

IBM SPSS Forecasting 32



Nota

Antes de utilizar esta información y el producto al que da soporte, lea la información del apartado “Avisos” en la página 57.

Información del producto

Esta edición se aplica a la versión 32, release 0, modificación 0 de IBM® SPSS Estadística y a todas las versiones y modificaciones posteriores hasta que se indique lo contrario en nuevas ediciones.

© Copyright International Business Machines Corporation .

Contenido

Capítulo 1. Previsiones.....	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 previsiones.....	3
Modelizador de series temporales.....	3
Especificación de opciones para el modelizador experto.....	5
Modelos personalizados de suavizado exponencial.....	6
Modelos ARIMA personalizados.....	7
Resultado.....	10
Almacenamiento de predicciones de modelos y especificaciones de modelo.....	12
Opciones.....	13
Características adicionales del comando TSMODEL.....	14
Aplicar modelos de series temporales.....	14
Resultado.....	16
Almacenamiento de predicciones de modelos y especificaciones de modelo.....	19
Opciones.....	20
Características adicionales del comando TSAPPLY.....	20
Descomposición estacional.....	20
Descomposición estacional: Guardar.....	21
Características adicionales del comando SEASON.....	22
Diagramas espectrales.....	22
Características adicionales del comando SPECTRA.....	24
Modelos causales temporales.....	24
Obtener un modelo causal temporal.....	25
Series temporales a modelo.....	25
Observaciones	27
Intervalo de tiempo del análisis	28
Agregación y distribución	28
Valores que faltan.....	29
Opciones de datos generales	30
Opciones de generación generales	30
Series a mostrar	30
Opciones de resultados	31
Periodo de estimación	33
Predicción	34
Guardar	34
Resultados interactivos	35
Aplicación de modelos causales temporales.....	36
Aplicación de modelos causales temporales.....	36
Predicción de modelos causales temporales.....	36
Escenarios de modelos causales temporales.....	42
Modelos autorregresivos vectoriales.....	48
Variables.....	49
Estadísticos.....	49
Gráficos.....	50
Guardar.....	51
Opciones.....	51
Medidas de bondad de ajuste.....	52
Tipos de valores atípicos.....	53

Guía de gráficos ACF/PACF.....	53
Avisos.....	57
Marcas comerciales.....	58
Índice.....	61

Capítulo 1. Previsiones

Las características siguientes de predicciones están incluidas en SPSS Statistics edición Professional o la opción Predicciones.

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 mediciones tomadas a intervalos regulares.

Tabla 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	6	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

Datos basados en columna

Cada campo de serie temporal contiene los datos para una sola serie temporal. Esta estructura es la estructura tradicional de los datos de serie temporal, tal como se utiliza en el procedimiento del modelador de series temporales, el procedimiento de descomposición estacional y el procedimiento de diagramas espectrales. 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 series temporales, cada serie se deberá organizar en una columna en la hoja de cálculo. 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.

Datos multidimensionales

Para los datos multidimensionales, cada campo de serie temporal contiene los datos para varias series temporales. Las series temporales separadas, en un campo particular, se identifican mediante un conjunto de valores de campos categóricos a los que se hace referencia como campos de *dimensión*.

Por ejemplo, los datos de ventas para regiones y marcas diferentes se pueden almacenar en un único campo *ventas*, de forma que las dimensiones de este caso son *región* y *marca*. Cada combinación de *región* y *marca* identifica una serie temporal particular para *ventas*. Por ejemplo, en la tabla siguiente, los registros que tienen 'norte' para *región* y 'marcaX' para *marca* definen un serie temporal única.

fecha	región	marca	ventas
01/01/2014	norte	brandX	82350
01/01/2014	norte	brandY	86380
01/01/2014	sur	brandX	91375
01/01/2014	sur	brandY	70320
01/02/2014	norte	brandX	83275
01/02/2014	norte	brandY	85260
01/02/2014	sur	brandX	94760
01/02/2014	sur	brandY	69870

Nota: Los datos que se importan desde cubos de OLAP como, por ejemplo, de IBM Cognos TM1, se representan como datos multidimensionales.

Transformaciones de los datos

Una serie de procedimientos de transformación de datos que se proporciona en el sistema básico es útil en el análisis de series temporales. Estas transformaciones solo se aplican a los datos basados en columna, donde cada campo de serie temporal contiene los datos para una sola serie temporal.

- El procedimiento Definir fechas (en el menú Datos) genera las variables de fecha que se utilizan para establecer la periodicidad y para distinguir entre periodos históricos, de validación y de previsión. 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 existente. 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 perdidos del 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 la *Guía del usuario del sistema básico* si desea información detallada sobre 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 generación de modelos. 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 previsiones

El módulo adicional Predicciones proporciona los procedimientos siguientes para llevar a cabo las tareas de creación de modelos y producción de predicciones:

- El procedimiento “[Modelizador de series temporales](#)” en la [página 3](#) 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 generación de modelos personalizados.
- El procedimiento “[Aplicar modelos de series temporales](#)” en la [página 14](#) 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.
- El procedimiento “[Modelos causales temporales](#)” en la [página 24](#) crea modelos de serie temporal autorregresivos para cada destino y determina automáticamente las mejores entrada que tienen una relación causal con el destino. El procedimiento produce resultado interactivo que se puede utilizar para explorar las relaciones causales. El procedimiento también puede generar predicciones, detectar valores atípicos y determinar las series que causan de forma más probable un valor atípico.
- El procedimiento “[Predicción de modelos causales temporales](#)” en la [página 36](#) aplica un modelo causal temporal al conjunto de datos activo. Puede utilizar este procedimiento para obtener predicciones para series para las que están disponibles datos más actuales, sin volver a crear los modelos. También puede utilizarlo para determinar series que causan con mayor probabilidad los valores atípicos que el procedimiento de Modelos causales temporales ha detectado.

Modelizador de series temporales

El procedimiento Modelador de series temporales estima modelos de suavizado exponencial, de media móvil integrada autorregresiva univariante (ARIMA) y ARIMA multivariante (o modelos de función de transferencia) para series temporales, y elabora previsiones. El procedimiento incluye un Modelizador experto que intenta identificar y estimar automáticamente el ARIMA de mejor ajuste o el modelo de suavizado exponencial para una o más series de variables dependientes, de esta forma se elimina la necesidad de identificar un modelo apropiado mediante el método de prueba y error. De forma alternativa, puede especificar un ARIMA personalizado o un modelo de suavizado exponencial.

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 Estadística, 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.

Estadísticas. Medidas de bondad de ajuste: R cuadrado estacionaria, R cuadrado (R^2), raíz de la media cuadrática de los errores (RMSE), media del error absoluto (MAE), error absoluto porcentual promedio (MAPE), error absoluto máximo (MaxAE), error porcentual absoluto máximo (MaxAPE) y criterio de información bayesiano (BIC) normalizado. Residuos: función de autocorrelación, función de autocorrelación parcial, Ljung-Box Q . 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 de la media cuadrática de los errores (RMSE), media del error absoluto (MAE), error absoluto porcentual promedio (MAPE), error absoluto máximo (MaxAE), error porcentual absoluto máximo (MaxAPE), criterio de información bayesiano (BIC) normalizado, diagramas de cajas de autocorrelaciones de residuos 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 forma más eficaz de transformar una serie no estacionaria en una estacionaria es a través de una transformación de diferencias, disponible en el Recuadro 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

1. En los menús seleccione:

Analizar > Previsiones > Crear modelos tradicionales...

Nota: Los campos resaltados en rojo son obligatorios. Los botones **Pegar** y **Aceptar** se activan después de introducir valores válidos en todos los campos obligatorios.

2. En la pestaña Variables, seleccione una o más variables dependientes para modelarlas.
3. En el cuadro desplegable Método, seleccione un método de modelado. Si desea modelar automáticamente, deje el método predeterminado 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:

4. Pulse en la pestaña **Opciones**.
5. 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 las predicciones, 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.
- 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. De forma predeterminada, 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. De forma predeterminada, el período de estimación incluye todos los casos del conjunto de datos activo. Para establecer el periodo 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, de forma predeterminada, 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

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 desde 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

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 valores atípicos de forma automática. De forma predeterminada, 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

Modelos personalizados de suavizado exponencial

Tipo de modelo. Sólo modelos de suavizado exponencial¹ 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 series en las que no hay tendencia ni 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.

¹ Gardner, E. S. 1985. Exponential smoothing: The state of the art. *Journal of Forecasting*, 4, 1-28.

- **Tendencia lineal Holt.** Este modelo es apropiado para series en las que existe una tendencia lineal y ninguna 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 Brown.** Este modelo es apropiado para series en las que existe una tendencia lineal y ninguna 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 series con una tendencia lineal que se está extinguiendo 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.
- **Estacional simple.** Este modelo es adecuado para series sin tendencia y con un efecto estacional 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, $p = 12$).
- **Aditivo de Winters.** Este modelo es adecuado para series con una 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 $p = 12$).
- **Multiplicativo de Winters.** Este modelo es adecuado para series con una 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 desde 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 modelador de series temporales le permite crear modelos ARIMA (Autoregresivo Integrated Moving Average) no estacionales o estacionales personalizados, también conocidos como Box-Jenkins²modelos-- 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.

² Box, G. E. P., G. M. Jenkins, y G. C. Reinsel. 1994. *Time series analysis: Forecasting and control*, 3ª ed. Englewood Cliffs, N.J.: Prentice Hall.

- 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

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 celdas de la columna estacional solo están habilitadas si se ha definido una periodicidad para el conjunto de datos activo (consulte "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 respecto al 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 desde 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 medio de la serie global es 0. La exclusión de la constante se recomienda cuando se aplica la diferenciación.

Funciones de transferencia en modelos ARIMA personalizados

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 prever 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 numerador, el modelo incluye los pedidos 2, 1 y 0. Si especifica 3 para el denominador, el modelo incluye los pedidos 3, 2 y 1. Las celdas de la columna estacional solo están habilitadas si se ha definido una periodicidad para el conjunto de datos activo (consulte "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 desde 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

La pestaña Valores atípicos proporciona las opciones siguientes para el manejo de valores atípicos³: detectarlos automáticamente, especificar puntos concretos como valores atípicos o no detectarlos o modelarlos.

