

# IBM SPSS Forecasting 19



*Note:* Before using this information and the product it supports, read the general information under Notices el p. 116.

This document contains proprietary information of SPSS Inc, an IBM Company. It is provided under a license agreement and is protected by copyright law. The information contained in this publication does not include any product warranties, and any statements provided in this manual should not be interpreted as such.

When you send information to IBM or SPSS, you grant IBM and SPSS a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

**© Copyright SPSS Inc. 1989, 2010.**

---

# Prefacio

IBM® SPSS® Statistics es un sistema global para el análisis de datos. El módulo adicional opcional Predicciones proporciona las técnicas de análisis adicionales que se describen en este manual. El módulo adicional Predicciones se debe utilizar con el sistema básico de SPSS Statistics y está completamente integrado en dicho sistema.

## ***Acerca de SPSS Inc., an IBM Company***

SPSS Inc., an IBM Company, es uno de los principales proveedores globales de software y soluciones de análisis predictivo. La gama completa de productos de la empresa (recopilación de datos, análisis estadístico, modelado y distribución) capta las actitudes y opiniones de las personas, predice los resultados de las interacciones futuras con los clientes y, a continuación, actúa basándose en esta información incorporando el análisis en los procesos comerciales. Las soluciones de SPSS Inc. tratan los objetivos comerciales interconectados en toda una organización centrándose en la convergencia del análisis, la arquitectura de TI y los procesos comerciales. Los clientes comerciales, gubernamentales y académicos de todo el mundo confían en la tecnología de SPSS Inc. como ventaja ante la competencia para atraer, retener y hacer crecer los clientes, reduciendo al mismo tiempo el fraude y mitigando los riesgos. SPSS Inc. fue adquirida por IBM en octubre de 2009. Para obtener más información, visite <http://www.spss.com>.

## ***Asistencia técnica***

El servicio de asistencia técnica está a disposición de todos los clientes de mantenimiento. Los clientes podrán ponerse en contacto con este servicio de asistencia técnica si desean recibir ayuda sobre la utilización de los productos de SPSS Inc. o sobre la instalación en alguno de los entornos de hardware admitidos. Para ponerse en contacto con el servicio de asistencia técnica, consulte el sitio web de SPSS Inc. en <http://support.spss.com> o encuentre a su representante local a través del sitio web <http://support.spss.com/default.asp?refpage=contactus.asp>. Tenga a mano su identificación, la de su organización y su contrato de asistencia cuando solicite ayuda.

## ***Servicio de atención al cliente***

Si tiene cualquier duda referente a la forma de envío o pago, póngase en contacto con su oficina local, que encontrará en el sitio Web en <http://www.spss.com/worldwide>. Recuerde tener preparado su número de serie para identificarse.

## ***Cursos de preparación***

SPSS Inc. ofrece cursos de preparación, tanto públicos como in situ. Todos los cursos incluyen talleres prácticos. Los cursos tendrán lugar periódicamente en las principales ciudades. Si desea obtener más información sobre estos cursos, póngase en contacto con su oficina local que encontrará en el sitio Web en <http://www.spss.com/worldwide>.

## ***Publicaciones adicionales***

Los documentos *SPSS Statistics: Guide to Data Analysis*, *SPSS Statistics: Statistical Procedures Companion* y *SPSS Statistics: Advanced Statistical Procedures Companion*, escritos por Marija Norušis y publicados por Prentice Hall, están disponibles y se recomiendan como material adicional. Estas publicaciones cubren los procedimientos estadísticos del módulo SPSS Statistics Base, el módulo Advanced Statistics y el módulo Regression. Tanto si da sus primeros pasos en el análisis de datos como si ya está preparado para las aplicaciones más avanzadas, estos libros le ayudarán a aprovechar al máximo las funciones ofrecidas por IBM® SPSS® Statistics. Si desea información adicional sobre el contenido de la publicación o muestras de capítulos, consulte el sitio web de la autora: <http://www.norusis.com>

---

# Contenido

## **Parte I: Manual del usuario**

### **1 Introducción a las series temporales 1**

Datos de serie temporal . . . . .	1
Transformaciones de los datos . . . . .	2
Períodos de estimación y de validación . . . . .	2
Generación de modelos y producción de predicciones . . . . .	2

### **2 Modelizador de series temporales 4**

Especificación de opciones para el modelizador experto . . . . .	8
Selección de modelo y especificación de eventos . . . . .	8
Tratamiento de valores atípicos con el modelizador experto. . . . .	10
Modelos personalizados de suavizado exponencial . . . . .	11
Modelos ARIMA personalizados . . . . .	12
Especificación de modelo para modelos ARIMA personalizados . . . . .	13
Funciones de transferencia en modelos ARIMA personalizados. . . . .	15
Valores atípicos en modelos ARIMA personalizados . . . . .	17
Resultados . . . . .	18
Tablas de estadísticos y predicciones . . . . .	19
Gráficos. . . . .	21
Limitación de resultados para los modelos de mejor o peor ajuste . . . . .	23
Almacenamiento de predicciones de modelos y especificaciones de modelo . . . . .	24
Opciones . . . . .	26
Funciones adicionales del comando TSMODEL. . . . .	28

### **3 Aplicar modelos de series temporales 29**

Resultados . . . . .	32
Tablas de estadísticos y predicciones . . . . .	32
Gráficos. . . . .	34
Limitación de resultados para los modelos de mejor o peor ajuste . . . . .	36

Almacenamiento de predicciones de modelos y especificaciones de modelo . . . . .	37
Opciones . . . . .	39
Funciones adicionales del comando TSAPPLY . . . . .	40
<b>4 Descomposición estacional</b>	<b>41</b>
Descomposición estacional: Guardar . . . . .	42
Funciones adicionales del comando SEASON. . . . .	43
<b>5 Diagramas espectrales</b>	<b>44</b>
Funciones adicionales del comando SPECTRA . . . . .	46
 <b>Parte II: Ejemplos</b>	
<b>6 Predicción masiva con el modelizador experto</b>	<b>49</b>
Examen de los datos. . . . .	49
Ejecución del análisis . . . . .	51
Gráficos de resumen del modelo. . . . .	57
Predicciones del modelo . . . . .	58
Resumen . . . . .	59
<b>7 Revisión masiva de predicciones mediante la aplicación de modelos guardados</b>	<b>60</b>
Ejecución del análisis . . . . .	60
Estadísticos de ajuste del modelo . . . . .	63
Predicciones del modelo . . . . .	64
Resumen . . . . .	64

## **8 *Uso del modelizador experto para determinar predictores significativos*** **65**

Gráficos de los datos . . . . .	65
Ejecución del análisis . . . . .	67
Gráfico de serie . . . . .	73
Tabla de descripción del modelo . . . . .	73
Tabla de estadísticos del modelo . . . . .	74
Tabla de parámetros del modelo ARIMA . . . . .	74
Resumen . . . . .	75

## **9 *Experimentación con predictores mediante la aplicación de modelos guardados*** **76**

Ampliación de la serie de predictores . . . . .	76
Modificación de valores predictores en el período de predicción . . . . .	80
Ejecución del análisis . . . . .	82

## **10 *Descomposición estacional*** **87**

Eliminación de la estacionalidad de los datos de ventas . . . . .	87
Determinación y configuración de la periodicidad . . . . .	87
Ejecución del análisis . . . . .	91
Descripción del resultado . . . . .	92
Resumen . . . . .	94
Procedimientos relacionados . . . . .	94

## **11 *Diagramas espectrales*** **95**

Uso de gráficos espectrales para verificar expectativas de periodicidad . . . . .	95
Ejecución del análisis . . . . .	95
Descripción del periodograma y la densidad espectral . . . . .	97
Resumen . . . . .	99
Procedimientos relacionados . . . . .	99

## ***Apéndices***

<b><i>A</i></b>	<b><i>Medidas de la bondad de ajuste</i></b>	<b><i>100</i></b>
<b><i>B</i></b>	<b><i>Tipos de valores atípicos</i></b>	<b><i>101</i></b>
<b><i>C</i></b>	<b><i>Guía de los gráficos de la función de autocorrelación simple (FAS) y parcial (FAP)</i></b>	<b><i>102</i></b>
<b><i>D</i></b>	<b><i>Archivos muestrales</i></b>	<b><i>106</i></b>
<b><i>E</i></b>	<b><i>Notices</i></b>	<b><i>116</i></b>
	<b><i>Bibliografía</i></b>	<b><i>118</i></b>
	<b><i>Índice</i></b>	<b><i>120</i></b>

***Parte I:***  
***Manual del usuario***



# Introducción a las series temporales

Una **serie temporal** es un conjunto de observaciones que se obtiene midiendo una variable única de manera regular a lo largo de un período de tiempo. Por ejemplo, en una serie de los datos de inventario, las observaciones pueden representar los niveles diarios de inventario durante varios meses. Una serie que muestra las cuotas de mercado de un producto puede consistir en las cuotas de mercado semanales registradas durante varios años. Una serie de las cifras de ventas totales puede consistir en una observación mensual durante muchos años. Lo que estos ejemplos tienen en común es que se ha observado alguna variable a intervalos conocidos y regulares a lo largo de un cierto período de tiempo. Por lo tanto, la forma de los datos para una serie temporal habitual es una secuencia o lista de observaciones única que representa medidas tomadas a intervalos regulares.

Tabla 1-1  
*Serie temporal de inventario diario*

Hora	Semana	Día	Nivel de inventario
$t_1$	1	Lunes	160
$t_2$	1	Martes	135
$t_3$	1	Miércoles	129
$t_4$	1	Jueves	122
$t_5$	1	Viernes	108
$t_6$	2	Lunes	150
		...	
$t_{60}$	12	Viernes	120

Una de las razones más importantes para realizar el análisis de las series temporales es intentar predecir los valores futuros de la serie. Un modelo de la serie que explique los valores pasados también puede predecir si aumentarán o disminuirán los próximos valores y en qué medida lo harán. La capacidad de realizar dichas predicciones correctamente es muy importante para cualquier negocio o disciplina científica.

## Datos de serie temporal

Al definir los datos de la serie temporal para utilizarlos con el módulo adicional Predicciones, cada serie se corresponde con una variable distinta. Por ejemplo, para definir una serie temporal en el Editor de datos, pulse en la pestaña Vista de variables e introduzca un nombre de variable en cualquier fila vacía. Cada observación de una serie temporal se corresponde con un caso (una fila del Editor de datos).

Si abre una hoja de cálculo que contiene datos de serie temporal, cada serie debe estar organizada en una columna de la base de datos. Si tiene ya una hoja de cálculo con una serie temporal organizada en filas, puede abrirla de todas formas y usar Transponer desde el menú Datos para voltear las filas en columnas.

## ***Transformaciones de los datos***

Disponer de un cierto número de procedimientos de transformación de datos que ofrece el sistema básico es útil para el análisis de series temporales.

- El procedimiento Definir fechas (del menú Datos) genera variables de fecha que se utilizan para establecer la periodicidad y distinguir entre períodos históricos, de validación y de predicción. Predicciones está diseñado para trabajar con las variables creadas por el procedimiento Definir fechas.
- El procedimiento Crear serie temporal (del menú Transformar) crea nuevas variables de series temporales como funciones de variables de series temporales existentes. Se incluyen aquí funciones que utilizan observaciones vecinas para el suavizado, el promedio y la diferenciación.
- El procedimiento Reemplazar valores perdidos (del menú Transformar) reemplaza los valores perdidos del sistema y los valores definidos como perdidos por el usuario por estimaciones basadas en uno de varios métodos. Los valores perdidos al principio o fin de una serie no suponen un problema especial; sencillamente acortan la longitud útil de la serie. Las discontinuidades que aparecen en mitad de una serie (datos *incrustados* perdidos) pueden ser un problema mucho más grave.

Consulte el *Manual del usuario del sistema básico* para obtener información detallada referente a las transformaciones de datos para series temporales.

## ***Períodos de estimación y de validación***

Normalmente resulta útil dividir la serie temporal en un período de *estimación* o *histórico* y un período de *validación*. Se desarrolla un modelo basándose en las observaciones realizadas en el período de estimación (histórico) y, a continuación, se prueba para ver si funciona bien en el período de validación. Al forzar que el modelo realice predicciones para los puntos que ya conoce (los puntos del período de validación), obtendrá una idea de cómo realiza el modelo las predicciones.

Los casos del período de validación se suelen denominar casos reservados porque no se incluyen en el proceso de creación del modelo. El período de estimación consta de los casos seleccionados actualmente en el conjunto de datos activo. Todos los casos posteriores al último caso seleccionado pueden utilizarse como casos reservados. Cuando considere que el modelo realiza un trabajo adecuado de predicción, puede volver a definir el período de estimación para incluir los casos reservados y, a continuación, crear el modelo final.

## ***Generación de modelos y producción de predicciones***

El módulo adicional Predicción proporciona dos procedimientos para llevar a cabo las tareas de crear modelos y producir predicciones.

- El procedimiento [Modelizador de series temporales](#) crea modelos de series temporales y genera predicciones. Incluye un modelizador experto que determina automáticamente el mejor modelo para cada serie temporal. Para analistas experimentados que buscan un mayor grado de control, también ofrece herramientas para la creación de modelos personalizados.
- El procedimiento [Aplicar modelos de series temporales](#) aplica modelos de series temporales creados por el modelizador de series temporales al conjunto de datos activo. Esto permite obtener predicciones para series que cuentan con datos nuevos o revisados sin tener que volver a generar los modelos. Si existe un motivo para pensar que un modelo ha cambiado, se puede volver a generar mediante el modelizador de series temporales.

# ***Modelizador de series temporales***

El procedimiento Modelizador de series temporales estima el modelo de suavizado exponencial, el modelo autorregresivo integrado de media móvil (ARIMA) univariante y los modelos ARIMA (o modelos de función de transferencia) multivariantes para series temporales, y genera predicciones. Este procedimiento incluye un modelizador experto que identifica y estima automáticamente el modelo ARIMA o de suavizado exponencial que mejor se ajuste para una o más series de variables dependientes, lo que elimina la necesidad de identificar un modelo adecuado mediante ensayo y error. Otra posibilidad consiste en especificar un modelo ARIMA o de suavizado exponencial personalizado.

**Ejemplo.** Imagine que es un director de producto responsable de la predicción de venta de unidades e ingresos del próximo mes para 100 productos diferentes, y tiene poca o ninguna experiencia en modelar series temporales. Los datos históricos de ventas de unidades para los 100 productos se almacenan en una única hoja de cálculo de Excel. Después de abrir la hoja de cálculo en IBM® SPSS® Statistics, debe utilizar el modelizador experto y solicitar predicciones de aquí a un mes. El modelizador experto busca el mejor modelo de ventas de unidades para cada uno de los productos y utiliza dicho modelo para generar las predicciones. Puesto que el modelizador experto puede trabajar con varias series de entradas, sólo tiene que ejecutar el procedimiento una vez si desea obtener predicciones para todos los productos. Si opta por guardar las predicciones en el conjunto de datos activo, puede exportar fácilmente los resultados a Excel.

**Estadísticos.** Medidas de bondad de ajuste:  $R$  cuadrado estacionaria,  $R$  cuadrado ( $R^2$ ), raíz del error cuadrático promedio (RMSE), error absoluto promedio (MAE), error absoluto porcentual promedio (MAPE), error absoluto máximo (MaxAE), error absoluto máximo porcentual (MaxAPE) y criterio de información bayesiano (BIC) normalizado. Residuos: función de autocorrelación, función de autocorrelación parcial y  $Q$  de Ljung-Box. Para modelos ARIMA: órdenes ARIMA para variables dependientes, órdenes de la función de transferencia para variables independientes y estimaciones de valores atípicos. Además, se utilizan estimaciones de parámetros de suavizado para modelos de suavizado exponencial.

**Diagramas.** Gráficos de resumen en todos los modelos: histogramas de  $R$  cuadrado estacionaria,  $R$  cuadrado ( $R^2$ ), raíz del error cuadrático promedio (RMSE), error absoluto promedio (MAE), error absoluto porcentual promedio (MAPE), error absoluto máximo (MaxAE), error absoluto máximo porcentual (MaxAPE), criterio de información bayesiano (BIC) normalizado, diagramas de caja de autocorrelaciones residuales y autocorrelaciones parciales. Resultados para modelos individuales: valores de predicción, valores ajustados, valores observados, límites de confianza superiores e inferiores, autocorrelaciones residuales y autocorrelaciones parciales.

## ***Modelizador de series temporales: Consideraciones sobre los datos***

**Datos.** La variable dependiente y todas las variables independientes deben ser numéricas.

**Supuestos.** La variable dependiente y todas las variables independientes se tratan como series temporales, lo que significa que cada caso representa un punto del tiempo, los casos son consecutivos separados por un intervalo de tiempo constante.

- **Estacionariedad.** Para los modelos ARIMA personalizados, la serie temporal que se va a modelar debe ser estacionaria. La manera más eficaz de transformar una serie no estacionaria en una estacionaria es mediante una transformación diferencia, disponible en el cuadro de diálogo Crear serie temporal.
- **Predicciones.** Para generar predicciones mediante modelos con variables independientes (predictoras), el conjunto de datos activo debe contener valores de estas variables para todos los casos del período de predicción. Además, las variables independientes no pueden contener ningún valor perdido en el período de estimación.

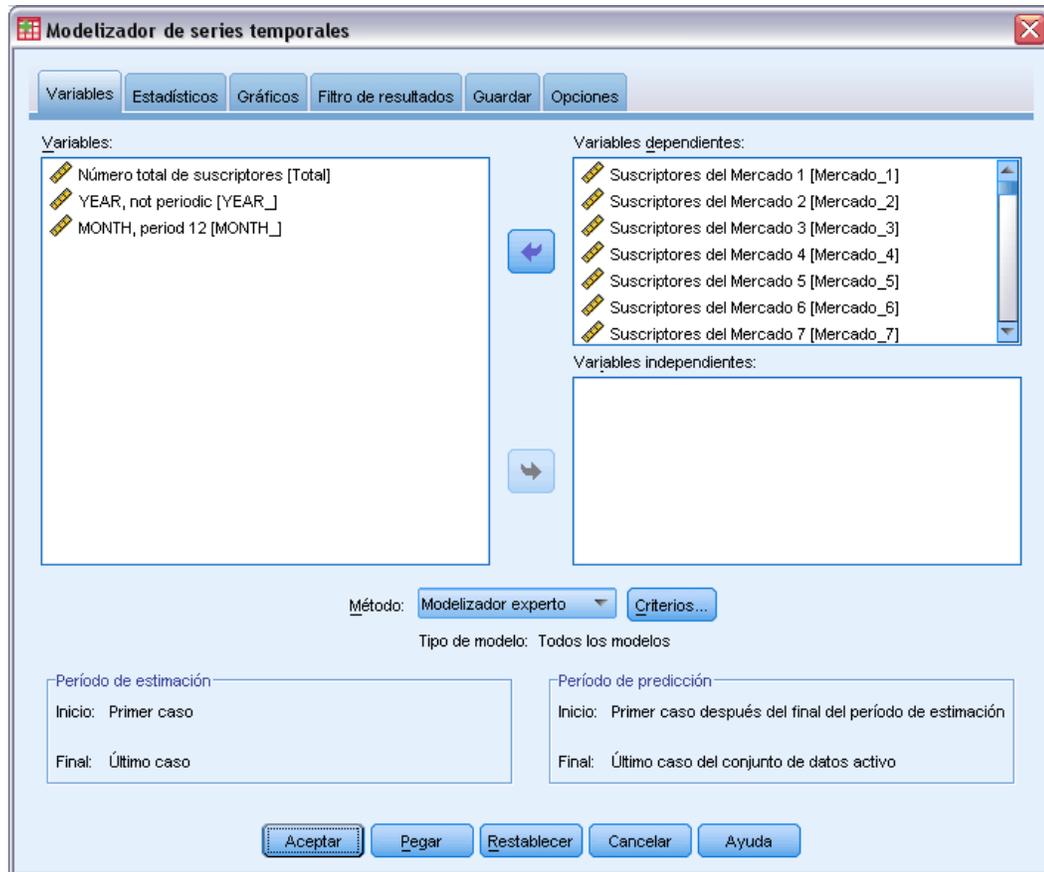
### ***Definición de fechas***

Aunque no es obligatorio, se recomienda el uso del cuadro de diálogo Definir fechas para especificar la fecha asociada al primer caso y el intervalo de tiempo entre los casos sucesivos. Esto se debe hacer antes de utilizar el modelizador de series temporales y ofrece como resultado un conjunto de variables que etiquetan la fecha asociada a cada caso. De este modo se establece también una periodicidad asumida de los datos (por ejemplo, una periodicidad de 12 si el intervalo de tiempo entre los casos sucesivos es un mes). Esta periodicidad es necesaria si desea crear modelos estacionales. Si no desea crear modelos estacionales y no necesita etiquetas de fecha en los resultados, puede omitir el cuadro de diálogo Definir fechas. La etiqueta asociada a cada caso es sencillamente el número de caso.

### ***Para utilizar el modelizador de series temporales***

- ▶ Elija en los menús:  
Analizar > Forecasting > Crear modelos...

Figura 2-1  
Modelizador de series temporales, pestaña Variables



- ▶ En la pestaña Variables, seleccione una o más variables dependientes para modelarlas.
- ▶ En el cuadro desplegable Método, seleccione un método de modelado. Si desea modelar automáticamente, deje el método por defecto de Modelizador experto. De este modo se invoca el modelizador experto para determinar el modelo que mejor se ajusta a cada una de las variables dependientes.

Para generar predicciones:

- ▶ Pulse en la pestaña Opciones.
- ▶ Especifique el período de predicción. De este modo se genera un gráfico que incluye las predicciones y los valores observados.

Si lo desea, puede:

- Seleccionar una o más variables independientes. Las variables independientes se tratan de forma muy similar a las variables predictoras del análisis de regresión, aunque son opcionales. Se pueden incluir en modelos ARIMA, pero no en modelos de suavizado exponencial. Si especifica Modelizador experto como el método de modelado e incluye variables independientes, sólo se tienen en cuenta los modelos ARIMA.

- Pulsar en Criterios para especificar los detalles del modelado.
- [Guardar los pronósticos, los intervalos de confianza y los residuos de ruido.](#)
- [Guardar los modelos estimados en formato XML.](#) Los modelos guardados se pueden aplicar a datos nuevos o revisados para obtener predicciones actualizadas sin necesidad de volver a generar los modelos. Para ello, se utiliza el procedimiento [Aplicar modelos de series temporales](#).
- [Obtener estadísticos de resumen para todos los modelos estimados.](#)
- [Especificar las funciones de transferencia para las variables independientes de los modelos ARIMA personalizados.](#)
- [Activar la detección automática de valores atípicos.](#)
- [Modelar determinados puntos de tiempo como valores atípicos para los modelos ARIMA personalizados.](#)

### ***Métodos de modelado***

Los métodos de modelado disponibles son:

**Modelizador experto.** El modelizador experto busca automáticamente el modelo que mejor se ajusta a cada serie dependiente. Si se especifican variables independientes (predictoras), el modelizador experto selecciona, para su inclusión en los modelos ARIMA, las que tienen una relación estadísticamente significativa con la serie dependiente. Las variables del modelo se transforman cuando es necesario mediante una diferenciación y/o una raíz cuadrada o una transformación logarítmica natural. Por defecto, el modelizador experto tiene en cuenta tanto los modelos de suavizado exponencial como los modelos ARIMA. No obstante, puede limitar el modelizador experto para que busque sólo modelos ARIMA o sólo modelos de suavizado exponencial. Además, puede especificar la detección automática de valores atípicos.

**Suavizado exponencial.** Utilice esta opción para especificar un modelo de suavizado exponencial personalizado. Puede elegir entre una amplia variedad de modelos de suavizado exponencial que difieren en el tratamiento de la tendencia y la estacionalidad.

**ARIMA.** Utilice esta opción para especificar un modelo ARIMA personalizado. Esto implica la especificación explícita de órdenes autorregresivos y de media móvil además del grado de diferenciación. Puede incluir variables independientes (predictoras) y definir las funciones de transferencia para alguna o todas las variables. Además, puede especificar la detección automática de valores atípicos o especificar un conjunto explícito de valores atípicos.

### ***Períodos de estimación y predicción***

**Período de estimación.** El período de estimación define el conjunto de casos utilizados para determinar el modelo. Por defecto, el período de estimación incluye todos los casos del conjunto de datos activo. Para establecer el período de estimación, seleccione Basándose en el rango del tiempo o de los casos en el cuadro de diálogo Seleccionar casos. Según los datos disponibles, el período de estimación utilizado por el procedimiento puede variar según la variable dependiente y, por consiguiente, ser distinto del valor mostrado. Para una variable dependiente determinada, el período de estimación real es el período restante tras la eliminación de todos los valores perdidos contiguos de la variable al principio o al final del período especificado.

**Período de predicción.** El período de predicción empieza en el primer caso después del período de estimación y, por defecto, llega hasta el último caso del conjunto de datos activo. Puede establecer el final del período de predicción en la pestaña [Opciones](#).

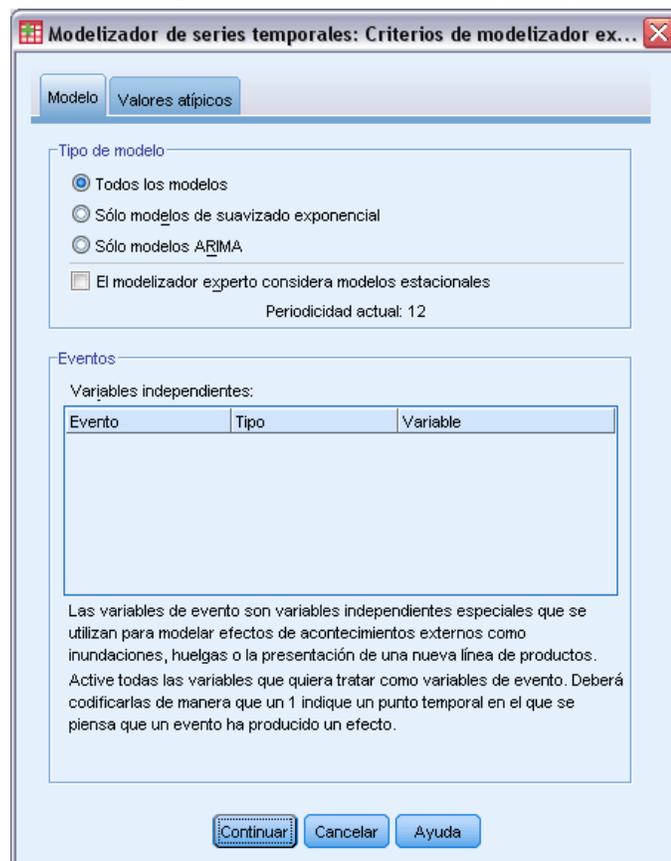
## Especificación de opciones para el modelizador experto

El modelizador experto proporciona una serie de opciones para restringir el conjunto de modelos candidatos mediante la especificación del tratamiento de los valores atípicos y la inclusión de variables de evento.

### Selección de modelo y especificación de eventos

Figura 2-2

Cuadro de diálogo Criterios de modelizador experto, pestaña Modelo



La pestaña Modelo le permite especificar los tipos de modelo que tiene en cuenta el modelizador experto y las variables de evento.

**Tipo de modelo.** Se encuentran disponibles las siguientes opciones:

- **Todos los modelos.** El modelizador experto tiene en cuenta tanto los modelos ARIMA como los modelos de suavizado exponencial.

- **Sólo modelos de suavizado exponencial.** El modelizador experto sólo tiene en cuenta los modelos de suavizado exponencial.
- **Sólo modelos ARIMA.** El modelizador experto sólo tiene en cuenta los modelos ARIMA.

**El modelizador experto considera modelos estacionales.** Esta opción sólo está activada si se ha definido una periodicidad para el conjunto de datos activo. Si esta opción está seleccionada (marcada), el modelizador experto tiene en cuenta los modelos tanto estacionales como no estacionales. Si esta opción no está seleccionada, el modelizador experto sólo tiene en cuenta los modelos no estacionales.

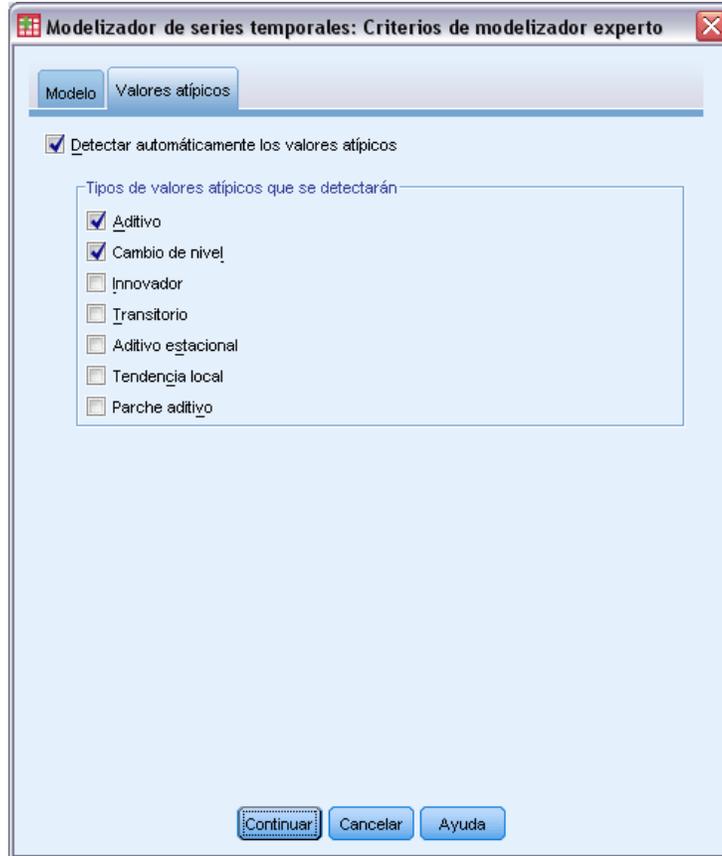
**Periodicidad actual.** Indica la periodicidad (si la hay) definida actualmente para el conjunto de datos activo. La periodicidad actual se expresa como un número entero (por ejemplo, 12 para la periodicidad anual, donde cada caso representa un mes). El valor *Ninguna* se muestra si no se ha establecido ninguna periodicidad. Los modelos estacionales requieren una periodicidad. Puede establecer la periodicidad en el cuadro de diálogo Definir fechas.

**Eventos.** Seleccione las variables independientes que se deben tratar como variables de evento. En las variables de evento, los casos con un valor igual a 1 indican las ocasiones en que se prevé que la serie dependiente se ve afectada por el evento. Los valores distintos de 1 indican que no se produce ningún efecto.

## Tratamiento de valores atípicos con el modelizador experto

Figura 2-3

Cuadro de diálogo Criterios de modelizador experto, pestaña Valores atípicos



La pestaña Valores atípicos le permite elegir la detección automática de valores atípicos además del tipo de valores atípicos que se va a detectar.

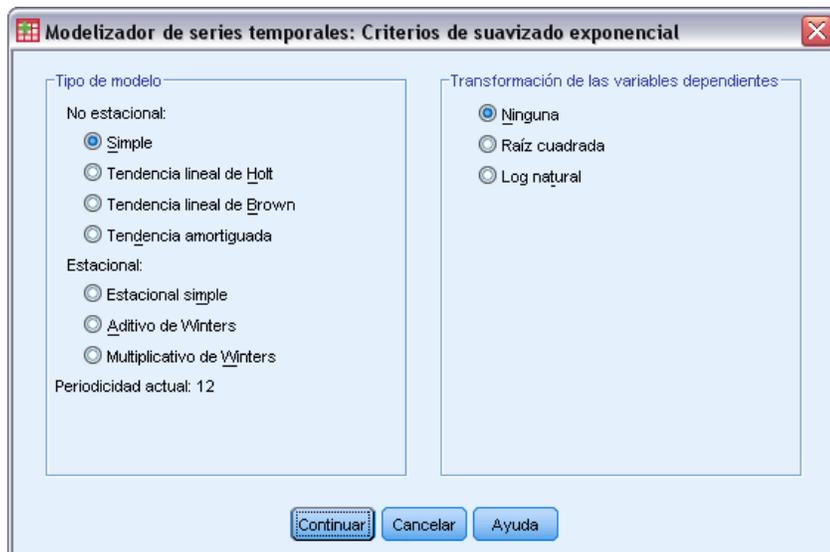
**Detectar automáticamente los valores atípicos.** Por defecto, no se realiza la detección automática de valores atípicos. Seleccione (marque) esta opción para realizar una detección automática de valores atípicos y, a continuación, seleccione uno o más de los siguientes tipos de valores atípicos:

- Aditivo
- Cambio de nivel
- Innovador
- Transitorio
- Aditivo estacional
- Tendencia local
- Parche aditivo

Si desea obtener más información, consulte el tema [Tipos de valores atípicos](#) en el apéndice B el p. 101.

## Modelos personalizados de suavizado exponencial

Figura 2-4  
Cuadro de diálogo Criterios de suavizado exponencial



**Tipo de modelo.** Los modelos de suavizado exponencial (Gardner, 1985) se clasifican como estacionales o no estacionales. Los modelos estacionales sólo están disponibles si se ha definido una periodicidad para el conjunto de datos activo (véase “Periodicidad actual” a continuación).

- **Simple.** Este modelo es adecuado para las series en las que no existe tendencia o estacionalidad. Su único parámetro de suavizado es el nivel. El suavizado exponencial simple es el más similar a un modelo ARIMA con cero órdenes de autorregresión, un orden de diferenciación, un orden de media móvil y sin constante.
- **Tendencia lineal de Holt.** Este modelo es adecuado para las series con una tendencia lineal y sin estacionalidad. Sus parámetros de suavizado son el nivel y la tendencia, y sus valores no se restringen mutuamente. El modelo de Holt es más general que el modelo de Brown pero puede llevar más tiempo de computación con series largas. El modelo de suavizado exponencial de Holt es muy similar a un modelo ARIMA con cero órdenes de autorregresión, dos órdenes de diferenciación y dos órdenes de media móvil.
- **Tendencia lineal de Brown.** Este modelo es adecuado para las series con una tendencia lineal y sin estacionalidad. Sus parámetros de suavizado son el nivel y la tendencia, que se asumen iguales. Por ello, el modelo de Brown es un caso especial del modelo de Holt. El modelo de suavizado exponencial de Brown es muy similar a un modelo ARIMA con cero órdenes de autorregresión, dos órdenes de diferenciación y dos órdenes de media móvil, con el coeficiente para el segundo orden de media móvil igual al cuadrado de la mitad del coeficiente de primer orden.
- **Tendencia amortiguada.** Este modelo es adecuado para las series con una tendencia lineal que va desapareciendo y sin estacionalidad. Sus parámetros de suavizado son el nivel, la tendencia y la amortiguación de la tendencia. El suavizado exponencial amortiguado es muy similar a un modelo ARIMA con un orden de autorregresión, un orden de diferenciación y dos órdenes de media móvil.

- **Simple estacional.** Este modelo es adecuado para series con tendencia y un efecto estacional que es constante a lo largo del tiempo. Sus parámetros de suavizado son el nivel y la estación. El modelo de suavizado exponencial simple estacional es muy similar a un modelo ARIMA con cero órdenes de autorregresión, un orden de diferenciación, un orden de diferenciación estacional y órdenes de media móvil 1,  $p$  y  $p + 1$ , donde  $p$  es el número de períodos contenidos en un intervalo estacional (para los datos mensuales,  $p = 12$ ).
- **De Winters aditivo.** Este modelo es adecuado para las series con tendencia lineal y un efecto estacional que no depende del nivel de la serie. Sus parámetros de suavizado son el nivel, la tendencia y la estación. El modelo de suavizado exponencial aditivo de Winters es muy similar a un modelo ARIMA con cero órdenes de autorregresión, un orden de diferenciación, un orden de diferenciación estacional y  $p + 1$  órdenes de media móvil, donde  $p$  es el número de períodos contenidos en un intervalo estacional (para datos mensuales  $p = 12$ ).
- **De Winters multiplicativo.** Este modelo es adecuado para las series con tendencia lineal y un efecto estacional que depende del nivel de la serie. Sus parámetros de suavizado son el nivel, la tendencia y la estación. El modelo de suavizado exponencial multiplicativo de Winters no es similar a ningún modelo ARIMA.

**Periodicidad actual.** Indica la periodicidad (si la hay) definida actualmente para el conjunto de datos activo. La periodicidad actual se expresa como un número entero (por ejemplo, 12 para la periodicidad anual, donde cada caso representa un mes). El valor *Ninguna* se muestra si no se ha establecido ninguna periodicidad. Los modelos estacionales requieren una periodicidad. Puede establecer la periodicidad en el cuadro de diálogo Definir fechas.

**Transformación de las variables dependientes.** Puede especificar una transformación para que se lleve a cabo sobre cada variable dependiente antes de su modelado.

- **Ninguno.** No se lleva a cabo ninguna transformación.
- **Raíz cuadrada.** Transformación de raíz cuadrada.
- **Log natural.** Transformación de logaritmo natural.

## ***Modelos ARIMA personalizados***

El modelizador de series temporales le permite generar modelos ARIMA (modelo autorregresivo integrado de media móvil) estacionales y no estacionales personalizados (también conocidos como modelos Box-Jenkins (Box, Jenkins, y Reinsel, 1994)) con o sin un conjunto fijo de variables predictoras. Puede definir funciones de transferencia para algunas o todas las variables predictoras y especificar la detección automática de valores atípicos o especificar un conjunto explícito de valores atípicos.

