.0757964	-4.11470	0.000062
AR(2)	-0.364052	0.0773783	-4.70483	0.000006
SAR(1)	0.562169	0.0739487	7.60215	0.000000
SAR(2)	0.334525	0.0752590	4.44499	0.000016

De aquí que el modelo para los datos será, de acuerdo con estos valores:

$$(1 - 0.5622B^{12} - 0.3345B^{24}) \times (1 + 0.3312B + 0.3641B^2)Z_t = e_t \quad (7.1)$$

Donde:  $Z_t = \Delta \ln(Y_t)$ .

### 7.8 Pronósticos

En cuanto a los pronósticos, es notable el cambio en la tendencia. En lugar del crecimiento sostenido anterior, el modelo muestra un patrón a la baja. Las líneas hacia arriba y abajo del pronóstico son los límites del intervalo de confianza, es decir, se prevé que el valor real de las remesas esté entre estos dos valores, con un 95 por ciento de confianza. En los datos de 1995 a 2008 se observa un ajuste excelente del modelo propuesto.

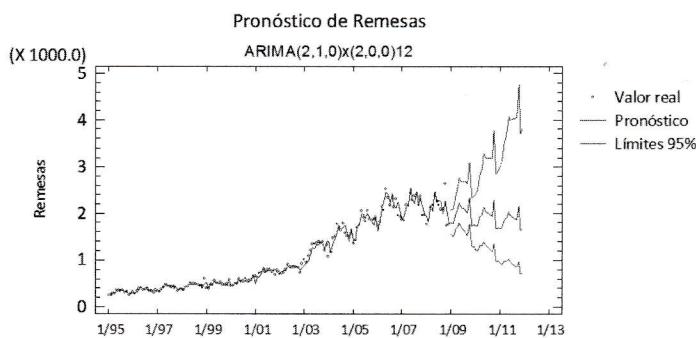


Figura 7.7: Datos y pronósticos de las remesas.

Por otro lado, en la Figura 7.8 se observan únicamente los pronósticos con sus correspondientes intervalos de confianza del 95 por ciento. En este caso

son notables los valores altos para los meses de mayo (día de las madres) y octubre, con un valor mínimo en noviembre de cada año.



Figura 7.8: Pronósticos de las remesas.

## 7.9 Discusión

El modelo obtenido puede considerarse apropiado, ya que sus residuales se comportan como ruido blanco, cumple satisfactoriamente con las pruebas estadísticas y, durante el periodo de validación, genera pronósticos ajustados a los datos reales. Además, todos los coeficientes son estadísticamente significativos.

Como es evidente en la Figura 7.1, los datos muestran heteroscedasticidad que se estabilizó a través de la transformación logaritmo natural, así como una tendencia dividida en tres intervalos con distinto comportamiento. Un periodo de crecimiento moderado de 1995 a 2000; otro con incremento más pronunciado de 2001 a 2005; y finalmente el periodo 2006 a 2008 muestra inicialmente un comportamiento similar al periodo anterior, hasta el año 2007, en el cual la pendiente parece tender a cero o inclusive hacerse negativa.

De acuerdo con lo anterior, el modelo pronostica, para los años de 2009 a 2011, una pendiente negativa, con intervalos de confianza crecientes, debidos a la heteroscedasticidad ya mencionada.

De manera consistente con los datos, este modelo predice valores altos de las remesas para los meses de mayo y octubre. El mes de mayo se explica porque hay un envío importante debido a los festejos del día de la Madre. El incremento de octubre es consistente en los últimos cuatro años, pero su origen no es claro y requiere de más investigación. Una hipótesis es que esté relacionado con la fiesta de San Judas Tadeo, muy venerado en México, que se celebra el 28 de octubre. Por otra parte, los meses de noviembre y diciembre se aprecia un decremento sistemático, muy posiblemente asociado con el hecho de que son temporadas en las que muchos de los migrantes regresan a México a pasar las fiestas de navidad y traen con ellos los recursos, en lugar de enviarlos por los medios tradicionales.

## 7.9 Discusión

---

El pronóstico parece apoyar entonces la hipótesis de que las remesas que ingresan a nuestro país disminuirán en cantidad, frecuencia y monto durante los próximos años. Esto debe tomarse en cuenta ya que han sido consideradas como un ingreso sustancial que repercute tanto en las familias de los migrantes (efecto microeconómico), como en el país que en años recientes se ha sustentado de manera significativa en ellas (efecto macroeconómico). Esto coincide con Cortina *et al.* [21] quienes consideran “probable que el flujo de remesas a México decrezca durante esta década” y que “el gobierno de México cometería un error si supone que los flujos de remesas continuarán creciendo”.

Ante esta situación, el propio Banco de México señala que “México es uno de los países con mayor flujo de emigrantes del mundo” y que “esto refleja la necesidad de llevar a cabo cambios estructurales todavía pendientes en nuestro país, que permitan generar mayores oportunidades de trabajo productivo y bien remunerado” [6]. Asimismo, indica que las remesas familiares no son una fuente de ingresos sostenible a largo plazo, en virtud de que las segundas y terceras generaciones de mexicanos se han establecido allá como familias completas, además de que el flujo de emigrantes puede variar debido a razones diversas como el reforzamiento de la seguridad en las fronteras y el endurecimiento de las políticas de contratación de migrantes.