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

Detectar valores atípicos de forma automática. 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

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 con soporte son: aditivo (valor predeterminado), 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.

Resultado

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 previsiones

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 estacionario
- *R* cuadrado
- Raíz de la media cuadrática de los errores
- Media del error porcentual absoluto

³ Pena, D., G. C. Tiao, y R. S. Tsay, eds. 2001. *A course in time series analysis*. New York: John Wiley and Sons.

- Media del error absoluto
- Error porcentual absoluto máximo
- Error absoluto máximo
- BIC normalizado

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 estadísticas de resumen y percentiles para R cuadrado estacionario, R cuadrado, error cuadrático promedio de raíz, error absoluto porcentual promedio, error absoluto promedio, error absoluto porcentual máximo, error absoluto máximo y Criterio de información bayesiano normalizado.
- **Función de autocorrelación simple (FAS) de residuo.** Tabla de estadísticos de resumen y percentiles para las autocorrelaciones de los residuos de todos los modelos estimados.
- **Función de autocorrelación parcial (FAP) de residuo.** 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) de residuo.** Muestra una tabla con las autocorrelaciones de residuos por retardo para cada modelo estimado. La tabla incluye los intervalos de confianza para las autocorrelaciones.
- **Función de autocorrelación parcial (FAP) de residuo.** 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 previsión se establece en la pestaña Opciones.

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 estacionario
- R cuadrado
- Raíz de la media cuadrática de los errores
- Media del error porcentual absoluto
- Media del error absoluto
- Error porcentual absoluto máximo
- Error absoluto máximo
- BIC normalizado
- Función de autocorrelación simple (FAS) de residuo

- Función de autocorrelación parcial (FAP) de residuo

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 previsión.
- **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 previsión.
- **Intervalos de confianza de los valores ajustados.** Son los intervalos de confianza para el período de estimación.

Función de autocorrelación simple (FAS) de residuo. Muestra un gráfico con las autocorrelaciones de residuos para cada modelo estimado.

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

Limitación de resultados para los modelos de mejor o peor ajuste

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. De forma predeterminada, 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 predeterminado es R cuadrado estacionaria.

Almacenamiento de predicciones de modelos y especificaciones de modelo

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. De forma predeterminada, 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 predeterminados. 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 previsiones .

- **Archivo XML.** Las especificaciones de modelo se guardan en un archivo XML que se puede utilizar con IBM SPSS aplicaciones.
- **Archivo PMML.** Las especificaciones de modelo se guardan en un archivo XML compatible con PMML que se puede utilizar con aplicaciones compatibles con PMML, incluidas aplicaciones IBM SPSS.

Opciones

La pestaña Opciones le permite establecer el período de previsión, 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. De forma predeterminada, el final del periodo de estimación es el último caso del conjunto de datos activos, 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 previsiones 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 perdidos del usuario. Estas opciones controlan el tratamiento de los valores perdidos del usuario.

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

- **Tratar como válidos.** Los valores perdidos del 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 del 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 previsión, el procedimiento genera una advertencia y realiza previsiones en la medida de lo posible.

Ancho de intervalo de confianza (%). Los intervalos de confianza se calculan para las predicciones del modelo y las autocorrelaciones residuales. Puede especificar cualquier valor positivo menor que 100. De forma predeterminada, 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 exclusivos compuestos por un prefijo personalizable y un sufijo entero. Puede escribir un prefijo o dejar el valor predeterminado 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.

Características 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).

Consulte la *Referencia de sintaxis de comandos* para obtener información completa de la sintaxis.

Aplicar modelos de series temporales

El procedimiento Aplicar modelos de series temporales carga los 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 utilizando 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 prevén 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.

Estadísticas. Medidas de bondad de ajuste: R cuadrado estacionaria, R cuadrado (R^2), raíz de la media cuadrática de los errores (RMSE), media del error absoluto (MAE), error absoluto porcentual promedio (MAPE), error absoluto máximo (MaxAE), error porcentual absoluto máximo (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 de la media cuadrática de los errores (RMSE), media del error absoluto (MAE), error absoluto porcentual promedio (MAPE), error absoluto máximo (MaxAE), error porcentual absoluto máximo (MaxAPE), criterio de información bayesiano (BIC) normalizado, diagramas de cajas de autocorrelaciones de residuos 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

1. En los menús seleccione:

Analizar > Previsiones > Aplicar modelos tradicionales...

2. 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 las predicciones, 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 que se visualizan en el resultado y se utilizan para filtrar modelos (el de mejor o peor ajuste) se toman del archivo del modelo y reflejan los datos utilizados cuando se desarrolló cada modelo (o la última actualización). 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. De forma predeterminada, 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 predeterminado), 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.

Resultado

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 previsiones

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 estacionario
- R cuadrado
- Raíz de la media cuadrática de los errores
- Media del error porcentual absoluto
- Media del error absoluto
- Error porcentual absoluto máximo
- Error absoluto máximo
- BIC normalizado

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 estadísticas de resumen y percentiles para R cuadrado estacionario, R cuadrado, error cuadrático promedio de raíz, error absoluto porcentual promedio, error absoluto promedio, error absoluto porcentual máximo, error absoluto máximo y Criterio de información bayesiano normalizado.
- **Función de autocorrelación simple (FAS) de residuo.** 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 estiman de nuevo los parámetros de modelo (**Estimar de nuevo a partir de los datos** en la pestaña Modelos).
- **Función de autocorrelación parcial (FAP) de residuo.** 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 estiman de nuevo 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) de residuo.** Muestra una tabla con las autocorrelaciones de residuos por retardo para cada modelo estimado. La tabla incluye los intervalos de confianza para las autocorrelaciones. Esta tabla sólo está disponible si se estiman de nuevo los parámetros de modelo (**Estimar de nuevo a partir de los datos** en la pestaña Modelos).
- **Función de autocorrelación parcial (FAP) de residuo.** 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 estiman de nuevo los parámetros de modelo (**Estimar de nuevo a partir de los datos** en la pestaña Modelos).

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

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 previsiones).

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 estacionario
- *R* cuadrado
- Raíz de la media cuadrática de los errores
- Media del error porcentual absoluto
- Media del error absoluto
- Error porcentual absoluto máximo
- Error absoluto máximo
- BIC normalizado
- Función de autocorrelación simple (FAS) de residuo
- Función de autocorrelación parcial (FAP) de residuo

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 previsión.
- **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 previsión.
- **Intervalos de confianza de los valores ajustados.** Son los intervalos de confianza para el período de estimación.

Función de autocorrelación simple (FAS) de residuo. Muestra un gráfico con las autocorrelaciones de residuos para cada modelo estimado.

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

Limitación de resultados para los modelos de mejor o peor ajuste

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. De forma predeterminada, 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 predeterminado es R cuadrado estacionaria.

Almacenamiento de predicciones de modelos y especificaciones de modelo

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. De forma predeterminada, 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 predeterminados. 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.

Las especificaciones de modelo **Exportar archivo 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).

- **Archivo XML.** Las especificaciones de modelo se guardan en un archivo XML que se puede utilizar con IBM SPSS aplicaciones.
- **Archivo PMML.** Las especificaciones de modelo se guardan en un archivo XML compatible con PMML que se puede utilizar con aplicaciones compatibles con PMML, incluidas aplicaciones IBM SPSS.

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 perdidos del usuario. Estas opciones controlan el tratamiento de los valores perdidos del usuario.

- **Tratar como no válidos.** Los valores perdidos del usuario reciben el mismo tratamiento que los valores perdidos del sistema.
- **Tratar como válidos.** Los valores perdidos del 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 del 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 previsión, el procedimiento genera una advertencia y realiza previsiones en la medida de lo posible.

Ancho de intervalo de confianza (%). Los intervalos de confianza se calculan para las predicciones del modelo y las autocorrelaciones residuales. Puede especificar cualquier valor positivo menor que 100. De forma predeterminada, 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).

Características adicionales del comando TSAPPLY

Las características 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.

Consulte la *Referencia de sintaxis de comandos* para obtener información completa de la sintaxis.

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 mediciones 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.

Estadísticas. El conjunto de factores estacionales.

Descomposición estacional: Consideraciones sobre los datos

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

1. En los menús seleccione:

Analizar > Predicciones > Descomposición estacional...

2. 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.* 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 iguales.* Las medias móviles se calculan con una amplitud igual a la periodicidad y con todos los puntos ponderados por igual. 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

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

- *Añadir a 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.
- *Reemplazar las existentes.* Las nuevas series creadas por la descomposición estacional se guardan como variables temporales 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 archivo de datos de trabajo.

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.

Características adicionales del comando SEASON

La sintaxis de comandos también le permite:

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

Consulte la *Referencia de sintaxis de comandos* para obtener información completa de la sintaxis.

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 necesiten conocer a la hora de evaluar las cifras actuales?

Estadísticas. 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.

Consideraciones sobre los datos de diagramas espectrales

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.

- *Estacionario.* Una condición que debe cumplir la serie temporal a la que se ajusta 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

1. En los menús seleccione:

Análisis > Serie temporal > Análisis espectral...

Nota: Los campos resaltados en rojo son obligatorios. Los botones **Pegar** y **Aceptar** se activan después de introducir valores válidos en todos los campos obligatorios.

2. Seleccione una o más variables de la lista disponible y muévalas a la lista de variables. Observe que la lista sólo incluye variables numéricas.
3. 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 $W_k = .54D_p(2 \pi f_k) + .23D_p(2 \pi f_k + \pi/p) + .23D_p(2 \pi f_k - \pi/p)$, para $k = 0, \dots, p$, donde p es la parte entera de la mitad del tramo y D_p es el kernel de Dirichlet de orden p .
- *Tukey*. Las ponderaciones son $W_k = 0.5D_p(2 \pi f_k) + 0.25D_p(2 \pi f_k + \pi/p) + 0.25D_p(2 \pi f_k - \pi/p)$, para $k = 0, \dots, p$, donde p es la parte entera de la mitad del tramo 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 del tramo y $F[p/2]$ es el kernel de Fejer de orden $p/2$.
- *Bartlett*. La forma de una ventana espectral para la cual las ponderaciones de la mitad superior de la ventana se calculan como $W_k = F_p(2 \pi f_k)$, para $k = 0, \dots, p$, donde p es la parte entera de la mitad del tramo y F_p es el kernel de Fejer de orden p . La mitad inferior es simétrica con la mitad superior.
- *Daniell (Unidad)*. La forma de una ventana espectral cuyas ponderaciones son todas iguales a 1.
- *Ninguno*. Sin suavizado. Si se selecciona esta opción, la estimación de la densidad espectral es la misma que el periodograma.