- Todas las variables independientes (predictoras) especificadas en la pestaña Variables se incluyen de forma explícita en el modelo. Esto contrasta con el uso del modelizador experto, donde las variables independientes sólo se incluyen si tienen una relación estadísticamente significativa con la variable dependiente.

## Especificación de modelo para modelos ARIMA personalizados

Figura 2-5  
Cuadro de diálogo Criterios ARIMA, pestaña Modelo

Modelizador de series temporales: Criterios de ARIMA

Modelo Valores atípicos

Órdenes ARIMA

Estructura:

	No estacional	Estacional
Autorregresivo (p)	0	0
Diferencia (d)	0	0
Media móvil (q)	0	0

Periodicidad actual: 12

Transformación

Ninguna  
 Raíz cuadrada  
 Log natural

Incluir constante en modelo

Continuar Cancelar Ayuda

La pestaña Modelo le permite especificar la estructura de un modelo ARIMA personalizado.

**Órdenes ARIMA.** Escriba valores para los distintos componentes ARIMA del modelo en las casillas correspondientes de la cuadrícula Estructura. Todos los valores deben ser enteros no negativos. Para los componentes autorregresivos y de media móvil, el valor representa el orden máximo. Todos los órdenes inferiores positivos se incluyen en el modelo. Por ejemplo, si especifica 2, el modelo incluye los órdenes 2 y 1. Las casillas de la columna Cíclico sólo se activan si se ha definido una periodicidad para el conjunto de datos activo (véase “Periodicidad actual” a continuación).

- **Autorregresivo (p).** Es el número de órdenes autorregresivos del modelo. Los órdenes autorregresivos especifican los valores previos de la serie utilizados para predecir los valores actuales. Por ejemplo, un orden autorregresivo igual a 2 especifica que se van a utilizar los valores de la serie correspondientes a dos períodos de tiempo del pasado para predecir el valor actual.
- **Diferencia (d).** Especifica el orden de diferenciación aplicado a la serie antes de estimar los modelos. La diferenciación es necesaria si hay tendencias (las series con tendencias suelen ser no estacionarias y el modelado de ARIMA asume la estacionariedad) y se utiliza para eliminar su efecto. El orden de diferenciación se corresponde con el grado de la tendencia de

la serie (la diferenciación de primer orden representa las tendencias lineales, la diferenciación de segundo orden representa las tendencias cuadráticas, etc.).

- **Media móvil (q).** Es el número de órdenes de media móvil presentes en el modelo. Los órdenes de media móvil especifican el modo en que se utilizan las desviaciones de la media de la serie para los valores previos con el fin de predecir los valores actuales. Por ejemplo, los órdenes de media móvil de 1 y 2 especifican que las desviaciones del valor medio de la serie de cada uno de los dos últimos períodos de tiempo se tienen en cuenta al predecir los valores actuales de la serie.

**Órdenes estacionales.** Los componentes estacionales autorregresivos, de media móvil y de diferenciación tienen la misma función que los componentes no estacionales correspondientes. No obstante, en el caso de los órdenes estacionales, los valores de la serie actual se ven afectados por los valores de la serie anterior separados por uno o más períodos estacionales. Por ejemplo, para los datos mensuales (período estacional de 12), un orden estacional de 1 significa que el valor de la serie actual se ve afectado por el valor de la serie 12 períodos antes del actual. Un orden estacional de 1 para los datos mensuales equivale a la especificación de un orden no estacional de 12.

**Periodicidad actual.** Indica la periodicidad (si la hay) definida actualmente para el conjunto de datos activo. La periodicidad actual se expresa como un número entero (por ejemplo, 12 para la periodicidad anual, donde cada caso representa un mes). El valor *Ninguna* se muestra si no se ha establecido ninguna periodicidad. Los modelos estacionales requieren una periodicidad. Puede establecer la periodicidad en el cuadro de diálogo Definir fechas.

**Transformación de las variables dependientes.** Puede especificar una transformación para que se lleve a cabo sobre cada variable dependiente antes de su modelado.

- **Ninguno.** No se lleva a cabo ninguna transformación.
- **Raíz cuadrada.** Transformación de raíz cuadrada.
- **Log natural.** Transformación de logaritmo natural.

**Incluir constante en el modelo.** La inclusión de una constante es estándar a menos que esté seguro de que el valor de la media global de la serie es 0. Se recomienda la exclusión de la constante si se aplica la diferenciación.

## Funciones de transferencia en modelos ARIMA personalizados

Figura 2-6  
Cuadro de diálogo Criterios ARIMA, pestaña Función de transferencia

Modelizador de series temporales: Criterios de ARIMA

Modelo Función de transferencia Valores atípicos

Variables independientes

Número total de suscri...

Órdenes de la función de transferencia

Estructura:

	No estacional	Estacional
Numerador	0	0
Denominador	0	0
Diferencia	0	0

Periodicidad actual: 12

Retardo: 0

Transformación

Ninguna  
 Raíz cuadrada  
 Log natural

Continuar Cancelar Ayuda

La pestaña Función de transferencia (que sólo aparece si se especifican variables independientes) le permite definir funciones de transferencia para algunas o todas las variables independientes especificadas en la pestaña Variables. Las funciones de transferencia le permiten especificar el modo en que se utilizan los valores anteriores de las variables independientes (predictoras) para predecir valores futuros de la serie dependiente.

**Órdenes de la función de transferencia.** Escriba valores para los distintos componentes de la función de transferencia en las casillas correspondientes de la cuadrícula Estructura. Todos los valores deben ser enteros no negativos. Para los componentes de numerador y denominador, el valor representa el orden máximo. Todos los órdenes inferiores positivos se incluyen en el modelo. Además, el orden 0 siempre se incluye para los componentes de numerador. Por ejemplo, si especifica 2 para el numerador, el modelo incluye los órdenes 2, 1 y 0. Si especifica 3 para el denominador, el modelo incluye los órdenes 3, 2 y 1. Las casillas de la columna Cíclico sólo se activan si se ha definido una periodicidad para el conjunto de datos activo (véase “Periodicidad actual” a continuación).

- **Numerador.** Es el orden de numerador de la función de transferencia. Especifica los valores previos de la serie independiente (predictora) seleccionada que se utilizan para predecir los valores actuales de la serie dependiente. Por ejemplo, un orden de numerador de 1 especifica que se utiliza el valor de una serie independiente de un período anterior (además del valor actual de la serie independiente) para predecir el valor actual de cada serie dependiente.

- **Denominador.** Es el orden de denominador de la función de transferencia. Especifica cómo se utilizan las desviaciones de la media de la serie para los valores previos de la serie independiente (predictora) seleccionada para predecir los valores actuales de la serie dependiente. Por ejemplo, un orden de denominador de 1 especifica que las desviaciones del valor medio de una serie independiente para un período de tiempo anterior se tienen en cuenta al predecir el valor actual de cada serie dependiente.
- **Diferencia.** Especifica el orden de diferenciación aplicado a la serie independiente (predictora) seleccionada antes de estimar los modelos. La diferenciación es necesaria si hay tendencias y se utiliza para eliminar su efecto.

**Órdenes estacionales.** Los componentes estacionales de numerador, denominador y diferenciación tienen la misma función que los componentes no estacionales correspondientes. No obstante, en el caso de los órdenes estacionales, los valores de la serie actual se ven afectados por los valores de la serie anterior separados por uno o más períodos estacionales. Por ejemplo, para los datos mensuales (período estacional de 12), un orden estacional de 1 significa que el valor de la serie actual se ve afectado por el valor de la serie 12 períodos antes del actual. Un orden estacional de 1 para los datos mensuales equivale a la especificación de un orden no estacional de 12.

**Periodicidad actual.** Indica la periodicidad (si la hay) definida actualmente para el conjunto de datos activo. La periodicidad actual se expresa como un número entero (por ejemplo, 12 para la periodicidad anual, donde cada caso representa un mes). El valor *Ninguna* se muestra si no se ha establecido ninguna periodicidad. Los modelos estacionales requieren una periodicidad. Puede establecer la periodicidad en el cuadro de diálogo Definir fechas.

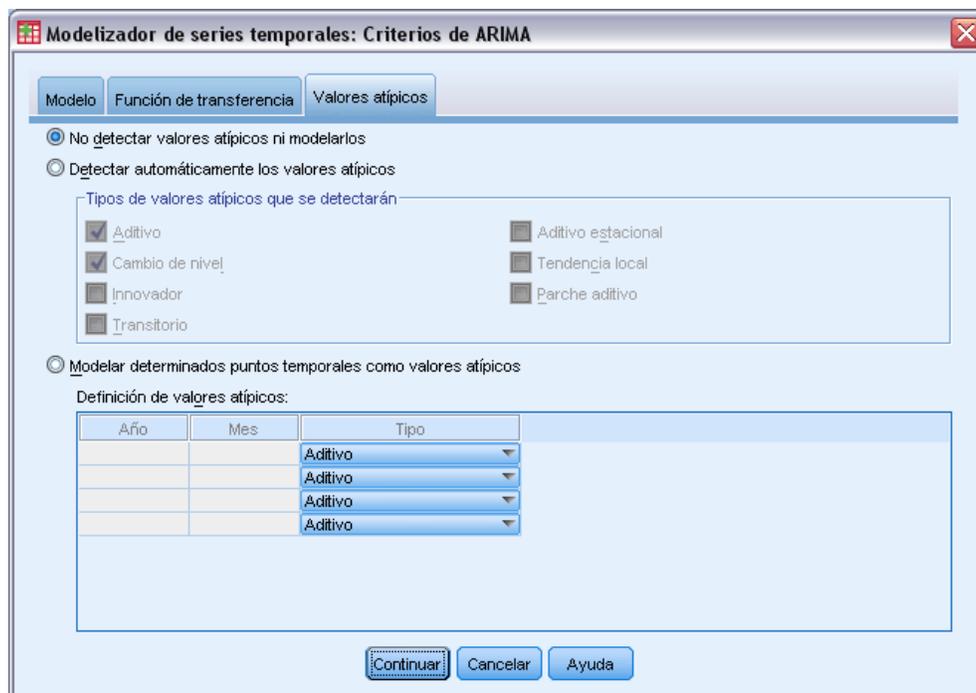
**Retardo.** Establecer un retardo provoca que la influencia de la variable independiente se retrase según el número de intervalos especificados. Por ejemplo, si el retardo se establece en 5, el valor de la variable independiente en el tiempo  $t$  no afecta a las predicciones hasta que han transcurrido cinco períodos ( $t + 5$ ).

**Transformación.** La especificación de una función de transferencia para un conjunto de variables independientes también incluye una transformación opcional que se puede aplicar a dichas variables.

- **Ninguno.** No se lleva a cabo ninguna transformación.
- **Raíz cuadrada.** Transformación de raíz cuadrada.
- **Log natural.** Transformación de logaritmo natural.

## Valores atípicos en modelos ARIMA personalizados

Figura 2-7  
Cuadro de diálogo Criterios ARIMA, pestaña Valores atípicos



La pestaña Valores atípicos ofrece las siguientes opciones para tratar los valores atípicos (Pena, Tiao, y Tsay, 2001): detección automática, especificación de puntos concretos como valores atípicos o no detectar ni modelar dichos valores.

**No detectar valores atípicos ni modelarlos.** Por defecto, los valores atípicos no se detectan ni modelan. Seleccione esta opción para desactivar la detección o el modelado de valores atípicos.

**Detectar automáticamente los valores atípicos.** Seleccione esta opción para realizar una detección automática de valores atípicos y seleccione uno o más de los siguientes tipos de valores atípicos:

- Aditivo
- Cambio de nivel
- Innovador
- Transitorio
- Aditivo estacional
- Tendencia local
- Parche aditivo

Si desea obtener más información, consulte el tema Tipos de valores atípicos en el apéndice B el p. 101.

**Modelar determinados puntos de tiempo como valores atípicos.** Seleccione esta opción para especificar puntos de tiempo concretos como valores atípicos. Utilice una fila independiente de la cuadrícula Definición de valores atípicos para cada valor atípico. Escriba valores para todas las casillas de una fila.

- **Tipo.** Es el tipo de valor atípico. Los tipos admitidos son: aditivo (por defecto), cambio de nivel, innovador, transitorio, aditivo estacional y tendencia local.

*Nota 1:* si no se ha definido ninguna especificación de fecha para el conjunto de datos activo, la cuadrícula Definición de valores atípicos muestra sólo la columna *Observación*. Para especificar un valor atípico, escriba el número de fila (tal como aparece en el Editor de datos) del caso correspondiente.

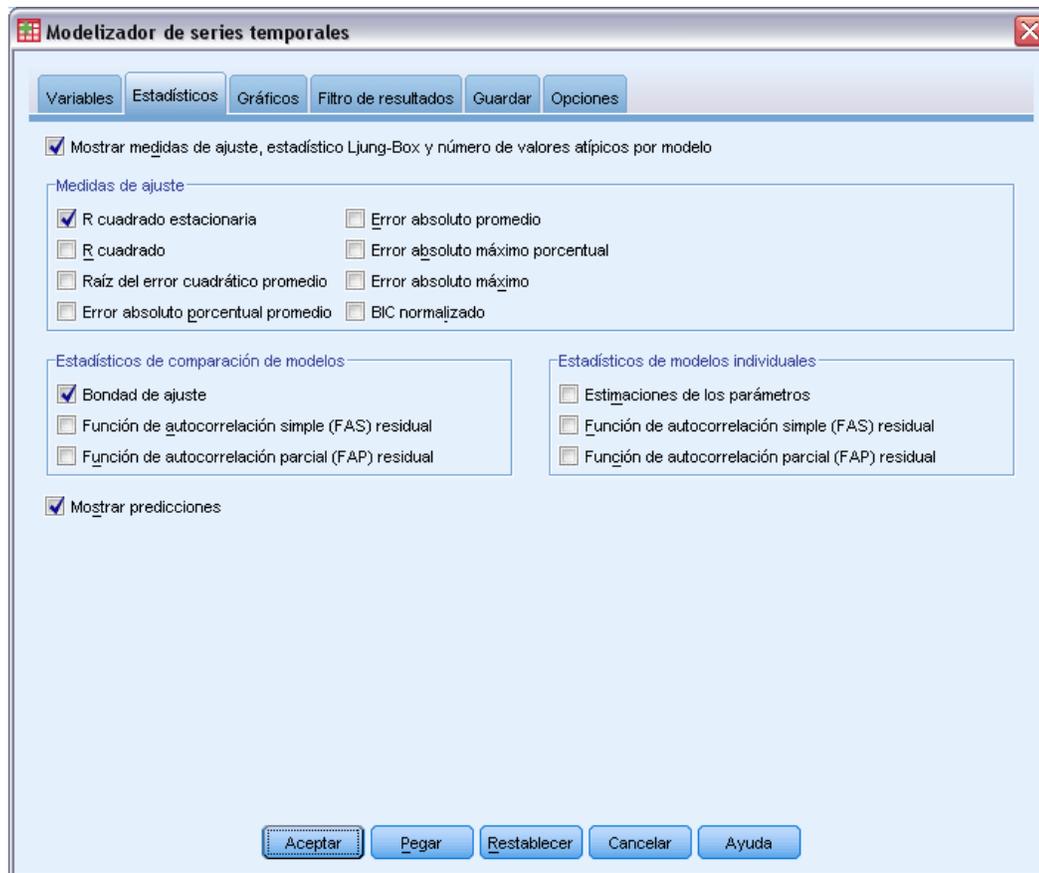
*Nota 2:* La columna *Ciclo* (si aparece) de la cuadrícula Definición de valores atípicos hace referencia al valor de la variable *CYCLE\_* del conjunto de datos activo.

## ***Resultados***

Los resultados disponibles incluyen tanto los de los modelos individuales como los calculados en todos los modelos. Los resultados de los modelos individuales se pueden limitar a un conjunto de modelos que mejor o peor se ajustan según los criterios especificados por el usuario.

## Tablas de estadísticos y predicciones

Figura 2-8  
Modelizador de series temporales, pestaña Estadísticos



La pestaña Estadísticos proporciona opciones para la presentación de tablas con los resultados del modelado.

### **Mostrar medidas de ajuste, estadístico Ljung-Box y número de valores atípicos por modelo.**

Seleccione (marque) esta opción para mostrar una tabla con las medidas de ajuste seleccionadas, el valor de Ljung-Box y el número de valores atípicos para cada modelo estimado.

**Medidas de ajuste.** Puede seleccionar una o más de las siguientes opciones para su inclusión en la tabla que contiene las medidas de ajuste para cada modelo estimado:

- $R$  cuadrado estacionaria
- $R$  cuadrado
- Raíz del error cuadrático promedio
- Error absoluto porcentual promedio
- Error absoluto promedio
- Error absoluto máximo porcentual

- Error absoluto máximo
- BIC normalizado

Si desea obtener más información, consulte el tema [Medidas de la bondad de ajuste en el apéndice A](#) el p. 100.

**Estadísticos de comparación de modelos.** Este grupo de opciones controla la visualización de las tablas que contienen los estadísticos calculados en todos los modelos estimados. Cada opción genera una tabla independiente. Puede seleccionar una o más de las siguientes opciones:

- **Bondad de ajuste.** Tabla de estadísticos de resumen y percentiles para  $R$  cuadrado estacionaria,  $R$  cuadrado, raíz del error cuadrático promedio, error absoluto porcentual promedio, error absoluto promedio, error absoluto máximo porcentual, error absoluto máximo y criterio de información bayesiano normalizado.
- **Función de autocorrelación simple (FAS) residual.** Tabla de estadísticos de resumen y percentiles para las autocorrelaciones de los residuos de todos los modelos estimados.
- **Función de autocorrelación parcial (FAP) residual.** Tabla de estadísticos de resumen y percentiles para las autocorrelaciones parciales de los residuos de todos los modelos estimados.

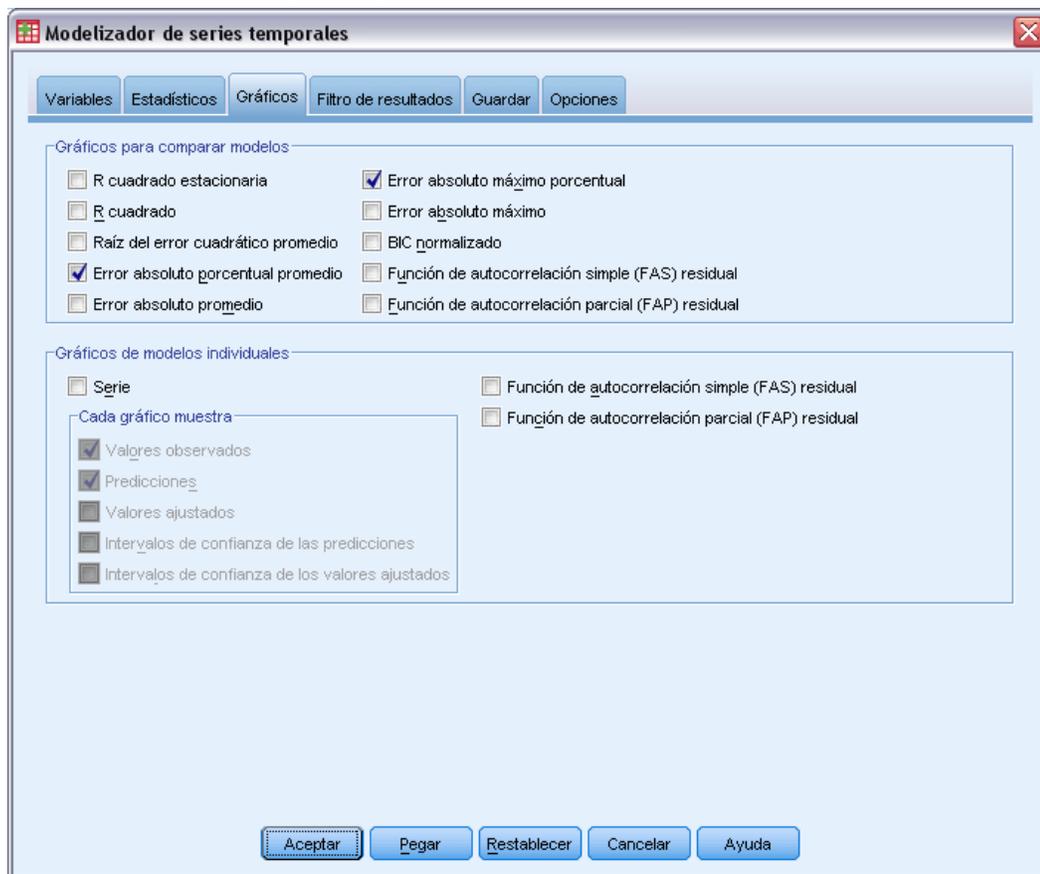
**Estadísticos de modelos individuales.** Este grupo de opciones controla la visualización de las tablas que contienen información detallada para cada modelo estimado. Cada opción genera una tabla independiente. Puede seleccionar una o más de las siguientes opciones:

- **Estimaciones de los parámetros.** Muestra una tabla de estimaciones de parámetros para cada modelo estimado. Se muestran tablas independientes para los modelos de suavizado exponencial y ARIMA. Si existen valores atípicos, las estimaciones de parámetros para dichos valores se muestran también en una tabla independiente.
- **Función de autocorrelación simple (FAS) residual.** Muestra una tabla con las autocorrelaciones residuales por retardo para cada modelo estimado. La tabla incluye los intervalos de confianza para las autocorrelaciones.
- **Función de autocorrelación parcial (FAP) residual.** Muestra una tabla con las autocorrelaciones parciales residuales por retardo para cada modelo estimado. La tabla incluye los intervalos de confianza para las autocorrelaciones parciales.

**Mostrar predicciones.** Muestra una tabla con las predicciones de modelo y los intervalos de confianza para cada modelo estimado. El período de predicción se establece en la pestaña Opciones.

## Gráficos

Figura 2-9  
Modelizador de series temporales, pestaña Gráficos



La pestaña Gráficos proporciona opciones para la presentación de gráficos con los resultados del modelado.

### **Gráficos para comparar modelos**

Este grupo de opciones controla la visualización de los gráficos que contienen los estadísticos calculados en todos los modelos estimados. Cada opción genera un gráfico independiente. Puede seleccionar una o más de las siguientes opciones:

- $R$  cuadrado estacionaria
- $R$  cuadrado
- Raíz del error cuadrático promedio
- Error absoluto porcentual promedio
- Error absoluto promedio
- Error absoluto máximo porcentual
- Error absoluto máximo

- BIC normalizado
- Función de autocorrelación simple (FAS) residual
- Función de autocorrelación parcial (FAP) residual

Si desea obtener más información, consulte el tema [Medidas de la bondad de ajuste en el apéndice A](#) el p. 100.

### **Gráficos de modelos individuales**

**Serie.** Seleccione (marque) esta opción para obtener los gráficos de los valores pronosticados para cada modelo estimado. Puede seleccionar una o más de las siguientes opciones para su inclusión en el gráfico:

- **Valores observados.** Son los valores observados de la serie dependiente.
- **Predicciones.** Son los valores pronosticados por el modelo para el período de predicción.
- **Valores ajustados.** Son los valores pronosticados por el modelo para el período de estimación.
- **Intervalos de confianza de las predicciones.** Son los intervalos de confianza para el período de predicción.
- **Intervalos de confianza de los valores ajustados.** Son los intervalos de confianza para el período de estimación.

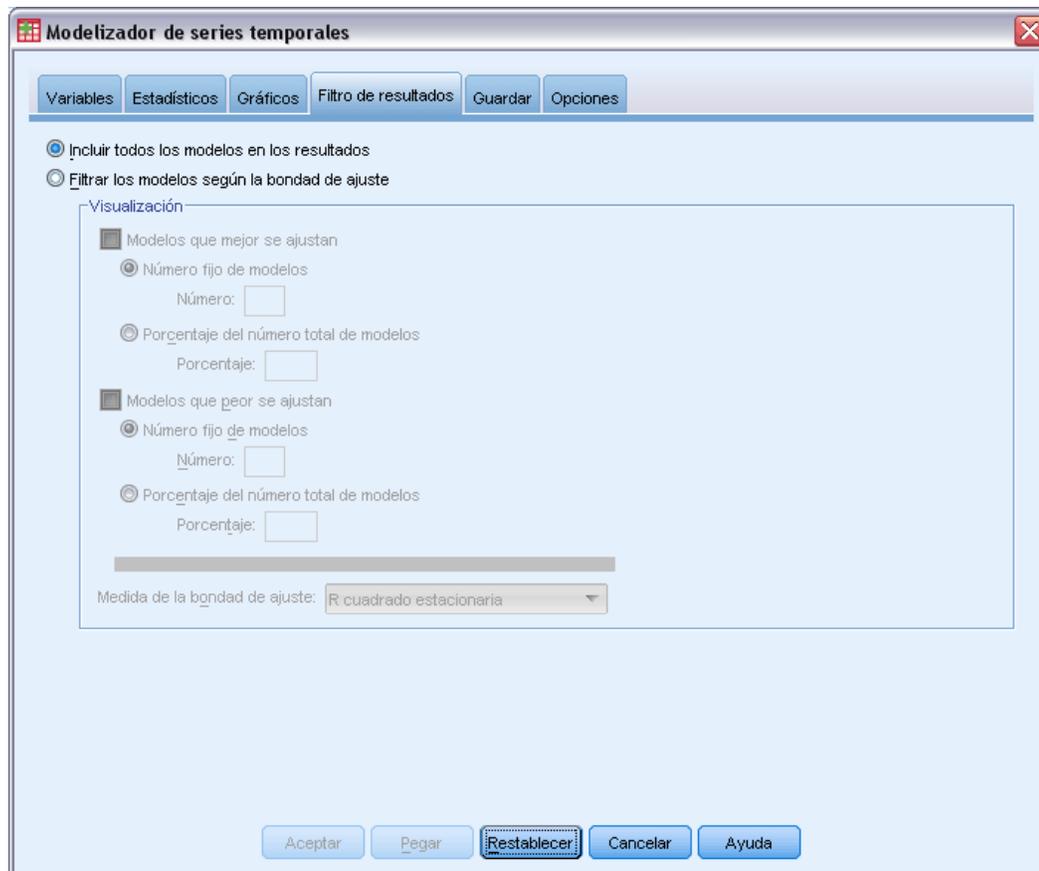
**Función de autocorrelación simple (FAS) residual.** Muestra un gráfico con las autocorrelaciones residuales para cada modelo estimado.

**Función de autocorrelación parcial (FAP) residual.** Muestra un gráfico con las autocorrelaciones parciales residuales para cada modelo estimado.

## Limitación de resultados para los modelos de mejor o peor ajuste

Figura 2-10

Modelizador de series temporales, pestaña Filtro de resultados



La pestaña Filtro de resultados proporciona opciones para restringir los resultados tanto tabulares como de gráficos a un subconjunto de modelos estimados. Puede optar por limitar los resultados a los modelos de mejor o peor ajuste según los criterios de ajuste proporcionados. Por defecto, todos los modelos estimados se incluyen en los resultados.

**Modelos que mejor se ajustan.** Seleccione (marque) esta opción para incluir los modelos que mejor se ajustan en los resultados. Seleccione una medida de bondad de ajuste y especifique el número de modelos que se van a incluir. La selección de esta opción no excluye la selección de los modelos que peor se ajustan. En este caso, los resultados incluyen tanto los modelos que peor se ajustan como los que mejor se ajustan.

- **Número fijo de modelos.** Especifica que los resultados se muestran para los  $n$  modelos que mejor se ajustan. Si este número es superior al número de modelos estimados, se muestran todos los modelos.
- **Porcentaje del número total de modelos.** Especifica que los resultados se muestran para los modelos con valores de bondad de ajuste en el porcentaje  $n$  superior de todos los modelos estimados.

**Modelos que peor se ajustan.** Seleccione (marque) esta opción para incluir los modelos que peor se ajustan en los resultados. Seleccione una medida de bondad de ajuste y especifique el número de modelos que se van a incluir. La selección de esta opción no excluye la selección de los modelos que mejor se ajustan. En este caso, los resultados incluyen tanto los modelos que mejor se ajustan como los que peor se ajustan.

- **Número fijo de modelos.** Especifica que los resultados se muestran para los  $n$  modelos que peor se ajustan. Si este número es superior al número de modelos estimados, se muestran todos los modelos.
- **Porcentaje del número total de modelos.** Especifica que los resultados se muestran para los modelos con valores de bondad de ajuste en el porcentaje  $n$  inferior de todos los modelos estimados.

**Medida de la bondad de ajuste.** Seleccione la medida de bondad de ajuste que se va a utilizar para filtrar los modelos. El valor por defecto es  $R$  cuadrado estacionaria.

## Almacenamiento de predicciones de modelos y especificaciones de modelo

Figura 2-11  
Modelizador de series temporales, pestaña Guardar

Modelizador de series temporales

Variables Estadísticos Gráficos Filtro de resultados Guardar Opciones

Guardar variables

Variables:

Descripción	Guardar	Prefijo del nombre de variable
Valores pronosticados	<input checked="" type="checkbox"/>	Pronosticado
Límites de confianza inferiores	<input type="checkbox"/>	LCI
Límites de confianza superiores	<input type="checkbox"/>	LCS
Residuos de ruido	<input type="checkbox"/>	ResiduoN

Por cada elemento seleccionado, se guardará una variable para cada variable dependiente.

Exportar archivo de modelo

Archivo: C:\models.xml Examinar...

Aceptar Pegar Restablecer Cancelar Ayuda

La pestaña Guardar le permite guardar predicciones del modelo como variables nuevas en el conjunto de datos activo y guardar las especificaciones del modelo en un archivo externo en formato XML.

**Guardar variables.** Puede guardar predicciones del modelo, intervalos de confianza y residuos como variables nuevas en el conjunto de datos activo. Cada serie dependiente genera su propio conjunto de variables nuevas y cada variable nueva contiene valores para los períodos de estimación y predicción. Se añaden casos nuevos si el período de predicción se amplía más allá de la duración de la serie de variables dependiente. Para guardar las variables nuevas, seleccione la casilla de verificación Guardar asociada a cada variable. Por defecto, no se guarda ninguna variable nueva.

- **Valores pronosticados.** Son los valores pronosticados por el modelo.
- **Límites de confianza inferiores.** Son los límites de confianza inferiores para los valores pronosticados.
- **Límites de confianza superiores.** Son los límites de confianza superiores para los valores pronosticados.
- **Residuos de ruido.** Son los residuos del modelo. Si se realizan transformaciones de la variable dependiente (por ejemplo, logaritmo natural), estos son los residuos de la serie transformada.
- **Prefijo del nombre de variable.** Especifique los prefijos que se van a utilizar para los nombres de variables nuevas o deje los prefijos por defecto. Los nombres de variables se componen del prefijo, el nombre de la variable dependiente asociada y un identificador del modelo. El nombre de la variable se amplía en caso de ser necesario para evitar conflictos de nombres de variables. El prefijo debe cumplir las reglas de nombres de variables válidos.

**Exportar archivo de modelo.** Las especificaciones de modelo para todos los modelos estimados se exportan al archivo especificado en formato XML. Los modelos guardados se pueden utilizar para obtener predicciones actualizadas, basadas en los datos más recientes, mediante el procedimiento [Aplicar modelos de series temporales](#).

## Opciones

Figura 2-12  
Modelizador de series temporales, pestaña Opciones

The screenshot shows the 'Opciones' (Options) tab of the 'Modelizador de series temporales' (Time Series Modeler) software. The window title is 'Modelizador de series temporales'. The 'Opciones' tab is selected, showing settings for the prediction period, handling of missing values, confidence interval width, model identifier prefix, and maximum number of lags.

**Período de predicción**

Primer caso después del final del período de estimación hasta el último caso del conjunto de datos activo

Primer caso después del final del período de estimación hasta una fecha especificada

Fecha:

Año	Mes
2004	3

**Valores definidos como perdidos por el usuario**

Tratar como no válidos

Tratar como válidos

Ancho del intervalo de confianza (%):

Prefijo de los identificadores de modelo de los resultados:

Número máximo de retardos que se muestran en resultados de las FAS y FAP:

Buttons: Aceptar, Pegar, Restablecer, Cancelar, Ayuda

La pestaña Opciones le permite establecer el período de predicción, especificar el tratamiento de los valores perdidos, establecer el ancho del intervalo de confianza, especificar un prefijo personalizado para los identificadores de modelo y establecer el número de retardos mostrados para las autocorrelaciones.

**Período de predicción.** El período de predicción siempre empieza con el primer caso después del final del período de estimación (conjunto de casos utilizado para determinar el modelo) y se extiende hasta el último caso del conjunto de datos activo o hasta una fecha especificada por el usuario. Por defecto, el final del período de estimación es el último caso del conjunto de datos activo, aunque se puede cambiar en el cuadro de diálogo Seleccionar casos seleccionando Basándose en el rango del tiempo o de los casos.

- **Primer caso después del final del período de estimación hasta el último caso del conjunto de datos activo.** Seleccione esta opción si el final del período de estimación es anterior al último caso del conjunto de datos activo y desea obtener predicciones hasta el último caso. Esta opción se

suele utilizar para generar predicciones para un período de datos reservados, lo que permite la comparación de las predicciones del modelo con un subconjunto de los valores que existen.

■ **Primer caso después del final del período de estimación hasta una fecha especificada.**

Seleccione esta opción para especificar de forma explícita el final del período de predicción. Esta opción se suele utilizar para generar predicciones más allá del final de la serie actual. Escriba valores para todas las casillas de la cuadrícula Fecha.

Si no se ha definido ninguna especificación de fecha para el conjunto de datos activo, la cuadrícula Fecha muestra sólo la columna *Observación*. Para especificar el final del período de predicción, escriba el número de fila (tal como aparece en el Editor de datos) del caso correspondiente.

La columna *Ciclo* (si aparece) de la cuadrícula Fecha hace referencia al valor de la variable *CYCLE\_* del conjunto de datos activo.

**Valores definidos como perdidos por el usuario.** Estas opciones controlan el tratamiento de los valores perdidos definidos por el usuario.

■ **Tratar como no válidos.** Los valores definidos como perdidos por el usuario reciben el mismo tratamiento que los valores perdidos del sistema.

■ **Tratar como válidos.** Los valores perdidos definidos por el usuario se tratan como datos válidos.

**Política de valores perdidos.** Las siguientes reglas se aplican al tratamiento de los valores perdidos (incluye los valores perdidos del sistema y los valores perdidos definidos por el usuario tratados como no válidos) durante el procedimiento de modelado:

- Los casos con valores perdidos de una variable dependiente que se producen durante el período de estimación se incluyen en el modelo. El tratamiento específico del valor perdido depende del método de estimación.
- Se genera una advertencia si una variable independiente tiene valores perdidos en el período de estimación. Para el modelizador experto, los modelos que implican la variable independiente se estiman sin la variable. Para los modelos ARIMA personalizados, no se estiman los modelos que implican la variable independiente.
- Si una variable independiente tiene valores perdidos en el período de predicción, el procedimiento genera una advertencia y realiza predicciones en la medida de lo posible.

**Ancho del intervalo de confianza (%).** Los intervalos de confianza se calculan para las predicciones del modelo y las autocorrelaciones residuales. Puede especificar cualquier valor positivo inferior a 100. Por defecto, se utiliza un intervalo de confianza del 95 %.

**Prefijo de los identificadores de modelo de los resultados.** Cada variable dependiente especificada en la pestaña Variables genera un modelo estimado independiente. Los modelos se distinguen mediante nombres únicos compuestos por un prefijo personalizable y un sufijo entero. Puede escribir un prefijo o dejar el valor por defecto de *Modelo*.

**Número máximo de retardos que se muestran en resultados de las FAS y FAP.** Puede establecer el número máximo de retardos que se muestran en las tablas y en los gráficos de autocorrelaciones y autocorrelaciones parciales.

## ***Funciones adicionales del comando TSMODEL***

Se puede personalizar el modelado de series temporales si se pegan las selecciones en una ventana de sintaxis y se edita la sintaxis del comando `TSMODEL` resultante. El lenguaje de sintaxis de comandos permite:

- Especificar el período estacional de los datos (con la palabra clave `SEASONLENGTH` del subcomando `AUXILIARY`). De este modo se anula la periodicidad actual (si existe) para el conjunto de datos activo.
- Especificar retardos no consecutivos para los componentes `ARIMA` personalizados y de función de transferencia (con los subcomandos `ARIMA` y `TRANSFERFUNCTION`). Por ejemplo, puede especificar un modelo `ARIMA` personalizado con retardos autorregresivos de órdenes 1, 3 y 6 o una función de transferencia con retardos de numerador de órdenes 2, 5 y 8.
- Proporcionar más de un conjunto de especificaciones de modelado (por ejemplo, método de modelado, órdenes `ARIMA`, variables independientes, etc.) para una sola ejecución del procedimiento del modelizador de series temporales (con el subcomando `MODEL`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

# Aplicar modelos de series temporales

El procedimiento Aplicar modelos de series temporales carga modelos de series temporales existentes desde un archivo externo y los aplica al conjunto de datos activo. Puede utilizar este procedimiento para obtener predicciones para series que cuentan con datos nuevos o revisados, sin tener que volver a generar los modelos. Los modelos se generan mediante el procedimiento [Modelizador de series temporales](#).

**Ejemplo.** Imagine que es el director de inventario de un importante minorista y se encarga de 5.000 productos. Ha utilizado el modelizador experto para crear modelos que predican las ventas de cada producto para un plazo de tres meses. Su almacén de datos se actualiza cada mes con datos de ventas reales que desea utilizar para generar predicciones actualizadas mensuales. El procedimiento Aplicar modelos de series temporales le permite hacerlo utilizando los modelos originales, con sólo volver a estimar los parámetros de los modelos para que tengan en cuenta los datos nuevos.