A estas preocupaciones debe agregarse la posibilidad de que muchos migrantes regresen al país, para encontrar un mercado de trabajo contraído:

La disminución de las transferencias impactará el consumo privado. Se pronostica, además, un retorno masivo de inmigrantes; el gobierno de la ciudad de México prevé el regreso de al menos 30 mil inmigrantes más en los próximos meses, lo que incrementará la presión sobre el ya débil mercado laboral. La información más reciente del Centro Estadounidense de Estudios de Inmigración indica una importante disminución del número de inmigrantes latinos en el país durante mayo-enero, resultado, en su mayoría, de la decisión de los propios indocumentados de abandonar el país [55].

El problema de la migración mexicana a los Estados Unidos es complejo y multifactorial. Inclusive la aplicación del término inmigrante a los mexicanos que llegan a este país ha sido cuestionado con diversos argumentos, como los que plantea Hernández Zamora [28]: “Nosotros no cruzamos la frontera; la frontera nos cruzó a nosotros” o “¿Quiénes son los nativos americanos? ¿Los pueblos nativos de América o los migrantes que vinieron de Europa?”

Algunos autores como McKenzie y Rapoport [34] consideran que la migración tiende a reducir la desigualdad económica, en la medida en que se crean y refuerzan las redes sociales de mexicanos migrantes en los Estados Unidos. Sin embargo, parece más apropiada la posición de Aragónés *et al.* [2] en cuanto a que “los aportes de las remesas no deben destinarse a infraestructura, ésa es la obligación del gasto público del gobierno, no de los salarios que obtienen los trabajadores en Estados Unidos”. En este mismo sentido, Canales y Armas [17] señalan que “las remesas en ningún caso son la solu-

ción a este abandono del Estado y al fracaso de la iniciativa privada: son su consecuencia, el único recurso que les queda a los migrantes y sus familias.”

La posible baja en las remesas apoya la tesis de que el modelo de desarrollo económico basado en estos recursos no sólo distorsiona la noción misma de desarrollo, sino que “esconde las causas de fondo de la migración bajo el espejismo de una economía ficticia e insustentable creado por la reciente dependencia de las remesas” [60].

La situación actual de crisis económica mundial y de recesión creciente se yergue como una amenaza real para los migrantes mexicanos que trabajan en los Estados Unidos y para los familiares que dependen de ellos en México. Esto hace que se requieran hoy, más que nunca, de investigaciones que permitan entender mejor este problema tan complejo y eventualmente ofrezcan información confiable a quienes desde sus posiciones de poder político o económico, puedan tomar las mejores decisiones en beneficio de los trabajadores migrantes, de sus familias y de nuestra nación.

Para México las remesas han representado una gallina de los huevos de oro, pues han permitido subsanar la falta de empleos bien remunerados del país y la asistencia deficitaria del gobierno mexicano en muchos rubros como vivienda, salud y educación. Sin embargo, los pronósticos hacen prever una “baja de postura” para los próximos años, y no hay datos que indiquen que éste sea un fenómeno reversible.

## 7.10 Actividades sugeridas

1. Realiza una búsqueda en el sitio del ISI *Web of Knowledge* (<http://isiknowledge.com>) con las palabras clave (ARIMA OR (Box AND Jenkins)) en el campo *Topic*. Reporta el número de artículos recabados en esta búsqueda y señala tu opinión al respecto.
2. Realiza una búsqueda en el sitio de REDALYC, ubicado en <http://redalyc.uaemex.mx/>, con las palabras clave (ARIMA OR (Box AND Jenkins)) en el campo que existe para ello. Reporta el número de artículos recabados en esta búsqueda y señala tu opinión al respecto.
3. Realiza una búsqueda en el sitio de ScIELO (<http://www.scielo.org.mx/scielo.php>) con las palabras clave (ARIMA OR (Box AND Jenkins)) en el campo *Título*. Reporta el número de artículos recabados en esta búsqueda y señala tu opinión al respecto.
4. Recupera algún artículo de investigación original realizado con la metodología de Box-Jenkins, en alguno de los sitios mencionados arriba. Léelo con cuidado. Identifica cada una de sus partes y escribe un resumen,

## CONCLUSIÓN

---

tus conclusiones y comentarios al respecto. Comparte la información con tus compañeros.

5. Escribe un artículo de investigación semejante al del ejemplo de este texto. Envíalo a una revista científica para su publicación.

# Conclusión

En el paso decisivo del descubrimiento se suspenden los códigos  
del razonamiento disciplinado, tal como ocurre en sueños y ensueños,  
en vuelos alocados del pensamiento, cuando las corrientes de ideas  
pueden desviarse de su cauce, por la gravedad emocional,  
en una forma que aparenta no seguir reglas.

ARTHUR KOESTLER

En este texto se han expuesto los conceptos principales relacionados con los pronósticos y las series de tiempo. Se detalló la metodología de Box-Jenkins y la forma de encontrar el modelo más apropiado para cualquier conjunto de datos cuantitativos observados. La mayor ventaja de este enfoque es que sólo se requiere de un conjunto de sesenta o más observaciones para obtener, en la mayoría de los casos, un buen pronóstico.