Expandir. Rango de valores consecutivos a través del cual 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.

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*. Gráfico sin suavizar de amplitud espectral (trazado en una escala logarítmica) con respecto a la frecuencia o 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, que es una medida de la correlación de los componentes de frecuencia fuera de fase de dos series temporales. 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.
- *Densidad espectral de potencia*. Densidad espectral de potencia (o densidad de potencia o espectro de densidad de potencia o espectro de potencia) de la señal describe la potencia presente en la señal en función de la frecuencia, por unidad de frecuencia.
- *coherencia*. La coherencia es un estadístico que se puede utilizar para examinar la relación entre dos señales o conjuntos de datos en el proceso de señales. Se utiliza comúnmente para estimar la transferencia de potencia entre la entrada y la salida de un sistema lineal.
- *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 de una de las series. Cada una de las dos series tiene su propio valor de ganancia.

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

Características adicionales del comando SPECTRA

La sintaxis de comandos también le permite:

- 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.

Consulte la *Referencia de sintaxis de comandos* para obtener información completa de la sintaxis.

Modelos causales temporales

El modelado causal temporal intenta descubrir relaciones causales clave en datos de series temporales. En el modelado causal temporal, se especifica un conjunto de series objetivo y un conjunto de entradas candidato de dichos objetivos. El procedimiento crea un modelo de serie temporal autorregresivo para cada objetivo y solo incluye las entradas que tienen una relación causal con el objetivo. Este enfoque difiere del modelado de serie temporal tradicional donde debe especificar explícitamente los predictores para una serie de objetivos. Puesto que el modelado causal temporal normalmente implica crear modelos para varias series temporal relacionadas, se hace referencia al resultado como un *sistema de modelo*.

En el contexto del modelado causal temporal, el término *causal* hace referencia a la causalidad Granger. En una serie temporal X se indica que la "causa Granger" provoca otra serie temporal Y si realiza una regresión para Y en términos de valores pasados de ambos resultados, X e y, en un modelo mejor para Y que realiza una regresión solo en los valores pasados en Y.

Ejemplos

Los tomadores de decisiones empresariales pueden utilizar el modelado causal temporal para descubrir relaciones causales en un conjunto grande de métricas basadas en tiempo que describen el negocio. El análisis puede revelar unas pocas entradas controlables, que tienen el mayor impacto en los indicadores de rendimiento clave.

Los gestores de sistemas de TI grandes pueden utilizar el modelado causal temporal para detectar anomalías en un conjunto grande de métricas operativas interrelacionadas. El modelo causal permite ir más allá de la detección de anomalías y descubrir las causas principales más probables de las anomalías.

Requisitos de campo

Debe haber al menos un objetivo. De forma predeterminada, no se utilizan los campos con un papel predefinido de Ninguno.

Estructura de datos

El modelado causal temporal soporta dos tipos de estructuras de datos.

Datos basados en columna

Para datos basados en columna, cada campo de serie temporal contiene los datos para una sola serie temporal. Esta estructura es la estructura tradicional de los datos de serie temporal, tal como se utiliza en el modelador de series temporales.

Datos multidimensionales

Para los datos multidimensionales, cada campo de serie temporal contiene los datos para varias series temporales. Las series temporales separadas, en un campo particular, se identifican mediante un conjunto de valores de campos categóricos a los que se hace referencia como campos de *dimensión*. Por ejemplo, los datos de ventas para dos canales de ventas diferentes (minorista y web) se podrían almacenar en un único campo *ventas*. Un campo de dimensión que se llama *canal*, con los valores 'minorista' y 'web', identifica los registros que están asociados a cada uno de los dos canales de ventas.

Obtener un modelo causal temporal

Esta función requiere la opción de predicción de estadísticas.

En los menús, seleccione:

Analizar > Previsiones > Crear modelos causales temporales...

1. Si las observaciones se definen mediante un campo de fecha/hora, especifique el campo.
2. Si los datos son multidimensionales, especifique los campos de dimensión que identifican las series temporales.
 - El orden especificado de los campos de dimensión define el orden en el cual aparecen en todos los diálogos y resultados posteriores. Utilice los botones de flecha hacia arriba y abajo para reordenar los campos de dimensión.
 - Puede especificar subconjuntos de valores de dimensión que limitan el análisis a valores particulares de los campos de dimensión. Por ejemplo, si tiene dimensiones para región y marca, puede limitar el análisis a una región específica. Los subconjuntos de dimensiones se aplican a todos los campos de métrica utilizados en el análisis.
 - También puede personalizar el análisis especificando valores de dimensión por rol en el nivel de campo de métrica. Por ejemplo, si tiene una dimensión para el canal de ventas (con valores 'minorista' y 'web') y métricas para ventas y publicidad en estos canales, puede especificar la publicidad web como una entrada para ambas ventas, minorista y web. De forma predeterminada, este tipo de personalización está habilitado y limitado a seleccionar desde un número especificado máximo de valores distintos (250 de forma predeterminada) de cada campo de dimensión.
3. Pulse **Continuar**.

Nota: Los pasos 1, 2 y 3 no se aplican si el conjunto de datos activo tiene una especificación de fecha. Las especificaciones de fecha se crean desde el cuadro de diálogo Definir fechas o el comando DATE .
4. Pulse **Campos** para especificar las series temporales para incluir en el modelo y para especificar cómo se definen las observaciones. Al menos, se debe especificar un campo como objetivo o como entrada y, también, objetivo.
5. Pulse **Especificaciones de datos** para especificar valores opcionales que incluyen el intervalo de tiempo para los valores de análisis, agregación y distribución y el manejo de valores que faltan.
6. Pulse **Opciones de generación** para definir el periodo de estimación, especifique el contenido del resultado y especifique los valores de generación como, por ejemplo, el número máximo de entradas por objetivo.
7. Pulse **Opciones de modelo** para solicitar predicciones, guardar predicciones y exportar el sistema del modelo a un archivo externo.
8. Pulse en **Ejecutar** para ejecutar el procedimiento.

Series temporales a modelo

En la pestaña Campos, utilice el valor de **Serie temporal** para especificar la serie para incluir en el sistema de modelo.

Para datos basados en columnas, el término *serie* tiene el mismo significado que el término *campo*. Para datos multidimensionales, los campos que contienen series temporales se denominan campos *de métrica*. Una serie temporal, para datos multidimensionales, se define mediante un campo de métrica y

un valor para cada uno de los campos de dimensión. Las consideraciones siguientes se aplican a ambos datos, basados en columna y multidimensionales.

- Las series que se han especificado como entradas candidato o como objetivo y entrada se consideran para la inclusión en el modelo para cada objetivo. El modelo para cada objetivo siempre incluye valores retardados del propio objetivo.
- Las series que se especifican como entradas forzadas siempre se incluyen en el modelo para cada objetivo.
- Al menos, se debe especificar una serie como un objetivo o como un objetivo y también entrada.
- Cuando está seleccionado **Utilizar papeles definidos previamente**, los campos que tienen un papel de Entrada se establecen como entradas candidato. Ningún rol predefinido se correlaciona con una entrada forzada.

Datos multidimensionales

Para datos multidimensionales, especifique campos de métrica y papeles asociados en una cuadrícula, donde cada fila de la cuadrícula especifica una única métrica y un único papel. De forma predeterminada, el sistema del modelo incluye series para todas las combinaciones de los campos de dimensión para cada fila de la cuadrícula. Por ejemplo, si hay dimensiones para *región* y *marca*, de forma predeterminada, especificar la métrica *ventas* como un objetivo significa que hay una serie de objetivos de ventas individual para cada combinación de *región* y *marca*.

Para cada fila de la cuadrícula, puede personalizar el conjunto de valores para cualquiera de los campos de dimensión pulsando el botón de puntos suspensivos para una dimensión. Esta acción abre el subdiálogo Seleccionar valores de dimensión. También puede añadir, suprimir o copiar filas de cuadrícula.

La columna **Recuento de series** muestra el número de conjuntos de valores de dimensión que se han especificado actualmente para la métrica asociada. El valor visualizado puede ser mayor que el número real de series (una serie por conjunto). Esta condición se produce cuando algunas de las combinaciones especificadas de valores de dimensión no corresponden a las series incluidas por la métrica asociada.

Seleccionar valores de dimensión

Para datos multidimensionales, puede personalizar los análisis especificando qué valores de dimensión aplicar a un campo de métrica particular con un papel determinado. Por ejemplo, si *ventas* es un campo de métrica y *canal* es una dimensión con valores 'minorista' y 'web', puede especificar que las ventas 'web' es una entrada y que las ventas 'minorista' es un objetivo. También puede especificar subconjuntos de dimensiones que se aplican a todos los campos de métrica utilizados en el análisis. Por ejemplo, si *región* es un campo de dimensión que indica una región geográfica, puede limitar el análisis a regiones particulares.

Todos los valores

Especifica que todos los valores del campo de dimensión actual se incluyen. Esta opción es la predeterminada.

Seleccionar valores que se van a incluir o excluir

Utilice esta opción para especificar el conjunto de valores para el campo de dimensión actual. Cuando está seleccionado **Incluir** para el **Modo**, solo se incluyen los valores que se han especificado en la lista **Valores seleccionados**. Cuando está seleccionado **Excluir** para el **Modo**, se incluyen todos los valores distintos a los valores que se han especificado en la lista **Valores seleccionados**.

Puede filtrar el conjunto de valores entre los que elegir. Los valores que cumplen la condición de filtro aparecen en la pestaña **Coincidentes** y los valores que no cumplen la condición de filtro aparecen en la pestaña **No coincidentes** de la lista **Valores no seleccionados**. La pestaña **Todos** lista todos los valores no seleccionados, independientemente de cualquier condición de filtro.