**Estadísticos.** Medidas de bondad de ajuste:  $R$  cuadrado estacionaria,  $R$  cuadrado ( $R^2$ ), raíz del error cuadrático promedio (RMSE), error absoluto promedio (MAE), error absoluto porcentual promedio (MAPE), error absoluto máximo (MaxAE), error absoluto máximo porcentual (MaxAPE) y criterio de información bayesiano (BIC) normalizado. Residuos: función de autocorrelación, función de autocorrelación parcial y  $Q$  de Ljung-Box.

**Diagramas.** Gráficos de resumen en todos los modelos: histogramas de  $R$  cuadrado estacionaria,  $R$  cuadrado ( $R^2$ ), raíz del error cuadrático promedio (RMSE), error absoluto promedio (MAE), error absoluto porcentual promedio (MAPE), error absoluto máximo (MaxAE), error absoluto máximo porcentual (MaxAPE), criterio de información bayesiano (BIC) normalizado, diagramas de caja de autocorrelaciones residuales y autocorrelaciones parciales. Resultados para modelos individuales: valores de predicción, valores ajustados, valores observados, límites de confianza superiores e inferiores, autocorrelaciones residuales y autocorrelaciones parciales.

## **Aplicar modelos de series temporales: Consideraciones sobre los datos**

**Datos.** Las variables (dependientes e independientes) a las que se aplican los modelos deben ser numéricas.

**Supuestos.** Los modelos se aplican a las variables del conjunto de datos activo con los mismos nombres que las variables especificadas en el modelo. Todas estas variables se tratan como series temporales, lo que significa que cada caso representa un punto del tiempo, los casos son consecutivos separados por un intervalo de tiempo constante.

- **Predicciones.** Para generar predicciones mediante modelos con variables independientes (predictoras), el conjunto de datos activo debe contener valores de estas variables para todos los casos del período de predicción. Si se vuelven a estimar los parámetros del modelo, las variables independientes no pueden contener ningún valor perdido en el período de estimación.

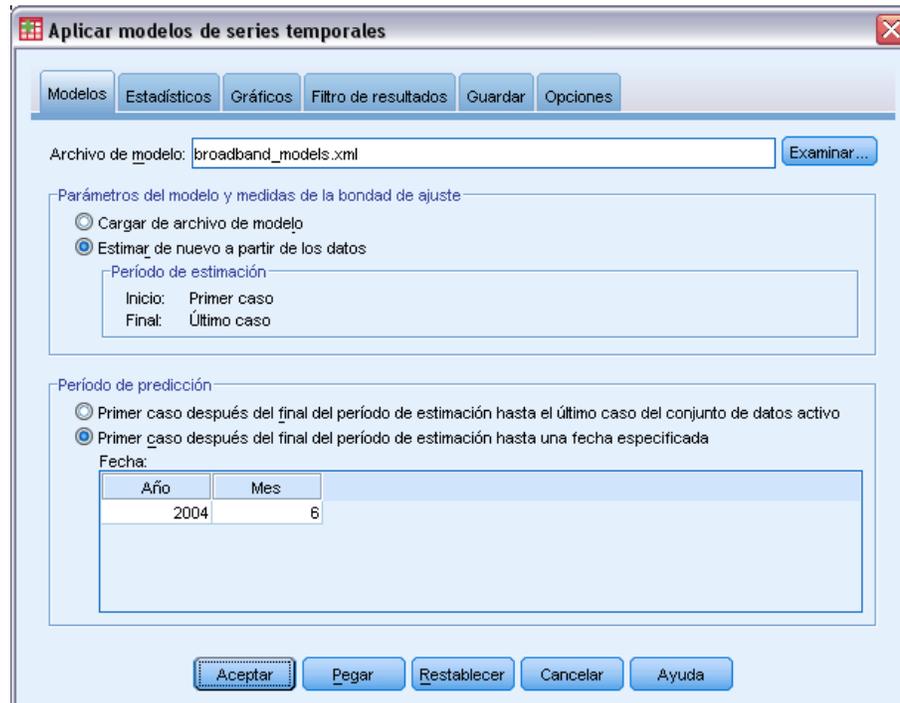
### Definición de fechas

El procedimiento Aplicar modelos de series temporales requiere que la periodicidad, si existe, del conjunto de datos activo coincida con la periodicidad de los modelos que se van a aplicar. Si sólo va a generar predicciones con el mismo conjunto de datos (quizá con datos nuevos o revisados) que el utilizado para generar el modelo, esta condición se cumple. Si no se da ninguna periodicidad para el conjunto de datos activo, tiene la posibilidad de desplazarse hasta el cuadro de diálogo Definir fechas para crear una. No obstante, si los modelos se crean sin especificar ninguna periodicidad, el conjunto de datos activo no puede tener tampoco ninguna periodicidad.

### Para aplicar modelos

- Elija en los menús:  
Analizar > Predicción > Aplicar modelos...

Figura 3-1  
Aplicar modelos de series temporales, pestaña Modelos



- Escriba la especificación de archivo para un archivo de modelo o pulse en Examinar y seleccione un archivo de modelo (los archivos de modelo se crean con el procedimiento [Modelizador de series temporales](#)).

Si lo desea, puede:

- Volver a estimar los parámetros del modelo con los datos del conjunto de datos activo. Las predicciones se crean con los parámetros que se han vuelto a estimar.
- [Guardar los pronósticos, los intervalos de confianza y los residuos de ruido.](#)
- [Guardar los modelos que se han vuelto a estimar en formato XML.](#)

### ***Parámetros del modelo y medidas de la bondad de ajuste***

**Cargar de archivo de modelo.** Las predicciones se generan con los parámetros de modelo del archivo de modelo sin volver a estimar dichos parámetros. [Las medidas de bondad de ajuste](#) mostradas en los resultados y utilizadas para filtrar los modelos (de mejor o peor ajuste) proceden del archivo de modelo y reflejan los datos utilizados en el desarrollo (o última actualización) de cada modelo. Con esta opción, las predicciones no tienen en cuenta los datos históricos (para las variables tanto dependientes como independientes) del conjunto de datos activo. Debe elegir Estimar de nuevo a partir de los datos si desea que los datos históricos influyan en las predicciones. Además, las predicciones no tienen en cuenta los valores de la serie dependiente en el período de predicción, aunque sí tienen en cuenta los valores de las variables independientes en el período de predicción. Si tiene más valores actuales de la serie dependiente y desea incluirlos en las predicciones, debe realizar otra estimación y ajustar el período de estimación para incluir dichos valores.

**Estimar de nuevo a partir de los datos.** Los parámetros del modelo se vuelven a estimar con los datos del conjunto de datos activo. La nueva estimación de los parámetros del modelo no tiene ningún efecto en la estructura del modelo. Por ejemplo, un modelo ARIMA(1,0,1) no varía, pero los parámetros autorregresivos y de media móvil se vuelven a estimar. La nueva estimación no tiene como resultado la detección de nuevos valores atípicos. Los valores atípicos, si los hay, siempre proceden del archivo de modelo.

- **Período de estimación.** El período de estimación define el conjunto de casos utilizados para volver a estimar los parámetros del modelo. Por defecto, el período de estimación incluye todos los casos del conjunto de datos activo. Para establecer el período de estimación, seleccione Basándose en el rango del tiempo o de los casos en el cuadro de diálogo Seleccionar casos. Según los datos disponibles, el período de estimación utilizado por el procedimiento puede variar según el modelo y, por consiguiente, ser distinto del valor mostrado. Para un modelo determinado, el período de estimación real es el período restante tras eliminar de todos los valores perdidos contiguos, de la variable dependiente del modelo, que aparecen al principio o al final del período especificado.

### ***Período de predicción***

El período de predicción de cada modelo siempre empieza con el primer caso después del final del período de estimación y se extiende hasta el último caso del conjunto de datos activo o hasta una fecha especificada por el usuario. Si no se vuelve a estimar ningún parámetro (valor por defecto), el período de estimación para cada modelo es el conjunto de casos utilizado en el desarrollo (o última actualización) de cada modelo.

- **Primer caso después del final del período de estimación hasta el último caso del conjunto de datos activo.** Seleccione esta opción si el final del período de estimación es anterior al último caso del conjunto de datos activo y desea obtener predicciones hasta el último caso.
- **Primer caso después del final del período de estimación hasta una fecha especificada.** Seleccione esta opción para especificar de forma explícita el final del período de predicción. Escriba valores para todas las casillas de la cuadrícula Fecha.

Si no se ha definido ninguna especificación de fecha para el conjunto de datos activo, la cuadrícula Fecha muestra sólo la columna *Observación*. Para especificar el final del período de predicción, escriba el número de fila (tal como aparece en el Editor de datos) del caso correspondiente.

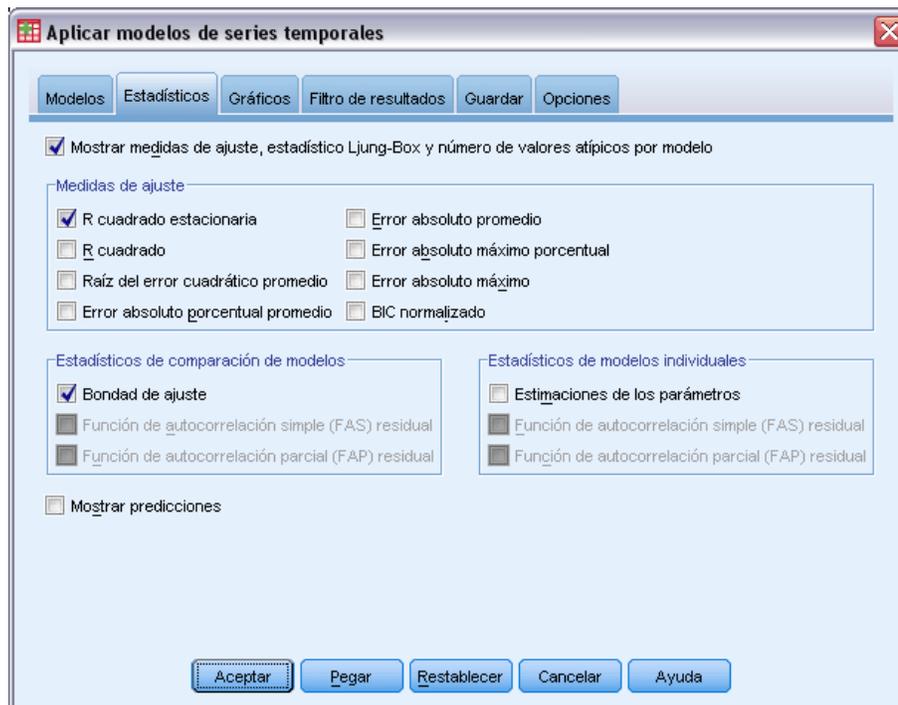
La columna *Ciclo* (si aparece) de la cuadrícula Fecha hace referencia al valor de la variable *CYCLE\_* del conjunto de datos activo.

## Resultados

Los resultados disponibles incluyen tanto los de los modelos individuales como los de todos los modelos. Los resultados de los modelos individuales se pueden limitar a un conjunto de modelos que mejor o peor se ajustan según los criterios especificados por el usuario.

## Tablas de estadísticos y predicciones

Figura 3-2  
Aplicar modelos de series temporales, pestaña Estadísticos



La pestaña Estadísticos proporciona opciones para mostrar tablas de estadísticos de ajuste del modelo, parámetros de modelo, funciones de autocorrelación y predicciones. A menos que se vuelvan a estimar los parámetros del modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos), los valores de las medidas de ajuste, los valores de Ljung-Box y los parámetros de modelo mostrados proceden del archivo de modelo y reflejan los datos utilizados en el desarrollo (o última actualización) de cada modelo. La información sobre los valores atípicos siempre procede del archivo de modelo.

**Mostrar medidas de ajuste, estadístico Ljung-Box y número de valores atípicos por modelo.**

Seleccione (marque) esta opción para mostrar una tabla con las medidas de ajuste seleccionadas, el valor de Ljung-Box y el número de valores atípicos para cada modelo.

**Medidas de ajuste.** Puede seleccionar una o más de las siguientes opciones para su inclusión en la tabla que contiene las medidas de ajuste para cada modelo:

- *R* cuadrado estacionaria
- *R* cuadrado
- Raíz del error cuadrático promedio
- Error absoluto porcentual promedio
- Error absoluto promedio
- Error absoluto máximo porcentual
- Error absoluto máximo
- BIC normalizado

Si desea obtener más información, consulte el tema [Medidas de la bondad de ajuste en el apéndice A el p. 100](#).

**Estadísticos de comparación de modelos.** Este grupo de opciones controla la visualización de las tablas que contienen los estadísticos de todos los modelos. Cada opción genera una tabla independiente. Puede seleccionar una o más de las siguientes opciones:

- **Bondad de ajuste.** Tabla de estadísticos de resumen y percentiles para *R* cuadrado estacionaria, *R* cuadrado, raíz del error cuadrático promedio, error absoluto porcentual promedio, error absoluto promedio, error absoluto máximo porcentual, error absoluto máximo y criterio de información bayesiano normalizado.
- **Función de autocorrelación simple (FAS) residual.** Tabla de estadísticos de resumen y percentiles para las autocorrelaciones de los residuos de todos los modelos estimados. Esta tabla sólo está disponible si se vuelven a estimar los parámetros de modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos).
- **Función de autocorrelación parcial (FAP) residual.** Tabla de estadísticos de resumen y percentiles para las autocorrelaciones parciales de los residuos de todos los modelos estimados. Esta tabla sólo está disponible si se vuelven a estimar los parámetros de modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos).

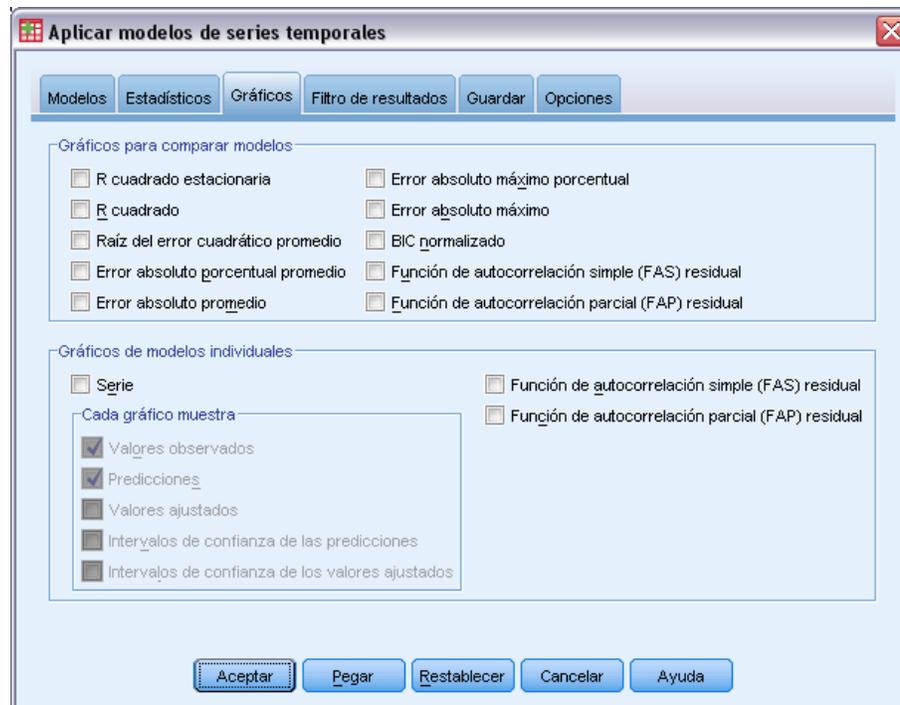
**Estadísticos de modelos individuales.** Este grupo de opciones controla la visualización de las tablas que contienen información detallada sobre cada modelo. Cada opción genera una tabla independiente. Puede seleccionar una o más de las siguientes opciones:

- **Estimaciones de los parámetros.** Muestra una tabla de estimaciones de parámetros para cada modelo. Se muestran tablas independientes para los modelos de suavizado exponencial y ARIMA. Si existen valores atípicos, las estimaciones de parámetros para dichos valores se muestran también en una tabla independiente.
- **Función de autocorrelación simple (FAS) residual.** Muestra una tabla con las autocorrelaciones residuales por retardo para cada modelo estimado. La tabla incluye los intervalos de confianza para las autocorrelaciones. Esta tabla sólo está disponible si se vuelven a estimar los parámetros de modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos).
- **Función de autocorrelación parcial (FAP) residual.** Muestra una tabla con las autocorrelaciones parciales residuales por retardo para cada modelo estimado. La tabla incluye los intervalos de confianza para las autocorrelaciones parciales. Esta tabla sólo está disponible si se vuelven a estimar los parámetros de modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos).

**Mostrar predicciones.** Muestra una tabla con las predicciones de modelo y los intervalos de confianza para cada modelo.

## Gráficos

Figura 3-3  
Aplicar modelos de series temporales, pestaña Gráficos



La pestaña Gráficos proporciona opciones para mostrar gráficos de estadísticos de ajuste del modelo, funciones de autocorrelación y valores de la serie (incluidas las predicciones).

### **Gráficos para comparar modelos**

Este grupo de opciones controla la visualización de los gráficos que contienen los estadísticos en todos los modelos. A menos que se vuelvan a estimar los parámetros del modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos), los valores mostrados proceden del archivo de modelo y reflejan los datos utilizados en el desarrollo (o última actualización) de cada modelo. Además, los gráficos de autocorrelación sólo están disponibles si se vuelven a estimar los parámetros de modelo. Cada opción genera un gráfico independiente. Puede seleccionar una o más de las siguientes opciones:

- *R* cuadrado estacionaria
- *R* cuadrado
- Raíz del error cuadrático promedio
- Error absoluto porcentual promedio
- Error absoluto promedio
- Error absoluto máximo porcentual
- Error absoluto máximo
- BIC normalizado
- Función de autocorrelación simple (FAS) residual
- Función de autocorrelación parcial (FAP) residual

Si desea obtener más información, consulte el tema [Medidas de la bondad de ajuste en el apéndice A el p. 100](#).

### **Gráficos de modelos individuales**

**Serie.** Seleccione (marque) esta opción para obtener los gráficos de los valores pronosticados para cada modelo. Los valores observados, los valores ajustados, los intervalos de confianza y las autocorrelaciones sólo están disponibles si se vuelven a estimar los parámetros de modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos). Puede seleccionar una o más de las siguientes opciones para su inclusión en el gráfico:

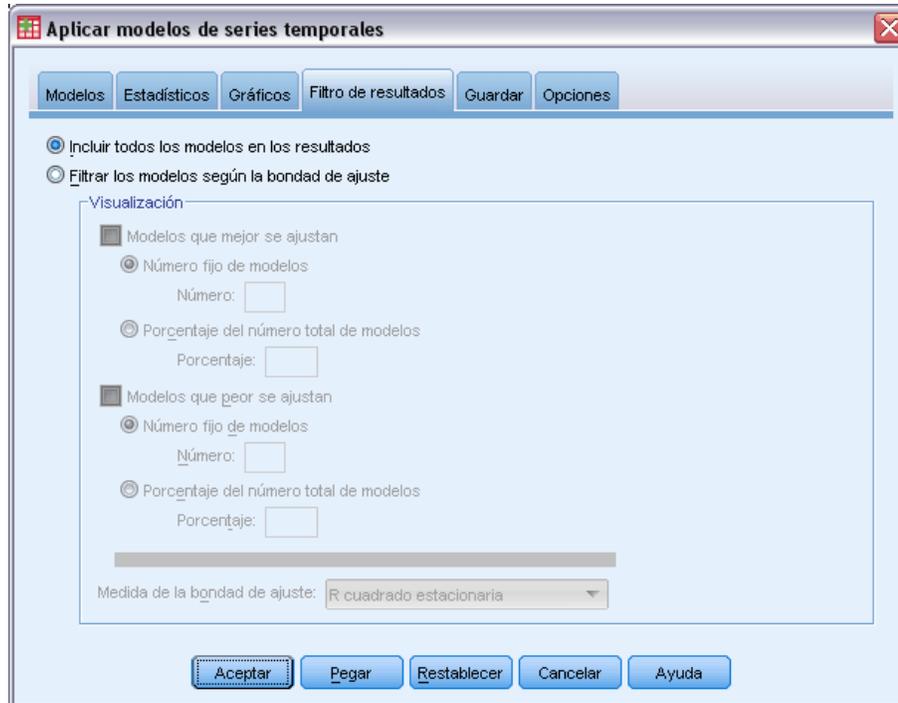
- **Valores observados.** Son los valores observados de la serie dependiente.
- **Predicciones.** Son los valores pronosticados por el modelo para el período de predicción.
- **Valores ajustados.** Son los valores pronosticados por el modelo para el período de estimación.
- **Intervalos de confianza de las predicciones.** Son los intervalos de confianza para el período de predicción.
- **Intervalos de confianza de los valores ajustados.** Son los intervalos de confianza para el período de estimación.

**Función de autocorrelación simple (FAS) residual.** Muestra un gráfico con las autocorrelaciones residuales para cada modelo estimado.

**Función de autocorrelación parcial (FAP) residual.** Muestra un gráfico con las autocorrelaciones parciales residuales para cada modelo estimado.

## Limitación de resultados para los modelos de mejor o peor ajuste

Figura 3-4  
Aplicar modelos de series temporales, pestaña Filtro de resultados



La pestaña Filtro de resultados proporciona opciones para restringir los resultados tanto tabulares como de gráficos a un subconjunto de modelos. Puede optar por limitar los resultados a los modelos de mejor o peor ajuste según los criterios de ajuste proporcionados. Por defecto, todos los modelos se incluyen en los resultados. A menos que se vuelvan a estimar los parámetros del modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos), los valores de las medidas de ajuste utilizados para filtrar los modelos proceden del archivo de modelo y reflejan los datos utilizados en el desarrollo (o última actualización) de cada modelo.

**Modelos que mejor se ajustan.** Seleccione (marque) esta opción para incluir los modelos que mejor se ajustan en los resultados. Seleccione una medida de bondad de ajuste y especifique el número de modelos que se van a incluir. La selección de esta opción no excluye la selección de los modelos que peor se ajustan. En este caso, los resultados incluyen tanto los modelos que peor se ajustan como los que mejor se ajustan.

- **Número fijo de modelos.** Especifica que los resultados se muestran para los  $n$  modelos que mejor se ajustan. Si este número es superior al número de modelos, se muestran todos los modelos.
- **Porcentaje del número total de modelos.** Especifica que los resultados se muestran para los modelos con valores de bondad de ajuste en el porcentaje  $n$  superior de todos los modelos.

**Modelos que peor se ajustan.** Seleccione (marque) esta opción para incluir los modelos que peor se ajustan en los resultados. Seleccione una medida de bondad de ajuste y especifique el número de modelos que se van a incluir. La selección de esta opción no excluye la selección de los

modelos que mejor se ajustan. En este caso, los resultados incluyen tanto los modelos que mejor se ajustan como los que peor se ajustan.

- **Número fijo de modelos.** Especifica que los resultados se muestran para los  $n$  modelos que peor se ajustan. Si este número es superior al número de modelos, se muestran todos los modelos.
- **Porcentaje del número total de modelos.** Especifica que los resultados se muestran para los modelos con valores de bondad de ajuste en el porcentaje  $n$  inferior de todos los modelos.

**Medida de la bondad de ajuste.** Seleccione la medida de bondad de ajuste que se va a utilizar para filtrar los modelos. El valor por defecto es  $R$  cuadrado estacionaria.

## Almacenamiento de predicciones de modelos y especificaciones de modelo

Figura 3-5  
Aplicar modelos de series temporales, pestaña Guardar



La pestaña Guardar le permite guardar predicciones del modelo como variables nuevas en el conjunto de datos activo y guardar las especificaciones del modelo en un archivo externo en formato XML.

**Guardar variables.** Puede guardar predicciones del modelo, intervalos de confianza y residuos como variables nuevas en el conjunto de datos activo. Cada modelo genera su propio conjunto de variables nuevas. Se añaden casos nuevos si el período de predicción se amplía más allá de la duración de la serie de variables dependiente asociada al modelo. A menos que se vuelvan a estimar los parámetros de modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos), los valores pronosticados y los límites de confianza sólo se crean para el período de predicción.

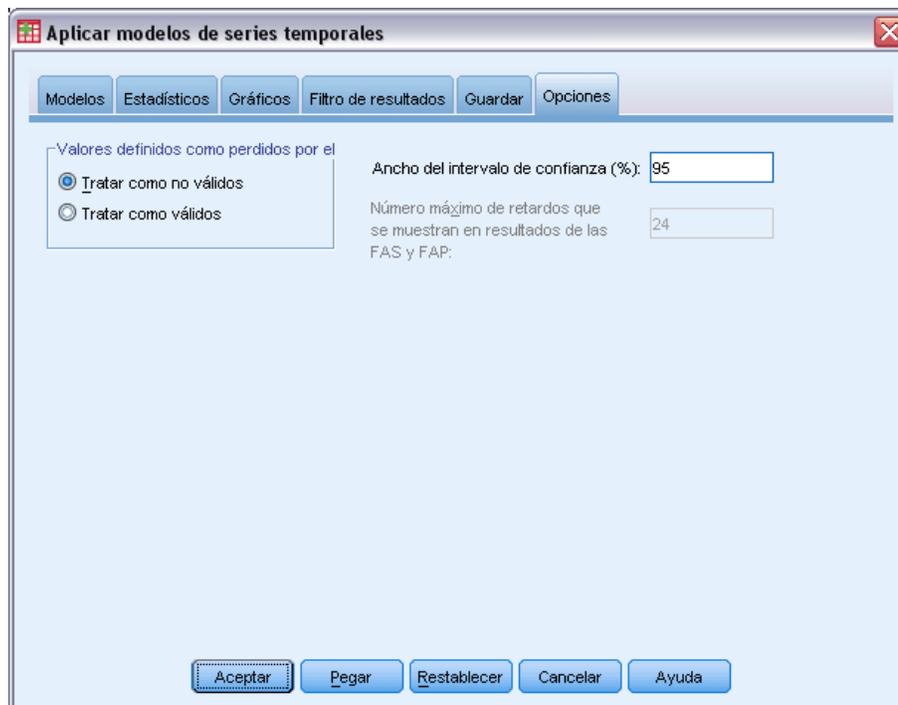
Para guardar las variables nuevas, seleccione la casilla de verificación Guardar asociada a cada variable. Por defecto, no se guarda ninguna variable nueva.

- **Valores pronosticados.** Son los valores pronosticados por el modelo.
- **Límites de confianza inferiores.** Son los límites de confianza inferiores para los valores pronosticados.
- **Límites de confianza superiores.** Son los límites de confianza superiores para los valores pronosticados.
- **Residuos de ruido.** Son los residuos del modelo. Si se realizan transformaciones de la variable dependiente (por ejemplo, logaritmo natural), estos son los residuos de la serie transformada. Esta opción sólo está disponible si se vuelven a estimar los parámetros de modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos).
- **Prefijo del nombre de variable.** Especifique los prefijos que se van a utilizar para los nombres de variables nuevas o deje los prefijos por defecto. Los nombres de variables se componen del prefijo, el nombre de la variable dependiente asociada y un identificador del modelo. El nombre de la variable se amplía en caso de ser necesario para evitar conflictos de nombres de variables. El prefijo debe cumplir las reglas de nombres de variables válidos.

**Exportar archivo de modelo que contiene los parámetros que se han estimado de nuevo.** Las especificaciones de modelo, que contienen los parámetros y estadísticos de ajuste que se han vuelto a estimar, se exportan al archivo especificado en formato XML. Esta opción sólo está disponible si se vuelven a estimar los parámetros de modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos).

## Opciones

Figura 3-6  
Aplicar modelos de series temporales, pestaña Opciones



La pestaña Opciones le permite especificar el tratamiento de los valores perdidos, establecer el ancho del intervalo de confianza y establecer el número de retardos mostrados para las autocorrelaciones.

**Valores definidos como perdidos por el usuario.** Estas opciones controlan el tratamiento de los valores perdidos definidos por el usuario.

- **Tratar como no válidos.** Los valores definidos como perdidos por el usuario reciben el mismo tratamiento que los valores perdidos del sistema.
- **Tratar como válidos.** Los valores perdidos definidos por el usuario se tratan como datos válidos.

**Política de valores perdidos.** Las siguientes reglas se aplican al tratamiento de los valores perdidos (incluye los valores perdidos del sistema y los valores perdidos definidos por el usuario tratados como no válidos):

- Los casos con valores perdidos de una variable dependiente que se producen durante el período de estimación se incluyen en el modelo. El tratamiento específico del valor perdido depende del método de estimación.
- Para los modelos ARIMA, se genera una advertencia si un predictor tiene valores perdidos en el período de estimación. Los modelos que implican el uso de dicho predictor no se vuelven a estimar.
- Si una variable independiente tiene valores perdidos en el período de predicción, el procedimiento genera una advertencia y realiza predicciones en la medida de lo posible.

**Ancho del intervalo de confianza (%).** Los intervalos de confianza se calculan para las predicciones del modelo y las autocorrelaciones residuales. Puede especificar cualquier valor positivo inferior a 100. Por defecto, se utiliza un intervalo de confianza del 95 %.

**Número máximo de retardos que se muestran en resultados de las FAS y FAP.** Puede establecer el número máximo de retardos que se muestran en las tablas y en los gráficos de autocorrelaciones y autocorrelaciones parciales. Esta opción sólo está disponible si se vuelven a estimar los parámetros de modelo (Estimar de nuevo a partir de los datos en la pestaña Modelos).

## ***Funciones adicionales del comando TSAPPLY***

Las funciones adicionales están disponibles si pega las selecciones en una ventana de sintaxis y edita la sintaxis del comando `TSAPPLY` resultante. El lenguaje de sintaxis de comandos permite:

- Especificar que sólo se va a aplicar un subconjunto de modelos de un archivo de modelo al conjunto de datos activo (con las palabras clave `DROP` y `KEEP` del subcomando `MODEL`).
- Aplicar modelos de dos o más archivos de modelo a los datos (con el subcomando `MODEL`). Por ejemplo, un archivo de modelo puede contener modelos para la serie que representa las ventas de unidades y otro puede contener modelos para la serie que representa los ingresos.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

# Descomposición estacional

El procedimiento Descomposición estacional descompone una serie en sus componentes estacionales, un componente de ciclo y tendencia combinado, así como un componente de “error”. El procedimiento es una implementación del método del censo I (Census Method I), también conocido como el método de la razón sobre la media móvil.

**Ejemplo.** Un científico está interesado en analizar mensualmente las medidas del nivel de ozono en una estación meteorológica particular. El objetivo es determinar si hay alguna tendencia en los datos. A fin de descubrir la tendencia real, el científico primero necesita tener en cuenta la variación de las lecturas debido a efectos estacionales. El procedimiento Descomposición estacional puede utilizarse para eliminar cualquier variación estacional sistemática. A continuación, se realiza el análisis de tendencia en una serie corregida estacionalmente.

**Estadísticos.** El conjunto de factores estacionales.

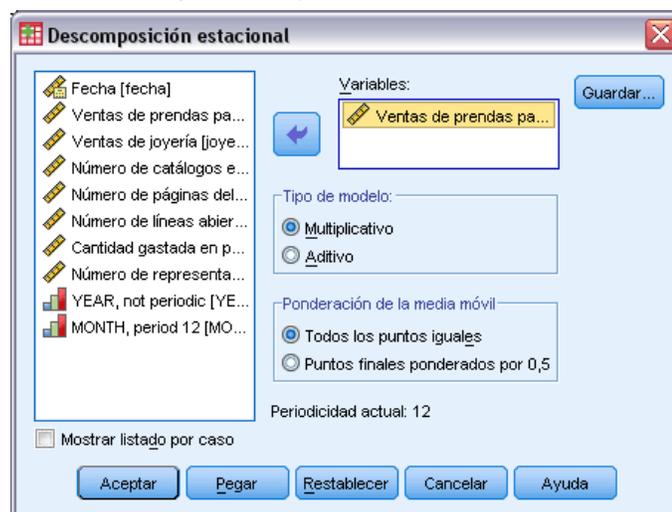
**Datos.** Las variables deben ser numéricas.

**Supuestos.** Las variables no deben contener datos perdidos incrustados. Debe definirse al menos un componente de fecha periódico.

## Estimación de los factores estacionales

- Elija en los menús:  
Analizar > Forecasting > Descomposición estacional...

Figura 4-1  
Cuadro de diálogo Descomposición estacional



- Seleccione una o más variables de la lista disponible y muévalas a la lista de variables. Observe que la lista sólo incluye variables numéricas.

**Tipo de modelo.** El procedimiento Descomposición estacional ofrece dos métodos diferentes para modelar los factores estacionales: multiplicativo o aditivo.

- **Multiplicativo.** En la estacionalidad multiplicativa, el componente estacional es un factor por el que se multiplica la serie corregida estacionalmente para dar lugar a la serie original. En la práctica, componentes estacionales proporcionales al nivel general de la serie. Las observaciones sin variación estacional tendrán un componente estacional de 1.
- **Aditivo.** Las correcciones estacionales se añaden a la serie corregida estacionalmente para obtener los valores observados. Estas correcciones pretenden eliminar de la serie el efecto estacional, para poder estudiar otras características de interés que puedan estar "enmascaradas" por el componente estacional. En la práctica, los componentes estacionales no dependen del nivel general de la serie. Las observaciones sin variación estacional tendrán un componente estacional de 0.

**Ponderación de la media móvil.** Las opciones de Ponderación de la media móvil permiten especificar la manera de tratar la serie al calcular las medias móviles. Estas opciones sólo están disponibles si la periodicidad de la serie es par. Si la periodicidad es impar, todos los puntos son ponderados por igual.

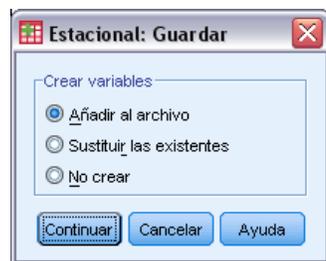
- **Todos los puntos son iguales.** Las medias móviles se calculan con una amplitud igual a la periodicidad y con todos los puntos ponderados igualmente. Si la periodicidad es impar, siempre se utiliza este método.
- **Puntos finales ponderados por 0,5.** Las medias móviles de las series con una periodicidad par se calculan con una duración igual a la periodicidad más 1, y con los puntos finales de la duración ponderados por 0,5.

Si lo desea, puede:

- Pulsar en Guardar para especificar el modo en que las nuevas variables deberían guardarse.

## ***Descomposición estacional: Guardar***

Figura 4-2  
Cuadro de diálogo Estacional: Guardar



**Crear variables** Permite elegir el modo de tratar las variables nuevas.

- **Añadir al archivo.** Las nuevas series creadas por la Descomposición estacional se guardan como variables regulares en el conjunto de datos activo. Los nombres de variable están constituidos por un prefijo de tres letras, un carácter de subrayado y un número.

- **Sustituir las existentes.** Las nuevas series creadas por la Descomposición estacional se guardan como variables regulares en el conjunto de datos activo. Al mismo tiempo, se elimina cualquier variable temporal existente creada por los procedimientos de Predicciones. Los nombres de variable están constituidos por un prefijo de tres letras, un símbolo de almohadilla (#) y un número.
- **No crear.** Las nuevas series no se añaden al conjunto de datos activo.

#### ***Nombres de variables nuevas***

El procedimiento Descomposición estacional genera cuatro variables nuevas (series) con los siguientes prefijos de tres letras para cada serie especificada:

**SAF.** *Factores de corrección estacional.* Estos valores indican el efecto de cada período en el nivel de la serie.

**SAS.** *Serie corregida estacionalmente.* Son los valores obtenidos después de eliminar la variación estacional de una serie.

**STC.** *Componentes de tendencia-ciclo suavizado.* Estos valores muestran la tendencia y el comportamiento cíclico de la serie.

**ERR.** *Valores de residuo o "error".* Son los valores que permanecen después de eliminar los componentes estacionales, de tendencia y ciclo de la serie.

## ***Funciones adicionales del comando SEASON***

Con el lenguaje de sintaxis de comandos también podrá:

- Especificar cualquier tipo de periodicidad dentro del comando SEASON en lugar de seleccionar las alternativas que ofrece el procedimiento Definir fechas.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

# ***Diagramas espectrales***

El procedimiento Diagramas espectrales se utiliza para identificar el comportamiento periódico en la serie temporal. En lugar de analizar la variación desde un punto temporal al siguiente, analiza la variación de las series como un todo en componentes periódicos de frecuencias diferentes. Las series suaves tienen componentes periódicos más fuertes a bajas frecuencias; la variación aleatoria (“ruido blanco”) dispersa la variación a través de todas las frecuencias.

Las series que incluyen datos perdidos no se pueden analizar con este procedimiento.

**Ejemplo.** La tasa a la que se construyen las casas nuevas es un barómetro importante del estado de la economía. Los comienzos de los datos para las viviendas muestran generalmente un componente estacional fuerte. Pero, ¿hay ciclos más largos presentes en los datos que los analistas necesitan conocer a la hora de evaluar las cifras actuales?

**Estadísticos.** Las transformaciones de seno y coseno, el valor del periodograma y la estimación de densidad espectral para cada frecuencia o componente periódico. Cuando se selecciona el análisis bivariado: las partes reales e imaginarias del periodograma cruzado, la densidad coespectral, el espectro de cuadratura, la ganancia, la coherencia cuadrada y el espectro de fase para cada frecuencia o componente periódico.