Por supuesto, la metodología de Box-Jenkins no es el único enfoque que existe para estimar valores futuros, ni es el mejor en todos los casos. Sin embargo, suele arrojar pronósticos bastante precisos y cercanos a la realidad cuando los fenómenos presentan cierta estabilidad a lo largo del tiempo. Es un modelo matemático sencillo, que no tiene la capacidad de prever situaciones complejas o sorpresivas. Sin embargo, puede combinarse con metodologías de corte cualitativo, para ampliar el panorama y tomar en cuenta aspectos externos a los datos que pudieran influir en ellos. Asimismo, existe el llamado “análisis de intervención”, cuyo objetivo es determinar si existe un cambio en la tendencia de las observaciones, ocasionada por algún hecho o intervención. Por ejemplo, puede evaluarse si la contaminación en una zona es menor después de aplicar un programa de tipo “hoy no circula” o “verificación de autos”. También podría analizarse el desempeño académico antes y después de introducir un ambiente virtual de aprendizaje. En fin, puede aplicarse a cualquier conjunto de datos en los que se presente un cambio significativo.

Contar con buenas estimaciones acerca de lo que sucederá más adelante puede permitir a las personas, empresas, organizaciones y países, tomar las mejores decisiones. A su vez, las decisiones acertadas promoverán avances sociales, económicos, personales, entre otros. De ahí la importancia de la aplicación de métodos objetivos en la elaboración de pronósticos.

## CONCLUSIÓN

---

Si bien la metodología puede parecer laboriosa, hoy en día se cuenta con muchas aplicaciones de software estadístico que permiten que el analista se centre en la selección de un modelo apropiado para los datos. Desarrollos como *Statgraphics* o SAS ofrecen la opción de probar y seleccionar, de manera automática, el “mejor” modelo para pronosticar. Esto se hace a través de criterios que pueden fijarse de antemano. Tal vez el lector se pregunte por qué se revela hasta ahora este pequeño secreto. En realidad, la forma de explotar al máximo este tipo de apoyos computacionales es contar con una comprensión clara acerca del funcionamiento de un método, su forma de uso, sus restricciones y deficiencias, así como sus ventajas.

Por otro lado, existen herramientas de software gratuito de alta calidad, a disposición de los interesados en pronosticar. Ejemplo de ellas son la versión *Statgraphics Online* y el software libre R. Así que no es estrictamente necesario invertir en estos recursos y el analista puede dirigir sus esfuerzos al desempeño intelectual.

Por todo lo anterior, en este texto pretende se espera brindar una guía práctica para la elaboración de pronósticos con modelos ARIMA. En particular, se presentan ejemplos que pertenecen al contexto de nuestro país y que muestran problemas actuales, como la inflación, el desempleo y el turismo. Asimismo, se proponen algunas actividades que no sólo tienen el propósito de reforzar el aprendizaje, sino de motivar al estudiante y hacerlo partícipe de la solución de problemas concretos.

Es importante destacar también que cada vez se cuenta con más datos y fuentes confiables en nuestro país. Por ejemplo, el Banco de México y el Instituto Nacional de Estadística y Geografía (INEGI), así como algunas de las secretarías de estado y otros organismos, ofrecen sitios web con información actualizada. Además, por lo general los datos se presentan en formatos manejables, como archivos de *Excel*, que permiten su manipulación y exportación al software estadístico.

También debe señalarse que tanto los servicios de información que involucran un costo, como el ISI *Web of Knowledge*, como los de acceso abierto, tales como *Google Académico*, REDALYC y SCIELO, ofrecen la posibilidad de localizar y recuperar artículos de investigación de gran calidad. Entre ellos existen muchos que presentan resultados obtenidos a través de la aplicación de la metodología de Box-Jenkins, de manera que cualquier persona interesada puede revisar la forma en que este método se utiliza en otros países y en diversos ámbitos del conocimiento.

Así pues, hoy en día el interesado en pronosticar a través de datos cuantitativos tiene opciones y facilidades para hacerlo. Sin embargo, no debe olvidarse que los métodos matemáticos funcionan de manera insensible a la realidad. Es decir, es indispensable que el analista tome en cuenta el contexto y toda la información disponible sobre el problema que estudia. En algunas ocasiones será necesario entrevistar a expertos e involucrados. A veces una serie de tiempo puede ser afectada por eventos locales, nacionales o internacionales. De ahí que sea conveniente, en la medida de lo posible, combinar el método matemático cuantitativo con alguna técnica de tipo cualitativo.

## CONCLUSIÓN

Quien aplique en este sentido las matemáticas tiene la responsabilidad de aprovechar sus conocimientos para contribuir en la prevención de los acontecimientos futuros, de tal manera que puedan tomarse las decisiones que sean siempre de mayor beneficio para las comunidades. En todos los casos, habrá de combinar las herramientas matemáticas y tecnológicas con el sentido común, la experiencia, el conocimiento de las características peculiares de los fenómenos analizados, la capacidad de dialogar y preguntar, la ética y los principios, para lograr los resultados óptimos.