- Puede utilizar asteriscos (*) para indicar caracteres comodín cuando especifique un filtro.
- Para borrar el filtro actual, especifique un valor vacío para el término de búsqueda en el diálogo Filtrar valores visualizados.

Observaciones

En la pestaña Campos, utilice el valor **Observaciones** para especificar los campos que definen las observaciones.

Nota: Si el conjunto de datos activo tiene una especificación de fecha, las observaciones se definen mediante la especificación de fecha y no se pueden modificar en el procedimiento de modelado causal temporal. Las especificaciones de fecha se crean desde el cuadro de diálogo Definir fechas o el comando DATE .

Observaciones que se definen por fecha/horas

Puede especificar que las observaciones se definen mediante un campo con formato de fecha, hora o fecha/hora, o mediante un campo de cadena que representa una fecha/hora. Los campos de cadena pueden representar una fecha en el formato AAAA-MM-DD, un hora en formato HH:MM:SS o una hora/fecha en el formato AAAA-MM-DD HH:MM:SS. Los ceros iniciales se pueden omitir en la representación de la cadena. Por ejemplo, la cadena 2014-9-01 es equivalente a 2014-09-01.

Además del campo que define las observaciones, seleccione el intervalo de tiempo apropiado que describe las observaciones. En función del intervalo de tiempo especificado, también puede especificar otros valores como, por ejemplo, el intervalo entre observaciones (incremento) o el número de días por semana. Las consideraciones siguientes se aplican al intervalo de tiempo:

- Utilice el valor **Irregular** cuando las observaciones se asignan de forma irregular en el tiempo como, por ejemplo, la hora a la que se procesa un pedido de compra. Si está seleccionado **Irregular**, debe especificar el intervalo de tiempo que se utiliza para el análisis, desde la configuración de **Intervalo de tiempo** en la pestaña Especificaciones de datos.
- Cuando las observaciones representan una fecha y una hora y el intervalo de tiempo es horas, minutos o segundos, utilice **Horas del día**, **Minutos del día** o **Segundos del día**. Cuando las observaciones representan un periodo de tiempo (duración) sin referencia a una fecha y el intervalo de tiempo es horas, minutos o segundos, utilice **Horas (no periódico)**, **Minutos (no periódico)** o **Segundos (no periódico)**.
- Basándose en el intervalo de tiempo seleccionado, el procedimiento puede detectar observaciones que faltan. La detección de observaciones que faltan es necesaria porque el procedimiento presupone que todas las observaciones se han espaciado de forma uniforme en el tiempo y que no falta ninguna observación. Por ejemplo, si el intervalo de tiempo es días y la fecha 2014-10-27 viene seguida por 2014-10-29, falta una observación para 2014-10-28. Los valores se imputan para cualquier observación que falta. La configuración para manejar los valores que faltan se puede especificar en la pestaña Especificaciones de datos.
- El intervalo de tiempo especificado permite al procedimiento detectar varias observaciones en el mismo intervalo de tiempo que se deben agregar juntas y alinear observaciones en un límite de intervalo como, por ejemplo, el primer día del mes, para garantizar que las observaciones se espacian de forma uniforme. Por ejemplo, si el intervalo de tiempo es Meses, varias fechas del mismo mes se pueden agregar juntas. Se hace referencia a este tipo de agregación como *agrupación*. De forma predeterminada, las observaciones se suman cuando se agrupan. Puede especificar un método diferente para la agrupación, como la media de las observaciones, desde la configuración **Agregación y distribución** en la pestaña Especificaciones de datos.
- Para algunos intervalos de tiempo, los valores adicionales pueden definir saltos en los intervalos normales espaciados de forma uniforme. Por ejemplo, si el intervalo de tiempo es Días, pero solo son válidos los fines de semana, puede especificar que hay cinco días en una semana y que la semana empieza el lunes.

Observaciones que se definen mediante periodos o periodos cíclicos

Las observaciones se pueden definir mediante uno o más campos de enteros que representan periodos o ciclos repetitivos de periodos, hasta un número arbitrario de niveles de ciclo. Con esta estructura, puede describir series de observaciones que no caben en uno de los intervalos de tiempo estándar. Por ejemplo, un año fiscal con solo 10 meses se puede describir con un campo de ciclo que representa años y un campo de periodo que representa meses, donde la longitud de un ciclo es 10.

Los campos que especifican periodos cíclicos definen una jerarquía de niveles periódicos, donde el nivel más bajo se define mediante el campo **Periodo**. El siguiente nivel más alto se especifica mediante un campo de ciclo cuyo nivel es 1, seguido de un campo de ciclo cuyo nivel es 2 y, así, sucesivamente. Los valores de campo para cada nivel, excepto para el más alto, deben ser periódicos con respecto al siguiente nivel superior. Los valores para el nivel superior no pueden ser periódicos. Por ejemplo, en el caso de un año fiscal de 10 meses, los meses son periódicos dentro de los años y los años no son periódicos.

- La longitud de un ciclo en un nivel particular es la periodicidad del siguiente nivel inferior. Para el ejemplo del año fiscal, solo hay un nivel de ciclo y la longitud del ciclo es 10 porque el siguiente nivel inferior representa meses y hay 10 meses en el año fiscal especificado.
- Especifique el valor inicial para cualquier campo periódico que no comience en 1. Este valor es necesario para detectar valores perdidos. Por ejemplo, si un campo periódico empieza en el 2, pero el valor inicial se especifica como 1, el procedimiento supone que falta un valor para el primer periodo de cada ciclo de dicho campo.

Observaciones que se definen mediante el orden de registro

Para datos basados en columna, puede especificar que las observaciones se definen mediante el orden de registro, de forma que el primer registro representa la primera observación, el segundo registro representa la segunda observación y, así, sucesivamente. Se supone que los registros representan observaciones que están espaciadas de forma uniforme en el tiempo.

Intervalo de tiempo del análisis

El intervalo de tiempo que se utiliza para el análisis puede diferir del intervalo de tiempo de las observaciones. Por ejemplo, si el intervalo de tiempo de las observaciones es Días, podría elegir Meses para el intervalo de tiempo del análisis. Los datos se añaden después a partir de datos de diarios a mensuales antes de que se genere el modelo. También puede elegir distribuir los datos de un intervalo de tiempo más largo a uno más corto. Por ejemplo, si las observaciones son trimestrales, puede distribuir los datos de datos trimestrales a datos mensuales.

Las opciones disponibles para el intervalo de tiempo en el que se realiza el análisis dependen de cómo se han definido las observaciones y del intervalo de tiempo de estas observaciones. En concreto, cuando las observaciones se han definido mediante periodos cíclicos o cuando una especificación de fecha se ha definido para el conjunto de datos activo, solo se da soporte a la agregación. En este caso, el intervalo de tiempo del análisis debe ser mayor o igual que el intervalo de tiempo de las observaciones.

El intervalo de tiempo para el análisis se especifica desde la configuración de **Intervalo de tiempo** en la pestaña Especificaciones de datos. El método mediante el cual los datos se agregan o distribuyen se especifica en la configuración de **Agregación y distribución** en la pestaña Especificaciones de datos.

Agregación y distribución

Funciones de agregación

Cuando el intervalo de tiempo que se utiliza para el análisis es más largo que el intervalo de tiempo de las observaciones, se agregan los datos de entrada. Por ejemplo, la agregación se realiza cuando el intervalo de tiempo de las observaciones es Días y el intervalo de tiempo para el análisis es Meses. Están disponibles las funciones de agregación siguientes: media, suma, moda, mín o máx.

Funciones de distribución

Cuando el intervalo de tiempo que se utiliza para el análisis es más corto que el intervalo de tiempo de las observaciones, se distribuyen los datos de entrada. Por ejemplo, la distribución se realiza cuando el intervalo de tiempo de las observaciones es Trimestres y el intervalo de tiempo para el análisis es Meses. Están disponibles las funciones de distribución siguientes: media o suma.

Funciones de agrupación

La agrupación se aplica cuando las observaciones se definen mediante fecha/horas y se producen varias observaciones en el mismo intervalo de tiempo. Por ejemplo, si el intervalo de tiempo de las observaciones es Meses, se agrupan varias fechas del mismo mes y se asocian al mes en el cual se producen. Están disponibles las funciones de agregación siguientes: media, suma, modo, mín o

máx. La agrupación siempre se realiza cuando las observaciones se definen mediante fecha/horas y el intervalo de tiempo de las observaciones se especifica como Irregular.

Nota: Aunque la agrupación es una forma de agregación, se realiza antes de manejar los valores que faltan, mientras que la agregación formal se realiza después de que se manejen los valores que faltan. Cuando el intervalo de tiempo de las observaciones se especifica como irregular, la agregación solo se realiza con la función de agrupación.

Agregar observaciones del día al día anterior

Especifica si las observaciones con horas que cruzan un límite de día se agregan a los valores del día anterior. Por ejemplo para las observaciones por hora con un día de ocho horas que empieza a las 20:00, este valor especifica si las observaciones entre 00:00 y 04:00 se incluyen en los resultados agregados para el día anterior. Este valor se aplica solo si el intervalo de tiempo de las observaciones es Horas por día, Minutos por día o Segundos por día y el intervalo de tiempo para el análisis es Días.

Valores personalizados para campos especificados

Puede especificar funciones de agregación, distribución y agrupación en un campo por campo. Estos valores alteran temporalmente los valores predeterminados para las funciones de agregación, distribución y agrupación.

Valores que faltan

Los valores que faltan en los datos de entrada se sustituyen con un valor imputado. Están disponibles los métodos de sustitución siguientes:

Interpolación lineal

Sustituye los valores perdidos utilizando una interpolación lineal. Se utilizan para la interpolación el último valor válido antes del valor perdido y el primer valor válido después del valor perdido. Si la primera o la última observación de la serie tiene un valor que falta, se utilizan los dos valores que no faltan más cercanos al principio o al final de la serie.

Media de series

Sustituye los valores perdidos con la media de la serie completa.

Media de puntos cercanos

Sustituye los valores perdidos por la media de los valores válidos circundantes. La amplitud de los puntos adyacentes es el número de valores válidos anteriores y posteriores del valor que falta que se utilizan para calcular la media.

Mediana de puntos adyacentes

Sustituye los valores perdidos por la mediana de los valores válidos circundantes. La amplitud de los puntos adyacentes es el número de valores válidos antes y después que el valor que falta que se utilizan para calcular la mediana.