**Diagramas.** Para el análisis univariado y bivariado: periodograma y densidad espectral. Para el análisis bivariado: la coherencia cuadrada, el espectro de cuadratura, la amplitud cruzada, la densidad coespectral, el espectro de fase y la ganancia.

**Datos.** Las variables deben ser numéricas.

**Supuestos.** Las variables no deben contener datos perdidos incrustados. La serie temporal que analizar debe ser estacionaria y cualquier media distinta de cero debe eliminarse de la serie.

- **Estacionaria.** Condición que deben satisfacer las series temporales a las que se quiere ajustar un modelo ARIMA. Las series MA puras serán estacionarias. Sin embargo, las series AR y ARMA pueden no serlo. Una serie estacionaria tiene una media constante y una varianza constante a lo largo del tiempo.

## ***Obtener análisis espectral***

- ▶ En los menús, seleccione:  
Análisis > Serie temporal > Análisis espectral...

Figura 5-1  
Cuadro de diálogo Diagramas espectrales



- ▶ Seleccione una o más variables de la lista disponible y muévalas a la lista de variables. Observe que la lista sólo incluye variables numéricas.
- ▶ Seleccione una de las opciones de Ventana espectral para elegir el modo de suavizar el periodograma con objeto de obtener una estimación de la densidad espectral. Las opciones de suavizado disponibles son Tukey-Hamming, Tukey, Parzen, Bartlett, Daniell (Unidad) y Ninguno.
  - **Tukey-Hamming.** Las ponderaciones son  $W_k = 0,54 D_p(2 \pi f_k) + 0,23 D_p(2 \pi f_k + \pi/p) + 0,23 D_p(2 \pi f_k - \pi/p)$ , para  $k = 0, \dots, p$ ; donde  $p$  es la parte entera de la mitad de la amplitud y  $D_p$  es el kernel de Dirichlet de orden  $p$ .
  - **Tukey.** Las ponderaciones son  $W_k = 0,5 D_p(2 \pi f_k) + 0,25 D_p(2 \pi f_k + \pi/p) + 0,25 D_p(2 \pi f_k - \pi/p)$ , para  $k = 0, \dots, p$ , donde  $p$  es la parte entera de la mitad de la amplitud, y  $D_p$  es el kernel de Dirichlet de orden  $p$ .
  - **Parzen.** Las ponderaciones son  $W_k = 1/p(2 + \cos(2 \pi f_k)) (F[p/2](2 \pi f_k))^{**2}$ , para  $k=0, \dots, p$ , donde  $p$  es la parte entera de la mitad de la amplitud y  $F[p/2]$  es el kernel de Fejer de orden  $p/2$ .
  - **Bartlett.** Forma de una ventana espectral para la que las ponderaciones ( $W_k$ ) de la mitad superior se calculen mediante la fórmula  $W_k = F_p(2 \pi f_k)$ , para  $k=0, \dots, p$ , donde  $p$  es la parte entera de la mitad de la amplitud, y  $F_p$  es el kernel de Fejer de orden  $p$ . La mitad inferior es simétrica a la mitad superior.
  - **Daniell (unidad).** La forma de una ventana espectral cuyas ponderaciones son todas iguales a 1.
  - **Ninguno.** No hay suavizado. Si se selecciona esta opción, la estimación de la densidad espectral es la misma que el periodograma.

**Amplitud.** Rango de valores consecutivos a lo largo de los cuales se lleva a cabo el suavizado. Generalmente, se utiliza un entero impar. Las amplitudes grandes suavizan más que las amplitudes pequeñas el diagrama de densidad espectral.

**Centrar las variables.** Corrige las series para que tengan una media de cero antes de calcular el espectro y eliminar el término mayor que puede estar asociado a la media de la serie.

**El análisis bivariado — la primera variable con cada uno.** Si se han seleccionado dos o más variables, puede seleccionar esta opción para solicitar análisis espectrales bivariados.

- La primera variable de la lista de variables se trata como la variable independiente y el resto de las variables se tratan como variables dependientes.
- Cada serie posterior a la primera se analiza con la primera serie independientemente de las demás series seleccionadas. También se realizan análisis univariados de cada serie.

**Gráfico.** El periodograma y la densidad espectral están disponibles tanto para el análisis univariado como bivariado. El resto de elecciones sólo están disponibles para el análisis bivariado.

- **Periodograma.** Representación no suavizada de la amplitud espectral (representada en escala logarítmica) respecto a la frecuencia o respecto al periodo. La variación de baja frecuencia es característica de series suaves. La variación distribuida de manera uniforme en todas las frecuencias indica "ruido blanco".
- **Coherencia cuadrada.** El producto de las ganancias de las dos series.
- **Espectro de cuadratura.** La parte imaginaria del periodograma cruzado la cual es una medida de la correlación de los componentes de la frecuencia fuera de fase de las dos series. Los componentes están fuera de fase en  $\pi/2$  radianes.
- **Amplitud cruzada.** La raíz cuadrada de la suma de la densidad coespectral al cuadrado y el espectro de cuadratura al cuadrado.
- **Densidad espectral.** Periodograma que se ha suavizado para eliminar la variación irregular.
- **Densidad coespectral.** La parte real del periodograma cruzado que es una medida de la correlación de los componentes de la frecuencia en-fase de las dos series.
- **Espectro de fase.** Medida del grado en que cada componente de la frecuencia de una serie precede o sigue a la otra.
- **Ganancia.** El cociente de dividir la amplitud cruzada por la densidad espectral para una de las series. Cada una de las dos series tiene su propio valor de ganancia.

**Por la frecuencia.** Todos los gráficos son generados por la frecuencia, con un rango desde la frecuencia 0 (el término constante o medio) hasta la frecuencia 0,5 (el término para un ciclo de dos observaciones).

**Por el período.** Todos los gráficos se generan por período, con un rango desde 2 (el término para un ciclo de dos observaciones) hasta un período igual al número de observaciones (el término constante o medio). El período se representa en escala logarítmica.

## ***Funciones adicionales del comando SPECTRA***

Con el lenguaje de sintaxis de comandos también podrá:

- Guardar variables calculadas de análisis espectral en el conjunto de datos activo para su uso posterior.
- Especificar ponderaciones personalizadas para la ventana espectral.

- Generar diagramas por frecuencia y período.
- Imprimir un listado completo de cada valor que aparece en el diagrama.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

# ***Parte II: Ejemplos***

# ***Predicción masiva con el modelizador experto***

Un analista que trabaja para un proveedor de banda ancha a nivel nacional debe generar predicciones de las suscripciones de usuarios para predecir la utilización de la banda ancha. Las predicciones se deben realizar para cada uno de los 85 mercados locales que conforman la base de suscriptores nacional. Los datos históricos mensuales se encuentran en *broadband\_1.sav*. [Si desea obtener más información, consulte el tema Archivos muestrales en el apéndice D en IBM SPSS Forecasting 19.](#)

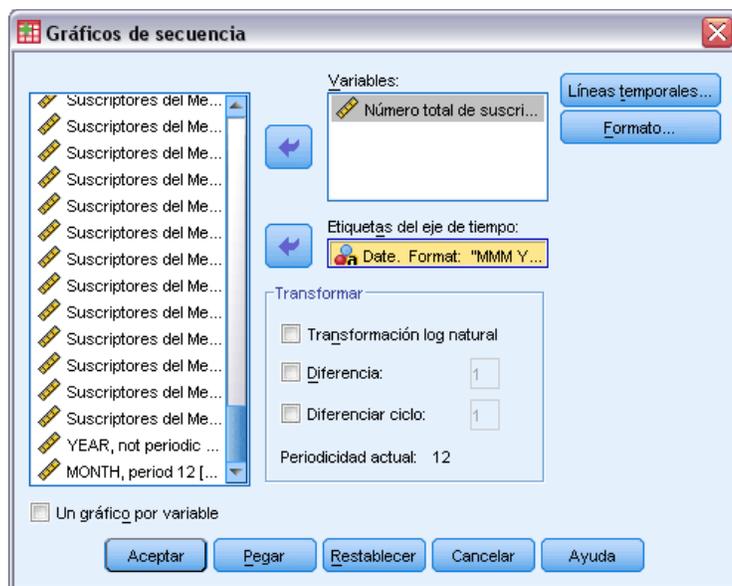
En este ejemplo, utilizará el modelizador experto para generar predicciones de los siguientes tres meses para cada uno de los 85 mercados locales y guardará los modelos generados en un archivo XML externo. Una vez que haya finalizado, podrá continuar con el siguiente ejemplo, [Revisión masiva de predicciones mediante la aplicación de modelos guardados en el capítulo 7 el p. 60](#), que aplica los modelos guardados a un conjunto de datos actualizado para ampliar las predicciones a otros tres meses sin tener que volver a generar los modelos.

## ***Examen de los datos***

Siempre es conveniente conocer la naturaleza de los datos antes de generar un modelo. ¿Los datos muestran variaciones estacionales? Aunque el modelizador experto encontrará de forma automática el mejor modelo estacional o no estacional para cada serie, a menudo se pueden obtener resultados más rápidos si se limita la búsqueda a los modelos no estacionales cuando la estacionalidad no está presente en los datos. Sin examinar los datos para cada uno de los 85 mercados locales, podemos obtener una imagen aproximada al realizar un gráfico del número total de personas suscritas en todos los mercados.

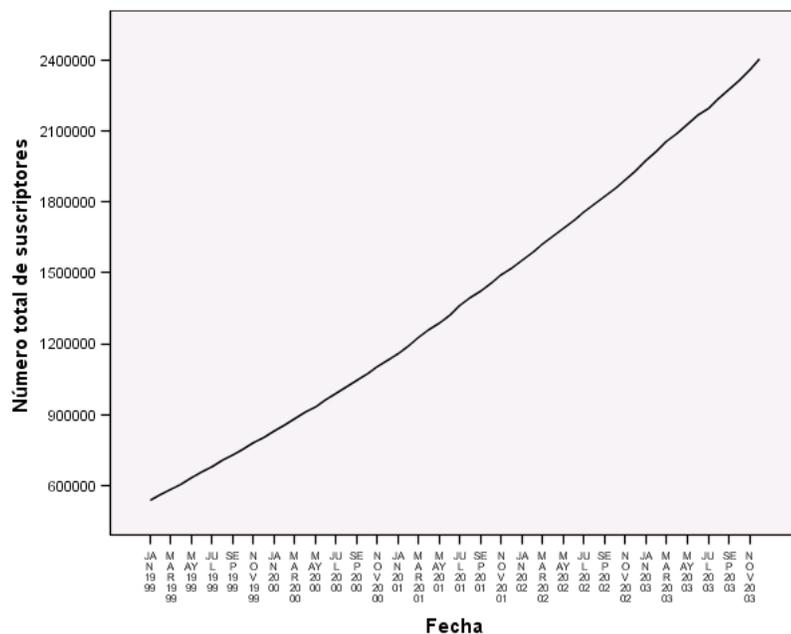
- ▶ Seleccione en los menús:  
Analizar > Predicciones > Gráficos de secuencia...

Figura 6-1  
Cuadro de diálogo Gráficos de secuencia



- ▶ Seleccione *Número total de suscriptores* y muévelo a la lista Variables.
- ▶ Seleccione *Date* y muévelo al cuadro Etiquetas del eje de tiempo.
- ▶ Pulse en Aceptar.

Figura 6-2  
*Número total de suscriptores de banda ancha en todos los mercados*



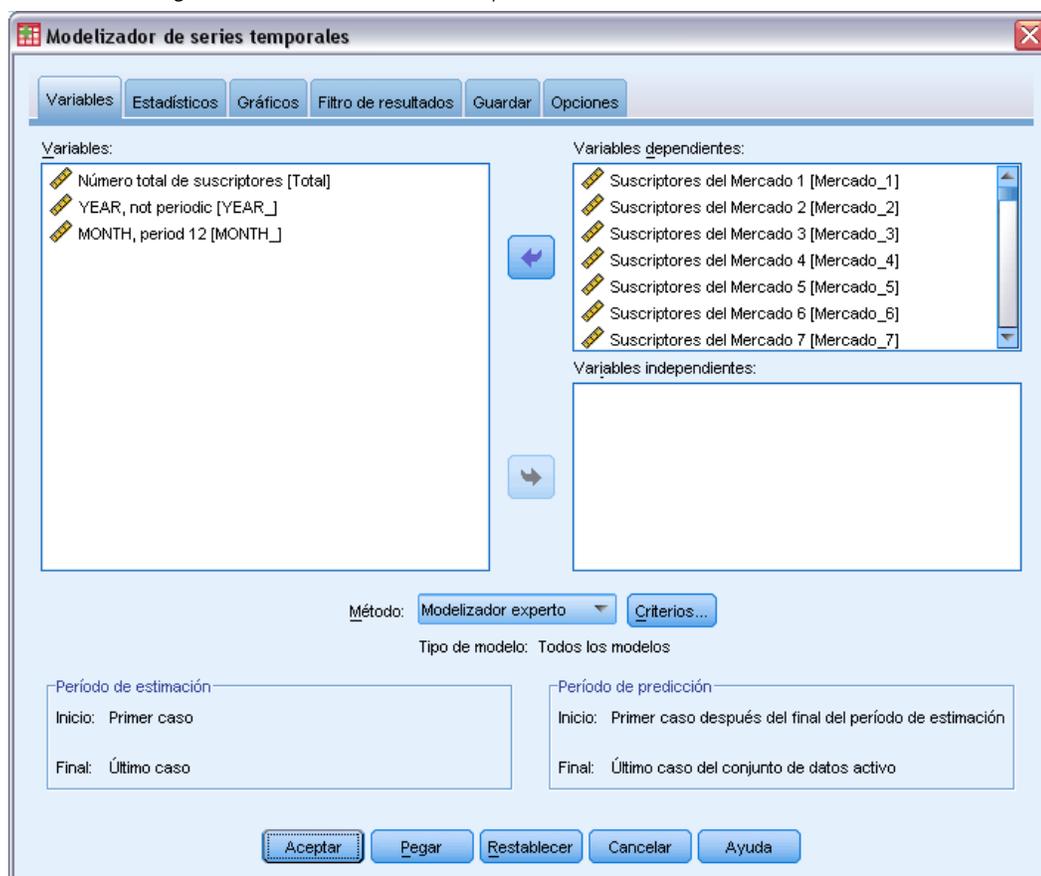
La serie muestra una tendencia ascendente muy suave sin indicios de variaciones estacionales. Puede haber series individuales con estacionalidad, aunque parece que dicha estacionalidad no es una característica prominente de los datos en general. Por supuesto, debe inspeccionar cada una de las series antes de descartar los modelos estacionales. A continuación, puede separar las series que muestren estacionalidad y realizar sus modelos independientemente. En este caso, la inspección de las 85 series mostraría que ninguna incluye estacionalidad.

## Ejecución del análisis

Para utilizar el modelizador experto:

- Seleccione en los menús:  
Analizar > Predicciones > Crear modelos...

Figura 6-3  
Cuadro de diálogo Modelizador de series temporales



- Seleccione desde *Suscriptores del mercado 1* hasta *Suscriptores del mercado 85* para las variables dependientes.

- Compruebe que Modelizador experto está seleccionado en la lista desplegable Método. El modelizador experto buscará automáticamente el modelo que mejor se ajuste a cada serie de las variables dependientes.

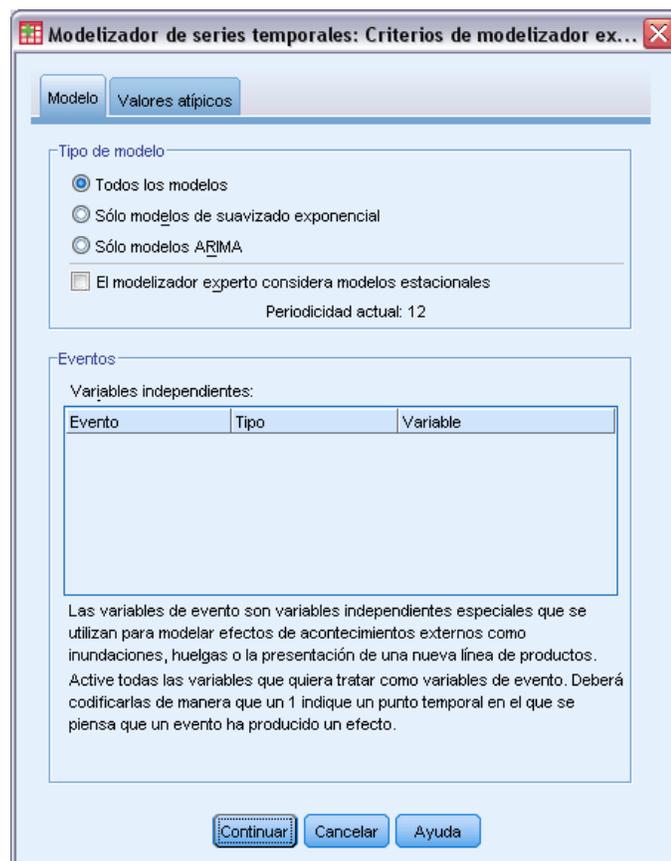
El conjunto de casos utilizados para estimar el modelo se denomina **período de estimación**. Por defecto, incluye todos los casos del conjunto de datos activo. Puede establecer el período de estimación seleccionando Basándose en el rango del tiempo o de los casos en el cuadro de diálogo Seleccionar casos. Para este ejemplo, utilizaremos el período por defecto.

Observe también que el período de predicción por defecto empieza después del período de estimación y llega hasta el último caso del conjunto de datos activo. Si realiza una predicción que va más allá del último caso, deberá ampliar el período de predicción. Esta acción se realiza desde la pestaña Opciones, como se verá en este ejemplo.

- Pulse en Criterios.

Figura 6-4

Cuadro de diálogo Criterios de modelizador experto, pestaña Modelo



- Anule la selección de El modelizador experto considera modelos estacionales del grupo Tipo de modelo.

Aunque los datos son mensuales y la periodicidad actual es 12, se ha visto que los datos no muestran estacionalidad, por lo que no es necesario tener en cuenta modelos estacionales. Esto reduce el espacio de modelos que busca el modelizador experto y puede reducir el tiempo de cálculo considerablemente.

- ▶ Pulse en Continuar.
- ▶ Pulse en la pestaña Opciones del cuadro de diálogo Modelizador de series temporales.

Figura 6-5  
Modelizador de series temporales, pestaña Opciones

Modelizador de series temporales

Variables Estadísticos Gráficos Filtro de resultados Guardar Opciones

Período de predicción

Primer caso después del final del período de estimación hasta el último caso del conjunto de datos activo

Primer caso después del final del período de estimación hasta una fecha especificada

Fecha:

Año	Mes
2004	3

Valores definidos como perdidos por el usuario

Tratar como no válidos

Tratar como válidos

Ancho del intervalo de confianza (%): 95

Prefijo de los identificadores de modelo de los resultados: Modelo

Número máximo de retardos que se muestran en resultados de las FAS y FAP: 24

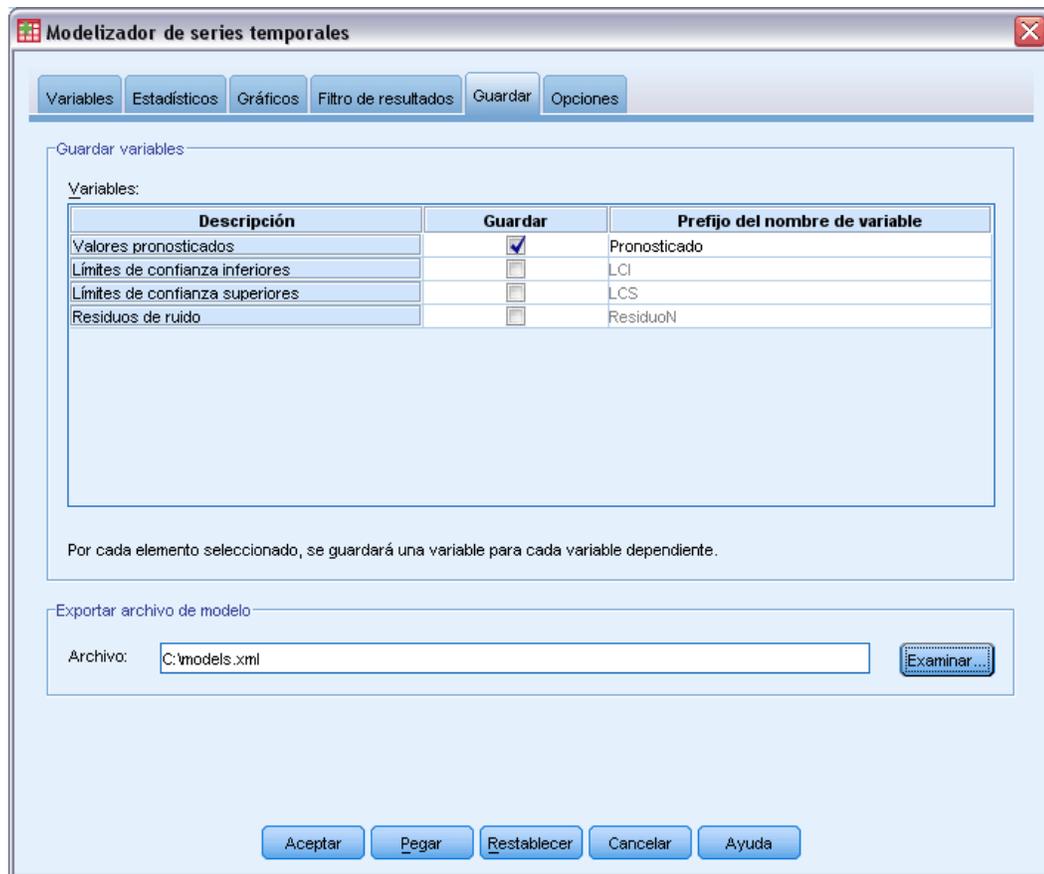
Aceptar Pegar Restablecer Cancelar Ayuda

- ▶ Seleccione Primer caso después del final del período de estimación hasta una fecha especificada en el grupo Período de predicción.
- ▶ En la cuadrícula Fecha, escriba 2004 para el año y 3 para el mes.

El conjunto de datos contiene datos desde enero de 1999 hasta diciembre de 2003. Con la configuración actual, el período de predicción será de enero de 2004 a marzo de 2004.

- ▶ Pulse en la pestaña Guardar.

Figura 6-6  
Modelizador de series temporales, pestaña Guardar

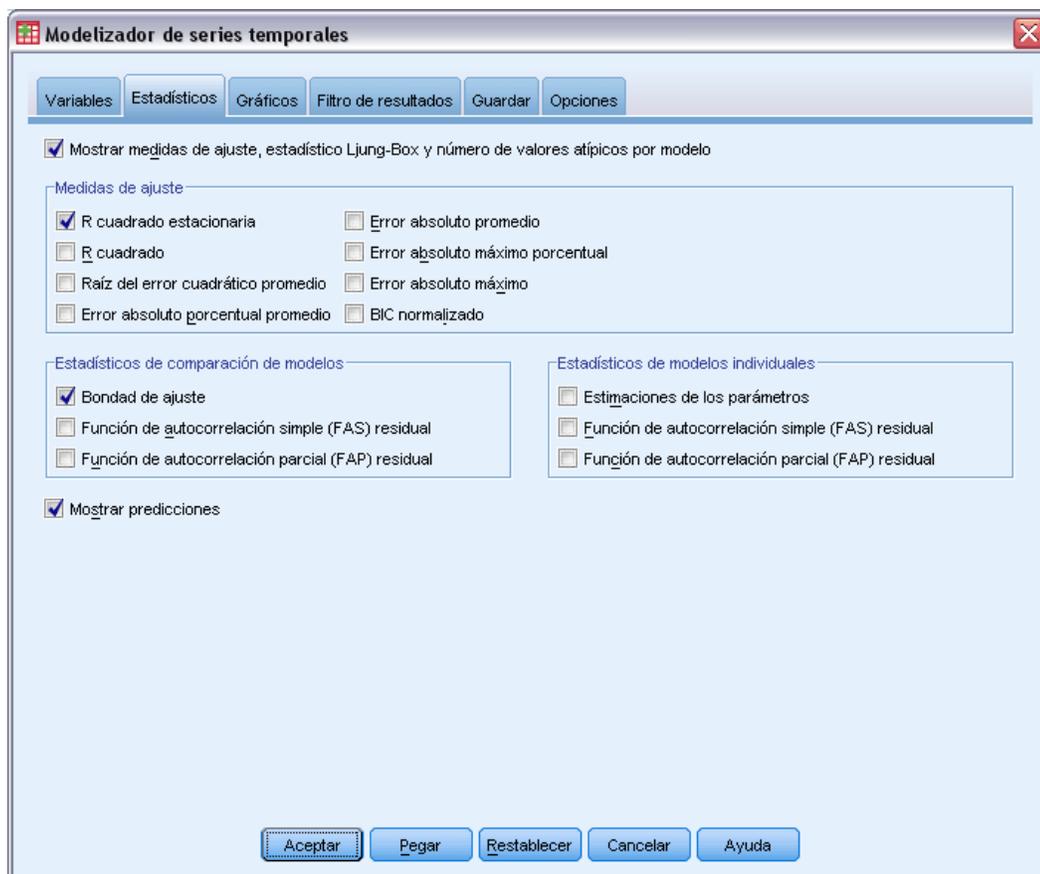


- ▶ Seleccione (marque) la entrada de valores pronosticados en la columna *Guardar* y deje el valor por defecto *Pronosticado* como Prefijo del nombre de variable.

Las predicciones del modelo se guardan como variables nuevas en el conjunto de datos activo, utilizando el prefijo *Pronosticado* para los nombres de variable. También puede guardar las especificaciones de cada uno de los modelos en un archivo XML externo. Así podrá volver a utilizar los modelos para ampliar las predicciones a medida que disponga de más datos.

- ▶ Pulse en el botón Examinar de la pestaña Guardar.  
Aparecerá el cuadro de diálogo estándar para guardar un archivo.
- ▶ Desplácese hasta la carpeta donde desea guardar el archivo de modelo en formato XML, escriba un nombre de archivo y pulse en Guardar.
- ▶ Pulse en la pestaña Estadísticos.

Figura 6-7  
Modelizador de series temporales, pestaña Estadísticos



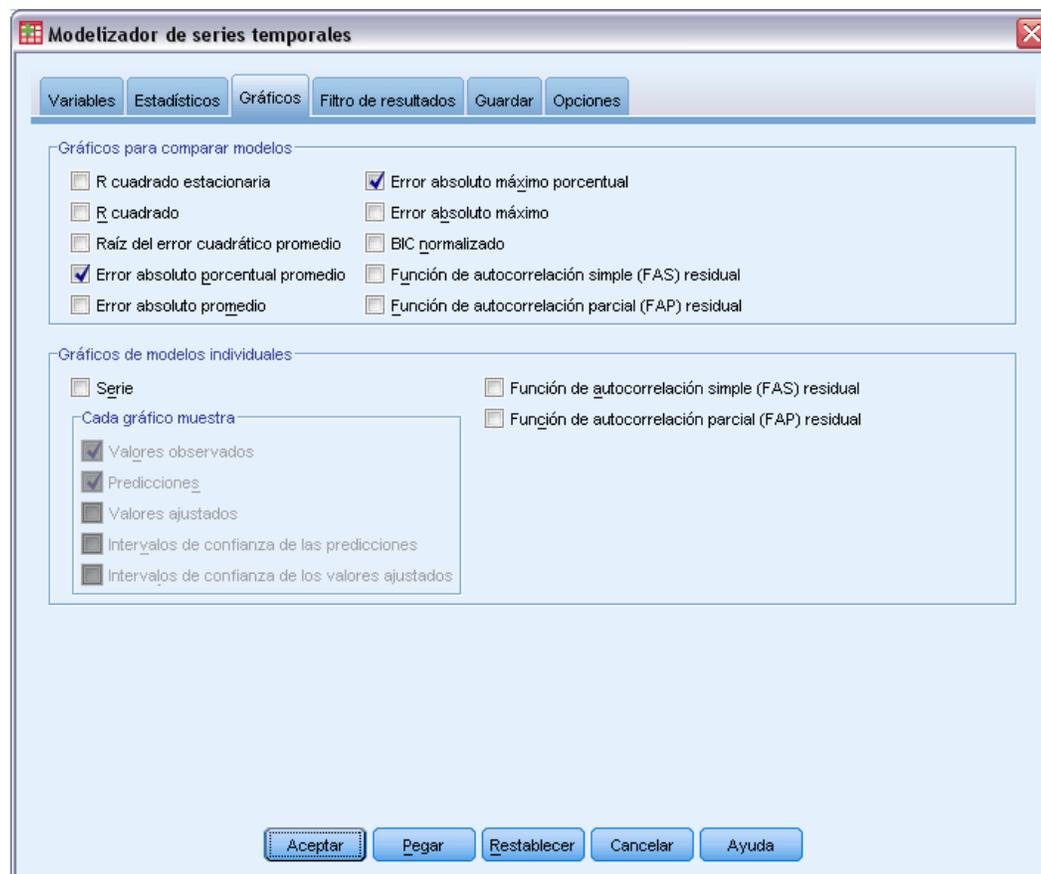
- Seleccione Mostrar predicciones.

Esta opción genera una tabla de valores de predicción para cada serie de las variables dependientes y proporciona otra opción, distinta de guardar las predicciones como variables nuevas, para obtener estos valores.

La selección por defecto de Bondad de ajuste (en el grupo Estadísticos de comparación de modelos) genera una tabla con estadísticos de ajuste, como  $R$ -cuadrado, error absoluto porcentual promedio y BIC normalizado, calculados en todos los modelos. Proporciona un resumen conciso sobre cómo se ajustan los modelos a los datos.

- Pulse en la pestaña Gráficos.

Figura 6-8  
Modelizador de series temporales, pestaña Gráficos



- Anule la selección de Serie en el grupo Gráficos de modelos individuales.

Esto suprime la generación de gráficos de la serie para cada uno de los modelos. En este ejemplo, nos interesa más guardar las predicciones como nuevas variables que generar gráficos de las mismas.

El grupo Gráficos para comparar modelos proporciona varios gráficos (en forma de histogramas) de estadísticos de ajuste calculados en todos los modelos.

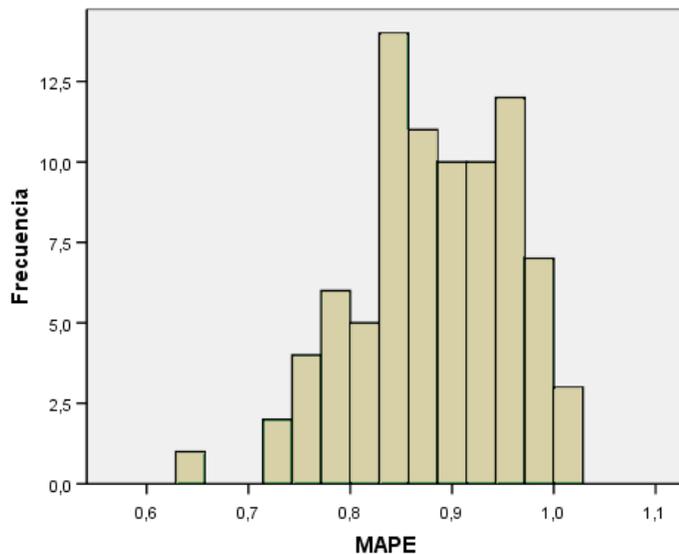
- Seleccione Error absoluto porcentual promedio y Error absoluto máximo porcentual en el grupo Gráficos para comparar modelos.

El error absoluto porcentual es una medida que indica cuánto varía una serie dependiente con respecto al nivel pronosticado por el modelo. Al examinar la media y el máximo en todos los modelos, puede obtener una indicación de la incertidumbre de las predicciones. Además, es recomendable observar los gráficos de resumen de los errores porcentuales en lugar de los errores absolutos, ya que las series dependientes representan los números de suscriptores para mercados de tamaños distintos.

- Pulse en Aceptar en el cuadro de diálogo Modelizador de series temporales.

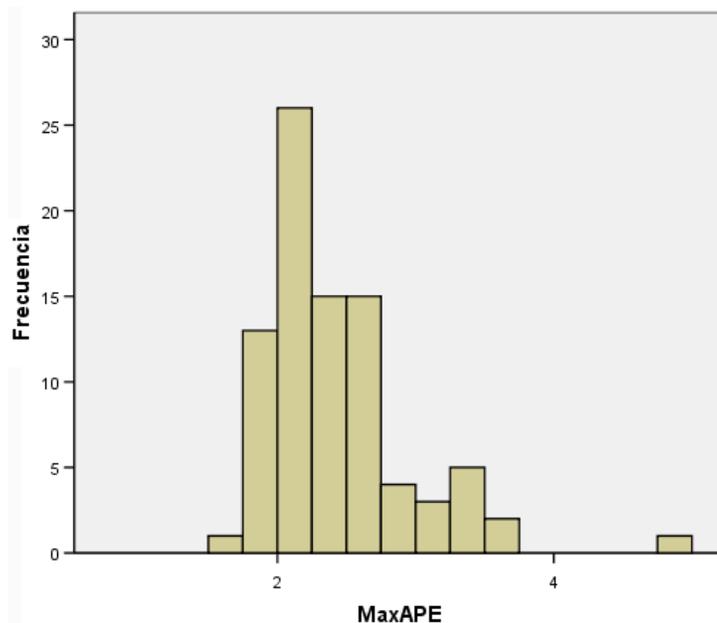
## Gráficos de resumen del modelo

Figura 6-9  
Histograma del error absoluto porcentual promedio



Este histograma muestra el error absoluto porcentual promedio (MAPE) de todos los modelos. Muestra una incertidumbre media de alrededor del 1% en todos los modelos.

Figura 6-10  
Histograma del error absoluto máximo porcentual



Este histograma muestra el error absoluto máximo porcentual (MaxAPE) de todos los modelos y es útil para imaginar el peor escenario para las predicciones. Muestra que el mayor error porcentual de cada modelo se encuentra en el rango del 1% a 5%. ¿Representan estos valores una

cantidad de incertidumbre aceptable? En esta situación, el sentido empresarial entra en juego, ya que el riesgo aceptable varía de un problema a otro.

## Predicciones del modelo

Figura 6-11  
Variables nuevas que contienen predicciones del modelo

YEAR_	MONTH_	DATE_	Predicted_Market_1_Model_1	Predicted_Market_2_Model_2
2003	10	OCT 2003	11820	51084
2003	11	NOV 2003	11857	51273
2003	12	DEC 2003	11687	53082
2004	1	JAN 2004	11503	54893
2004	2	FEB 2004	11447	55856
2004	3	MAR 2004	11390	56704

El Editor de datos muestra las nuevas variables que contienen las predicciones del modelo. Aunque sólo se muestran dos, hay 85 variables nuevas, una por cada una de las 85 series dependientes. Los nombres de variable constan de un prefijo por defecto *Pronosticado*, seguido del nombre de la variable dependiente asociada (por ejemplo, *Mercado\_1*), seguido de un identificador de modelo (por ejemplo, *Modelo\_1*).

Se han añadido tres nuevos casos que contienen las predicciones desde enero de 2004 a marzo de 2004 al conjunto de datos, junto con etiquetas de datos generadas automáticamente. Cada una de las variables nuevas contiene las predicciones del modelo para el período de estimación, (enero de 1999 a diciembre de 2003), lo que permite observar cómo se ajustan los modelos a los valores conocidos.

Figura 6-12  
Tabla de predicciones

Modelo		ENE 2004	FEB 2004	MAR 2004
Suscriptores del Mercado 1-Modelo_1	Previsión	11503	11447	11390
	LCS	11686	11767	11870
	LCI	11321	11126	10910
Suscriptores del Mercado 2-Modelo_2	Previsión	54893	55856	56704
	LCS	55632	57195	58575
	LCI	54154	54518	54832
Suscriptores del Mercado 3-Modelo_3	Previsión	59656	59305	58954
	LCS	60457	60753	61158
	LCI	58856	57857	56750
Suscriptores del Mercado 4-Modelo_4	Previsión	18235	18424	18628
	LCS	18413	18731	19121
	LCI	18058	18116	18136

También puede crear una tabla con los valores de predicción. La tabla consta de los valores de predicción del período de predicción pero, a diferencia de las variables nuevas que contienen las predicciones del modelo, no incluye los valores pronosticados para el período de estimación. Los resultados se organizan por modelo y se identifican por nombre del modelo, que consta del nombre (o la etiqueta) de la variable dependiente asociada seguida de un identificador de modelo, como los nombres de las variables nuevas que contienen las predicciones del modelo. La tabla también incluye los límites de confianza superiores (LCS) y los límites de confianza inferiores (LCI) para los valores de predicción (95% por defecto).

Ya ha visto dos maneras para obtener los valores de predicción: guardar las predicciones como variables nuevas en el conjunto de datos activo y crear una tabla de predicciones. Con cada manera, dispondrá de distintas opciones para exportar las predicciones (por ejemplo, a una hoja de cálculo de Excel).

## ***Resumen***

Ha aprendido a utilizar el modelizador experto para generar predicciones para series múltiples y ha guardado los modelos resultantes en un archivo XML externo. En el siguiente ejemplo, aprenderá a ampliar las predicciones a medida que disponga de más datos, sin tener que volver a generar los modelos, mediante el procedimiento Aplicar modelos de series temporales.

# ***Revisión masiva de predicciones mediante la aplicación de modelos guardados***

Ha utilizado el modelizador de series temporales para crear modelos para los datos de las series temporales y crear predicciones iniciales basadas en los datos disponibles. Desea volver a utilizar estos modelos para ampliar las predicciones a medida que dispone de más datos actuales, por lo que ha guardado los modelos en un archivo externo. Ya puede aplicar los modelos guardados.