# Bibliografía

- [1] ALTAMIRANO, T.: «Transnacionalismo, remesas y economía doméstica.» *Cuadernos Electrónicos de Filosofía del Derecho*, 2004, (10), pp. 1–31.
- [2] ARAGONÉS, ANA MARÍA; SALGADO, U. y RÍOS, E.: «¿A quién benefician las remesas?» *Economía UNAM*, 2008, p. 37.
- [3] ARMSTRONG, JOHN SCOTT: *Long Range Forecasting. From crystal ball to computers*. Wiley, USA, 1978.
- [4] AYDEMIR, A. y BORJAS, G. J.: «Cross-country variation in the impact of international migration: Canada, Mexico, and the United States». *Journal of the European Economic Association*, 2007, 5(4)(4), pp. 663–708.
- [5] BANCO DE MÉXICO: «Glosario». En línea, 2008. Fecha de consulta: 03/04 2009.  
<http://www.banxico.gob.mx/tipo/didactico/glosario.pdf>
- [6] ———: «Remesas». En línea, 2008. Fecha de consulta: 03/11 2009.  
<http://www.banxico.gob.mx/audiencias/estudiantes/Remesas.pdf>
- [7] ———: «Balanza de pagos». En línea, 2009. Fecha de consulta: 02/21/2009.  
<http://www.banxico.gob.mx/polmoneinflacion/estadisticas/>
- [8] ———: «Colección numismática». Sitio web, 2009. Fecha de consulta: 2009/12/29.  
<http://www.banxico.org.mx/billetes-y-monedas/informacion-general/colección-numismática/página-colección-numismática.html>
- [9] BARTLETT, M. S.: «On the Theoretical Specification and Sampling Properties of Autocorrelated Time-Series». *Supplement to the Journal of the Royal Statistical Society*, 1946, 8(1), pp. 27–41.
- [10] BOVAS, ABRAHAM y LEDOLTER, JOHANNES: *Statistical Methods for Forecasting*. Wiley Series in Probability and Mathematic Statistics. Wiley, New York, 1983.
- [11] BOWERMANN, B. L. y O'CONNELL, R. T.: *Time Series and Forecasting*. Duxbury Press, Boston, USA, 1979.
- [12] BOX, GEORGE E. P. y JENKINS, GWILYM: *Time series analysis. Forecasting and control*. Holden Day, 1976.

## BIBLIOGRAFÍA

---

- [13] Box, GEORGE E. P. y PIERCE, D. A.: «Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models». *Journal of the American Statistical Association*, 1970, **65(332)**, pp. 1509–1526.
- [14] Box, GEORGE E. P y TIAO, G. C.: «Intervention Analysis with Applications to Economic and Environmental Problems». *Journal of the American Statistical Association*, 1975, **70(349)**, pp. 70–79.
- [15] BROCKWELL, P.J. y DAVIS, R.A.: *Time series: theory and methods*. Springer, 1991.
- [16] BURDEN, R. L. y FAIRES, J. D.: *Análisis Numérico*. Thomson Learning Ibero, 2001.
- [17] CANALES, A. y ARMAS, I.: «Remesas e Inversión Productiva en Comunidades de Alta Migración a Estados Unidos. El Caso de Teocaltiche, Jalisco». *Migraciones internacionales*, 2004, **2(3)**, pp. 142–172.
- [18] CHATFIELD, C.: *The analysis of time series: an introduction*. CRC Press, 2004.
- [19] CONSEJO NACIONAL DE POBLACIÓN: «Población residente en Estados Unidos», 2007. Fecha de consulta: 03/14/2009.  
<http://www.conapo.gob.mx/>
- [20] ——: «Series sobre Migración Internacional 2008», 2008. Fecha de consulta: 03/14/2009.  
<http://www.conapo.gob.mx/>
- [21] CORTINA, J.; DE LA GARZA, R. y OCHOA-REZA, E.: «Remesas: límites al optimismo». *Foreign Affairs en Español*, 2005, **34**, p. 101.
- [22] DELGADO WISE, R. y FAVELA GAVIA, D. M.: *Nuevas tendencias y desafíos de la migración internacional México-Estados Unidos*. UNAM - Universidad Autónoma de Zacatecas - Miguel Ángel Porruá, 2004.
- [23] DURY, J. y MASSEY, D.: *Crossing the Border: Research from the Mexican Migration Project*. Russell Sage Foundation Publications, 2006.
- [24] DUTTA, M.: *Econometric Models*. South Western Publishing, USA, 1975.
- [25] ENDERS, W.: *Applied Econometric Time Series*. Wiley New York, 1995.
- [26] GONZÁLEZ VIDEGARAY, MARICARMEN: *Modelos de Decisión con Procesos Estocásticos II. Metodología de Box-Jenkins*. UNAM - ENEP Acatlán, México, 1990.
- [27] HARVEY, A. C.: *Time Series Models*. Phillip Allan Publishers, England, 1981.
- [28] HERNÁNDEZ ZAMORA, GREGORIO: «¿Somos “Migrantes”?» *La Jornada Migración*, 2008.