Tendencia lineal

Esta opción utiliza todas las observaciones que no faltan en la serie para ajustar un modelo de regresión lineal simple, que se utiliza para imputar los valores que faltan.

Otros valores:

Porcentaje máximo de valores perdidos (%)

Especifica el porcentaje máximo de valores perdidos que están permitidos para cualquier serie. Las series con más valores perdidos que el máximo especificado se excluyen del análisis.

Valores perdidos del usuario

Esta opción especifica si los valores perdidos del usuario se tratan como datos válidos y, por lo tanto, se incluyen en la serie. De forma predeterminada, los valores que faltan del usuario se excluyen y se tratan como valores que faltan del sistema, que se imputan después.

Opciones de datos generales

Número máximo de valores distintos por campo de dimensión

Este valor se aplica a datos multidimensionales y especifica el número máximo de valores distintos que están permitidos para cualquier campo de dimensión. De forma predeterminada, este límite se establece en 10000, pero se puede aumentar a un número arbitrariamente grande.

Opciones de generación generales

Ancho de intervalo de confianza (%)

Este valor controla los intervalos de confianza para ambos parámetros, de previsiones y de modelo. Puede especificar cualquier valor positivo menor que 100. De forma predeterminada, se utiliza un intervalo de confianza del 95%.

Número máximo de entradas para cada objetivo

Este valor especifica el número máximo de entradas que están permitidas en el modelo para cada objetivo. Puede especificar un entero dentro del rango de 1 a 20. El modelo para cada objetivo siempre incluye valores retardados de sí mismo, de modo que establecer este valor en 1 especifica que la entrada solo es el objetivo propio.

Tolerancia del modelo

Este valor controla el proceso iterativo que se utiliza para determinar el mejor conjunto de entradas para cada objetivo. Se puede especificar cualquier valor mayor que cero. El valor predeterminado es 0.001.

Umbral de valores atípicos (%)

Una observación se marca como valor atípico si la probabilidad, tal como se calcula a partir del modelo, de que sea un valor atípico excede este umbral. Puede especificar un valor dentro del rango de 50 a 100.

Número de retardos de cada entrada

Este valor especifica el número de términos de retardo para cada entrada en el modelo para cada objetivo. De forma predeterminada, el número de términos de retardo se determina automáticamente a partir del intervalo de tiempo utilizado para el análisis. Por ejemplo, si el intervalo de tiempo es de meses (con un incremento de un mes), el número de retardos es 12. De forma opcional, se puede especificar explícitamente el número de retardos. El valor especificado debe ser un entero en el rango de 1 a 20.

Series a mostrar

Estas opciones especifican las series (de objetivos o entradas) para los cuales se visualiza el resultado. El contenido del resultado para las series especificadas se determina mediante el valor de **Opciones de resultado**.

Mostrar objetivos asociados con modelos que mejor se ajustan

De forma predeterminada, se visualiza el resultado para los objetivos que están asociados a los 10 modelos de mejor ajuste, según lo que determina el valor R cuadrado. Puede especificar un número fijo diferente de modelos de mejor ajuste o puede especificar un porcentaje de modelos de mejor ajuste. También puede elegir entre las siguientes medidas de bondad de ajuste:

R cuadrado

Medida de la bondad de ajuste de un modelo lineal; en ocasiones recibe el nombre de coeficiente de determinación. Es la proporción de la variación de la variable objetivo explicada por el modelo. Es un valor en un rango de 0 a 1. Los valores pequeños indican que el modelo no se ajusta bien a los datos.

Raíz de la media cuadrática de los errores porcentuales

Una medida de cuánto difieren los valores pronosticados por el modelo de los valores observados de las series. Es independiente de las unidades utilizadas y, por tanto, se puede utilizar para comparar series con distintas unidades.

Raíz de la media cuadrática de los errores

La raíz cuadrada de la media cuadrática de los errores. Una medida de cuánto se desvía la serie dependiente del nivel pronosticado por el modelo, expresado en las mismas unidades que la serie dependiente.

BIC

Criterio de información Bayesian. Una medida para seleccionar y comparar modelos basados en el logaritmo de probabilidad -2 reducida. Los valores menores indican modelos mejores. El BIC también "penaliza" modelos sobreparametrizados (modelos complejos con un gran número de entradas, por ejemplo), pero de forma más estricta que el AIC.

AIC

Criterio de información Akaike. Una medida para seleccionar y comparar modelos basados en el logaritmo de probabilidad -2 reducida. Los valores menores indican modelos mejores. El AIC "penaliza" modelos sobreparametrizados (modelos complejos con un gran número de entradas, por ejemplo).

Especifique series individuales

Puede especificar series individuales para los cuales desea un resultado.

- Para los datos basados en columna, especifique los campos que contienen las series que desea. El orden de los campos especificados define el orden en el cual aparecen en el resultado.
- Para los datos multidimensionales, especifique una serie particular añadiendo una entrada a la cuadrícula para el campo de métrica que contiene la serie. Especifique los valores de los campos de dimensión que definen las series.
 - Puede especificar el valor para cada campo de dimensión directamente en la cuadrícula o puede seleccionar en la lista de valores de dimensión disponibles. Para seleccionar en la lista de valores de dimensión disponibles, pulse el botón de puntos suspensivos en la casilla para la dimensión que desee. Esta acción abre el subdiálogo Seleccionar valor de dimensión.
 - Puede buscar en la lista de valores de dimensión, en el subdiálogo Seleccionar valor de dimensión, pulsando el icono de los prismáticos y especificando un término de búsqueda. Los espacios se tratan como parte del término de búsqueda. Los asteriscos (*) en el término de búsqueda no indican caracteres comodín.
 - El orden de las series de la cuadrícula define el orden en el cual aparecen en el resultado.

Para ambos datos, los que se basan en columna y los multidimensionales, el resultado está limitado a 30 series. Este límite incluye series individuales (entradas u objetivos) que especifique y los objetivos que están asociados a los modelos de mejor ajuste. Las series especificadas de forma individual tienen prioridad sobre los objetivos que están asociados a los modelos de mejor ajuste.

Opciones de resultados

Estas opciones especifican el contenido del resultado. Las opciones del grupo **Objetivos de resultado** generan el resultado para los objetivos que están asociados a los modelos de mejor ajuste en los valores **Series para visualizar**. Las opciones del grupo **Resultado para series** generan el resultado para las series individuales que se han especificado en los valores **Series para visualizar**.

Sistema global de modelo

Muestra una representación gráfica de las relaciones causales entre series en el sistema del modelo. Las tablas tanto en las estadísticas de ajuste del modelo y los valores atípicos para los objetivos visualizado se incluyen como parte del elemento de resultado. Cuando esta opción está seleccionada en el grupo **Resultado para series**, se crea un elemento de resultado separado para cada serie individual que se ha especificado en los valores de **Series para visualizar**.

Las relaciones causales entre las series tienen un nivel de significación asociado, donde un nivel de significación más pequeño indica un conexión más significativa. Puede elegir ocultar las relaciones con un nivel de significación que es mayor que un valor especificado.

Estadísticos de ajuste del modelo y valores atípicos

Las tablas de estadísticas y valores atípicos del ajuste del modelo para las series de objetivos que se han seleccionado para su visualización. Estas tablas contienen la misma información que las tablas en la visualización de Sistema global de modelo. Estas tablas soportan todas las funciones estándar para tablas dinámicas y edición.

Efectos del modelo y parámetros del modelo

Tablas de pruebas de efectos del modelo y parámetros de modelo para las series de objetivos que se han seleccionado para su visualización. Las pruebas de efectos del modelo incluyen la estadística F y un valor de significación asociado para cada entrada incluida en el modelo.

Diagrama de impacto

Muestra una representación gráfica de las relaciones causales entre una serie de interés y otras series a las que afecta o que le afectan. Las series que afectan a las series de interés se denominan *causas*. La selección de **Efectos** genera un diagrama de impacto inicializado para mostrar efectos. La selección de **Causas** genera un diagrama de impacto inicializado para mostrar causas. La selección de **Causas y efectos** genera dos diagramas de impacto independientes, uno inicializado para las causas y otro inicializado para los efectos. Puede alternar de forma interactiva entre causas y efectos en el elemento de resultado que muestra el diagrama de impacto.

Puede especificar el número de niveles de causas o efectos para mostrar, donde el primer nivel es simplemente la serie de interés. Cada nivel adicional muestra más causas o efectos indirectas de las series de interés. Por ejemplo, el tercer nivel de la visualización de efectos está formado por la series que contienen series del segundo nivel como una entrada directa. Las series del tercer nivel son afectadas directamente por las series de interés puesto que las series de interés es una entrada directa en las series del segundo nivel.

Gráfico de series

Gráficos de valores observados y pronosticados para las series de objetivos que se han seleccionado para la visualización. Cuando se solicitan previsiones, el gráfico también muestra los valores pronosticados y los intervalos de confianza para las previsiones.

Gráfico de residuos

Gráficos de los residuos del modelo para las series de objetivos que se han seleccionado para la visualización.

Entradas superiores

Gráficos para cada objetivo visualizado, a lo largo del tiempo, junto con las 3 entradas superiores del objetivo. Las entradas superiores son las entradas con el valor de significación menor. Para acomodar distintas escalas para las entradas y el objetivo, el eje Y representa la puntuación Z para cada serie.

Tabla de predicciones

Tablas de valores pronosticados e intervalos de confianza de estas previsiones para las series de objetivos que se han seleccionado para la visualización.

Análisis de la causa principal del valor atípico

Determina qué series tienen la mayor probabilidad de ser la causa de cada valor atípico en una serie de interés. El análisis de causa principal de valor atípico se realiza para cada serie de objetivos que se incluye en la lista de series individuales en el valor **Series para visualizar**.

Resultado

Tabla y gráfico de valores atípicos interactivos

Tabla y gráfico de valores atípicos y causas principales de estos valores atípicos para cada serie de interés. La tabla contiene una sola fila para cada valor atípico. El gráfico es un diagrama de impactos. Seleccionar una fila en la tabla resalta la ruta, en el diagrama de impactos, de la serie de interés en la serie que es más probable que provoque el valor atípico asociado.