Este ejemplo es una ampliación natural del anterior, [Predicción masiva con el modelizador experto en el capítulo 6 el p. 49](#), aunque también puede utilizarse independientemente. En este escenario, imagine que es un analista para un proveedor de banda ancha a nivel nacional que debe realizar predicciones mensuales de suscripciones de usuarios para cada uno de los 85 mercados locales. Ya ha utilizado el modelizador experto para crear modelos y hacer una predicción de tres meses. El almacén de datos se ha actualizado con los datos reales para el período de predicción original, así que desea utilizar esos datos para ampliar el horizonte otros tres meses.

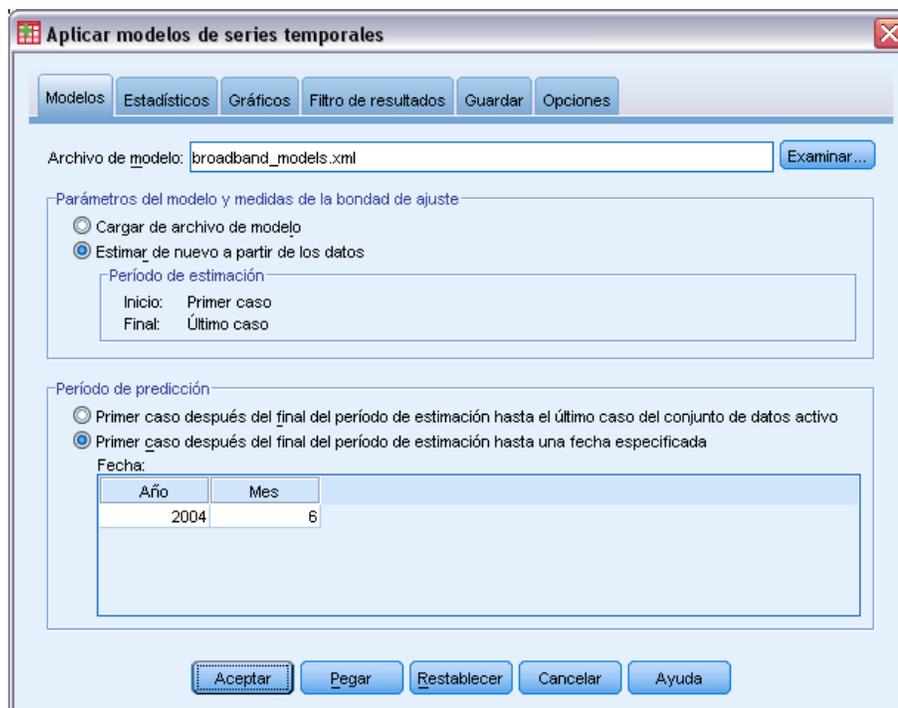
Los datos históricos mensuales actualizados se encuentran en *broadband\_2.sav* y los modelos guardados se encuentran en *broadband\_models.xml*. [Si desea obtener más información, consulte el tema Archivos muestrales en el apéndice D en IBM SPSS Forecasting 19](#). Por supuesto, si trabajó en el ejemplo anterior y guardó su propio archivo de modelo, puede utilizar dicho archivo en lugar de *broadband\_models.xml*.

## ***Ejecución del análisis***

Para aplicar modelos:

- ▶ Elija en los menús:  
Analizar > Predicciones > Aplicar modelos...

Figura 7-1  
Cuadro de diálogo Aplicar modelos de series temporales



- ▶ Pulse en Examinar, desplácese al archivo *broadband\_models.xml* y selecciónelo (o elija su propio archivo de modelo guardado del ejemplo anterior). [Si desea obtener más información, consulte el tema Archivos muestrales en el apéndice D en IBM SPSS Forecasting 19.](#)
- ▶ Seleccione Estimar de nuevo a partir de los datos.  
Para incorporar nuevos valores de las series temporales a las predicciones, el procedimiento Aplicar modelos de series temporales deberá volver a estimar los parámetros del modelo. Sin embargo, la estructura de los modelos permanece intacta, por lo que el tiempo de cálculo para repetir la estimación es mucho menor que el tiempo de cálculo original para generar los modelos.  
El conjunto de casos que se utilizan para repetir la estimación debe incluir los datos nuevos. Esto se garantizará si utiliza el período de estimación por defecto de Primer caso a Último caso. Si necesita establecer el período de estimación en un valor distinto del valor por defecto, puede hacerlo seleccionando Basándose en el rango del tiempo o de los casos en el cuadro de diálogo Seleccionar casos.
- ▶ Seleccione Primer caso después del final del período de estimación hasta una fecha especificada en el grupo Período de predicción.
- ▶ En la cuadrícula Fecha, escriba 2004 para el año y 6 para el mes.  
El conjunto de datos contiene datos desde enero de 1999 hasta marzo de 2004. Con la configuración actual, el período de predicción será de abril de 2004 a junio de 2004.
- ▶ Pulse en la pestaña Guardar.

Figura 7-2  
Aplicar modelos de series temporales, pestaña Guardar

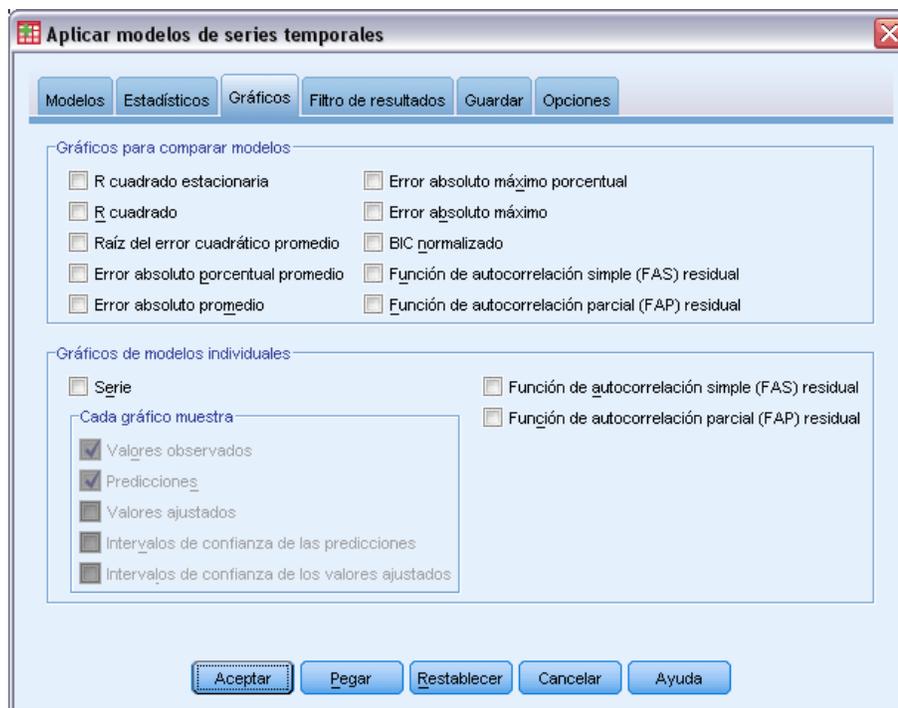


- Seleccione (marque) la entrada para Valores pronosticados en la columna *Guardar* y deje el valor por defecto *Pronosticado* como el prefijo del nombre de variable.

Las predicciones del modelo se guardarán como variables nuevas en el conjunto de datos activo, utilizando el prefijo *Pronosticado* para los nombres de variable.

- Pulse en la pestaña Gráficos.

Figura 7-3  
Aplicar modelos de series temporales, pestaña Gráficos



- Anule la selección de Serie en el grupo Gráficos de modelos individuales.

Esto suprime la generación de gráficos de la serie para cada uno de los modelos. En este ejemplo, nos interesa más guardar las predicciones como nuevas variables que generar gráficos de las mismas.

- Pulse en Aceptar en el cuadro de diálogo Aplicar modelos de series temporales.

## Estadísticos de ajuste del modelo

Figura 7-4  
Tabla Ajuste del modelo

Estadístico de ajuste	Media	ET	Mínimo	Máximo	Percentil						
					5	10	25	50	75	90	95
R-cuadrado estacionaria	,181	,141	-4,2E-015	,609	-8,2E-016	-4,4E-016	,070	,193	,252	,359	,453
R-cuadrado	,999	,000	,998	1,000	,999	,999	,999	,999	,999	,999	1,000
RMSE	187,766	149,959	45,048	764,698	52,558	59,848	93,155	138,913	205,742	433,883	493,904
MAPE	,886	,074	,869	1,026	,754	,792	,832	,894	,948	,976	,996
MaxAPE	2,446	,512	1,742	4,373	1,840	1,952	2,103	2,321	2,663	3,313	3,676
MAE	146,373	113,749	35,573	612,989	41,209	48,742	73,553	111,569	162,306	331,207	364,582
MaxAE	499,714	434,112	116,966	2143,993	131,261	144,918	226,960	345,221	574,071	1165,207	1587,793
BIC normalizado	10,086	1,343	7,749	13,412	8,049	8,285	9,193	9,968	10,717	12,210	12,494

La tabla Ajuste del modelo proporciona estadísticos de ajuste calculados en todos los modelos. Proporciona un resumen conciso sobre cómo se ajustan a los datos los modelos con parámetros que se han vuelto a estimar. Para cada estadístico, la tabla proporciona la media, el error típico (ET), el valor mínimo y máximo de todos los modelos. Además, contiene valores de los

percentiles que proporcionan información sobre la distribución del estadístico en los modelos. Para cada percentil, ese porcentaje de modelos tiene un valor del estadístico de ajuste por debajo del valor establecido. Por ejemplo, el 95% de los modelos tiene un valor de MaxAPE (error absoluto máximo porcentual) inferior a 3,676.

Aunque se muestran distintos estadísticos, nos centraremos en dos: MAPE (error absoluto porcentual promedio) y MaxAPE (error absoluto máximo porcentual). El error absoluto porcentual es una medida que indica cómo varía una serie dependiente respecto al nivel pronosticado por el modelo y proporciona una indicación de la incertidumbre de las predicciones. El error absoluto porcentual promedio varía desde un mínimo de 0,669% hasta un máximo de 1,026% en todos los modelos. El error absoluto máximo porcentual varía de 1,742% a 4,373% en todos los modelos. Por tanto, la incertidumbre media de la predicción de cada modelo ronda el 1% y la incertidumbre máxima está en torno al 2,5% (el valor medio de MaxAPE), siendo el peor escenario aproximadamente del 4%. El grado de riesgo que esté dispuesto a aceptar determinará si estos valores representan un nivel de incertidumbre aceptable.

## Predicciones del modelo

Figura 7-5  
Variables nuevas que contienen predicciones del modelo

YEAR_	MONTH_	DATE_	Predicted_Market_1_Model_1	Predicted_Market_2_Model_2
2004	1	JAN 2004	11513	54947
2004	2	FEB 2004	11806	56810
2004	3	MAR 2004	11950	57344
2004	4	APR 2004	12312	59631
2004	5	MAY 2004	12501	60717
2004	6	JUN 2004	12689	61659

El Editor de datos muestra las nuevas variables que contienen las predicciones del modelo. Aunque sólo se muestran dos, hay 85 variables nuevas, una por cada una de las 85 series dependientes. Los nombres de variable constan de un prefijo por defecto *Pronosticado*, seguido del nombre de la variable dependiente asociada (por ejemplo, *Mercado\_1*), seguido de un identificador de modelo (por ejemplo, *Modelo\_1*).

Se han añadido tres nuevos casos que contienen las predicciones desde abril de 2004 a junio de 2004 al conjunto de datos, junto con etiquetas de datos generadas automáticamente.

## Resumen

Ha aprendido cómo aplicar modelos guardados para ampliar las predicciones anteriores cuando hay más datos actuales disponibles. Además, lo ha hecho sin volver a generar los modelos. Por supuesto, si tiene motivos para pensar que un modelo ha cambiado, debería volver a generarlo mediante el procedimiento Modelizador de series temporales.

## ***Uso del modelizador experto para determinar predictores significativos***

Una compañía de venta por catálogo, interesada en desarrollar un modelo de predicción, ha recopilado datos de las ventas mensuales de prendas para hombre junto con varias series que podrían utilizarse para explicar parte de la variación en las ventas. Los posibles predictores incluyen el número de catálogos enviados por correo, el número de páginas del catálogo, el número de líneas telefónicas abiertas para realizar pedidos, el capital invertido en publicidad impresa y el número de representantes del servicio de atención al cliente. ¿Alguno de estos predictores es útil para la predicción?

En este ejemplo, utilizará el modelizador experto con todos los predictores candidatos para buscar el mejor modelo. Como el modelizador experto sólo selecciona los predictores que tengan una relación estadísticamente significativa con la serie dependiente, detectará los predictores que son útiles y obtendrá un modelo para realizar predicciones con ellos. Una vez que haya finalizado, podrá continuar con el siguiente ejemplo, [Experimentación con predictores mediante la aplicación de modelos guardados en el capítulo 9 el p. 76](#), que investiga el efecto sobre las ventas de distintos escenarios de predictores con el modelo generado en este ejemplo.

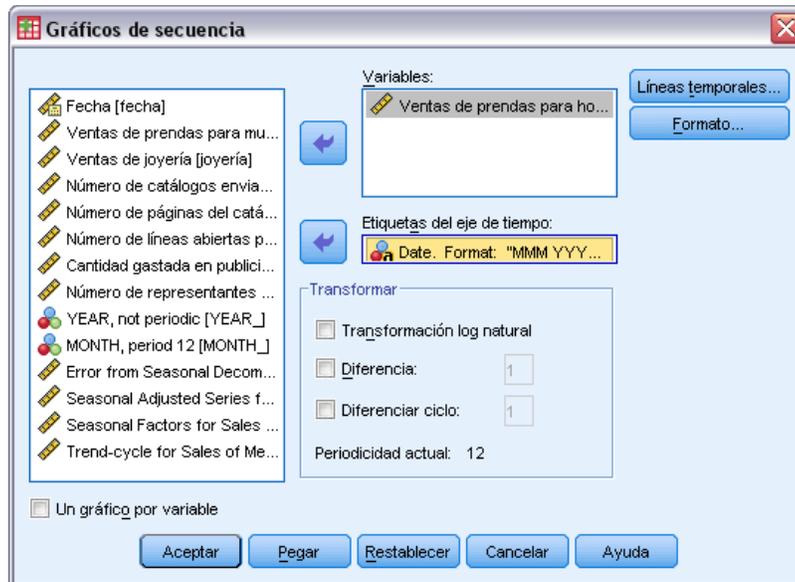
Los datos para el ejemplo actual se encuentran en *catalog\_seasfac.sav*. Si desea obtener más información, consulte el tema [Archivos muestrales en el apéndice D en IBM SPSS Forecasting 19](#).

### ***Gráficos de los datos***

Siempre es aconsejable realizar gráficos de los datos, especialmente si sólo trabaja con una serie:

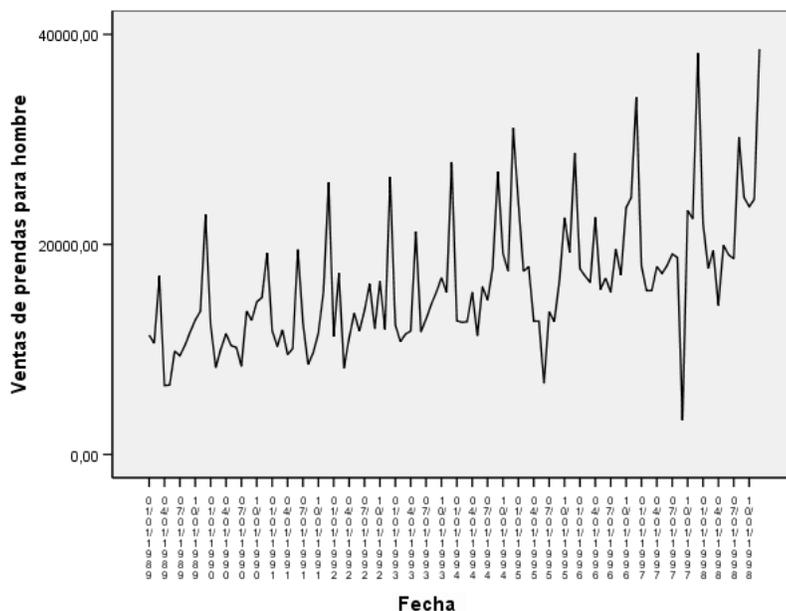
- ▶ Seleccione en los menús:  
Analizar > Predicciones > Gráficos de secuencia...

Figura 8-1  
Cuadro de diálogo Gráficos de secuencia



- ▶ Seleccione *Ventas de prendas para hombre* y muévelo a la lista Variables.
- ▶ Seleccione *Date* y muévelo al cuadro Etiquetas del eje de tiempo.
- ▶ Pulse en Aceptar.

Figura 8-2  
Ventas de prendas para hombre (en dólares estadounidenses)



La serie muestra muchos picos, muchos de los cuales parecen estar espaciados uniformemente, así como una tendencia ascendente clara. Los picos espaciados uniformemente sugieren la presencia de un componente periódico en la serie temporal. Dada la naturaleza estacional de las ventas, que suelen registrar máximos durante la temporada vacacional, no sería raro encontrar un componente estacional anual en los datos.

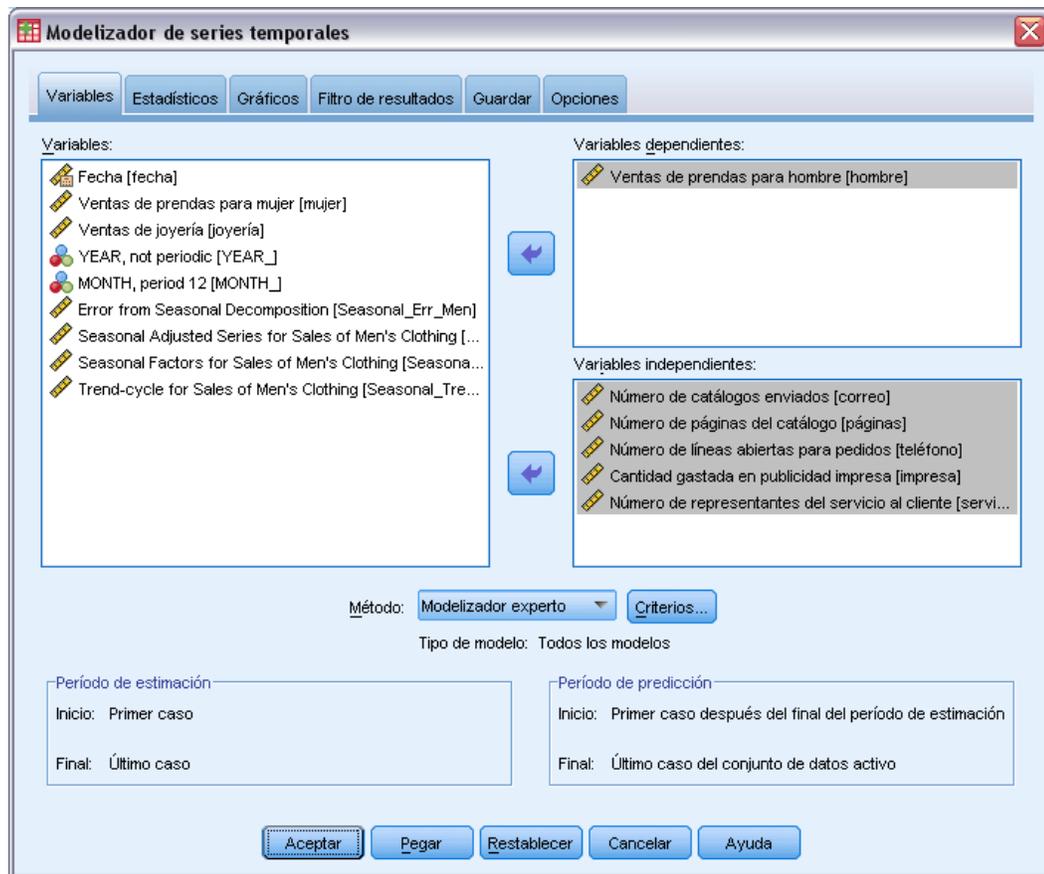
También existen picos que no parecen formar parte del patrón estacional y que representan desviaciones significativas de los puntos de datos adyacentes. Estos puntos pueden ser valores atípicos, que el modelizador experto puede y debe tratar.

## Ejecución del análisis

Para utilizar el modelizador experto:

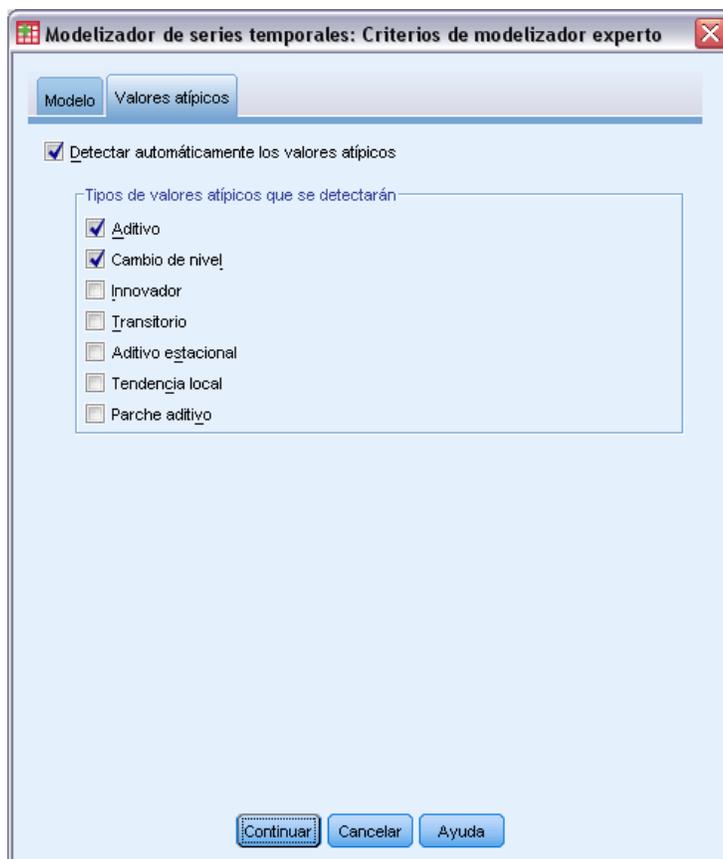
- ▶ Seleccione en los menús:  
Analizar > Predicciones > Crear modelos...

Figura 8-3  
Cuadro de diálogo Modelizador de series temporales



- ▶ Seleccione *Ventas de prendas para hombre* para la variable dependiente.
- ▶ Seleccione desde *Número de catálogos enviados* hasta *Número de representantes del servicio al cliente* para las variables independientes.
- ▶ Compruebe que *Modelizador experto* está seleccionado en la lista desplegable *Método*. El modelizador experto buscará automáticamente el modelo estacional o no estacional que mejor se ajuste a la serie de la variable dependiente.
- ▶ Pulse en *Criterios* y, a continuación, pulse en la pestaña *Valores atípicos*.

Figura 8-4  
Cuadro de diálogo Criterios de modelizador experto, pestaña Valores atípicos

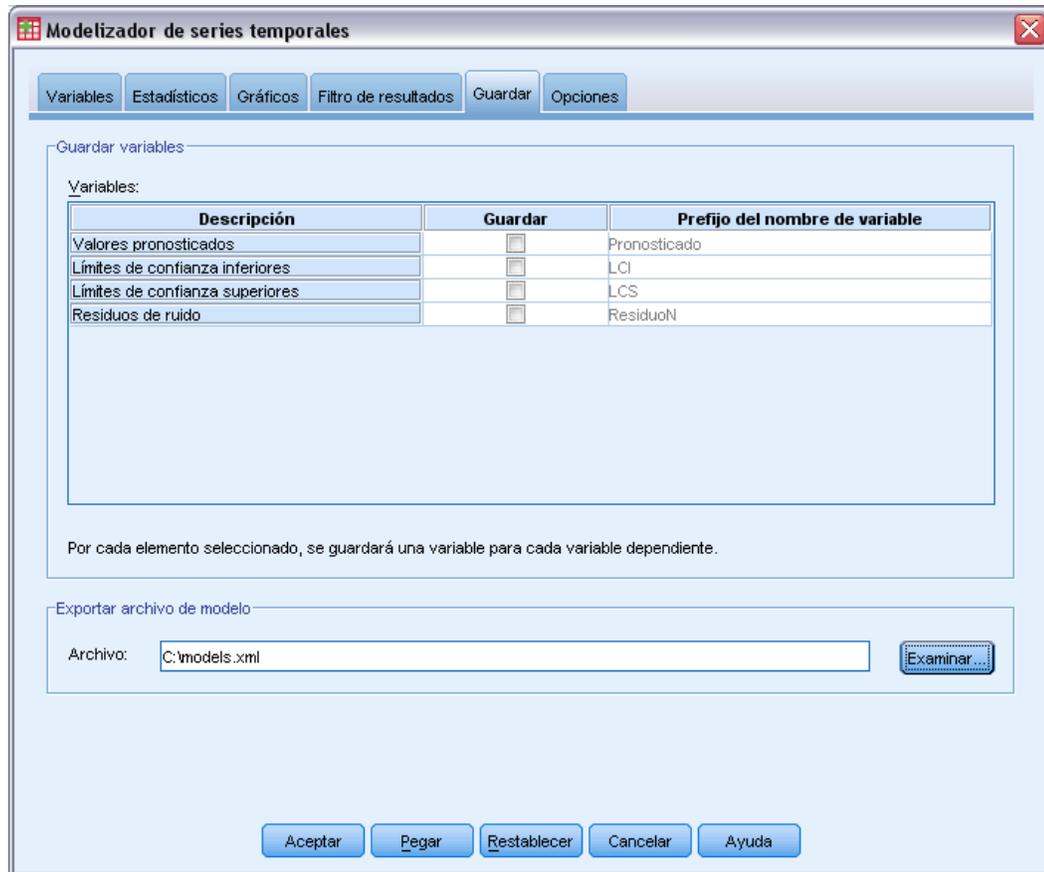


- ▶ Seleccione Detectar automáticamente los valores atípicos y deje las selecciones por defecto para los tipos de valores atípicos que se van a detectar.

La inspección visual de los datos sugiere que puede haber valores atípicos. Con las opciones actuales, el modelizador experto buscará los tipos de valores atípicos más comunes e incorporará cualquier valor atípico al modelo final. La detección de valores atípicos puede incrementar significativamente el tiempo de cálculo que necesita el modelizador experto, por lo que esta característica se debe utilizar con cierta discreción, especialmente al realizar modelos de muchas series a la vez. Por defecto, los valores atípicos no se detectan.

- ▶ Pulse en Continuar.
- ▶ Pulse en la pestaña Guardar del cuadro de diálogo Modelizador de series temporales.

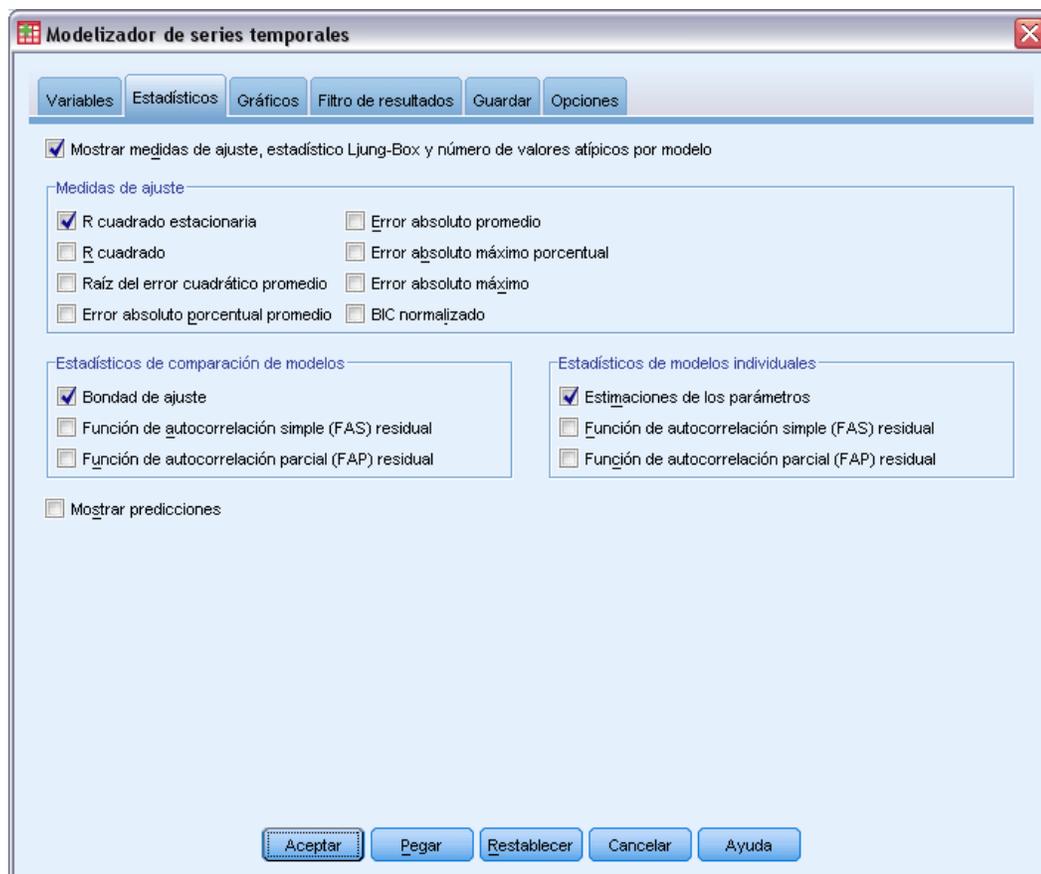
Figura 8-5  
Modelizador de series temporales, pestaña Guardar



Vamos a guardar el modelo estimado en un archivo XML externo para poder experimentar con distintos valores de los predictores, utilizando el procedimiento Aplicar modelos de series temporales, sin tener que volver a generar el modelo.

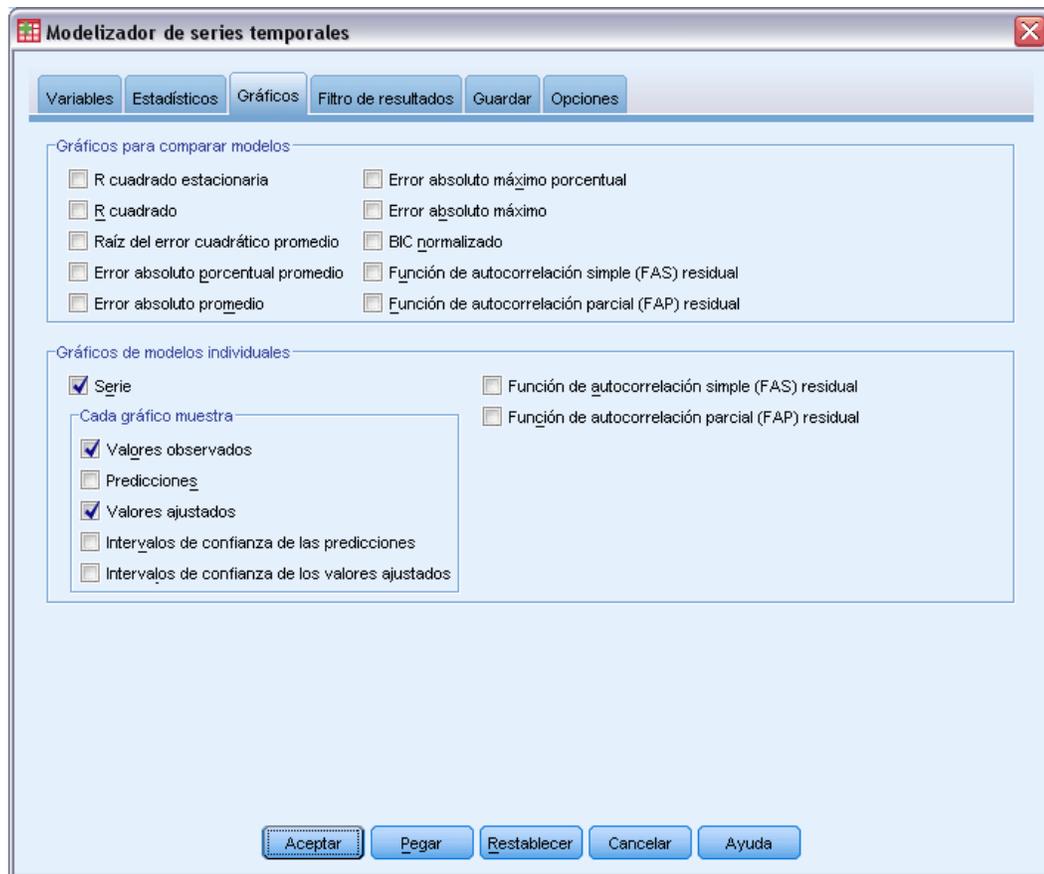
- ▶ Pulse en el botón Examinar de la pestaña Guardar.  
Aparecerá el cuadro de diálogo estándar para guardar un archivo.
- ▶ Desplácese hasta la carpeta donde desea guardar el archivo de modelo en formato XML, escriba un nombre de archivo y pulse en Guardar.
- ▶ Pulse en la pestaña Estadísticos.

Figura 8-6  
Modelizador de series temporales, pestaña Estadísticos



- ▶ Seleccione Estimaciones de los parámetros.  
Esta opción genera una tabla que muestra todos los parámetros, incluidos los predictores significativos, para el modelo elegido por el modelizador experto.
- ▶ Pulse en la pestaña Gráficos.

Figura 8-7  
Modelizador de series temporales, pestaña Gráficos



- Anule la selección de Predicciones.

En el ejemplo actual, sólo nos interesa determinar los predictores significativos y generar un modelo. No realizaremos ninguna predicción.

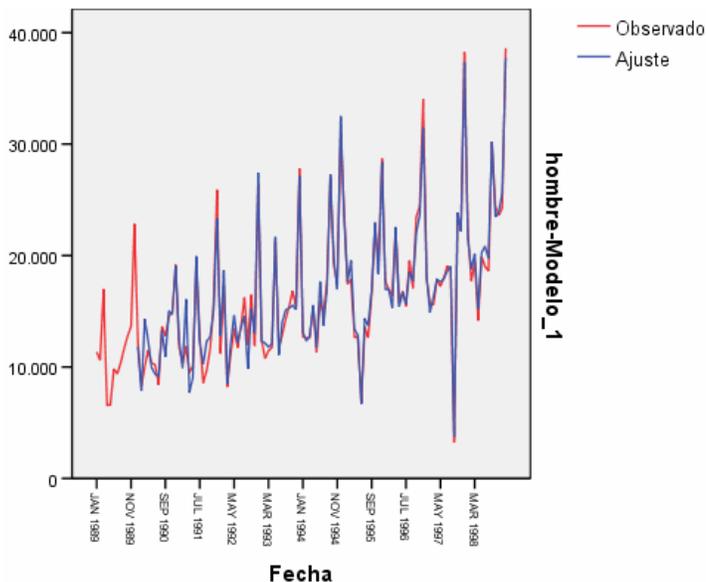
- Seleccione Valores ajustados.

Esta opción muestra los valores pronosticados en el período utilizado para estimar el modelo. Este período se denomina **período de estimación** e incluye todos los casos del conjunto de datos activo para este ejemplo. Estos valores proporcionan una indicación de cómo se ajusta el modelo a los valores observados, por lo que se denominan **valores ajustados**. El gráfico resultante estará compuesto por los valores observados y los valores ajustados.

- Pulse en Aceptar en el cuadro de diálogo Modelizador de series temporales.

## Gráfico de serie

Figura 8-8  
Valores pronosticados y observados



Los valores pronosticados muestran un buen acuerdo con los valores observados, lo que indica que el modelo tiene una capacidad predictora satisfactoria. Observe cómo predice el modelo los picos estacionales. Además, funciona correctamente a la hora de capturar la tendencia ascendente de los datos.

## Tabla de descripción del modelo

Figura 8-9  
Tabla de descripción del modelo

ID del modelo	Ventas de prendas para hombre	Modelo_1	Tipo de modelo
			ARIMA(0,0,0)(0,1,0)

La tabla de descripción del modelo contiene una entrada para cada modelo estimado e incluye un identificador de modelo y el tipo de modelo. El identificador de modelo consta del nombre (o la etiqueta) de la variable dependiente asociada y un nombre asignado por el sistema. En el ejemplo actual, la variable es *Ventas de prendas para hombre* y el nombre asignado por el sistema es *Modelo\_1*.

El modelizador de series temporales soporta tanto el modelo de suavizado exponencial como los modelos ARIMA. Los tipos de modelos de suavizado exponencial se enumeran por sus nombres más habituales, como Holt y Aditivo de Winters. Los tipos de modelos ARIMA se enumeran mediante la notación estándar de ARIMA  $(p,d,q)(P,D,Q)$ , donde  $p$  es el orden de autorregresión,  $d$  es el orden de la diferenciación (o la integración) y  $q$  es el orden de media móvil y  $(P,D,Q)$  son sus homólogos estacionales.

El modelizador experto ha determinado que las ventas de prendas para hombre se describen mejor mediante un modelo ARIMA estacional con un orden de diferenciación. La naturaleza estacional del modelo explica los picos estacionales que vimos en el gráfico de la serie, y el orden simple de diferenciación refleja la tendencia ascendente que era evidente en los datos.