## BIBLIOGRAFÍA

---

- [29] INSTITUTO NACIONAL DE ESTADÍSTICA Y GEORGRAFÍA: «Tasa de Desocupación Abierta en México», 2009. Fecha de consulta: 2009/12/20.  
<http://www.inegi.gob.mx>
- [30] LUDLOW WIECHERS, JORGE: *Modelos, Pronósticos y Volatilidad de las Series de Tiempo Generadas en la Bolsa Mexicana de Valores*. Economía. UAM Azcapotzalco, 1997.
- [31] MAKRIDAKIS, SPYROS y WHEELWRIGHT, STEVEN C.: *Forecasting Methods and Applications*. Wiley, 1983.
- [32] ———: *Forecasting for Management*. Wiley, USA, 1985.
- [33] MARRONI, M. G.: «Migrantes mexicanas en los escenarios familiares de las comunidades de origen: amor, desamor y dolor». *Estudios Sociológicos*, 2006, **24**, p. 72.
- [34] MCKENZIE, D. y RAPOPORT, H.: «Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico». *Journal of Development Economics*, 2007, **84**(1), pp. 1–24.
- [35] MENTZER, JOHN T. y BIENSTOCK, CAROL C.: *Sales Forecasting Management*. SAGE Publications, 1998.
- [36] ORGANIZACIÓN DE LAS NACIONES UNIDAS: «Declaración Universal de los Derechos humanos. Adoptada y proclamada por la Resolución de la Asamblea General 217 A (iii) del 10 de diciembre de 1948, Organización de las Naciones Unidas». En línea, 1998. Fecha de consulta: 2008/03/05.  
<http://www.un.org/spanish/aboutun/rights.htm>
- [37] PALIT, AJOY K. y POPOVI, DOBRIVOJE: *Computational Intelligence in Time Series Forecasting: theory and engineering applications*. Birkhäuser, 2005. Definición.
- [38] PEÑA, D. y BOX, GEORGE E. P.: «Identifying a Simplifying Structure in Time Series». *Journal of the American Statistical Association*, 1987, **82**(399), pp. 836–843.
- [39] PERCIVAL, DONALD B. y WALDEN, ANDREW T.: *Wavelet Methods for Time Series Analysis*. Cambridge University Press, 2006.
- [40] PHILLIPS, TONY: «Long Range Solar Forecast», 2006. Fecha de consulta: 2009/12/28.  
[http://science.nasa.gov/headlines/y2006/10may\\_longrange.htm](http://science.nasa.gov/headlines/y2006/10may_longrange.htm)
- [41] PINDYCK, R. y RUBINFELD, D.: *Econometric Models and Economic Forecasts*. McGraw-Hill, USA, 1981.
- [42] QUANTITATIVE MICRO SOFTWARE: «EViews». Software, 2009. Fecha de consulta: 2009/12/28.  
<http://www.eviews.com/>
- [43] QUENOUILLE, M. H.: «Approximate Tests of Correlation in Time-Series». *Journal of the Royal Statistical Society. Series B (Methodological)*, 1949, **11**(1), pp. 68–84.

## BIBLIOGRAFÍA

---

- [44] R FOUNDATION: «The R Project for Statistical Computing». Software libre, 2009. Fecha de consulta: 2009/12/28.  
<http://www.r-project.org/>
- [45] SAS INSTITUTE: «Econometrics & Time Series Analysis with SAS/ETS® Software». Software, 2009. Fecha de consulta: 2009/12/28.  
<http://www.sas.com/technologies/analyticforecasting/ets/>
- [46] SHARMA, J. N.: *Numerical methods for engineers and scientists*. Alpha Science International Ltd., 2004.
- [47] SISTEMA NACIONAL DE INFORMACIÓN ESTADÍSTICA Y GEOGRÁFICA: «Banco de Información Económica - Tasa de desocupación mensual abierta», 2009. Fecha de consulta: 2009/12/20.  
<http://dgcnesyp.inegi.org.mx/cgi-win/bdieintsi.exe/>
- [48] ———: «Banco de Información Económica - Turismo receptivo (miles de personas)», 2009. Fecha de consulta: 2009/12/20.  
<http://dgcnesyp.inegi.org.mx/cgi-win/bdieintsi.exe/Consultar>
- [49] SPSS INC.: «SPSS 15.0 para Windows». Software, 2006.  
<http://www.spss.com/>
- [50] STATGRAPHICS.NET: «Statgraphics Online». Software, 2010.  
<http://www.statgraphicsonline.com/>
- [51] STATPOINT: «Statgraphics Centurion XV Versión 15.1.02». Software, 2006. Fecha de consulta: 2009/12/20.  
<http://statgraphics.com/>
- [52] TAYLOR, STEPHEN J.: *Modelling financial time series*. World Scientific, 2007. Definición.
- [53] THE MATHWORKS: «System Identification Toolbox». En línea, 2009. Fecha de consulta: 02/21 2009.  
<http://www.mathworks.com/>
- [54] TIAO, G. C. y BOX, GEORGE E. P.: «Modeling Multiple Times Series with Applications». *Journal of the American Statistical Association*, 1981, **76(376)**, pp. 802–816.
- [55] UNIT, ECONOMIST INTELLIGENCE: «Más desempleo, menos remesas». *La Jornada*, 2008.
- [56] VANDAELE, WALTER: *Applied Time Series and Box-Jenkins Models*. Academic Press, USA, 1983.
- [57] VARGAS-SILVA, C.: «The Tale of Three Amigos: Remittances, Exchange Rates, and Money Demand in Mexico.» *Review of Development Economics*, 2009, **13(1)**, pp. 1–14.
- [58] WEI, WILLIAM S.: *Time Series Analysis. Univariate and Multivariate Methods*. Addison Wesley, USA, 1990.