Tabla dinámica de valores atípicos

Tabla de valores atípico y causas principales de estos valores atípico para cada series de interés. Esta tabla contiene la misma información que la tabla en la visualización interactiva. Esta tabla soporta todas las funciones estándar para tablas dinámicas y de edición.

Niveles causales

Puede especificar el número de niveles para incluir en la búsqueda para las causas principales. El concepto de niveles que se utiliza aquí es el mismo que el que se describe para diagramas de impactos.

Ajuste de modelo en todos los modelos

Histograma del ajuste del modelo para todos los modelos y para las estadísticas de ajuste seleccionadas. Están disponibles las estadísticas de ajuste siguientes:

R cuadrado

Medida de la bondad de ajuste de un modelo lineal; en ocasiones recibe el nombre de coeficiente de determinación. Es la proporción de la variación de la variable objetivo explicada por el modelo. Es un valor en un rango de 0 a 1. Los valores pequeños indican que el modelo no se ajusta bien a los datos.

Raíz de la media cuadrática de los errores porcentuales

Una medida de cuánto difieren los valores pronosticados por el modelo de los valores observados de las series. Es independiente de las unidades utilizadas y, por tanto, se puede utilizar para comparar series con distintas unidades.

Raíz de la media cuadrática de los errores

La raíz cuadrada de la media cuadrática de los errores. Una medida de cuánto se desvía la serie dependiente del nivel pronosticado por el modelo, expresado en las mismas unidades que la serie dependiente.

BIC

Criterio de información Bayesian. Una medida para seleccionar y comparar modelos basados en el logaritmo de probabilidad -2 reducida. Los valores menores indican modelos mejores. El BIC también "penaliza" modelos sobreparametrizados (modelos complejos con un gran número de entradas, por ejemplo), pero de forma más estricta que el AIC.

AIC

Criterio de información Akaike. Una medida para seleccionar y comparar modelos basados en el logaritmo de probabilidad -2 reducida. Los valores menores indican modelos mejores. El AIC "penaliza" modelos sobreparametrizados (modelos complejos con un gran número de entradas, por ejemplo).

Valores atípicos en el tiempo

Diagrama de barras del número de valores atípicos, en todos los objetivos, para cada intervalo de tiempo en el periodo de estimación.

Transformaciones de series

Tabla de cualquier transformación que se ha aplicado a la serie en el sistema del modelo. Las posibles transformaciones son una imputación, agregación y distribución de un valor que falta.

Periodo de estimación

De forma predeterminada, el período de estimación inicia la hora de la primera observación y finaliza la hora de la última observación en todas las series.

Por hora de inicio y de finalización

Puede especificar el inicio y la finalización del período de estimación o puede especificar simplemente el inicio o solo el final. Si omite el inicio o la finalización del período de estimación, se utiliza el valor predeterminado.

- Si las observaciones se han definido mediante un campo de fecha/hora, especifique los valores para el inicio y el final con el mismo formato que se utiliza para el campo de fecha/hora.
- Para las observaciones que se han definido mediante periodos cíclicos, especifique un valor para cada uno de los campos de periodos cíclicos. Cada campo se visualiza en una columna separada.
- Si hay una especificación de fecha en vigor para el conjunto de datos activo, debe especificar un valor para cada componente (como Mes) de la especificación de fecha. Cada componente se visualiza en una columna separada.

- Cuando las observaciones se definen mediante el orden de registro, el inicio y la finalización del periodo de estimación se definen mediante el número de fila (tal como se visualiza en el editor de datos) del caso relevante.

Por los primeros o los últimos intervalos de tiempo

Define el período de estimación como un número especificado de intervalos de tiempo que empiezan en el primer intervalo de tiempo o finalizan el último intervalo de tiempo en los datos, con un desplazamiento opcional. En este contexto, el intervalo de tiempo hace referencia al intervalo de tiempo del análisis. Por ejemplo, asuma que las observaciones son mensuales, pero el intervalo de tiempo del análisis es trimestres. Especificar **Último** y un valor de 24 para el **Número de intervalos de tiempo** significa los últimos 24 trimestres.

De forma opcional, puede excluir un número especificado de intervalos de tiempo. Por ejemplo, especificar los últimos 24 intervalos de tiempo y 1 para el número para excluir significa que el periodo de estimación consta de los 24 intervalos que preceden al último.

Predicción

La opción de **Extender registros en el futuro** define el número de intervalos de tiempo para prever más allá del final del periodo de estimación. El intervalo de tiempo en este caso es el intervalo de tiempo del análisis, que se especifica en la pestaña Especificaciones de datos. Cuando se solicitan predicciones, se generan modelos autorregresivos automáticamente para cualquier serie de entrada que tampoco son objetivos. Estos modelos se utilizan para generar valores para estas series de entrada en el periodo de predicción.

Guardar

Opciones de destino

Puede guardar ambas cosas, las transformaciones de los datos (como agregación o imputación de valores que faltan) y las variables nuevas (especificadas en el valor **Guardar objetivos**) en un archivo de datos o un nuevo conjunto de datos de IBM SPSS Estadística en la sesión actual. Los valores fecha/hora de los datos guardados se alinean con el comienzo de cada intervalo de tiempo, como el primero del mes, y representan el intervalo de tiempo del análisis para el sistema modelo. Puede guardar cualquiera de las variables nuevas en el conjunto de datos activo sólo si las observaciones se definen mediante una especificación de fecha, una orden de registro y los datos no están agregados.

Guardar objetivos

Puede guardar predicciones de modelo, intervalos de confianza y residuos como variables nuevas. Cada objetivo especificado para guardar genera su propio conjunto de variables nuevas y cada variable nueva contiene valores para ambos periodos, el de estimación y el de predicción. Para obtener valores pronosticados, intervalos de confianza y residuos de ruido, puede especificar el nombre raíz que se va a utilizar como el prefijo para las variables nuevas. El nombre de variable completo es la concatenación del nombre raíz y el nombre del campo que contiene las series de objetivos. El nombre raíz debe cumplir las reglas para los nombres de variable válidos. El nombre de la variable se amplía en caso de ser necesario para evitar conflictos de nombres de variables.

Indicar casos que contienen previsiones

Crea una variable que indica si un registro contiene datos de previsión. Puede especificar el nombre de variable. El valor predeterminado es *ForecastIndicator*.

Objetivos para guardar

Especifica si se crean variables nuevas para todas las series de objetivos en el sistema del modelo o solo las series objetivos que se han especificado en los valores **Series para visualizar**.

Exportar sistema de modelo

Guardar el sistema del modelo en un archivo de archivado comprimido (archivo .zip). El procedimiento de la predicción del modelo causal temporal puede utilizar el archivo del sistema de modelo para obtener predicciones actualizadas o para generar cualquiera de los resultados disponibles. El procedimiento de escenarios de modelos causales temporales también lo puede utilizar para ejecutar el análisis de escenarios.

Resultados interactivos

El resultado del modelo causal temporal incluye un número de objetos de resultado interactivo. Las características interactivas están disponibles si se activa (se pulsa dos veces) el objeto del visor de resultados.

Sistema global de modelo

Muestra las relaciones causales entre las series en el sistema de modelo. Todas las líneas que conectan un destino concreto con sus entradas tienen el mismo color. El grosor de la línea indica la importancia de la conexión causal, donde las líneas más gruesas representan una conexión más importante. Las entradas que no sean también destinos se indican con un cuadrado negro.

- Puede mostrar relaciones para los modelos superiores, un serie especificada, todas las series o modelos sin entradas. Los modelos superiores son los modelos que cumplen los criterios que se han especificado para modelos de mejor ajuste en la configuración de **Series para visualizar**.
- Puede generar diagramas de impactos para una o más series seleccionando los nombres de serie en el gráfico, pulsando con el botón derecho del ratón y, después, eligiendo **Crear diagrama de impactos** en el menú contextual.
- Puede elegir ocultar las relaciones causales que tienen un nivel de significación que es mayor que un valor especificado. Los nivel de significación más pequeños indican una relación causal más significativa.
- Puede mostrar las relaciones de una serie concreta seleccionando el nombre de la serie en el diagrama, pulsando con el botón derecho del ratón y seleccionando a continuación **Resaltar relaciones para series** en el menú contextual.

Diagrama de impacto

Muestra una representación gráfica de las relaciones causales entre una serie de interés y otras series a las que afecta o que le afectan. Las series que afectan a las series de interés se denominan *causas*.

- Puede cambiar la serie de interés especificando el nombre de la serie que desea. Si pulsa dos veces cualquier nodo en el diagrama de impactos cambia la serie de interés por la serie asociada a dicho nodo.
- Puede conmutar la visualización entre causas y efectos y puede cambiar el número de niveles de causas o efectos para mostrar.
- Si pulsa una vez en cualquier nodo se abre un diagrama de secuencia detallada para la serie que está asociada al nodo.

Análisis de la causa principal del valor atípico

Determina qué series tienen la mayor probabilidad de ser la causa de cada valor atípico en una serie de interés.

- Puede mostrar la causa principal para cualquier valor atípico seleccionando la fila para el valor atípico en la tabla Valores atípicos. También puede mostrar la causa principal pulsando el icono para el valor atípico en el diagrama de la secuencia.
- Si pulsa una vez en cualquier nodo se abre un diagrama de secuencia detallada para la serie que está asociada al nodo.

Calidad global de modelo

Histograma del ajuste del modelo para todos los modelos, para una estadística de ajuste determinada. Pulsar una barra en el diagrama de barras filtra el gráfico de puntos de forma que solo muestra los modelos que están asociados a la barra seleccionada. Puede encontrar el modelo para una serie de objetivos concretos en el gráfico de puntos especificando el nombre de la serie.

Distribución de valores atípicos

Diagrama de barras del número de valores atípicos, en todos los objetivos, para cada intervalo de tiempo en el periodo de estimación. Al pulsar una barra del diagrama de barras se filtra el gráfico de puntos, de manera que sólo muestra los valores atípicos que están asociados a la barra seleccionada.

Nota:

- Si guarda un documento de resultados que contiene resultados interactivos del modelado causal temporal y desea conservar las características interactivas, asegúrese de que **Almacenar información del modelo necesaria con el documento de resultados** está seleccionado en el diálogo Guardar resultados como.
- Algunas características interactivas necesitan que el conjunto de datos activo sean los mismos datos que se han utilizado para construir el sistema del modelo causal temporal.