## Tabla de estadísticos del modelo

Figura 8-10  
tabla de estadísticos del modelo

Modelo	Número de predictores	Estadísticos de ajuste del modelo	Ljung-Box Q(18)			Número de valores atípicos
		R-cuadrado estacionaria	Estadísticos	GL	Sig.	
Ventas de prendas para hombre-Modelo_1	2	,948	7,589	18	,984	9

La tabla de estadísticos del modelo proporciona información de resumen y estadísticos de bondad de ajuste para cada modelo estimado. Los resultados de cada modelo se etiquetan con el identificador de modelo proporcionado en la tabla de descripción del modelo. Primero, observe que el modelo contiene dos de los cinco predictores candidatos que especificó en un principio. Por tanto, parece que el modelizador experto ha identificado dos variables independientes que pueden ser útiles para la realizar predicciones.

Aunque el modelizador de series temporales ofrece distintos estadísticos de bondad de ajuste, optamos sólo por el valor  $R$  cuadrado estacionaria. Este estadístico proporciona una estimación de la proporción de la variación total de la serie que se explica mediante el modelo y se prefiere al  $R$  cuadrado ordinario cuando existe una tendencia o un patrón estacional, como en este caso. Los valores más grandes de  $R$  cuadrado estacionario (hasta un valor máximo de 1) indican un ajuste mejor. Un valor de 0,948 indica que el modelo realiza un trabajo excelente al explicar la variación observada en las series.

El estadístico de Ljung-Box, también conocido como el estadístico Box-Pierce modificado, proporciona una indicación de si el modelo se ha especificado correctamente. Un valor de significación inferior a 0,05 implica que existe una estructura en la serie observada que el modelo no explica. El valor de 0,984 que se muestra aquí no es significativo, por lo que podemos estar seguros de que el modelo está correctamente especificado.

El modelizador experto detectó nueve puntos que se consideraban valores atípicos. Se ha creado un modelo de cada uno de estos puntos, por lo que no es necesario eliminarlos de la serie.

## Tabla de parámetros del modelo ARIMA

Figura 8-11  
tabla de parámetros del modelo ARIMA

				Estimación	ET	t	Sig.
Ventas de prendas para hombre-Modelo_1	Ventas de prendas para hombre	Sin transformación	Diferenciación estacional	1			
	Número de catálogos enviados	Sin transformación	Numerador Lag 0	1,549	,071	21,943	,000
			Diferenciación estacional	1			
	Número de líneas abiertas para pedidos	Sin transformación	Numerador Lag 0	315,262	15,298	20,607	,000
			Diferenciación estacional	1			

La tabla de parámetros del modelo ARIMA muestra valores para todos los parámetros del modelo, con una entrada para cada modelo estimado etiquetado por el identificador de modelo. En nuestro caso, mostrará una lista de todas las variables del modelo, incluida la variable dependiente y cualquier variable independiente que el modelizador experto consideró significativas. La tabla de estadísticos del modelo ha indicado que hay dos predictores significativos. La tabla de parámetros del modelo nos muestra que son *Número de catálogos enviados* y *Número de líneas abiertas para pedidos*.

## **Resumen**

Ha aprendido a utilizar el modelizador experto para generar un modelo e identificar predictores significativos y ha guardado el modelo resultante en un archivo externo. Ahora puede utilizar el procedimiento Aplicar modelos de series temporales para experimentar con escenarios alternativos para las series de predictores y ver cómo afectan las alternativas a la predicción de ventas.

# ***Experimentación con predictores mediante la aplicación de modelos guardados***

Ha utilizado el modelizador de series temporales para crear un modelo para los datos e identificar los predictores que pueden resultar útiles para realizar predicciones. Los predictores representan factores que se encuentran bajo su control, así, desea experimentar con sus valores en el período de predicción para ver cómo resultan afectadas las predicciones de la variable dependiente. Esta tarea se realiza fácilmente con el procedimiento Aplicar modelos de series temporales, utilizando el archivo de modelo que se crea con el procedimiento Modelizador de series temporales.

Este ejemplo es una ampliación natural del ejemplo anterior, [Uso del modelizador experto para determinar predictores significativos en el capítulo 8 el p. 65.](#), aunque también puede utilizarse independientemente. El escenario implica una compañía de venta por catálogo que ha reunido datos sobre las ventas mensuales de ropa para hombre desde enero de 1989 a diciembre de 1998, junto con varias series que se pensó que podían ser útiles como predictores de las ventas futuras. El modelizador experto ha determinado que sólo dos de los cinco predictores candidatos son significativos: el número de catálogos enviados y el número de líneas telefónicas abiertas para pedidos.

Al planificar la estrategia de ventas para el año siguiente, cuenta con recursos limitados para imprimir catálogos y mantener las líneas telefónicas abiertas para realizar los pedidos. El presupuesto para los tres primeros meses de 1999 permite contar con 2.000 catálogos adicionales o 5 líneas telefónicas adicionales respecto las proyecciones iniciales. ¿Qué opción generará un ingreso mayor de las ventas para este período de tres meses?

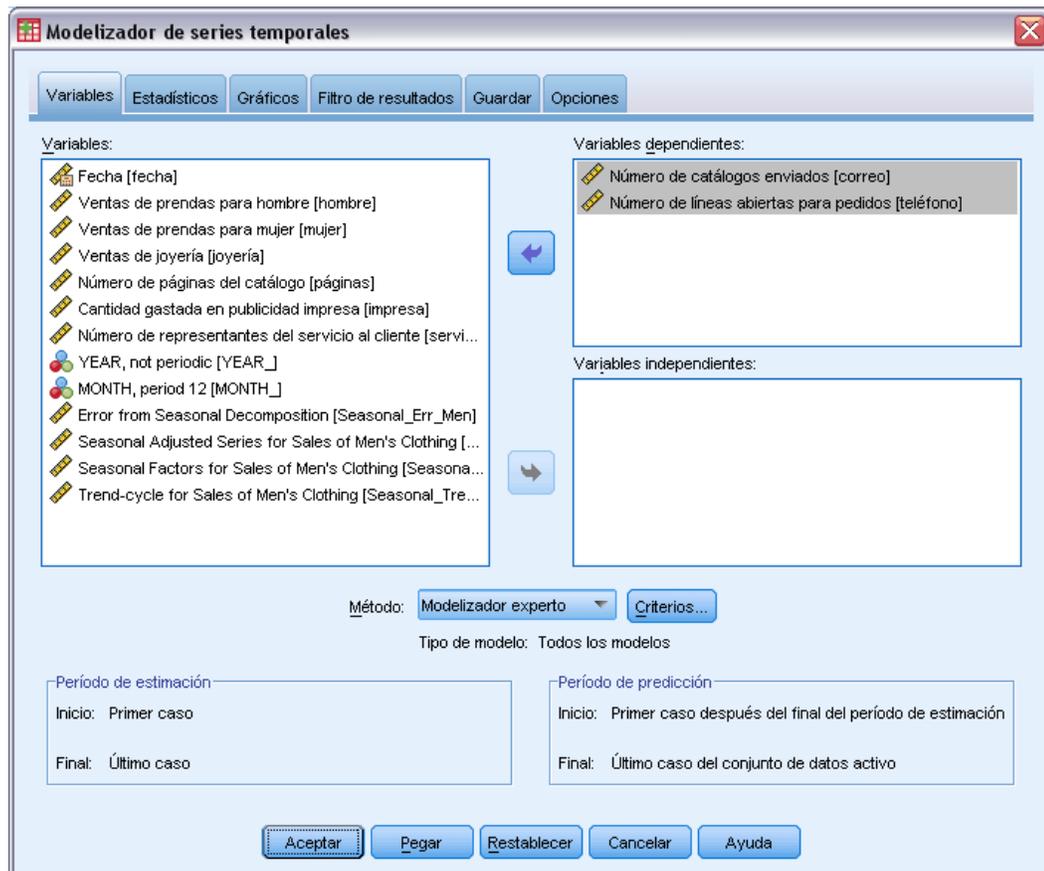
Los datos para este ejemplo se encuentran en *catalog\_seasfac.sav*, y *catalog\_model.xml* contienen el modelo de las ventas mensuales que se genera con el modelizador experto. [Si desea obtener más información, consulte el tema Archivos muestrales en el apéndice D en IBM SPSS Forecasting 19.](#) Por supuesto, si trabajó en el ejemplo anterior y guardó su propio archivo de modelo, puede utilizar dicho archivo en lugar de *catalog\_model.xml*.

## ***Ampliación de la serie de predictores***

Cuando crea predicciones para series dependientes con predictores, cada serie de predictores debe ampliarse en el período de predicción. A menos que sepa exactamente cuáles serán los valores futuros de los predictores, deberá estimarlos. Más tarde podrá modificar las estimaciones para probar los distintos escenarios de predictores. Las proyecciones iniciales se crean fácilmente mediante el modelizador experto.

- Elija en los menús:  
Analizar > Predicciones > Crear modelos...

Figura 9-1  
Cuadro de diálogo Modelizador de series temporales



- Seleccione *Número de catálogos enviados* y *Número de líneas abiertas para pedidos* para las variables dependientes.
- Pulse en la pestaña Guardar.

Figura 9-2  
Modelizador de series temporales, pestaña Guardar



- ▶ En la columna *Guardar*, seleccione (marque) la entrada para Valores pronosticados y deje el valor por defecto *Pronosticado* para el prefijo del nombre de variable.
- ▶ Pulse en la pestaña *Opciones*.

Figura 9-3  
Modelizador de series temporales, pestaña Opciones

Año	Mes
1999	3

- ▶ En el grupo Período de predicción, seleccione Primer caso después del final del período de estimación hasta una fecha especificada.

- ▶ En la cuadrícula Fecha, escriba 1999 para el año y 3 para el mes.

El conjunto de datos contiene datos desde enero de 1989 hasta diciembre de 1998, por lo que, con la configuración actual, el período de predicción comprenderá desde enero de 1999 hasta marzo de 1999.

- ▶ Pulse en Aceptar.

**Figura 9-4**  
Nuevas variables que contienen predicciones para la serie de predictores

	Pronosticado_correo_Modelo_1	Pronosticado_teléfono_Modelo_2	
121	11742	51	▲
122	11853	45	■
123	11965	45	▼

El Editor de datos muestra las nuevas variables *Pronosticado\_correo\_Modelo\_1* y *Pronosticado\_teléfono\_Modelo\_2*, que contienen los valores pronosticados por el modelo para el número de catálogos enviados por correo y el número de líneas telefónicas. Para ampliar la serie de predictores, sólo se necesitan los valores de enero de 1999 a marzo de 1999, que suponen los casos 121 al 123.

- ▶ Copie los valores de estos tres casos de *Pronosticado\_correo\_Modelo\_1* y añádalos a la variable *correo*.
- ▶ Repita este proceso para *Pronosticado\_teléfono\_Modelo\_2*, copiando los tres casos y añadiéndolos a la variable *teléfono*.

**Figura 9-5**  
Serie de predictores ampliada en el período de predicción

	correo	páginas	teléfono	impresa	servicio	YEAR_	MONTH_	DATE_
121	11742	.	51	.	.	1999	1	JAN 1999
122	11853	.	45	.	.	1999	2	FEB 1999
123	11965	.	45	.	.	1999	3	MAR 1999

Los predictores se han ampliado en el período de predicción.

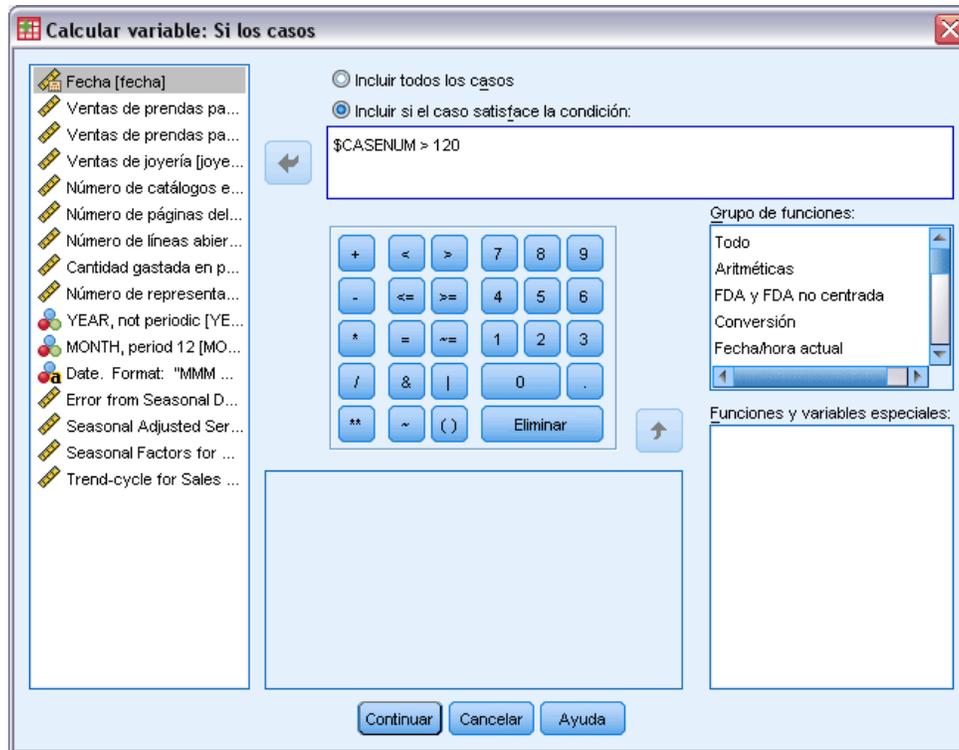
## **Modificación de valores predictores en el período de predicción**

La comprobación de los dos escenarios de enviar por correo más catálogos o proporcionar más líneas telefónicas requiere modificar las estimaciones para los predictores *correo* o *teléfono*, respectivamente. Como sólo se modifican los valores predictores para tres casos (meses), sería fácil escribir los nuevos valores directamente en las celdas adecuadas del Editor de datos. Sin embargo, por motivos didácticos, utilizaremos el cuadro de diálogo *Calcular variable*. Cuando tenga más variables para modificar, probablemente encuentre este cuadro de diálogo más cómodo.

- ▶ Elija en los menús:  
Transformar > Calcular variable...



Figura 9-7  
Cuadro de diálogo *Calcular variable: Si los casos*

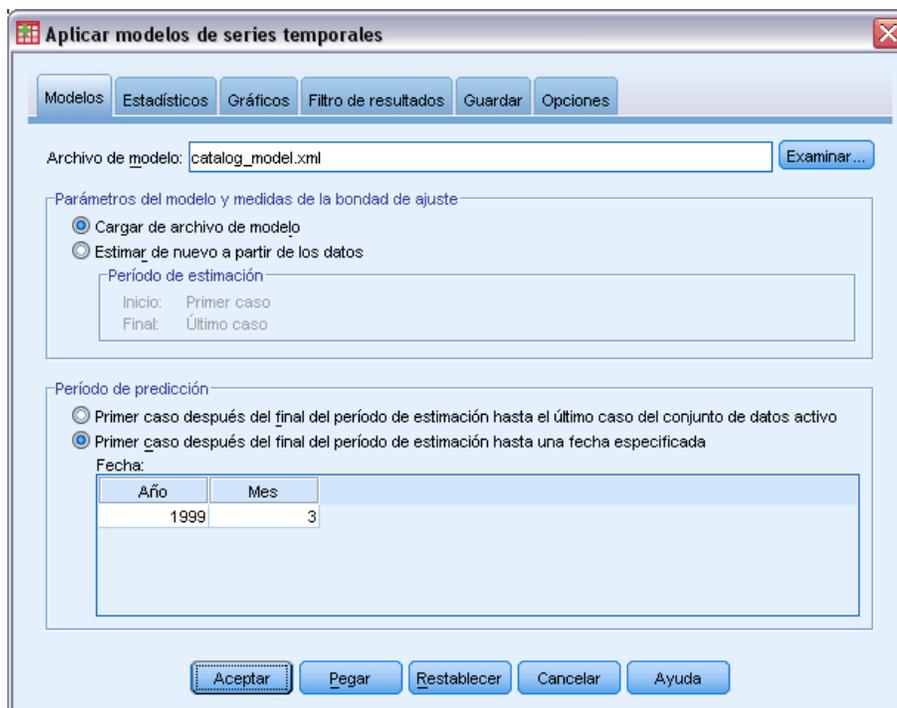


- ▶ Seleccione *Incluir si el caso satisface la condición*.
- ▶ En el cuadro de texto, escriba `$CASENUM > 120`.  
Esto limitará los cambios de la variable *correo* a los casos del período de predicción.
- ▶ Pulse en *Continuar*.
- ▶ Pulse en *Aceptar* en el cuadro de diálogo *Calcular variable* y pulse en *Aceptar* cuando se le pregunte si desea cambiar la variable existente.  
Esto incrementará los valores de *correo*, el número de catálogos enviados, en 2.000 para cada uno de los tres meses del período de predicción. Ahora los datos están preparados para probar el primer escenario y ya se puede ejecutar el análisis.

## Ejecución del análisis

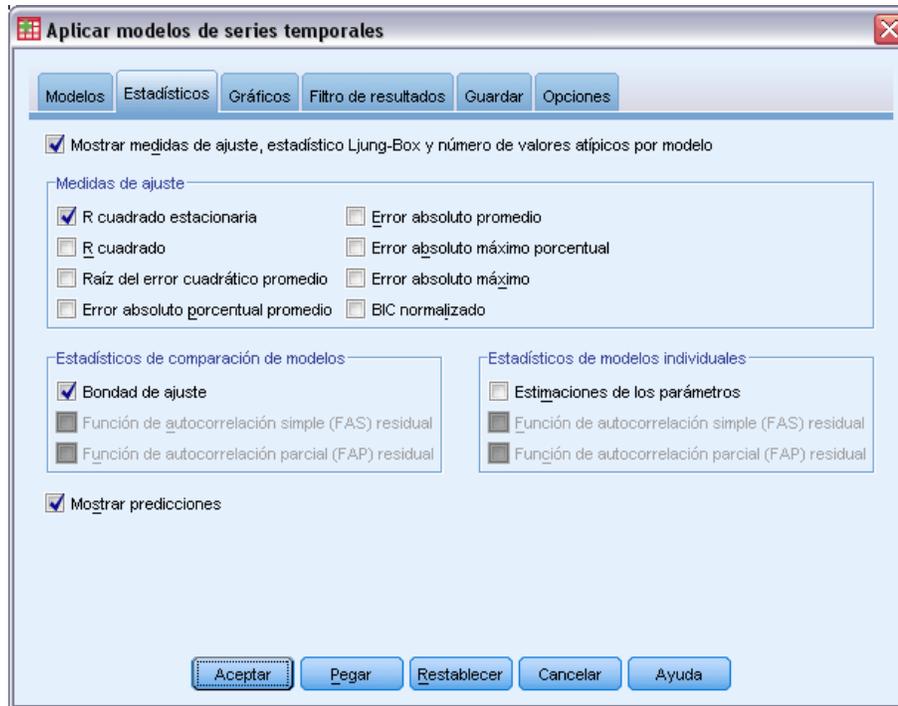
- ▶ Elija en los menús:  
Analizar > Predicciones > Aplicar modelos...

Figura 9-8  
Cuadro de diálogo Aplicar modelos de series temporales



- ▶ Pulse en Examinar, desplácese al archivo *catalog\_model.xml* y selecciónelo (o elija su propio archivo de modelo guardado del ejemplo anterior). [Si desea obtener más información, consulte el tema Archivos muestrales en el apéndice D en IBM SPSS Forecasting 19.](#)
- ▶ En el grupo Período de predicción, seleccione Primer caso después del final del período de estimación hasta una fecha especificada.
- ▶ En la cuadrícula Fecha, escriba 1999 para el año y 3 para el mes.
- ▶ Pulse en la pestaña Estadísticos.

Figura 9-9  
Aplicar modelos de series temporales, pestaña Estadísticos



- Seleccione Mostrar predicciones.

Se produce como resultado una tabla de valores de predicción para la variable dependiente.

- Pulse en Aceptar en el cuadro de diálogo Aplicar modelos de series temporales.

Figura 9-10  
Tabla de predicciones

Modelo		ENE 1999	FEB 1999	MAR 1999
Ventas de prendas para hombre-Modelo_1	Previsión	25279,91	22064,72	21580,96
	LCS	27591,62	24376,42	23892,66
	LCI	22968,21	19753,02	19269,25

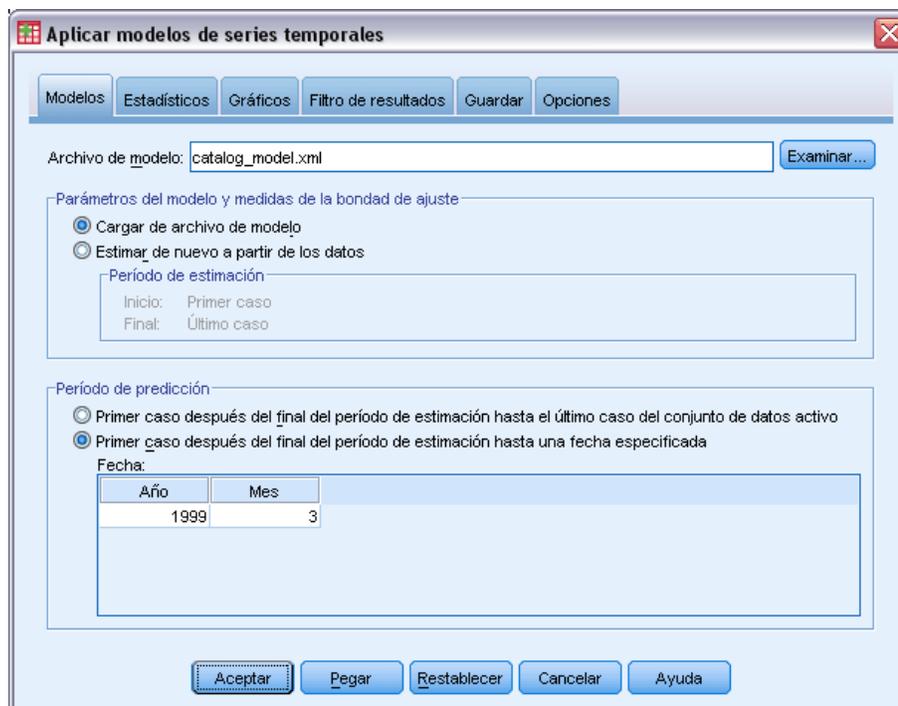
La tabla de predicciones contiene los valores pronosticados de la serie dependiente, teniendo en cuenta los valores de los dos predictores *correo* y *teléfono* del período de predicción. La tabla también incluye el límite de confianza superior (LCS) y el límite de confianza inferior (LCI) para las predicciones.

Ha generado la predicción de ventas para el escenario de enviar por correo 2.000 catálogos más cada mes. Ahora prepararemos los datos para el escenario de aumentar el número de líneas telefónicas, lo que supone restablecer la variable *correo* a los valores originales y aumentar la variable *teléfono* en 5. Puede restablecer *correo* copiando los valores de *Pronosticado\_correo\_Modelo\_1* en el período de predicción y pegándolos sobre los valores actuales de *correo* en el período de predicción. Además, puede aumentar el número de líneas telefónicas, 5 más para cada mes del período de predicción, directamente en el Editor de datos o mediante el cuadro de diálogo Calcular variable, como se hizo con el número de catálogos.

Para ejecutar el análisis, vuelva a abrir el cuadro de diálogo Aplicar modelos de series temporales como se indica a continuación:

- ▶ Pulse en el botón de la barra de herramientas Recuperar cuadros de diálogo.
- ▶ Seleccione Aplicar modelos de series temporales.

Figura 9-11  
Cuadro de diálogo Aplicar modelos de series temporales



- ▶ Pulse en Aceptar en el cuadro de diálogo Aplicar modelos de series temporales.

Figura 9-12  
Tablas de predicciones para los dos escenarios

#### Predicción con más catálogos

Modelo		ENE 1999	FEB 1999	MAR 1999
Ventas de prendas para hombre-Modelo_1	Previsión	25279,91	22064,72	21580,96
	LCS	27591,62	24376,42	23892,66
	LCI	22968,21	19753,02	19269,25

#### Predicción con más líneas telefónicas

Modelo		ENE 1999	FEB 1999	MAR 1999
Ventas de prendas para hombre-Modelo_1	Previsión	23757,25	20542,06	20058,29
	LCS	26068,95	22853,76	22370,00
	LCI	21445,55	18230,35	17746,59

Al mostrar las tablas de predicciones para ambos escenarios se desprende que, en cada uno de los tres meses pronosticados, al aumentar el número de catálogos enviados, se espera que se generen aproximadamente 1.500€ más en ventas que al aumentar el número de líneas telefónicas que se

abren para los pedidos. Según el análisis, parece más acertado asignar los recursos al envío de 2.000 catálogos adicionales.

# ***Descomposición estacional***

## ***Eliminación de la estacionalidad de los datos de ventas***

Una compañía de venta por catálogo está interesada en crear un modelo de la tendencia ascendente de las ventas en su línea de prendas para hombre en un conjunto de variables predictoras (como el número de catálogos enviados por correo y el número de líneas telefónicas abiertas para realizar pedidos). Para ello, la compañía ha recopilado las ventas mensuales de prendas para hombre durante un período de 10 años. Esta información se recoge en el archivo *catalog.sav*. [Si desea obtener más información, consulte el tema Archivos muestrales en el apéndice D en IBM SPSS Forecasting 19.](#)

Para realizar un análisis de tendencias, debe eliminar las variaciones estacionales presentes en los datos. Esta tarea se realiza fácilmente con el procedimiento Descomposición estacional.

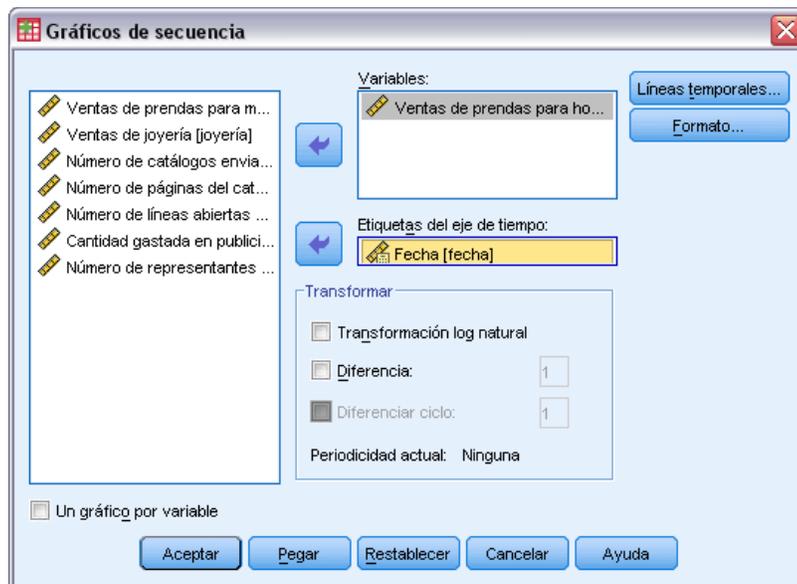
## ***Determinación y configuración de la periodicidad***

El procedimiento Descomposición estacional necesita la presencia de un componente de fecha periódico en el conjunto de datos activo; por ejemplo, una periodicidad anual de 12 (meses), una periodicidad semanal de 7 (días), y así sucesivamente. Es buena idea representar primero la serie temporal, ya que un gráfico de serie temporal normalmente ofrece una idea razonablemente buena acerca de la periodicidad subyacente.

Para obtener un gráfico de las ventas de prendas para hombre a lo largo del tiempo:

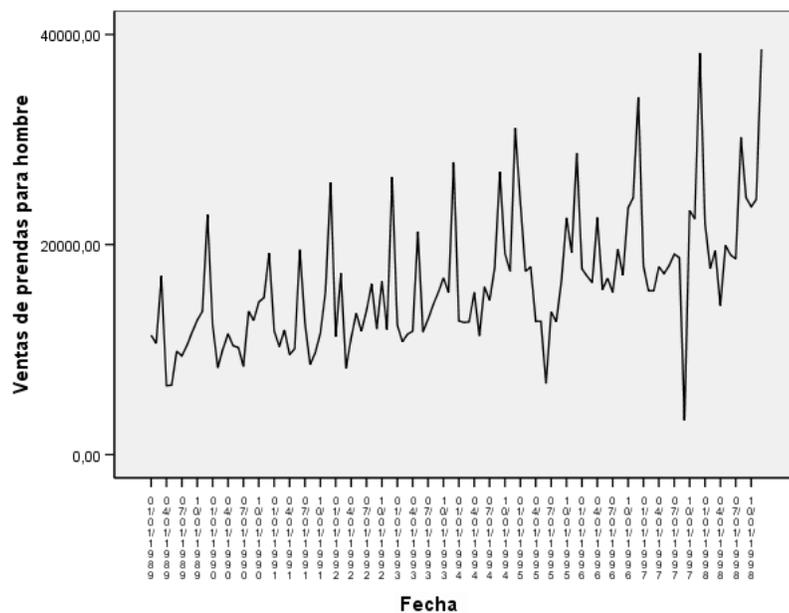
- ▶ Seleccione en los menús:  
Analizar > Predicciones > Gráficos de secuencia...

Figura 10-1  
Cuadro de diálogo Gráficos de secuencia



- ▶ Seleccione *Ventas de prendas para hombre* y muévelo a la lista Variables.
- ▶ Seleccione *Fecha* y muévela a la lista Etiquetas del eje de tiempo.
- ▶ Pulse en Aceptar.

Figura 10-2  
*Ventas de prendas para hombre (en dólares estadounidenses)*

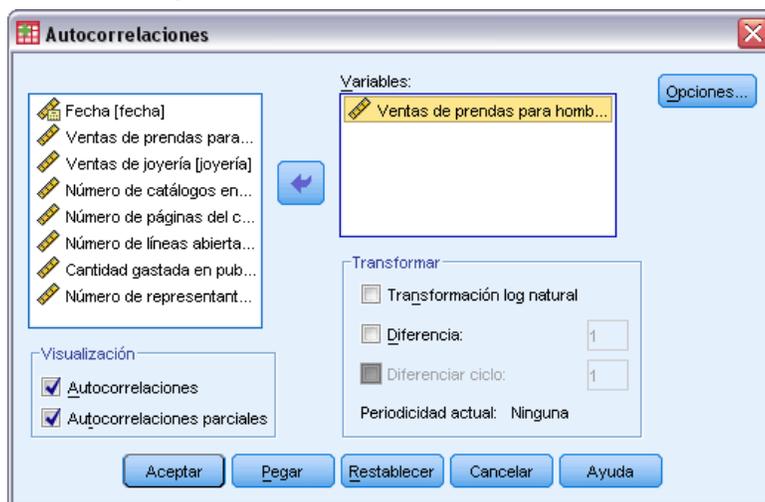


La serie muestra un número de picos, pero no parecen estar espaciados uniformemente. Este resultado sugiere que si la serie tiene un componente periódico, también dispone de fluctuaciones que no son periódicas, el caso típico de una serie temporal real. Aparte de las pequeñas fluctuaciones de escala, los picos significativos parecen estar separados por más de unos pocos meses. Dada la naturaleza estacional de las ventas, con máximos típicos durante la temporada estacional de diciembre, es probable que la serie temporal tenga una periodicidad anual. Observe también que las variaciones estacionales parecen crecer con la tendencia ascendente de la serie, sugiriendo que las variaciones estacionales pueden ser proporcionales al nivel de la serie, lo cual implica un modelo multiplicativo en lugar de uno aditivo.

Examinar las autocorrelaciones y autocorrelaciones parciales de una serie temporal proporciona una conclusión más cuantitativa acerca de la periodicidad subyacente.

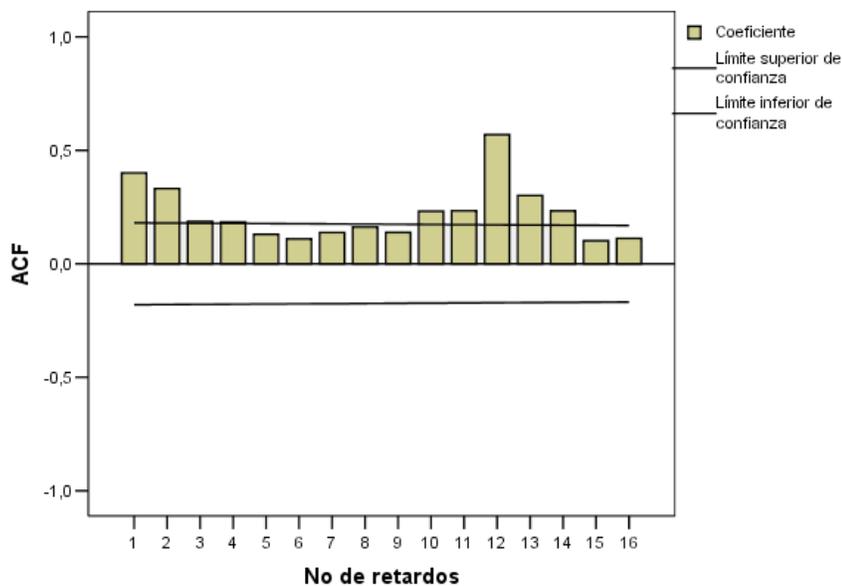
- Seleccione en los menús:  
Analizar > Predicciones > Autocorrelaciones...

Figura 10-3  
Cuadro de diálogo Autocorrelaciones



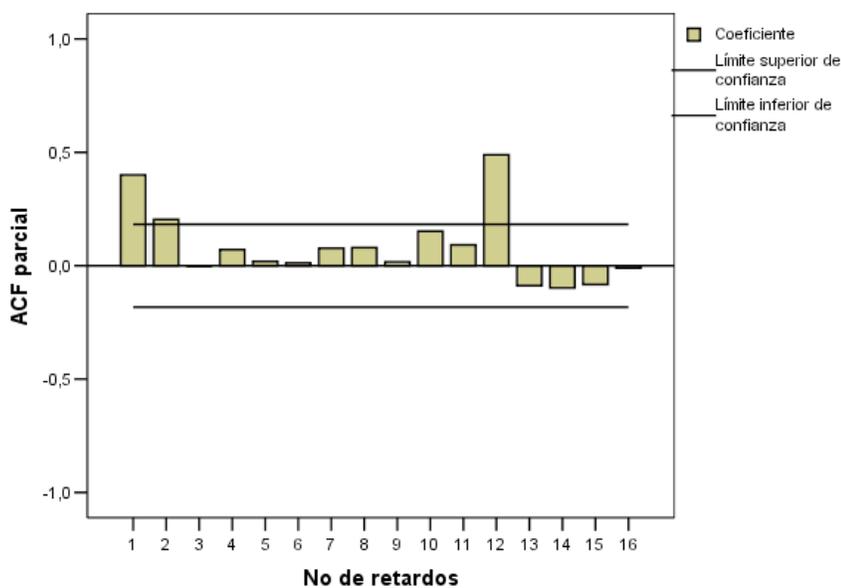
- Seleccione *Ventas de prendas para hombre* y muévelo a la lista Variables.
- Pulse en Aceptar.

Figura 10-4  
Gráfico de autocorrelación de prendas para hombre



La función de autocorrelación muestra un pico significativo en el retardo 1 con una cola exponencial larga, un patrón típico de las series temporales. El pico significativo en el retardo 12 sugiere la presencia de un componente estacional anual en los datos. El examen de la función de autocorrelación parcial permitirá obtener una conclusión más definitiva.

Figura 10-5  
Gráfico de autocorrelación parcial para hombres



El pico significativo en el retardo 12 en la función de autocorrelación parcial confirma la presencia de un componente estacional anual en los datos.

Para establecer la periodicidad anual:

- ▶ Seleccione en los menús:  
Datos > Definir fechas...

Figura 10-6  
Cuadro de diálogo Definir fechas



- ▶ Seleccione Años, meses en la lista Los casos son.
- ▶ Introduzca 1989 para el año y 1 para el mes.
- ▶ Pulse en Aceptar.

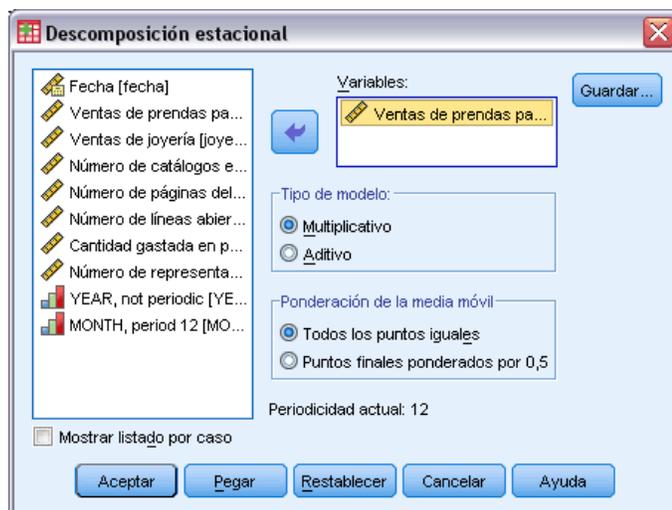
Esto establece la periodicidad en 12 y crea un conjunto de variables de fecha que están diseñados para trabajar con los procedimientos Predicción.

## ***Ejecución del análisis***

Para ejecutar el procedimiento Descomposición estacional:

- ▶ Seleccione en los menús:  
Analizar > Predicciones > Descomposición estacional...

Figura 10-7  
Cuadro de diálogo *Descomposición estacional*



- ▶ Pulse con el botón derecho del ratón en cualquier parte de la lista de variables de origen y seleccione en el menú contextual *Mostrar nombres de variable*.
- ▶ Seleccione *hombre* y muévela a la lista *Variables*.
- ▶ Seleccione *Multiplicativo* en el grupo *Tipo de modelo*.
- ▶ Pulse en *Aceptar*.