## BIBLIOGRAFÍA

---

- [59] WILKINSON, A.M.: *The scientist's handbook for writing papers and dissertations*. Prentice Hall Englewood Cliffs, NJ, 1991.
- [60] WISE, R. y COVARRUBIAS, H.: «La migración mexicana hacia Estados Unidos a la luz de la integración económica regional: nuevo dinamismo y paradojas». *Theomai*, 2006, p. 6.
- [61] WONG, R.; PALLONI, A. y SOLDÓ, B. J.: «Wealth in middle and old age in Mexico: The role of international migration». *International Migration Review*, 2007, **41(1)**, pp. 127–151.
- [62] XIE, ZHONGJIE: *Case studies in time series analysis*. World Scientific, 1992. Definición.

# Índice alfabético

- ACF, 26
- IMRYD, 11
- PACF, 26, 64
- actualización
  - adaptativa, 104
  - secuencial, 104
- amplitud, 27
- análisis de intervención, 135
- armónicas de Fourier, 31, 32
- autocorrelación parcial, 63
- Box-Jenkins
  - metodología, 23
- caminata aleatoria, 117
- choques aleatorios, 15
- ciclo, 19
- coeficientes de Fourier, 32
- componentes de una serie de tiempo,
  - 17
- condición de
  - estacionaridad, 66
  - invertibilidad, 74
- correlograma, 53
- costo del pronóstico, 99
- criterio 1, 125
- criterio 2, 125
- dato presente, 99
- datos
  - atípicos, 135
  - discrepantes, 135
  - influyentes, 135
  - faltantes, 135
  - históricos, 99
  - prehistóricos, 99
- decisiones, 13
- deflactar, 87
- diferencias
  - estacionales, 86, 91
  - finitas *véase* ordinarias 90
  - no estacionales *véase* ordinarias 90
  - ordinarias, 86, 90
- dualidad de modelos, 79
- ecuaciones de Yule-Walker, 71
- error de pronóstico, 99
- espectro lineal, 37
- estacionariedad, 49
  - en sentido amplio, 51
  - en sentido estricto, 49
- estadístico de Box-Pierce, 110
- estimación de parámetros del modelo, 117
- estimador
  - por intervalo, 98
  - puntual, 97
- estimador de la autocorrelación, 52
- estimador de la autocorrelación parcial, 64
- estimador de la media, 51
- estimador de la varianza, 52
- estimadores, 119
- extrapolación, 97
- fase, 27
- filtros, 23
- fines del modelado matemático, 9
- fluctuación aleatoria, 20
- frecuencia, 27
- frecuencias de Fourier, 32
- función

## ÍNDICE ALFABÉTICO

---

- de autocorrelación, 26, 52  
de autocorrelación muestral, 46  
de autocorrelación parcial, 26, 64  
de autocorrelación parcial muestral, 64  
de autocovarianza, 26  
de autocovarianza muestral, 46  
de media, 26  
de media muestral, 46  
de transferencia, 22  
de varianza, 26  
de varianza muestral, 46  
función de verosimilitud  
    condicional, 120  
    logarítmica condicional, 121  
función seno generalizada, 31  
golpes aleatorios, véase choques aleatorios  
heteroscedasticidad, 26  
hipótesis  
    alternativa, 108  
    nula, 108  
homoscedasticidad, 26  
horizonte de pronóstico, 99  
identificación del modelo, 108  
información  
    cuantitativa, 14  
    cuantitativa, 14  
inicialización de la serie de tiempo, 107  
integración, 91  
interpolación, 97  
linealización, 123  
longitud de periodo, 27  
método de máxima verosimilitud, 106  
método de mínimos cuadrados, 107  
mínimos cuadrados, 119  
matriz  
    de autocorrelaciones, 55  
    de varianzas y autocovarianzas, 55  
metodología de Box-Jenkins  
procedimiento, 105  
modelo  
    ARMA, 15  
    ARMA( $p, q$ ), 77, 78  
    autorregresivo, 65  
    de forma invertida, 61  
    de medias móviles, 72  
    general autorregresivo estacional, 142  
    general estacional, 146  
    general multiplicativo, 147  
    lineal general, 61  
    matemático, 9  
    mezclado ARMA( $p, q$ ), 77  
    no estacionario, 85  
    preliminar, 105  
    tentativo, 105  
modelos  
    ARIMA, 94  
    estacionales, 141  
    estocásticos, 10  
    integrados autorregresivos y de medias móviles, 94  
    multiplicativos, 141, 147  
modelos multivariados, 22  
modelos univariados, 22  
número máximo de iteraciones, 125  
nivel de significancia, 98  
notación abreviada, 78  
observaciones, 17  
operador  
    de salto hacia adelante, 121  
    de salto hacia atrás, 78  
orden  
    de la parte autorregresiva estacional, 142  
    de la parte de medias móviles estacionales, 142  
    de la parte estacional, 141  
    de las diferencias estacionales, 142  
P-value, 110  
parámetros, 119  
periodograma, 41