Aplicación de modelos causales temporales

Aplicación de modelos causales temporales

Hay disponibles dos procedimientos para aplicar modelos que se han creado con el procedimiento de modelado causal temporal. Ambos procedimientos requieren el archivo de sistema del modelo, que se puede guardar como parte del procedimiento de modelado causal temporal.

Predicción de modelos causales temporales

Puede utilizar este procedimiento para obtener predicciones para series para las que están disponibles datos más actuales, sin volver a crear los modelos. También puede generar cualquier resultado que esté disponible con el procedimiento de modelado causal temporal.

Escenarios de modelos causales temporales

Utilice este procedimiento para investigar cómo los valores específicos de una serie temporal particular, en un sistema del modelo, afectan a los valores pronosticados de la serie temporal que están relacionados de forma causal.

Predicción de modelos causales temporales

El procedimiento de predicción de modelos causales temporales carga un archivo de sistema de modelo que se ha creado mediante el procedimiento de modelado causal temporal y se aplica a los modelos en el conjunto de datos activo. Puede utilizar este procedimiento para obtener predicciones para series para las que están disponibles datos más actuales, sin volver a crear los modelos. También puede generar cualquier resultado que esté disponible con el procedimiento de modelado causal temporal.

Supuestos

- La estructura de los datos del conjunto de datos activo, ya se basen en columnas o multidimensionales, debe ser la misma estructura que se haya utilizado cuando se creó el sistema del modelo. Para datos multidimensionales, los campos de dimensión deben ser los mismos que se utilizan para crear el sistema del modelo. Además, los valores de dimensión que se han utilizado para crear el sistema del modelo deben existir en el conjunto de datos activo.
- Los modelos se aplican a campos del conjunto de datos activo con los mismos nombres que los campos especificados en el sistema del modelo.
- El campo, o campos, que definía las observaciones cuando se creó el sistema del modelo debe existir en el conjunto de datos activo. El intervalo de tiempo entre observaciones se supone que es el mismo que cuando se crearon los modelos. Si las observaciones se han definido mediante una especificación de fecha, debe existir la misma especificación de fecha en el conjunto de datos activo. Las especificaciones de fecha se crean desde el cuadro de diálogo Definir fechas o el comando DATE .
- El intervalo de tiempo del análisis y cualquier valor para la agregación, distribución y valores que faltan son iguales que cuando se crearon los modelos.

Utilizar predicciones de modelos causales temporales

Esta función requiere la opción de predicción de estadísticas.

En los menús, seleccione:

Analizar > Previsiones > Aplicar modelos causales temporales...

1. Especifique la especificación de archivo para un archivo de sistema de modelo o pulse **Examinar** y seleccione un archivo de sistema de modelo. Los archivos de sistema de modelo se crean con el procedimiento de modelado causal temporal.
2. Pulse la opción para volver a estimar modelos, crear predicciones y generar resultados.
3. Pulse **Continuar**.
4. Especifique si desea utilizar parámetros de modelo existentes o volver a estimar parámetros de modelo a partir de los datos del conjunto de datos activo.
5. Especifique hasta cuándo se realizarán predicciones en el futuro, o especifique que no hay predicciones.
6. Pulse **Opciones** para especificar el contenido del resultado.
7. Pulse **Guardar** para guardar predicciones y exporte el sistema del modelo actualizado a un archivo externo cuando se vuelvan a estimar parámetros de modelo.
8. Pulse en **Ejecutar** para ejecutar el procedimiento.

Parámetros y predicciones de modelo

Cargar de archivo de modelo

Las predicciones se generan mediante el uso de parámetros del archivo del sistema de modelo, y los datos del conjunto de datos activo, sin volver a estimar los parámetros del modelo. Las medidas de bondad de ajuste que se visualizan en el resultado y se utilizan para seleccionar los modelos de mejor ajuste se obtienen del archivo del sistema de modelo. Las medidas de ajuste reflejan los datos que se han utilizado cuando se ha creado cada modelo (o en la última actualización). Esta opción es apropiada para generar predicciones y resultados a partir de los datos que se han utilizado para crear el sistema de modelo.

Estimar de nuevo a partir de los datos

Los parámetros del modelo se vuelven a estimar utilizando los datos del conjunto de datos activo. Volver a estimar los parámetros del modelo no afecta a qué entradas se incluyen en el modelo para cada objetivo. Esta opción es apropiada cuando tiene datos nuevos más allá del período de estimación original y desea generar predicciones u otros resultados con los datos actualizados.

Todas las observaciones

Especifique que el período de estimación se inicia a la hora de la primera observación y finaliza a la hora de la última observación de todas las series.

Por hora de inicio y de finalización

Puede especificar el inicio y la finalización del período de estimación o puede especificar simplemente el inicio o solo el final. Si omite el inicio o la finalización del período de estimación, se utiliza el valor predeterminado.

- Si las observaciones se han definido mediante un campo de fecha/hora, especifique los valores para el inicio y el final con el mismo formato que se utiliza para el campo de fecha/hora.
- Para las observaciones que se han definido mediante periodos cíclicos, especifique un valor para cada uno de los campos de periodos cíclicos. Cada campo se visualiza en una columna separada.
- Si hay una especificación de fecha en vigor para el conjunto de datos activo, debe especificar un valor para cada componente (como Mes) de la especificación de fecha. Cada componente se visualiza en una columna separada.
- Cuando las observaciones se definen mediante el orden de registro, el inicio y la finalización del período de estimación se definen mediante el número de fila (tal como se visualiza en el editor de datos) del caso relevante.

Por los primeros o los últimos intervalos de tiempo

Define el período de estimación como un número especificado de intervalos de tiempo que empiezan en el primer intervalo de tiempo o finalizan el último intervalo de tiempo en los datos, con un desplazamiento opcional. En este contexto, el intervalo de tiempo hace referencia al

intervalo de tiempo del análisis. Por ejemplo, asuma que las observaciones son mensuales, pero el intervalo de tiempo del análisis es trimestres. Especificar **Último** y un valor de 24 para el **Número de intervalos de tiempo** significa los últimos 24 trimestres.

De forma opcional, puede excluir un número especificado de intervalos de tiempo. Por ejemplo, especificar los últimos 24 intervalos de tiempo y 1 para el número para excluir significa que el periodo de estimación consta de los 24 intervalos que preceden al último.

Extender registros en el futuro

Define el número de intervalos de tiempo para prever más allá del final del periodo de estimación. El intervalo de tiempo en este caso es el intervalo de tiempo del análisis. Cuando se solicitan predicciones, se generan modelos autorregresivos automáticamente para cualquier serie de entrada que tampoco son objetivos. Estos modelos se utilizan para generar valores para estas series de entrada en el periodo de predicción para obtener las predicciones de los objetivos de estas entradas.

Opciones generales

Ancho de intervalo de confianza (%)

Este valor controla los intervalos de confianza para ambos parámetros, de previsiones y de modelo. Puede especificar cualquier valor positivo menor que 100. De forma predeterminada, se utiliza un intervalo de confianza del 95%.

Umbral de valores atípicos (%)

Una observación se marca como valor atípico si la probabilidad, tal como se calcula a partir del modelo, de que sea un valor atípico excede este umbral. Puede especificar un valor dentro del rango de 50 a 100.

Serie que se va a visualizar

Estas opciones especifican las series (de objetivos o entradas) para los cuales se visualiza el resultado. El contenido del resultado para las series especificadas se determina mediante el valor de **Opciones de resultado**.

Mostrar objetivos asociados con modelos que mejor se ajustan

De forma predeterminada, se visualiza el resultado para los objetivos que están asociados a los 10 modelos de mejor ajuste, según lo que determina el valor R cuadrado. Puede especificar un número fijo diferente de modelos de mejor ajuste o puede especificar un porcentaje de modelos de mejor ajuste. También puede elegir entre las siguientes medidas de bondad de ajuste:

R cuadrado

Medida de la bondad de ajuste de un modelo lineal; en ocasiones recibe el nombre de coeficiente de determinación. Es la proporción de la variación de la variable objetivo explicada por el modelo. Es un valor en un rango de 0 a 1. Los valores pequeños indican que el modelo no se ajusta bien a los datos.

Raíz de la media cuadrática de los errores porcentuales

Una medida de cuánto difieren los valores pronosticados por el modelo de los valores observados de las series. Es independiente de las unidades utilizadas y, por tanto, se puede utilizar para comparar series con distintas unidades.

Raíz de la media cuadrática de los errores

La raíz cuadrada de la media cuadrática de los errores. Una medida de cuánto se desvía la serie dependiente del nivel pronosticado por el modelo, expresado en las mismas unidades que la serie dependiente.

BIC

Criterio de información Bayesian. Una medida para seleccionar y comparar modelos basados en el logaritmo de probabilidad -2 reducida. Los valores menores indican modelos mejores. El BIC también "penaliza" modelos sobreparametrizados (modelos complejos con un gran número de entradas, por ejemplo), pero de forma más estricta que el AIC.

AIC

Criterio de información Akaike. Una medida para seleccionar y comparar modelos basados en el logaritmo de probabilidad -2 reducida. Los valores menores indican modelos mejores. El AIC "penaliza" modelos sobreparametrizados (modelos complejos con un gran número de entradas, por ejemplo).

Especifique series individuales

Puede especificar series individuales para los cuales desea un resultado.

- Para los datos basados en columna, especifique los campos que contienen las series que desea. El orden de los campos especificados define el orden en el cual aparecen en el resultado.
- Para los datos multidimensionales, especifique una serie particular añadiendo una entrada a la cuadrícula para el campo de métrica que contiene la serie. Especifique los valores de los campos de dimensión que definen las series.
 - Puede especificar el valor para cada campo de dimensión directamente en la cuadrícula o puede seleccionar en la lista de valores de dimensión disponibles. Para seleccionar en la lista de valores de dimensión disponibles, pulse el botón de puntos suspensivos en la casilla para la dimensión que desee. Esta acción abre el subdiálogo Seleccionar valor de dimensión.
 - Puede buscar en la lista de valores de dimensión, en el subdiálogo Seleccionar valor de dimensión, pulsando el icono de los prismáticos y especificando un término de búsqueda. Los espacios se tratan como parte del término de búsqueda. Los asteriscos (*) en el término de búsqueda no indican caracteres comodín.
 - El orden de las series de la cuadrícula define el orden en el cual aparecen en el resultado.