### **Descripción del resultado**

El procedimiento *Descomposición estacional* crea cuatro variables nuevas para cada una de las variables originales analizadas por el procedimiento. Las variables nuevas se añaden al conjunto de datos activo por defecto. Las nuevas series tienen nombres que comienzan con los siguientes prefijos:

**SAF.** Factores de corrección estacionales, que representan la variación estacional. En el caso del modelo multiplicativo, el valor 1 representa la ausencia de variación estacional; mientras que en el caso del modelo aditivo, es el valor 0 el que representa la ausencia de variación estacional.

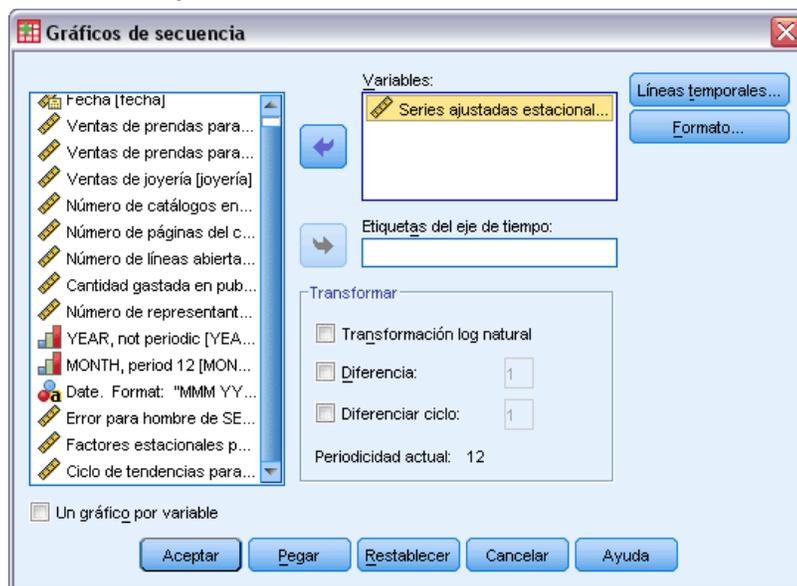
**SAS.** Serie corregida estacionalmente, representa la serie original con las variaciones estacionales eliminadas. Trabajar con una serie corregida estacionalmente, por ejemplo, permite aislar y analizar un componente de tendencia de forma independiente a cualquier componente estacional.

**STC.** Componente de tendencia y ciclo suavizado, que es una versión suavizada de la serie corregida estacionalmente que muestra los componentes de tendencia y ciclo.

**ERR.** El componente residual de la serie para una observación particular.

Para el caso actual, la serie corregida estacionalmente es la más apropiada, ya que representa la serie original con las variaciones estacionales eliminadas.

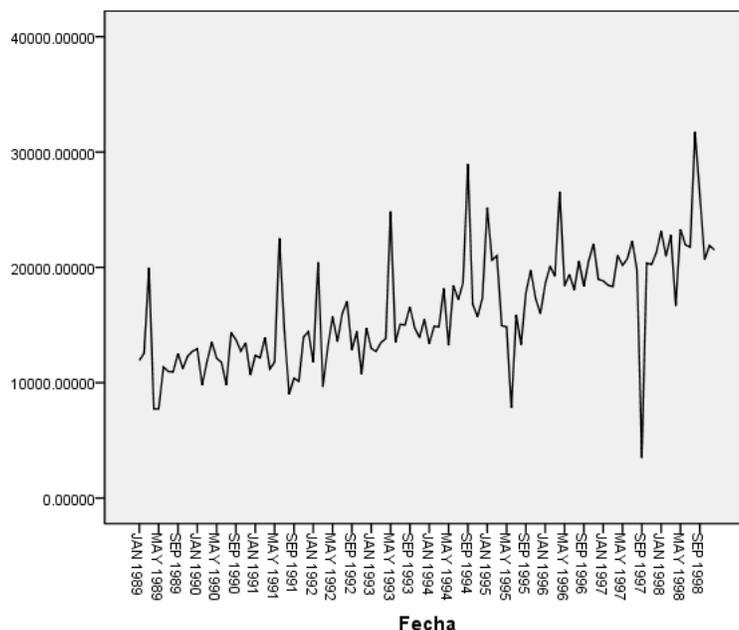
Figura 10-8  
Cuadro de diálogo Gráficos de secuencia



Para representar la serie corregida estacionalmente:

- ▶ Abra el cuadro de diálogo Gráficos de secuencia.
- ▶ Pulse en Restablecer para borrar cualquier selección anterior.
- ▶ Pulse con el botón derecho del ratón en cualquier parte de la lista de variables de origen y seleccione en el menú contextual Mostrar nombres de variable.
- ▶ Seleccione *SAS\_I* y muévala a lista Variables.
- ▶ Pulse en Aceptar.

Figura 10-9  
Serie corregida estacionalmente



La serie corregida estacionalmente muestra una clara tendencia ascendente. Se hacen patentes un número de picos, pero aparecen a intervalos aleatorios y no muestran evidencia de la existencia de un patrón anual.

## Resumen

Mediante el procedimiento Descomposición estacional, se han eliminado los componentes estacionales de una serie temporal periódica para generar una serie que sea más adecuada para el análisis de la tendencia. El examen de las autocorrelaciones y de las autocorrelaciones parciales de la serie temporal sirvió para determinar la periodicidad subyacente, en este caso, anual.

## Procedimientos relacionados

El procedimiento Descomposición estacional es útil para eliminar un único componente estacional de una serie temporal periódica.

- Para realizar un análisis detallado de la periodicidad de una serie temporal, que proporciona la función de correlación parcial, utilice el procedimiento Diagramas espectrales. Si desea obtener más información, consulte [el capítulo 11](#).

# ***Diagramas espectrales***

## ***Uso de gráficos espectrales para verificar expectativas de periodicidad***

La serie temporal que representa las ventas al por menor tienen, generalmente, una periodicidad anual subyacente debido al pico normal en las ventas durante la temporada vacacional. La creación de proyecciones de ventas implica construir un modelo de serie temporal, lo que significa que es necesario identificar todos los componentes periódicos. Puede que un gráfico de la serie temporal no descubra la periodicidad anual porque las series temporales contienen fluctuaciones aleatorias que suelen ocultar la estructura subyacente.

Los datos de ventas mensuales para una compañía de venta por catálogo están almacenados en *catalog.sav*. [Si desea obtener más información, consulte el tema Archivos muestrales en el apéndice D en IBM SPSS Forecasting 19](#). Antes de continuar con las proyecciones de ventas, desea confirmar que los datos de ventas muestran una periodicidad anual. Un gráfico de la serie temporal muestra que tiene muchos picos y que está espaciada de forma irregular, por lo que cualquier periodicidad subyacente no es evidente. Utilice el procedimiento Gráficos espectrales para identificar cualquier periodicidad que haya en los datos de venta.

### ***Ejecución del análisis***

Para ejecutar el procedimiento Gráficos espectrales:

- ▶ Elija en los menús:  
Analizar > Predicciones > Análisis espectral...

Figura 11-1  
Cuadro de diálogo Diagramas espectrales



- ▶ Seleccione *Ventas de prendas para hombre* y muévelo a la lista Variables.
- ▶ Seleccione Densidad espectral en el grupo Gráficos.
- ▶ Pulse en Aceptar.

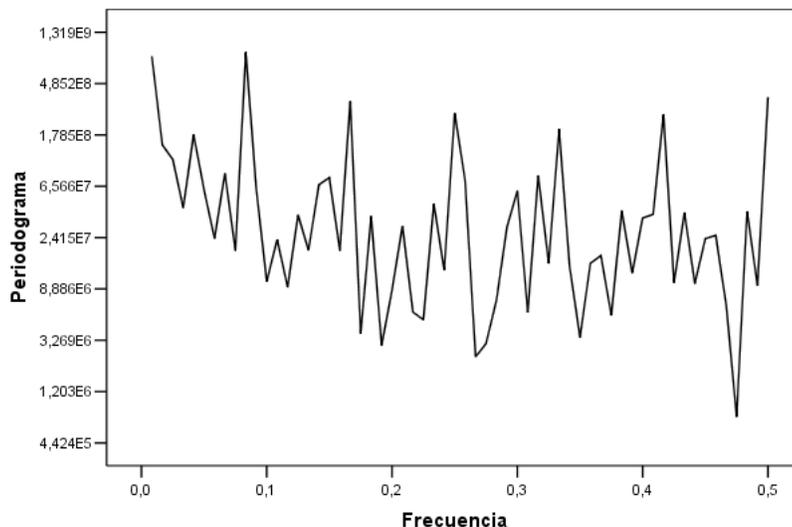
Estas selecciones generan la siguiente sintaxis de comandos:

```
* Spectral Analysis.
TSET PRINT=DEFAULT.
SPECTRA
  /VARIABLES=men
  /WINDOW=HAMMING (5)
  /CENTER
  /PLOT=PS BY FREQUENCY.
```

Tenga en cuenta que para obtener la tabla de estadísticas univariadas en los resultados, debe cambiar el comando TSET para leer TSET PRINT=DETAILED.

## Descripción del periodograma y la densidad espectral

Figura 11-2  
Periodograma



El gráfico del periodograma muestra una secuencia de picos que sobresalen del ruido de fondo, con el pico de frecuencia más bajo a una frecuencia un poco menor que 0,1. Si sospecha que los datos contienen un componente periódico anual, considere la contribución que un componente anual haría al periodograma. Cada uno de los puntos de datos de la serie temporal representa un mes, por lo que la periodicidad anual se corresponde con un período de 12 en el conjunto de datos actual. Como el período y la frecuencia son recíprocos entre sí, un período de 12 se corresponde con una frecuencia de  $1/12$  (o 0,083). Por lo tanto, un componente anual implica la aparición de un pico en el periodograma de 0,083, que parece coherente con la presencia del pico justo debajo de la frecuencia de 0,1.

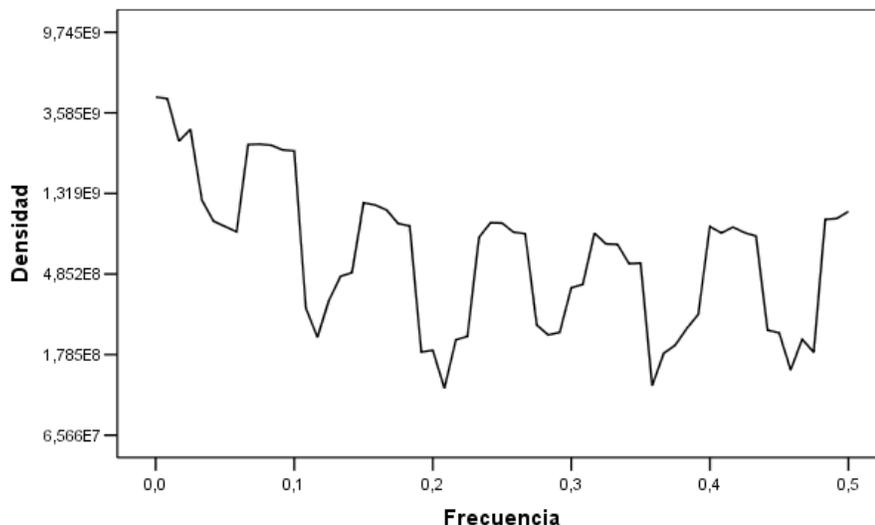
Figura 11-3  
Tabla de estadísticos univariados

Nombre de la serie: hombres

	Frecuencia	Transformación de seno	Transformación de coseno	Periodograma	Estimación de densidad espectral
1	,00000	,000	16242,813	,000	4,361E+09
2	,00833	-3696,643	370,153	828131182	4,278E+09
3	,01667	-1496,277	454,831	146743038	2,526E+09
4	,02500	-1336,400	252,087	110970821	2,921E+09
5	,03333	-662,146	529,734	43143315,6	1,210E+09
6	,04167	-1654,614	501,571	179359225	935824937
7	,05000	-784,814	-636,729	61281367,3	875492029
8	,05833	-335,646	532,062	23744855,0	820375352
9	,06667	-1094,178	-451,489	84064108,9	2,420E+09
10	,07500	264,554	492,876	18774933,5	2,429E+09
11	,08333	-3053,934	2370,483	896742149	2,401E+09
12	,09167	-978,882	-287,035	62435897,6	2,263E+09
13	,10000	-403,128	93,036	10270064,4	2,235E+09

La tabla de estadísticos univariados contiene los puntos de datos que se utilizan para representar el periodograma. Observe que para las frecuencias menores que 0,1, el mayor valor en la columna *Periodograma* aparece con una frecuencia de 0,08333, exactamente lo que esperaba encontrar si hubiera un componente periódico anual. Esta información confirma la identificación del pico de frecuencia menor con un componente periódico anual. Pero, ¿qué sucede con los demás picos a frecuencias mayores?

Figura 11-4  
*Densidad espectral*



Los picos restantes se analizan mejor con la función de densidad espectral, que es simplemente una versión suavizada del periodograma. El suavizado proporciona un medio de eliminar el ruido de fondo de un periodograma, lo que permite aislar más fácilmente la estructura subyacente.

La densidad espectral consta de cinco picos bien diferenciados que parecen estar uniformemente espaciados. El pico de frecuencia menor representa simplemente la versión suavizada del pico en 0,08333. Para comprender la significación de los cuatro picos de frecuencia mayores, recuerde que el periodograma se calcula creando un modelo de serie temporal como la suma de las funciones seno y coseno. Los componentes periódicos que tienen la forma de una función seno o coseno (sinusoidal) aparecen en el periodograma como picos únicos. Los componentes periódicos no sinusoidales aparecen como una serie de picos de diferentes alturas uniformemente espaciados, con el pico de frecuencia menor de la serie en la frecuencia de un componente periódico. Por lo tanto, los cuatro picos de frecuencia mayores de la densidad espectral simplemente indican que el componente periódico anual no es sinusoidal.

Ya se han explicado todas las partes de la estructura perceptible en el gráfico de densidad espectral y se ha llegado a la conclusión de que los datos contienen un único componente periódico con un período de 12 meses.

## **Resumen**

Mediante el procedimiento Gráficos espectrales, ha confirmado la existencia de un componente periódico anual de una serie temporal y ha comprobado que no existen más periodicidades significativas. Se ha demostrado que la densidad espectral es más útil que el periodograma para descubrir toda la estructura subyacente porque la densidad espectral suaviza las fluctuaciones causadas por el componente no periódico de los datos.

## **Procedimientos relacionados**

El procedimiento Gráficos espectrales es útil para identificar los componentes periódicos de una serie temporal.

- Para eliminar un componente periódico de una serie temporal ((por ejemplo, para realizar un análisis de tendencia), utilice el procedimiento descomposición estacional. Consulte [el capítulo 10](#) para obtener información.

## ***Medidas de la bondad de ajuste***

Esta sección proporciona definiciones de las medidas de bondad de ajuste utilizadas en el modelado de series temporales.

- **R-cuadrado estacionaria.** Una medida que compara la parte estacionaria del modelo con un modelo de promedio simple. Esta medida es preferible al R-cuadrado ordinario cuando existe tendencia o patrón estacional. R-cuadrado puede ser negativa con un rango desde menos infinito hasta 1. Los valores negativos significan que el modelo estudiado es peor que el modelo basal. Los valores positivos significan que el modelo estudiado es mejor que el modelo basal.
- **R cuadrado.** Una estimación de la proporción de la varianza total en la serie que explica el modelo. Esta medida es más útil cuando la serie es estacionaria. R-cuadrado puede ser negativa con un rango desde menos infinito hasta 1. Los valores negativos significan que el modelo estudiado es peor que el modelo basal. Los valores positivos significan que el modelo estudiado es mejor que el modelo basal.
- **RMSE.** Raíz del error cuadrático promedio. La raíz cuadrada del error cuadrático promedio. Una medida de cuánto se desvía la serie dependiente del nivel pronosticado por el modelo, expresado en las mismas unidades que la serie dependiente.
- **MAPE.** Error absoluto porcentual promedio. Medida de la desviación de la serie dependiente del nivel pronosticado por el modelo. Es independiente de las unidades utilizadas y se puede utilizar para comparar series con distintas unidades.
- **MAE.** Error absoluto promedio. Mide la desviación de la serie del nivel pronosticado por el modelo. El MAE se informa en las unidades originales de la serie.
- **MaxAPE.** Error absoluto máximo porcentual. El mayor error pronosticado, expresado como porcentaje. Esta medida es útil para imaginar el peor escenario de un caso en las predicciones.
- **MaxAE.** Error absoluto máximo. El mayor error pronosticado, expresado en las mismas unidades que la variable dependiente. Al igual que el MaxAPE, es útil para imaginar el peor escenario de los casos en la predicción. El error absoluto máximo y el error absoluto máximo porcentual pueden darse en distintos puntos de la serie. Por ejemplo, si el error absoluto de un valor de la serie grande es ligeramente mayor que el error absoluto de un valor de la serie pequeño. En ese caso el error absoluto máximo se obtendrá en el valor de la serie mayor y el error absoluto máximo porcentual corresponderá al valor de la serie menor.
- **BIC normalizado.** Criterio de información Bayesiano normalizado. Una medida general del ajuste global del modelo que intenta tener en cuenta la complejidad del modelo. Es una medida basada en el error cuadrático promedio que incluye una penalización para el número de parámetros presentes en el modelo y la longitud de la serie. La penalización elimina la ventaja de los modelos con mayor número de parámetros, haciendo que el estadístico sea fácil de comparar entre distintos modelos para la misma serie.

## ***Tipos de valores atípicos***

Esta sección proporciona definiciones de los tipos de valores atípicos utilizados en el modelado de series temporales.

- **Aditivo.** Un valor atípico que afecta a una sola observación. Por ejemplo, un error de codificación de los datos puede identificarse como un valor atípico aditivo.
- **Cambio de nivel.** Un valor atípico que desplaza todas las observaciones mediante una constante, comenzando en un punto concreto de la serie. Un cambio de nivel puede ser el resultado de un cambio de política.
- **De Innovación.** Un valor atípico que actúa como adición al término error en un punto particular de la serie. En las series estacionarias, un valor extremo de innovación afecta a varias observaciones. En las series no estacionarias, puede afectar a cada observación a partir de un punto particular de inicio en la serie.
- **Transitorio.** Un valor atípico cuyo impacto decae exponencialmente hacia cero.
- **Aditivo estacional.** Un valor atípico que afecta a una observación particular y a todas las observaciones siguientes separadas de ella por uno o más períodos estacionales. Todas las observaciones afectadas lo son de igual forma. Un valor atípico estacional puede ocurrir si, a partir de cierto año, las ventas son mayores cada enero.
- **Tendencia local.** Un valor atípico que da inicio a una tendencia local en un punto particular de la serie.
- **Parche aditivo.** Un grupo de dos o más valores atípicos aditivos consecutivos. La selección de este tipo de valores atípicos tiene como resultado la detección de los valores atípicos aditivos individuales en adición a las rachas compuestas por los mismos.



# ***Guía de los gráficos de la función de autocorrelación simple (FAS) y parcial (FAP)***

Los gráficos que se muestran aquí son aquéllos de procesos ARIMA puros o teóricos. Aquí se describen algunas normas generales para identificar el proceso:

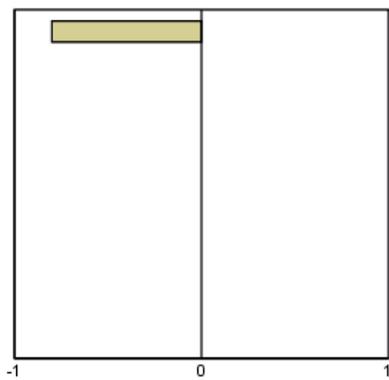
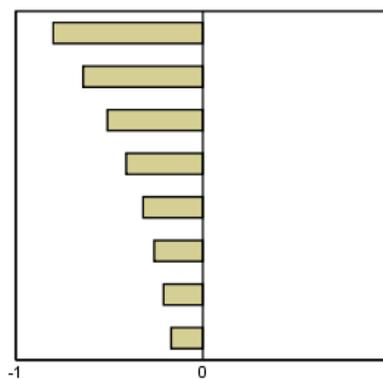
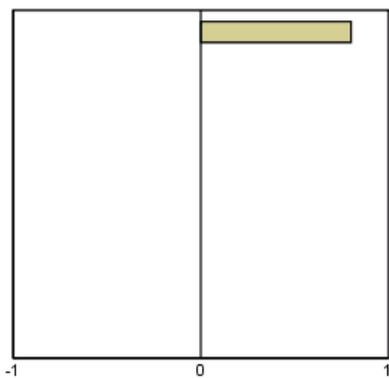
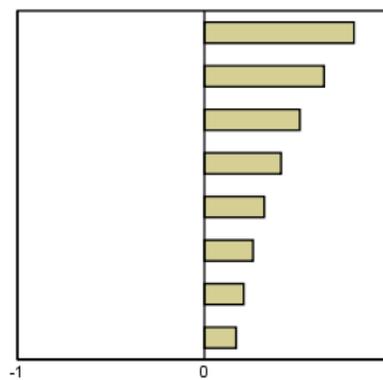
- Las series no estacionarias tienen una FAS que permanece significativa durante seis retrasos o más, en lugar de disminuir rápidamente a 0. Debe diferenciar dicha serie hasta que sea estacionaria antes de poder identificar el proceso.
- Los procesos autorregresivos tienen una FAS que disminuye exponencialmente y trazos de unión en el primer retardo o en más de la FAP. El número de trazos de unión indica el orden de la autorregresión.
- Los procesos de medias móviles tienen trazos de unión en el primer retardo o en más de la FAS y una FAP que disminuye exponencialmente. El número de trazos de unión indica el orden de la media móvil.
- Los procesos mixtos (ARMA) generalmente muestran las disminuciones exponenciales en las FAS y en las FAP.

En la etapa de identificación no es necesario preocuparse del signo de la FAS o FAP, o de la velocidad con la que se aproxima a 0 una FAS o FAP que disminuye exponencialmente. Estos valores dependen del signo y del valor real de los coeficientes AR y MA. En algunos casos, una FAS que disminuye exponencialmente alterna entre valores positivos y negativos.

*Los gráficos de FAS y FAP obtenidos a partir de datos reales no serán nunca tan claros como los gráficos que aquí se muestran. Debe aprender a escoger las partes esenciales de un gráfico. Compruebe siempre la FAS y la FAP de los residuos, en caso de que la interpretación sea errónea. Recuerde que:*

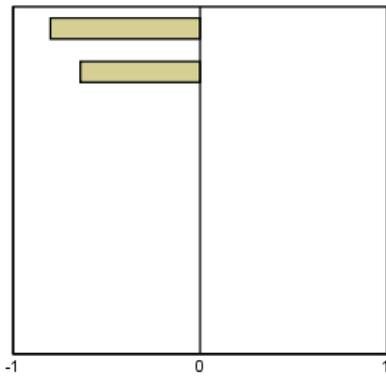
- Los procesos estacionales muestran estos patrones en los retardos estacionales (los múltiplos del período estacional).
- Si lo desea, puede tratar los valores no significativos como 0. Es decir, puede ignorar los valores contenidos en los intervalos de confianza en los gráficos. Sin embargo, no es necesario ignorarlos, especialmente si siguen el patrón de los valores estadísticamente significativos.
- Una autocorrelación ocasional será significativa estadísticamente sólo por casualidad. Puede ignorar una autocorrelación estadísticamente significativa si está aislada, preferiblemente en un retardo mayor, y si no se produce en un retardo estacional.

Consulte los textos que aparecen en los análisis ARIMA para obtener información más detallada de los gráficos de FAS y FAP.

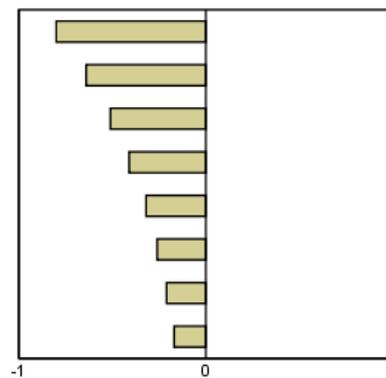
**ARIMA(0,0,1),  $\theta > 0$** **FAS****FAP****ARIMA(0,0,1),  $\theta < 0$** **FAS****FAP**

**ARIMA(0,0,2),  $\theta_1\theta_2 > 0$**

**FAS**

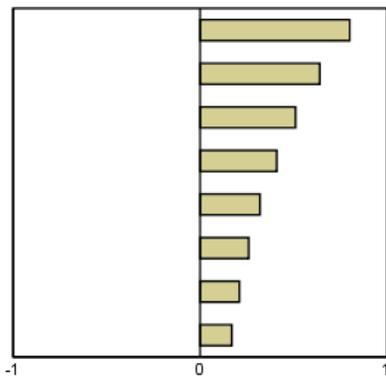


**FAP**

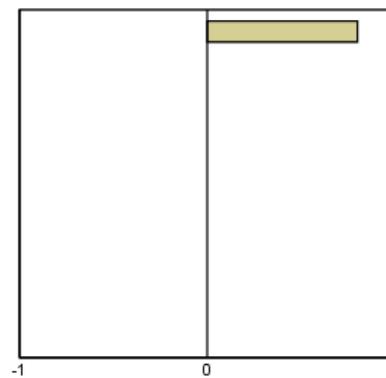


**ARIMA(1,0,0),  $\phi > 0$**

**FAS**

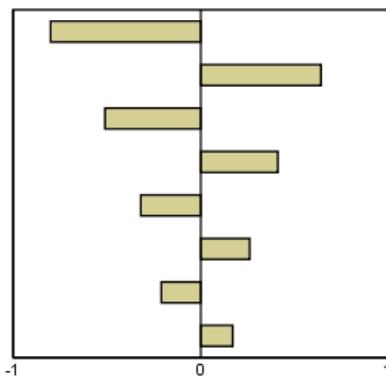


**FAP**

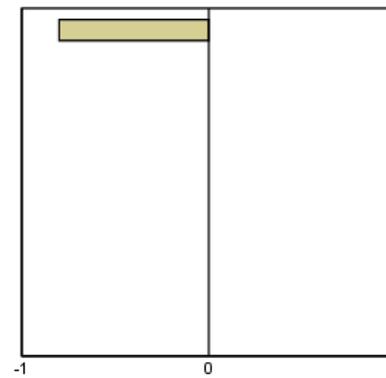


**ARIMA(1,0,0),  $\phi < 0$**

**FAS**

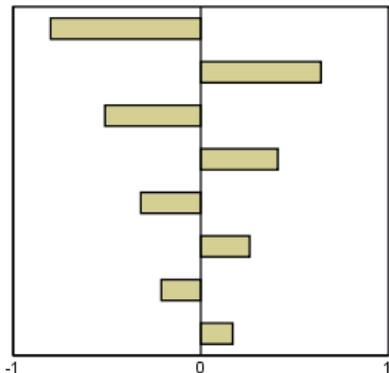


**FAP**

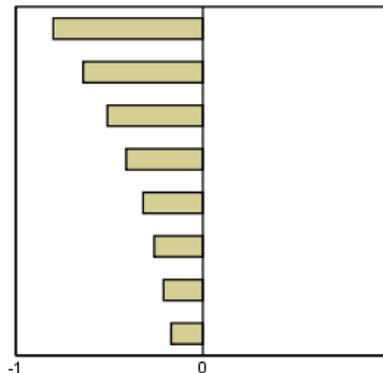


**ARIMA(1,0,1),  $\phi < 0, \theta > 0$**

**FAS**

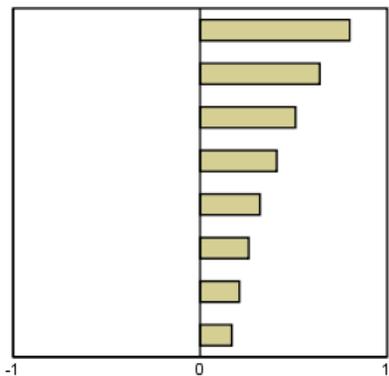


**FAP**

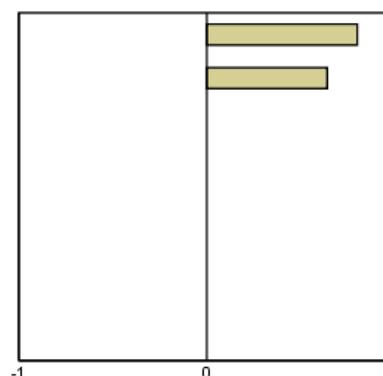


**ARIMA(2,0,0),  $\phi_1\phi_2 > 0$**

**FAS**

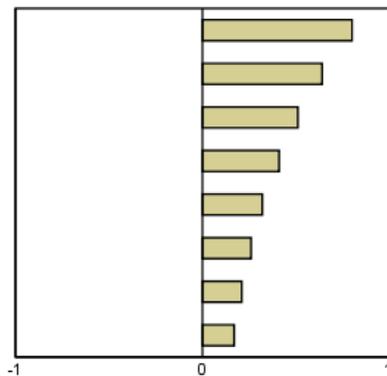


**FAP**



**ARIMA(0,1,0) (serie integrada)**

**FAS**



# Archivos muestrales

Los archivos muestrales instalados con el producto se encuentran en el subdirectorio *Samples* del directorio de instalación. Hay una carpeta independiente dentro del subdirectorio *Samples* para cada uno de los siguientes idiomas: Inglés, francés, alemán, italiano, japonés, coreano, polaco, ruso, chino simplificado, español y chino tradicional.

No todos los archivos muestrales están disponibles en todos los idiomas. Si un archivo muestral no está disponible en un idioma, esa carpeta de idioma contendrá una versión en inglés del archivo muestral.

## Descripciones

A continuación, se describen brevemente los archivos muestrales usados en varios ejemplos que aparecen a lo largo de la documentación.

- **accidents.sav.** Archivo de datos hipotéticos sobre una compañía de seguros que estudia los factores de riesgo de edad y género que influyen en los accidentes de automóviles de una región determinada. Cada caso corresponde a una clasificación cruzada de categoría de edad y género.
- **adl.sav.** Archivo de datos hipotéticos relativo a los esfuerzos para determinar las ventajas de un tipo propuesto de tratamiento para pacientes que han sufrido un derrame cerebral. Los médicos dividieron de manera aleatoria a pacientes (mujeres) que habían sufrido un derrame cerebral en dos grupos. El primer grupo recibió el tratamiento físico estándar y el segundo recibió un tratamiento emocional adicional. Tres meses después de los tratamientos, se puntuaron las capacidades de cada paciente para realizar actividades cotidianas como variables ordinales.
- **advert.sav.** Archivo de datos hipotéticos sobre las iniciativas de un minorista para examinar la relación entre el dinero invertido en publicidad y las ventas resultantes. Para ello, se recopilaron las cifras de ventas anteriores y los costes de publicidad asociados.
- **aflatoxin.sav.** Archivo de datos hipotéticos sobre las pruebas realizadas en las cosechas de maíz con relación a la aflatoxina, un veneno cuya concentración varía ampliamente en los rendimientos de cultivo y entre los mismos. Un procesador de grano ha recibido 16 muestras de cada uno de los 8 rendimientos de cultivo y ha medido los niveles de aflatoxinas en partes por millón (PPM).
- **aflatoxin20.sav.** Este archivo de datos contiene las medidas de aflatoxina de cada una de las 16 muestras de los rendimientos 4 y 8 procedentes del archivo de datos *aflatoxin.sav*.
- **anorectic.sav.** Mientras trabajaban en una sintomatología estandarizada del comportamiento anoréxico/bulímico, los investigadores (Van der Ham, Meulman, Van Strien, y Van Engeland, 1997) realizaron un estudio de 55 adolescentes con trastornos de la alimentación conocidos. Cada paciente fue examinado cuatro veces durante cuatro años, lo que representa un total

de 220 observaciones. En cada observación, se puntuó a los pacientes por cada uno de los 16 síntomas. Faltan las puntuaciones de los síntomas para el paciente 71 en el tiempo 2, el paciente 76 en el tiempo 2 y el paciente 47 en el tiempo 3, lo que nos deja 217 observaciones válidas.

- **autoaccidents.sav.** Archivo de datos hipotéticos sobre las iniciativas de un analista de seguros para elaborar un modelo del número de accidentes de automóvil por conductor teniendo en cuenta la edad y el género del conductor. Cada caso representa un conductor diferente y registra el sexo, la edad en años y el número de accidentes de automóvil del conductor en los últimos cinco años.
- **band.sav** Este archivo de datos contiene las cifras de ventas semanales hipotéticas de CD de música de una banda. También se incluyen datos para tres variables predictoras posibles.
- **bankloan.sav.** Archivo de datos hipotéticos sobre las iniciativas de un banco para reducir la tasa de moras de créditos. El archivo contiene información financiera y demográfica de 850 clientes anteriores y posibles clientes. Los primeros 700 casos son clientes a los que anteriormente se les ha concedido un préstamo. Al menos 150 casos son posibles clientes cuyos riesgos de crédito el banco necesita clasificar como positivos o negativos.
- **bankloan\_binning.sav.** Archivo de datos hipotéticos que contiene información financiera y demográfica sobre 5.000 clientes anteriores.
- **behavior.sav.** En un ejemplo clásico (Price y Bouffard, 1974), se pidió a 52 estudiantes que valoraran las combinaciones de 15 situaciones y 15 comportamientos en una escala de 10 puntos que oscilaba entre 0 = “extremadamente apropiado” y 9 = “extremadamente inapropiado”. Los valores promediados respecto a los individuos se toman como disimilaridades.
- **behavior\_ini.sav.** Este archivo de datos contiene una configuración inicial para una solución bidimensional de *behavior.sav*.
- **brakes.sav.** Archivo de datos hipotéticos sobre el control de calidad de una fábrica que produce frenos de disco para automóviles de alto rendimiento. El archivo de datos contiene las medidas del diámetro de 16 discos de cada una de las 8 máquinas de producción. El diámetro objetivo para los frenos es de 322 milímetros.
- **breakfast.sav.** En un estudio clásico (Green y Rao, 1972), se pidió a 21 estudiantes de administración de empresas de la Wharton School y sus cónyuges que ordenaran 15 elementos de desayuno por orden de preferencia, de 1 = “más preferido” a 15 = “menos preferido”. Sus preferencias se registraron en seis escenarios distintos, de “Preferencia global” a “Aperitivo, con bebida sólo”.
- **breakfast-overall.sav.** Este archivo de datos sólo contiene las preferencias de elementos de desayuno para el primer escenario, “Preferencia global”.
- **broadband\_1.sav** Archivo de datos hipotéticos que contiene el número de suscriptores, por región, a un servicio de banda ancha nacional. El archivo de datos contiene números de suscriptores mensuales para 85 regiones durante un período de cuatro años.
- **broadband\_2.sav** Este archivo de datos es idéntico a *broadband\_1.sav* pero contiene datos para tres meses adicionales.
- **car\_insurance\_claims.sav.** Un conjunto de datos presentados y analizados en otro lugar (McCullagh y Nelder, 1989) estudia las reclamaciones por daños en vehículos. La cantidad de reclamaciones media se puede modelar como si tuviera una distribución Gamma, mediante

una función de enlace inversa para relacionar la media de la variable dependiente con una combinación lineal de la edad del asegurado, el tipo de vehículo y la antigüedad del vehículo. El número de reclamaciones presentadas se puede utilizar como una ponderación de escalamiento.

- **car\_sales.sav.** Este archivo de datos contiene estimaciones de ventas, precios de lista y especificaciones físicas hipotéticas de varias marcas y modelos de vehículos. Los precios de lista y las especificaciones físicas se han obtenido de *edmunds.com* y de sitios de fabricantes.
- **car\_sales\_uprepared.sav.** Ésta es una versión modificada de *car\_sales.sav* que no incluye ninguna versión transformada de los campos.
- **carpet.sav** En un ejemplo muy conocido (Green y Wind, 1973), una compañía interesada en sacar al mercado un nuevo limpiador de alfombras desea examinar la influencia de cinco factores sobre la preferencia del consumidor: diseño del producto, marca comercial, precio, sello de *buen producto para el hogar* y garantía de devolución del importe. Hay tres niveles de factores para el diseño del producto, cada uno con una diferente colocación del cepillo del aplicador; tres nombres comerciales (*K2R*, *Glory* y *Bissell*); tres niveles de precios; y dos niveles (no o sí) para los dos últimos factores. Diez consumidores clasificaron 22 perfiles definidos por estos factores. La variable *Preferencia* contiene el rango de las clasificaciones medias de cada perfil. Las clasificaciones inferiores corresponden a preferencias elevadas. Esta variable refleja una medida global de la preferencia de cada perfil.
- **carpet\_prefs.sav** Este archivo de datos se basa en el mismo ejemplo que el descrito para *carpet.sav*, pero contiene las clasificaciones reales recogidas de cada uno de los 10 consumidores. Se pidió a los consumidores que clasificaran los 22 perfiles de los productos empezando por el menos preferido. Las variables desde *PREF1* hasta *PREF22* contienen los ID de los perfiles asociados, como se definen en *carpet\_plan.sav*.
- **catalog.sav** Este archivo de datos contiene cifras de ventas mensuales hipotéticas de tres productos vendidos por una compañía de venta por catálogo. También se incluyen datos para cinco variables predictoras posibles.
- **catalog\_seasfac.sav** Este archivo de datos es igual que *catalog.sav*, con la excepción de que incluye un conjunto de factores estacionales calculados a partir del procedimiento Descomposición estacional junto con las variables de fecha que lo acompañan.
- **cellular.sav.** Archivo de datos hipotéticos sobre las iniciativas de una compañía de telefonía móvil para reducir el abandono de clientes. Las puntuaciones de propensión al abandono de clientes se aplican a las cuentas, oscilando de 0 a 100. Las cuentas con una puntuación de 50 o superior pueden estar buscando otros proveedores.
- **ceramics.sav.** Archivo de datos hipotéticos sobre las iniciativas de un fabricante para determinar si una nueva aleación de calidad tiene una mayor resistencia al calor que una aleación estándar. Cada caso representa una prueba independiente de una de las aleaciones; la temperatura a la que registró el fallo del rodamiento.
- **cereal.sav.** Archivo de datos hipotéticos sobre una encuesta realizada a 880 personas sobre sus preferencias en el desayuno, teniendo también en cuenta su edad, sexo, estado civil y si tienen un estilo de vida activo o no (en función de si practican ejercicio al menos dos veces a la semana). Cada caso representa un encuestado diferente.
- **clothing\_defects.sav.** Archivo de datos hipotéticos sobre el proceso de control de calidad en una fábrica de prendas. Los inspectores toman una muestra de prendas de cada lote producido en la fábrica, y cuentan el número de prendas que no son aceptables.