## ÍNDICE ALFABÉTICO

- predicción, *véase* pronóstico  
principio  
    mejoramiento iterativo, 23  
    parsimonia, 23  
pronóstico, 14, 97  
    hacia atrás, 121  
pronósticos  
    límite de los, 101  
prueba *portmanteau*, 110  
prueba de hipótesis, 109
- random walk *véase* caminata aleatoria 117  
realización, 16, 45  
regresión lineal múltiple, 122  
residuales, 100, 119  
ruido blanco, 20
- serie de Fourier, 32  
serie de Taylor, 123  
serie de tiempo, 15  
    definición, 15  
    objetivos, 21  
        control, 22  
        descripción, 21  
        explicación, 21  
        pronóstico, 22  
series de tiempo  
    análisis de, 17  
significativa, 109  
suma de los cuadrados de los residuales, 120
- tendencia, 18  
    eliminación de, 88  
transformación de Box-Cox, 88  
transformaciones, 87
- validación del modelo, 128  
valor inicial, 123  
valores esperados  
    condicionales, 121  
    incondicionales, 121  
variable endógena, 23  
variable exógena, 23  
variables  
    dependientes, 123  
    endógenas, 123  
    exógenas, 123  
    independientes, 123  
variación estacional, 19  
    tratamiento de la, 91  
varianza  
    estabilización de, 87  
volatilidad, 87

# Índice de cuadros

1.1	Coeficientes de Fourier del Ejemplo 1. . . . .	33
1.2	Datos del Ejemplo 2. . . . .	34
1.3	Coeficientes de Fourier del Ejemplo 2. . . . .	35
1.4	Datos del Ejemplo 3. . . . .	35
1.5	Coeficientes de Fourier del Ejemplo 3. . . . .	36
1.6	Amplitudes y fases del Ejemplo 1. . . . .	38
1.7	Amplitudes y fases del Ejemplo 2. . . . .	39
1.8	Amplitudes y fases del Ejemplo 3. . . . .	40
2.1	Tasa de desocupación abierta mensual en México. . . . .	60
2.2	Dualidad y resumen de modelos ARMA . . . . .	79
6.1	Resumen de modelos ARIMA . . . . .	152
7.1	Parámetros del modelo ARIMA. . . . .	160

# Índice de figuras

1.1	La serie de tiempo. . . . .	16
1.2	Tasa de desocupación abierta mensual en México. . . . .	18
1.3	Turismo receptivo mensual en México. . . . .	19
1.4	Ciclos solares desde 1940. . . . .	20
1.5	Ruido blanco. . . . .	21
1.6	Enfoque de Box-Jenkins. . . . .	24
1.7	Metodología de Box-Jenkins. . . . .	25
1.8	Función seno. . . . .	28
1.9	Funciones seno con diferentes amplitudes. . . . .	28
1.10	Funciones seno con diferentes fases. . . . .	29
1.11	Funciones seno con longitud de 32 puntos. . . . .	30
1.12	Funciones seno con longitud de 16 puntos. . . . .	30
1.13	Ejemplo 1 - Función seno con $A = 1$ , $L = 16$ , $f = 1/16$ , $\phi = 0$ . . . . .	33
1.14	Ejemplo 2 - Función seno con $A = 1$ , $L = 2$ , $f = 8/16$ , $\phi = 0$ . . . . .	34
1.15	Ejemplo 3 - Función con valores aleatorios. . . . .	35
1.16	Cuadrantes del círculo trigonométrico. . . . .	37
1.17	Espectro lineal del Ejemplo 1. . . . .	38
1.18	Espectro lineal del Ejemplo 2. . . . .	39
1.19	Espectro lineal del Ejemplo 3. . . . .	40
1.20	Periodograma de valores aleatorios en SPSS. . . . .	41
1.21	Periodograma del Ejemplo 1. . . . .	42
1.22	Periodograma del Ejemplo 2. . . . .	42
1.23	Periodograma del Ejemplo 3. . . . .	43
1.24	Periodograma Integrado Ejemplo 3. . . . .	44
1.25	Periodograma Integrado Ejemplo 1. . . . .	44
1.26	Periodograma Integrado Ejemplo 2. . . . .	45
2.1	Moneda antigua de veinte centavos mexicanos. . . . .	50
2.2	Ejemplo de realización de un proceso Bernoulli. . . . .	50
2.3	Ejemplo de realización de un proceso binomial. . . . .	51
2.4	Proceso estocástico estacionario. . . . .	54
2.5	Proceso estocástico no estacionario. . . . .	54
2.6	Tasa de inflación mensual en México. . . . .	57
2.7	La tasa de inflación es estacionaria. . . . .	57
2.8	ACF de la tasa de inflación mensual en México. . . . .	58