- **coffee.sav.** Este archivo de datos pertenece a las imágenes percibidas de seis marcas de café helado (Kennedy, Riquier, y Sharp, 1996). Para cada uno de los 23 atributos de imagen de café helado, los encuestados seleccionaron todas las marcas que quedaban descritas por el atributo. Las seis marcas se denotan AA, BB, CC, DD, EE y FF para mantener la confidencialidad.
- **contacts.sav.** Archivo de datos hipotéticos sobre las listas de contactos de un grupo de representantes de ventas de ordenadores de empresa. Cada uno de los contactos está categorizado por el departamento de la compañía en el que trabaja y su categoría en la compañía. Además, también se registran los importes de la última venta realizada, el tiempo transcurrido desde la última venta y el tamaño de la compañía del contacto.
- **creditpromo.sav.** Archivo de datos hipotéticos sobre las iniciativas de unos almacenes para evaluar la eficacia de una promoción de tarjetas de crédito reciente. Para este fin, se seleccionaron aleatoriamente 500 titulares. La mitad recibieron un anuncio promocionando una tasa de interés reducida sobre las ventas realizadas en los siguientes tres meses. La otra mitad recibió un anuncio estacional estándar.
- **customer\_dbase.sav.** Archivo de datos hipotéticos sobre las iniciativas de una compañía para usar la información de su almacén de datos para realizar ofertas especiales a los clientes con más probabilidades de responder. Se seleccionó un subconjunto de la base de clientes aleatoriamente a quienes se ofrecieron las ofertas especiales y sus respuestas se registraron.
- **customer\_information.sav.** Archivo de datos hipotéticos que contiene la información de correo del cliente, como el nombre y la dirección.
- **customer\_subset.sav.** Un subconjunto de 80 casos de *customer\_dbase.sav*.
- **customers\_model.sav.** Este archivo contiene datos hipotéticos sobre los individuos a los que va dirigida una campaña de marketing. Estos datos incluyen información demográfica, un resumen del historial de compras y si cada individuo respondió a la campaña. Cada caso representa un individuo diferente.
- **customers\_new.sav.** Este archivo contiene datos hipotéticos sobre los individuos que son candidatos potenciales para una campaña de marketing. Estos datos incluyen información demográfica y un resumen del historial de compras de cada individuo. Cada caso representa un individuo diferente.
- **debate.sav.** Archivos de datos hipotéticos sobre las respuestas emparejadas de una encuesta realizada a los asistentes a un debate político antes y después del debate. Cada caso corresponde a un encuestado diferente.
- **debate\_aggregate.sav.** Archivo de datos hipotéticos que agrega las respuestas de *debate.sav*. Cada caso corresponde a una clasificación cruzada de preferencias antes y después del debate.
- **demo.sav.** Archivos de datos hipotéticos sobre una base de datos de clientes adquirida con el fin de enviar por correo ofertas mensuales. Se registra si el cliente respondió a la oferta, junto con información demográfica diversa.
- **demo\_cs\_1.sav.** Archivo de datos hipotéticos sobre el primer paso de las iniciativas de una compañía para recopilar una base de datos de información de encuestas. Cada caso corresponde a una ciudad diferente, y se registra la identificación de la ciudad, la región, la provincia y el distrito.
- **demo\_cs\_2.sav.** Archivo de datos hipotéticos sobre el segundo paso de las iniciativas de una compañía para recopilar una base de datos de información de encuestas. Cada caso corresponde a una unidad familiar diferente de las ciudades seleccionadas en el primer paso, y

se registra la identificación de la unidad, la subdivisión, la ciudad, el distrito, la provincia y la región. También se incluye la información de muestreo de las primeras dos etapas del diseño.

- **demo\_cs.sav.** Archivo de datos hipotéticos que contiene información de encuestas recopilada mediante un diseño de muestreo complejo. Cada caso corresponde a una unidad familiar distinta, y se recopila información demográfica y de muestreo diversa.
- **dmdata.sav.** Éste es un archivo de datos hipotéticos que contiene información demográfica y de compras para una empresa de marketing directo. *dmdata2.sav* contiene información para un subconjunto de contactos que recibió un envío de prueba, y *dmdata3.sav* contiene información sobre el resto de contactos que no recibieron el envío de prueba.
- **dietstudy.sav.** Este archivo de datos hipotéticos contiene los resultados de un estudio sobre la “dieta Stillman” (Rickman, Mitchell, Dingman, y Dalen, 1974). Cada caso corresponde a un sujeto distinto y registra sus pesos antes y después de la dieta en libras y niveles de triglicéridos en mg/100 ml.
- **dvdplayer.sav.** Archivo de datos hipotéticos sobre el desarrollo de un nuevo reproductor de DVD. El equipo de marketing ha recopilado datos de grupo de enfoque mediante un prototipo. Cada caso corresponde a un usuario encuestado diferente y registra información demográfica sobre los encuestados y sus respuestas a preguntas acerca del prototipo.
- **german\_credit.sav.** Este archivo de datos se toma del conjunto de datos “German credit” de las Repository of Machine Learning Databases (Blake y Merz, 1998) de la Universidad de California, Irvine.
- **grocery\_1month.sav.** Este archivo de datos hipotéticos es el archivo de datos *grocery\_coupons.sav* con las compras semanales “acumuladas” para que cada caso corresponda a un cliente diferente. Algunas de las variables que cambiaban semanalmente desaparecen de los resultados, y la cantidad gastada registrada se convierte ahora en la suma de las cantidades gastadas durante las cuatro semanas del estudio.
- **grocery\_coupons.sav.** Archivo de datos hipotéticos que contiene datos de encuestas recopilados por una cadena de tiendas de alimentación interesada en los hábitos de compra de sus clientes. Se sigue a cada cliente durante cuatro semanas, y cada caso corresponde a un cliente-semana distinto y registra información sobre dónde y cómo compran los clientes, incluida la cantidad que invierten en comestibles durante esa semana.
- **guttman.sav.** Bell (Bell, 1961) presentó una tabla para ilustrar posibles grupos sociales. Guttman (Guttman, 1968) utilizó parte de esta tabla, en la que se cruzaron cinco variables que describían elementos como la interacción social, sentimientos de pertenencia a un grupo, proximidad física de los miembros y grado de formalización de la relación con siete grupos sociales teóricos, incluidos multitudes (por ejemplo, las personas que acuden a un partido de fútbol), espectadores (por ejemplo, las personas que acuden a un teatro o de una conferencia), públicos (por ejemplo, los lectores de periódicos o los espectadores de televisión), muchedumbres (como una multitud pero con una interacción mucho más intensa), grupos primarios (íntimos), grupos secundarios (voluntarios) y la comunidad moderna (confederación débil que resulta de la proximidad cercana física y de la necesidad de servicios especializados).
- **health\_funding.sav.** Archivo de datos hipotéticos que contiene datos sobre inversión en sanidad (cantidad por 100 personas), tasas de enfermedad (índice por 10.000 personas) y visitas a centros de salud (índice por 10.000 personas). Cada caso representa una ciudad diferente.

- **hivassay.sav.** Archivo de datos hipotéticos sobre las iniciativas de un laboratorio farmacéutico para desarrollar un ensayo rápido para detectar la infección por VIH. Los resultados del ensayo son ocho tonos de rojo con diferentes intensidades, donde los tonos más oscuros indican una mayor probabilidad de infección. Se llevó a cabo una prueba de laboratorio de 2.000 muestras de sangre, de las cuales una mitad estaba infectada con el VIH y la otra estaba limpia.
- **hourlywagedata.sav.** Archivo de datos hipotéticos sobre los salarios por horas de enfermeras de puestos de oficina y hospitales y con niveles distintos de experiencia.
- **insurance\_claims.sav.** Éste es un archivo de datos hipotéticos sobre una compañía de seguros que desee generar un modelo para etiquetar las reclamaciones sospechosas y potencialmente fraudulentas. Cada caso representa una reclamación diferente.
- **insure.sav.** Archivo de datos hipotéticos sobre una compañía de seguros que estudia los factores de riesgo que indican si un cliente tendrá que hacer una reclamación a lo largo de un contrato de seguro de vida de 10 años. Cada caso del archivo de datos representa un par de contratos (de los que uno registró una reclamación y el otro no), agrupados por edad y sexo.
- **judges.sav.** Archivo de datos hipotéticos sobre las puntuaciones concedidas por jueces cualificados (y un aficionado) a 300 actuaciones gimnásticas. Cada fila representa una actuación diferente; los jueces vieron las mismas actuaciones.
- **kinship\_dat.sav.** Rosenberg y Kim (Rosenberg y Kim, 1975) comenzaron a analizar 15 términos de parentesco [tía, hermano, primos, hija, padre, nieta, abuelo, abuela, nieto, madre, sobrino, sobrina, hermana, hijo, tío]. Le pidieron a cuatro grupos de estudiantes universitarios (dos masculinos y dos femeninos) que ordenaran estos grupos según las similitudes. A dos grupos (uno masculino y otro femenino) se les pidió que realizaran la ordenación dos veces, pero que la segunda ordenación la hicieran según criterios distintos a los de la primera. Así, se obtuvo un total de seis “fuentes“. Cada fuente se corresponde con una matriz de proximidades de  $15 \times 15$  cuyas casillas son iguales al número de personas de una fuente menos el número de veces que se partitionaron los objetos en esa fuente.
- **kinship\_ini.sav.** Este archivo de datos contiene una configuración inicial para una solución tridimensional de *kinship\_dat.sav*.
- **kinship\_var.sav.** Este archivo de datos contiene variables independientes *sexo*, *gener(ación)*, y *grado* (de separación) que se pueden usar para interpretar las dimensiones de una solución para *kinship\_dat.sav*. Concretamente, se pueden usar para restringir el espacio de la solución a una combinación lineal de estas variables.
- **marketvalues.sav.** Archivo de datos sobre las ventas de casas en una nueva urbanización de Algonquin, Ill., durante los años 1999 y 2000. Los datos de estas ventas son públicos.
- **nhis2000\_subset.sav.** La National Health Interview Survey (NHIS, encuesta del Centro Nacional de Estadísticas de Salud de EE.UU.) es una encuesta detallada realizada entre la población civil de Estados Unidos. Las encuestas se realizaron en persona a una muestra representativa de las unidades familiares del país. Se recogió tanto la información demográfica como las observaciones acerca del estado y los hábitos de salud de los integrantes de cada unidad familiar. Este archivo de datos contiene un subconjunto de información de la encuesta de 2000. National Center for Health Statistics. National Health Interview Survey, 2000. Archivo de datos y documentación de uso público. [ftp://ftp.cdc.gov/pub/Health\\_Statistics/NCHS/Datasets/NHIS/2000/](ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/). Fecha de acceso: 2003.

- **ozono.sav.** Los datos incluyen 330 observaciones de seis variables meteorológicas para pronosticar la concentración de ozono a partir del resto de variables. Los investigadores anteriores (Breiman y Friedman, 1985), (Hastie y Tibshirani, 1990) han encontrado que no hay linealidad entre estas variables, lo que dificulta los métodos de regresión típica.
- **pain\_medication.sav.** Este archivo de datos hipotéticos contiene los resultados de una prueba clínica sobre medicación antiinflamatoria para tratar el dolor artrítico crónico. Resulta de particular interés el tiempo que tarda el fármaco en hacer efecto y cómo se compara con una medicación existente.
- **patient\_los.sav.** Este archivo de datos hipotéticos contiene los registros de tratamiento de pacientes que fueron admitidos en el hospital ante la posibilidad de sufrir un infarto de miocardio (IM o “ataque al corazón”). Cada caso corresponde a un paciente distinto y registra diversas variables relacionadas con su estancia hospitalaria.
- **patlos\_sample.sav.** Este archivo de datos hipotéticos contiene los registros de tratamiento de una muestra de pacientes que recibieron trombolíticos durante el tratamiento del infarto de miocardio (IM o “ataque al corazón”). Cada caso corresponde a un paciente distinto y registra diversas variables relacionadas con su estancia hospitalaria.
- **polishing.sav.** Archivo de datos “Nambeware Polishing Times” (Tiempo de pulido de metal) de la biblioteca de datos e historiales. Contiene datos sobre las iniciativas de un fabricante de cuberterías de metal (Nambe Mills, Santa Fe, N. M.) para planificar su programa de producción. Cada caso representa un artículo distinto de la línea de productos. Se registra el diámetro, el tiempo de pulido, el precio y el tipo de producto de cada artículo.
- **poll\_cs.sav.** Archivo de datos hipotéticos sobre las iniciativas de los encuestadores para determinar el nivel de apoyo público a una ley antes de una asamblea legislativa. Los casos corresponden a votantes registrados. Cada caso registra el condado, la población y el vecindario en el que vive el votante.
- **poll\_cs\_sample.sav.** Este archivo de datos hipotéticos contiene una muestra de los votantes enumerados en *poll\_cs.sav*. La muestra se tomó según el diseño especificado en el archivo de plan *poll\_csplan* y este archivo de datos registra las probabilidades de inclusión y las ponderaciones muestrales. Sin embargo, tenga en cuenta que debido a que el plan muestral hace uso de un método de probabilidad proporcional al tamaño (PPS), también existe un archivo que contiene las probabilidades de selección conjunta (*poll\_jointprob.sav*). Las variables adicionales que corresponden a los datos demográficos de los votantes y sus opiniones sobre la propuesta de ley se recopilaron y añadieron al archivo de datos después de tomar la muestra.
- **property\_assess.sav.** Archivo de datos hipotéticos sobre las iniciativas de un asesor del condado para mantener actualizada la evaluación de los valores de las propiedades utilizando recursos limitados. Los casos corresponden a las propiedades vendidas en el condado el año anterior. Cada caso del archivo de datos registra la población en que se encuentra la propiedad, el último asesor que visitó la propiedad, el tiempo transcurrido desde la última evaluación, la valoración realizada en ese momento y el valor de venta de la propiedad.
- **property\_assess\_cs.sav.** Archivo de datos hipotéticos sobre las iniciativas de un asesor de un estado para mantener actualizada la evaluación de los valores de las propiedades utilizando recursos limitados. Los casos corresponden a propiedades del estado. Cada caso del archivo de datos registra el condado, la población y el vecindario en el que se encuentra la propiedad, el tiempo transcurrido desde la última evaluación y la valoración realizada en ese momento.

- **property\_assess\_cs\_sample.sav** Este archivo de datos hipotéticos contiene una muestra de las propiedades recogidas en *property\_assess\_cs.sav*. La muestra se tomó en función del diseño especificado en el archivo de plan *property\_assess.csplan*, y este archivo de datos registra las probabilidades de inclusión y las ponderaciones muestrales. La variable adicional *Valor actual* se recopiló y añadió al archivo de datos después de tomar la muestra.
- **recidivism.sav**. Archivo de datos hipotéticos sobre las iniciativas de una agencia de orden público para comprender los índices de reincidencia en su área de jurisdicción. Cada caso corresponde a un infractor anterior y registra su información demográfica, algunos detalles de su primer delito y, a continuación, el tiempo transcurrido desde su segundo arresto, si ocurrió en los dos años posteriores al primer arresto.
- **recidivism\_cs\_sample.sav**. Archivo de datos hipotéticos sobre las iniciativas de una agencia de orden público para comprender los índices de reincidencia en su área de jurisdicción. Cada caso corresponde a un delincuente anterior, puesto en libertad tras su primer arresto durante el mes de junio de 2003 y registra su información demográfica, algunos detalles de su primer delito y los datos de su segundo arresto, si se produjo antes de finales de junio de 2006. Los delincuentes se seleccionaron de una muestra de departamentos según el plan de muestreo especificado en *recidivism\_cs.csplan*. Como este plan utiliza un método de probabilidad proporcional al tamaño (PPS), también existe un archivo que contiene las probabilidades de selección conjunta (*recidivism\_cs\_jointprob.sav*).
- **rfm\_transactions.sav**. Archivo de datos hipotéticos que contiene datos de transacciones de compra, incluida la fecha de compra, los artículos adquiridos y el importe de cada transacción.
- **salesperformance.sav**. Archivo de datos hipotéticos sobre la evaluación de dos nuevos cursos de formación de ventas. Sesenta empleados, divididos en tres grupos, reciben formación estándar. Además, el grupo 2 recibe formación técnica; el grupo 3, un tutorial práctico. Cada empleado se sometió a un examen al final del curso de formación y se registró su puntuación. Cada caso del archivo de datos representa a un alumno distinto y registra el grupo al que fue asignado y la puntuación que obtuvo en el examen.
- **satisf.sav**. Archivo de datos hipotéticos sobre una encuesta de satisfacción llevada a cabo por una empresa minorista en cuatro tiendas. Se encuestó a 582 clientes en total y cada caso representa las respuestas de un único cliente.
- **screws.sav** Este archivo de datos contiene información acerca de las características de tornillos, pernos, clavos y tacos (Hartigan, 1975).
- **shampoo\_ph.sav**. Archivo de datos hipotéticos sobre el control de calidad en una fábrica de productos para el cabello. Se midieron seis lotes de resultados distintos en intervalos regulares y se registró su pH. El intervalo objetivo es de 4,5 a 5,5.
- **ships.sav**. Un conjunto de datos presentados y analizados en otro lugar (McCullagh et al., 1989) sobre los daños en los cargueros producidos por las olas. Los recuentos de incidentes se pueden modelar como si ocurrieran con una tasa de Poisson dado el tipo de barco, el período de construcción y el período de servicio. Los meses de servicio agregados para cada casilla de la tabla formados por la clasificación cruzada de factores proporcionan valores para la exposición al riesgo.
- **site.sav**. Archivo de datos hipotéticos sobre las iniciativas de una compañía para seleccionar sitios nuevos para sus negocios en expansión. Se ha contratado a dos consultores para evaluar los sitios de forma independiente, quienes, además de un informe completo, han resumido cada sitio como una posibilidad “buena”, “media” o “baja”.

- **smokers.sav.** Este archivo de datos es un resumen de la encuesta sobre toxicomanía 1998 National Household Survey of Drug Abuse y es una muestra de probabilidad de unidades familiares americanas. (<http://dx.doi.org/10.3886/ICPSR02934>) Así, el primer paso de un análisis de este archivo de datos debe ser ponderar los datos para reflejar las tendencias de población.
- **stroke\_clean.sav.** Este archivo de datos hipotéticos contiene el estado de una base de datos médica después de haberla limpiado mediante los procedimientos de la opción Preparación de datos.
- **stroke\_invalid.sav.** Este archivo de datos hipotéticos contiene el estado inicial de una base de datos médica que incluye contiene varios errores de entrada de datos.
- **stroke\_survival.** Este archivo de datos hipotéticos registra los tiempos de supervivencia de los pacientes que finalizan un programa de rehabilitación tras un ataque isquémico. Tras el ataque, la ocurrencia de infarto de miocardio, ataque isquémico o ataque hemorrágico se anotan junto con el momento en el que se produce el evento registrado. La muestra está truncada a la izquierda ya que únicamente incluye a los pacientes que han sobrevivido al final del programa de rehabilitación administrado tras el ataque.
- **stroke\_valid.sav.** Este archivo de datos hipotéticos contiene el estado de una base de datos médica después de haber comprobado los valores mediante el procedimiento Validar datos. Sigue conteniendo casos potencialmente anómalos.
- **survey\_sample.sav.** Este archivo de datos contiene datos de encuestas, incluyendo datos demográficos y diferentes medidas de actitud. Se basa en un subconjunto de variables de NORC General Social Survey de 1998, aunque algunos valores de datos se han modificado y que existen variables ficticias adicionales se han añadido para demostraciones.
- **telco.sav.** Archivo de datos hipotéticos sobre las iniciativas de una compañía de telecomunicaciones para reducir el abandono de clientes en su base de clientes. Cada caso corresponde a un cliente distinto y registra diversa información demográfica y de uso del servicio.
- **telco\_extra.sav.** Este archivo de datos es similar al archivo de datos *telco.sav*, pero las variables de meses con servicio y gasto de clientes transformadas logarítmicamente se han eliminado y sustituido por variables de gasto del cliente transformadas logarítmicamente tipificadas.
- **telco\_missing.sav.** Este archivo de datos es un subconjunto del archivo de datos *telco.sav*, pero algunos valores de datos demográficos se han sustituido con valores perdidos.
- **testmarket.sav.** Archivo de datos hipotéticos sobre los planes de una cadena de comida rápida para añadir un nuevo artículo a su menú. Hay tres campañas posibles para promocionar el nuevo producto, por lo que el artículo se presenta en ubicaciones de varios mercados seleccionados aleatoriamente. Se utiliza una promoción diferente en cada ubicación y se registran las ventas semanales del nuevo artículo durante las primeras cuatro semanas. Cada caso corresponde a una ubicación semanal diferente.
- **testmarket\_1month.sav.** Este archivo de datos hipotéticos es el archivo de datos *testmarket.sav* con las ventas semanales “acumuladas” para que cada caso corresponda a una ubicación diferente. Como resultado, algunas de las variables que cambiaban semanalmente desaparecen y las ventas registradas se convierten en la suma de las ventas realizadas durante las cuatro semanas del estudio.
- **tree\_car.sav.** Archivo de datos hipotéticos que contiene datos demográficos y de precios de compra de vehículos.

- **tree\_credit.sav** Archivo de datos hipotéticos que contiene datos demográficos y de historial de créditos bancarios.
- **tree\_missing\_data.sav** Archivo de datos hipotéticos que contiene datos demográficos y de historial de créditos bancarios con un elevado número de valores perdidos.
- **tree\_score\_car.sav.** Archivo de datos hipotéticos que contiene datos demográficos y de precios de compra de vehículos.
- **tree\_textdata.sav.** Archivo de datos sencillos con dos variables diseñadas principalmente para mostrar el estado por defecto de las variables antes de realizar la asignación de nivel de medida y etiquetas de valor.
- **tv-survey.sav.** Archivo de datos hipotéticos sobre una encuesta dirigida por un estudio de TV que está considerando la posibilidad de ampliar la emisión de un programa de éxito. Se preguntó a 906 encuestados si verían el programa en distintas condiciones. Cada fila representa un encuestado diferente; cada columna es una condición diferente.
- **ulcer\_recurrence.sav.** Este archivo contiene información parcial de un estudio diseñado para comparar la eficacia de dos tratamientos para prevenir la reaparición de úlceras. Constituye un buen ejemplo de datos censurados por intervalos y se ha presentado y analizado en otro lugar (Collett, 2003).
- **ulcer\_recurrence\_recoded.sav.** Este archivo reorganiza la información de *ulcer\_recurrence.sav* para permitir modelar la probabilidad de eventos de cada intervalo del estudio en lugar de sólo la probabilidad de eventos al final del estudio. Se ha presentado y analizado en otro lugar (Collett et al., 2003).
- **verd1985.sav.** Archivo de datos sobre una encuesta (Verdegaal, 1985). Se han registrado las respuestas de 15 sujetos a 8 variables. Se han dividido las variables de interés en tres grupos. El conjunto 1 incluye *edad* y *ecivil*, el conjunto 2 incluye *mascota* y *noticia*, mientras que el conjunto 3 incluye *música* y *vivir*. Se escala *mascota* como nominal múltiple y *edad* como ordinal; el resto de variables se escalan como nominal simple.
- **virus.sav.** Archivo de datos hipotéticos sobre las iniciativas de un proveedor de servicios de Internet (ISP) para determinar los efectos de un virus en sus redes. Se ha realizado un seguimiento (aproximado) del porcentaje de tráfico de correos electrónicos infectados en sus redes a lo largo del tiempo, desde el momento en que se descubre hasta que la amenaza se contiene.
- **wheeze\_steubenville.sav.** Subconjunto de un estudio longitudinal de los efectos sobre la salud de la polución del aire en los niños (Ware, Dockery, Spiro III, Speizer, y Ferris Jr., 1984). Los datos contienen medidas binarias repetidas del estado de las sibilancias en niños de Steubenville, Ohio, con edades de 7, 8, 9 y 10 años, junto con un registro fijo de si la madre era fumadora durante el primer año del estudio.
- **workprog.sav.** Archivo de datos hipotéticos sobre un programa de obras del gobierno que intenta colocar a personas desfavorecidas en mejores trabajos. Se siguió una muestra de participantes potenciales del programa, algunos de los cuales se seleccionaron aleatoriamente para entrar en el programa, mientras que otros no siguieron esta selección aleatoria. Cada caso representa un participante del programa diferente.

# Notices

Licensed Materials – Property of SPSS Inc., an IBM Company. © Copyright SPSS Inc. 1989, 2010.

Patent No. 7,023,453

**The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:** SPSS INC., AN IBM COMPANY, PROVIDES THIS PUBLICATION “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. SPSS Inc. may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-SPSS and non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this SPSS Inc. product and use of those Web sites is at your own risk.

When you send information to IBM or SPSS, you grant IBM and SPSS a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Information concerning non-SPSS products was obtained from the suppliers of those products, their published announcements or other publicly available sources. SPSS has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-SPSS products. Questions on the capabilities of non-SPSS products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

## COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to SPSS Inc., for the purposes of developing,

using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. SPSS Inc., therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided “AS IS”, without warranty of any kind. SPSS Inc. shall not be liable for any damages arising out of your use of the sample programs.

### **Trademarks**

IBM, the IBM logo, and [ibm.com](http://www.ibm.com) are trademarks of IBM Corporation, registered in many jurisdictions worldwide. A current list of IBM trademarks is available on the Web at <http://www.ibm.com/legal/copytrade.shtml>.

SPSS is a trademark of SPSS Inc., an IBM Company, registered in many jurisdictions worldwide.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Java and all Java-based trademarks and logos are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

This product uses WinWrap Basic, Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com>.

Other product and service names might be trademarks of IBM, SPSS, or other companies.

Adobe product screenshot(s) reprinted with permission from Adobe Systems Incorporated.

Microsoft product screenshot(s) reprinted with permission from Microsoft Corporation.



---

# Bibliografía

- Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. Nueva York: Harper & Row.
- Blake, C. L., y C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mlern/MLRepository.html>.
- Box, G. E. P., G. M. Jenkins, y G. C. Reinsel. 1994. *Time series analysis: Forecasting and control*, 3rd ed. Englewood Cliffs, N.J.: Prentice Hall.
- Breiman, L., y J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, .
- Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.
- Gardner, E. S. 1985. Exponential smoothing: The state of the art. *Journal of Forecasting*, 4, .
- Green, P. E., y V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.
- Green, P. E., y Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.
- Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, .
- Hartigan, J. A. 1975. *Clustering algorithms*. Nueva York: John Wiley and Sons.
- Hastie, T., y R. Tibshirani. 1990. *Generalized additive models*. Londres: Chapman and Hall.
- Kennedy, R., C. Riquier, y B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing*, 5, .
- McCullagh, P., y J. A. Nelder. 1989. *Modelos lineales generalizados*, 2nd ed. Londres: Chapman & Hall.
- Pena, D., G. C. Tiao, y R. S. Tsay, eds. 2001. *A course in time series analysis*. Nueva York: John Wiley and Sons.
- Price, R. H., y D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, .
- Rickman, R., N. Mitchell, J. Dingman, y J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228, .
- Rosenberg, S., y M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, .
- Van der Ham, T., J. J. Meulman, D. C. Van Strien, y H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, .
- Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (en neerlandés)*. Leiden: Department of Data Theory, University of Leiden.

Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, y B. G. Ferris Jr.. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases*, 129, .

- análisis armónico, 44
- Aplicar modelos de series temporales, 29, 60, 76
  - estadístico de Box-Ljung, 32
  - estadísticos de bondad de ajuste, 32, 34, 63
  - estadísticos en todos los modelos, 32, 34, 63
  - función de autocorrelación parcial residual, 32, 34
  - función de autocorrelación residual, 32, 34
  - guardar modelos estimados de nuevo en XML, 37
  - guardar predicciones, 37, 62
  - intervalos de confianza, 34, 39
  - modelos de mejor o peor ajuste, 36
  - nombres de variables nuevas, 37, 64
  - parámetros de modelo, 32
  - período de estimación, 31
  - período de predicción, 31, 61, 83
  - predicciones, 32, 34, 84
  - tabla ajuste del modelo, 63
  - tabla de predicciones, 84
  - valores ajustados, 34
  - valores perdidos, 39
  - volver a estimar parámetros de modelo, 31, 61
- archivos de ejemplo
  - posición, 106
- autorregresión
  - modelos ARIMA, 13
  
- BIC normalizado (criterio de información bayesiano), 100
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
- bondad de ajuste
  - definiciones, 100
  - en Aplicar modelos de series temporales, 32, 34, 63
  - en modelizador de series temporales, 19, 21, 55
  
- casos reservados, 2
  
- datos históricos
  - en Aplicar modelos de series temporales, 34
  - en modelizador de series temporales, 21
- Descomposición estacional, 41–43
  - almacenamiento de nuevas variables, 42
  - cálculo de las medias móviles, 41
  - componente de fecha periódico, 87
  - crear variables, 42
  - modelos, 41
  - procedimientos relacionados, 94
  - supuestos, 41
  - variables nuevas, 92
- Diagramas espectrales, 44, 46
  - análisis espectral bivariado, 45
  - centrado de transformación, 46
  - densidad espectral, 97
  - periodograma, 97
  - procedimientos relacionados, 99
  - supuestos, 44
  - ventanas espectrales, 44
  
- error absoluto máximo, 100
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
- error absoluto máximo porcentual, 100
  - en Aplicar modelos de series temporales, 32, 34, 64
  - en modelizador de series temporales, 19, 21, 56
- error absoluto porcentual promedio, 100
  - en Aplicar modelos de series temporales, 32, 34, 64
  - en modelizador de series temporales, 19, 21, 56
- error absoluto promedio, 100
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
- estadístico de Box-Ljung
  - en Aplicar modelos de series temporales, 32
  - en modelizador de series temporales, 19, 74
- eventos, 9
  - en modelizador de series temporales, 8
  
- FAP
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
  - gráficos para procesos ARIMA puros, 102
- FAS
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
  - gráficos para procesos ARIMA puros, 102
- función de autocorrelación
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
  - gráficos para procesos ARIMA puros, 102
- función de autocorrelación parcial
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
  - gráficos para procesos ARIMA puros, 102
- funciones de transferencia, 15
  - órdenes de denominador, 15
  - órdenes de diferencia, 15
  - órdenes de numerador, 15
  - órdenes estacionales, 15
  - retardo, 15
  
- guardar
  - especificaciones de modelo en XML, 24

- modelos estimados de nuevo en XML, 37
- nombres de variables nuevas, 24, 37
- predicciones de modelo, 24, 37
- integración
  - modelos ARIMA, 13
- intervalos de confianza
  - en Aplicar modelos de series temporales, 34, 39
  - en modelizador de series temporales, 21, 26
- legal notices, 116
- MAE, 100
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
- MAPE, 100
  - en Aplicar modelos de series temporales, 32, 34, 64
  - en modelizador de series temporales, 19, 21, 56
- MaxAE, 100
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
- MaxAPE, 100
  - en Aplicar modelos de series temporales, 32, 34, 64
  - en modelizador de series temporales, 19, 21, 56
- media móvil
  - modelos ARIMA, 13
- Modelizador de series temporales, 4
  - ARIMA, 7, 13
  - estadístico de Box-Ljung, 19
  - estadísticos de bondad de ajuste, 19, 21, 55, 74
  - estadísticos en todos los modelos, 19, 21, 55, 57
  - eventos, 8
  - función de autocorrelación parcial residual, 19, 21
  - función de autocorrelación residual, 19, 21
  - funciones de transferencia, 15
  - guardar especificaciones de modelo en XML, 24, 54, 70
  - guardar predicciones, 24, 54
  - intervalos de confianza, 21, 26
  - Modelizador experto, 7, 49, 65
  - modelos de mejor o peor ajuste, 23
  - nombres de modelo, 26
  - nombres de variables nuevas, 24, 58
  - parámetros de modelo, 19, 71
  - periodicidad, 8, 11, 13, 15
  - período de estimación, 7, 52
  - período de predicción, 7, 26, 52–53
  - predicciones, 19, 21, 55
  - suavizado exponencial, 7, 11
  - tabla de descripción del modelo, 73
  - tabla de estadísticos del modelo, 74
  - tabla de parámetros del modelo ARIMA, 74
  - tabla de predicciones, 58
  - transformación de serie, 11, 13, 15
  - valores ajustados, 21, 72
  - valores atípicos, 10, 17, 69
  - valores perdidos, 26
- Modelizador experto, 7, 49
  - limitar espacio de modelo, 8, 52
  - valores atípicos, 10, 69
- modelo de suavizado exponencial amortiguado, 11
- modelo de suavizado exponencial de Brown, 11
- modelo de suavizado exponencial de Holt, 11
- modelo de suavizado exponencial estacional simple, 11
- modelo de suavizado exponencial simple, 11
- modelo de suavizado exponencial simple de Winters
  - aditivo, 11
  - multiplicativo, 11
- modelos
  - ARIMA, 7, 13
  - Modelizador experto, 7
  - suavizado exponencial, 7, 11
- modelos ARIMA, 7
  - constante, 13
  - funciones de transferencia, 15
  - órdenes autorregresivos, 13
  - órdenes de diferenciación, 13
  - órdenes de media móvil, 13
  - órdenes estacionales, 13
  - valores atípicos, 17
- modelos de suavizado exponencial, 7, 11
- nombres de modelo
  - en modelizador de series temporales, 26
- nombres de variable
  - en Aplicar modelos de series temporales, 37
  - en modelizador de series temporales, 24
- órdenes estacionales
  - modelos ARIMA, 13
- parámetros de modelo
  - en Aplicar modelos de series temporales, 32
  - en modelizador de series temporales, 19, 71
- periodicidad
  - en modelizador de series temporales, 8, 11, 13, 15
- período de estimación, 2
  - en Aplicar modelos de series temporales, 31
  - en modelizador de series temporales, 7, 52
- período de predicción
  - en Aplicar modelos de series temporales, 31, 61, 83
  - en modelizador de series temporales, 7, 26, 52–53
- período de validación, 2
- período histórico, 2
- predicciones
  - en Aplicar modelos de series temporales, 32, 34, 84
  - en modelizador de series temporales, 19, 21, 55
- $R^2$ , 100
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
- $R^2$  estacionaria, 100
  - en Aplicar modelos de series temporales, 32, 34

- en modelizador de series temporales, 19, 21, 74
- raíz del error cuadrático promedio, 100
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
- residuos
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
- RMSE, 100
  - en Aplicar modelos de series temporales, 32, 34
  - en modelizador de series temporales, 19, 21
  
- tabla ajuste del modelo
  - en Aplicar modelos de series temporales, 63
- tabla de descripción del modelo
  - en modelizador de series temporales, 73
- tabla de estadísticos del modelo
  - en modelizador de series temporales, 74
- tabla de parámetros del modelo ARIMA
  - en modelizador de series temporales, 74
- tabla de predicciones
  - en Aplicar modelos de series temporales, 84
  - en modelizador de series temporales, 58
- trademarks, 117
- transformación de raíz cuadrada
  - en modelizador de series temporales, 11, 13, 15
- transformación diferencia
  - modelos ARIMA, 13
- transformación diferencia estacional
  - modelos ARIMA, 13
- transformación logarítmica
  - en modelizador de series temporales, 11, 13, 15
- transformación logarítmica natural
  - en modelizador de series temporales, 11, 13, 15
  
- valor atípico aditivo, 101
  - en modelizador de series temporales, 10, 17
- valor atípico aditivo estacional, 101
  - en modelizador de series temporales, 10, 17
- valor atípico de cambio de nivel, 101
  - en modelizador de series temporales, 10, 17
- valor atípico de parche aditivo, 101
  - en modelizador de series temporales, 10, 17
- valor atípico de tendencia local, 101
  - en modelizador de series temporales, 10, 17
- valor atípico innovador, 101
  - en modelizador de series temporales, 10, 17
- valor atípico transitorio, 101
  - en modelizador de series temporales, 10, 17
- valores ajustados
  - en Aplicar modelos de series temporales, 34
  - en modelizador de series temporales, 21, 72
- valores atípicos
  - definiciones, 101
  - Modelizador experto, 10, 69
  - modelos ARIMA, 17
- valores perdidos
  - en Aplicar modelos de series temporales, 39
  - en modelizador de series temporales, 26
- volver a estimar parámetros de modelo
  - en Aplicar modelos de series temporales, 31, 61
  
- XML
  - guardar modelos de series temporales en XML, 24, 54, 70
  - guardar modelos estimados de nuevo en XML, 37