## ÍNDICE DE FIGURAS

---

2.9 Periodograma de la tasa de inflación mensual en México. . . . .	58
2.10 Periodograma integrado de la tasa de inflación mensual en México. . . . .	59
2.11 Tabla de periodograma para INPC (fragmento). . . . .	59
2.12 Tasa de desocupación abierta mensual. . . . .	61
2.13 Características de la tasa de desocupación. . . . .	62
2.14 ACF de un modelo AR(1) con coeficiente positivo. . . . .	68
2.15 ACF de un modelo AR(1) con coeficiente negativo. . . . .	68
2.16 PACF del modelo AR(1) con coeficiente positivo. . . . .	69
2.17 PACF del modelo AR(1) con coeficiente negativo. . . . .	69
2.18 ACF de un modelo MA(1) con coeficiente positivo. . . . .	74
2.19 PACF de un modelo MA(1) con coeficiente positivo. . . . .	74
2.20 ACF de un modelo MA(1) con coeficiente negativo. . . . .	75
2.21 PACF de un modelo MA(1) con coeficiente negativo. . . . .	75
2.22 Pantalla de captura en <i>Statgraphics</i> . . . . .	80
2.23 Elegir descripción de la serie de tiempo. . . . .	80
2.24 Selección de la serie de tiempo. . . . .	81
2.25 Paneles con características de la serie de tiempo. . . . .	82
2.26 Paneles con otras características de la serie de tiempo. . . . .	82
 3.1 Menú desplegable en <i>Statgraphics</i> . . . . .	88
3.2 Ventana para hacer transformaciones de datos. . . . .	89
3.3 Orden de la posible variación estacional. . . . .	89
3.4 Orden de las diferencias no estacionales. . . . .	92
3.5 Longitud de la variación estacional. . . . .	94
 4.1 Costo del error de pronóstico. . . . .	100
 5.1 Gráfica original del Ejemplo 1. . . . .	112
5.2 ACF del Ejemplo 1. . . . .	113
5.3 PACF del Ejemplo 1. . . . .	113
5.4 Periodograma integrado del Ejemplo 1. . . . .	114
5.5 Gráfica original del Ejemplo 2. . . . .	115
5.6 ACF del Ejemplo 2. . . . .	115
5.7 PACF del Ejemplo 2. . . . .	116
5.8 Periodograma integrado del Ejemplo 2. . . . .	116
5.9 Pronósticos en <i>Statgraphics</i> . . . . .	117
5.10 Especificación de modelos. . . . .	118
5.11 Opciones de estimación. . . . .	126
5.12 Ejemplo de modelo sobreestimado. . . . .	130
5.13 <i>Model Specification Options</i> . . . . .	131
5.14 Modelo sin el parámetro sobrante. . . . .	131
5.15 Comparación de modelos. . . . .	137
5.16 Comparación del periodo de estimación. . . . .	138
5.17 Comparación del periodo de validación. . . . .	138
5.18 Estimación del modelo ARMA(1,1,1). . . . .	138

## ÍNDICE DE FIGURAS

5.19 Pronósticos modelo ARIMA(0,1,1). . . . .	139
6.1 ACF y PACF del modelo SAR(1). . . . .	143
6.2 Periodograma integrado del modelo SAR(1). . . . .	144
6.3 ACF y PACF del modelo SMA(1). . . . .	145
6.4 Periodograma integrado del modelo SMA(1). . . . .	146
6.5 ACF y PACF del modelo SARMA(1,1). . . . .	147
6.6 ACF y PACF del modelo AR(1)SAR(1). . . . .	149
6.7 ACF y PACF del modelo MA(1)SMA(1). . . . .	150
7.1 Remesas familiares totales. . . . .	157
7.2 Logaritmo natural de las remesas. . . . .	158
7.3 Diferencia del logaritmo natural de las remesas. . . . .	158
7.4 ACF de la serie modificada. . . . .	159
7.5 PACF de la serie modificada. . . . .	159
7.6 Periodograma integrado de la serie modificada. . . . .	159
7.7 Datos y pronósticos de las remesas. . . . .	160
7.8 Pronósticos de las remesas. . . . .	161

*Pronosticos. Metodología de Box Jenkins* concluyó su edición el 1 de marzo del 2011 por la FES Acatlán y se terminó de imprimir el 6 de mayo del 2011 en los talleres de Tipos Futura, S. A. de C. V. Francisco González Bocanegra 47-B, Col. Ampliación Morelos, México, D.F. Se tiraron 300 ejemplares con interiores en papel Bond blanco de 97 gramos, y cubierta en cartulina sulfatada de 12 pts. El cuidado de la edición estuvo a cargo del Departamento de Tipografía y Diseño de la Unidad de Servicios Editoriales de la FES Acatlán. Corrección de estilo: Leticia Cuevas Muñoz.

La edición estuvo al cuidado de la Unidad de Servicios Editoriales de la FES Acatlán UNAM. La corrección de estilo la realizó Leticia Muñoz Cuevas



ISBN: 978-607-02-2223-8

9 786070 222238