Para ambos datos, los que se basan en columna y los multidimensionales, el resultado está limitado a 30 series. Este límite incluye series individuales (entradas u objetivos) que especifique y los objetivos que están asociados a los modelos de mejor ajuste. Las series especificadas de forma individual tienen prioridad sobre los objetivos que están asociados a los modelos de mejor ajuste.

Opciones de resultados

Estas opciones especifican el contenido del resultado. Las opciones del grupo **Objetivos de resultado** generan el resultado para los objetivos que están asociados a los modelos de mejor ajuste en los valores **Series para visualizar**. Las opciones del grupo **Resultado para series** generan el resultado para las series individuales que se han especificado en los valores **Series para visualizar**.

Sistema global de modelo

Muestra una representación gráfica de las relaciones causales entre series en el sistema del modelo. Las tablas tanto en las estadísticas de ajuste del modelo y los valores atípicos para los objetivos visualizado se incluyen como parte del elemento de resultado. Cuando esta opción está seleccionada en el grupo **Resultado para series**, se crea un elemento de resultado separado para cada serie individual que se ha especificado en los valores de **Series para visualizar**.

Las relaciones causales entre las series tienen un nivel de significación asociado, donde un nivel de significación más pequeño indica una conexión más significativa. Puede elegir ocultar las relaciones con un nivel de significación que es mayor que un valor especificado.

Estadísticos de ajuste del modelo y valores atípicos

Las tablas de estadísticas y valores atípicos del ajuste del modelo para las series de objetivos que se han seleccionado para su visualización. Estas tablas contienen la misma información que las tablas en la visualización de Sistema global de modelo. Estas tablas soportan todas las funciones estándar para tablas dinámicas y edición.

Efectos del modelo y parámetros del modelo

Tablas de pruebas de efectos del modelo y parámetros de modelo para las series de objetivos que se han seleccionado para su visualización. Las pruebas de efectos del modelo incluyen la estadística F y un valor de significación asociado para cada entrada incluida en el modelo.

Diagrama de impacto

Muestra una representación gráfica de las relaciones causales entre una serie de interés y otras series a las que afecta o que le afectan. Las series que afectan a las series de interés se denominan *causas*. La selección de **Efectos** genera un diagrama de impacto inicializado para mostrar efectos. La selección de **Causas** genera un diagrama de impacto inicializado para mostrar causas. La selección de **Causas y efectos** genera dos diagramas de impacto independientes, uno inicializado para las causas y otro inicializado para los efectos. Puede alternar de forma interactiva entre causas y efectos en el elemento de resultado que muestra el diagrama de impacto.

Puede especificar el número de niveles de causas o efectos para mostrar, donde el primer nivel es simplemente la serie de interés. Cada nivel adicional muestra más causas o efectos indirectas de las series de interés. Por ejemplo, el tercer nivel de la visualización de efectos está formado por la series que contienen series del segundo nivel como una entrada directa. Las series del tercer nivel son afectadas directamente por las series de interés puesto que las series de interés es una entrada directa en las series del segundo nivel.

Gráfico de series

Gráficos de valores observados y pronosticados para las series de objetivos que se han seleccionado para la visualización. Cuando se solicitan previsiones, el gráfico también muestra los valores pronosticados y los intervalos de confianza para las previsiones.

Gráfico de residuos

Gráficos de los residuos del modelo para las series de objetivos que se han seleccionado para la visualización.

Entradas superiores

Gráficos para cada objetivo visualizado, a lo largo del tiempo, junto con las 3 entradas superiores del objetivo. Las entradas superiores son las entradas con el valor de significación menor. Para acomodar distintas escalas para las entradas y el objetivo, el eje Y representa la puntuación Z para cada serie.

Tabla de predicciones

Tablas de valores pronosticados e intervalos de confianza de estas previsiones para las series de objetivos que se han seleccionado para la visualización.

Análisis de la causa principal del valor atípico

Determina qué series tienen la mayor probabilidad de ser la causa de cada valor atípico en una serie de interés. El análisis de causa principal de valor atípico se realiza para cada serie de objetivos que se incluye en la lista de series individuales en el valor **Series para visualizar**.

Resultado

Tabla y gráfico de valores atípicos interactivos

Tabla y gráfico de valores atípicos y causas principales de estos valores atípicos para cada serie de interés. La tabla contiene una sola fila para cada valor atípico. El gráfico es un diagrama de impactos. Seleccionar una fila en la tabla resalta la ruta, en el diagrama de impactos, de la serie de interés en la serie que es más probable que provoque el valor atípico asociado.

Tabla dinámica de valores atípicos

Tabla de valores atípico y causas principales de estos valores atípico para cada series de interés. Esta tabla contiene la misma información que la tabla en la visualización interactiva. Esta tabla soporta todas las funciones estándar para tablas dinámicas y de edición.

Niveles causales

Puede especificar el número de niveles para incluir en la búsqueda para las causas principales. El concepto de niveles que se utiliza aquí es el mismo que el que se describe para diagramas de impactos.

Ajuste de modelo en todos los modelos

Histograma del ajuste del modelo para todos los modelos y para las estadísticas de ajuste seleccionadas. Están disponibles las estadísticas de ajuste siguientes:

R cuadrado

Medida de la bondad de ajuste de un modelo lineal; en ocasiones recibe el nombre de coeficiente de determinación. Es la proporción de la variación de la variable objetivo explicada por el modelo. Es un valor en un rango de 0 a 1. Los valores pequeños indican que el modelo no se ajusta bien a los datos.

Raíz de la media cuadrática de los errores porcentuales

Una medida de cuánto difieren los valores pronosticados por el modelo de los valores observados de las series. Es independiente de las unidades utilizadas y, por tanto, se puede utilizar para comparar series con distintas unidades.

Raíz de la media cuadrática de los errores

La raíz cuadrada de la media cuadrática de los errores. Una medida de cuánto se desvía la serie dependiente del nivel pronosticado por el modelo, expresado en las mismas unidades que la serie dependiente.

BIC

Criterio de información Bayesian. Una medida para seleccionar y comparar modelos basados en el logaritmo de probabilidad -2 reducida. Los valores menores indican modelos mejores. El BIC también "penaliza" modelos sobreparametrizados (modelos complejos con un gran número de entradas, por ejemplo), pero de forma más estricta que el AIC.

AIC

Criterio de información Akaike. Una medida para seleccionar y comparar modelos basados en el logaritmo de probabilidad -2 reducida. Los valores menores indican modelos mejores. El AIC "penaliza" modelos sobreparametrizados (modelos complejos con un gran número de entradas, por ejemplo).

Valores atípicos en el tiempo

Diagrama de barras del número de valores atípicos, en todos los objetivos, para cada intervalo de tiempo en el periodo de estimación.

Transformaciones de series

Tabla de cualquier transformación que se ha aplicado a la serie en el sistema del modelo. Las posibles transformaciones son una imputación, agregación y distribución de un valor que falta.

Guardar

Guardar objetivos

Puede guardar predicciones de modelo, intervalos de confianza y residuos como variables nuevas. Cada objetivo especificado para guardar genera su propio conjunto de variables nuevas y cada variable nueva contiene valores para ambos periodos, el de estimación y el de predicción. Para obtener valores pronosticados, intervalos de confianza y residuos de ruido, puede especificar el nombre raíz que se va a utilizar como el prefijo para las variables nuevas. El nombre de variable completo es la concatenación del nombre raíz y el nombre del campo que contiene las series de objetivos. El nombre raíz debe cumplir las reglas para los nombres de variable válidos. El nombre de la variable se amplía en caso de ser necesario para evitar conflictos de nombres de variables.

Indicar casos que contienen previsiones

Crea una variable que indica si un registro contiene datos de previsión. Puede especificar el nombre de variable. El valor predeterminado es *ForecastIndicator*.

Objetivos para guardar

Especifica si se crean variables nuevas para todas las series de objetivos en el sistema del modelo o solo las series objetivos que se han especificado en los valores **Series para visualizar**.

Opciones de destino

Puede guardar ambas cosas, las transformaciones de los datos (como agregación o imputación de valores que faltan) y las variables nuevas (especificadas en el valor **Guardar objetivos**) en un archivo de datos o un nuevo conjunto de datos de IBM SPSS Estadística en la sesión actual. Los valores fecha/hora de los datos guardados se alinean con el comienzo de cada intervalo de tiempo, como el primero del mes, y representan el intervalo de tiempo del análisis para el sistema modelo. Puede guardar cualquiera de las variables nuevas en el conjunto de datos activo sólo si las observaciones se definen mediante una especificación de fecha, una orden de registro y los datos no están agregados.

Exportar sistema de modelo

Guardar el sistema del modelo en un archivo de archivado comprimido (archivo .zip). El procedimiento puede reutilizar el archivo del sistema de modelo. El procedimiento Escenarios de modelos causales temporales también puede utilizarlo para ejecutar un análisis del escenario. Esta opción solo está disponible cuando se vuelven a calcular los parámetros del modelo.

Escenarios de modelos causales temporales

El procedimiento de escenarios de modelos causales temporales ejecuta escenarios definidos por el usuario para un sistema de modelo causal temporal, con los datos del conjunto de datos activo. Un *escenario* se define mediante una serie temporal, a la que se hace referencia como *serie raíz* y un conjunto de valores definidos por el usuario para dicha serie a lo largo de un rango de tiempo especificado. Los valores especificados se utilizan para generar predicciones para las series temporales afectadas por la serie raíz. El procedimiento requiere un archivo de sistema de modelo que ha creado el procedimiento de modelado causal temporal. Se presupone que el conjunto de datos activo son los mismos datos que se han utilizado para crear el archivo del sistema de modelo.

Ejemplo

Mediante el uso del procedimiento de modelado causal temporal, un tomador de decisiones empresariales ha descubierto una métrica clave afecta a un número de indicadores de rendimiento importantes. La métrica se puede controlar, de forma que el tomador de decisiones desea investigar el efecto de distintos conjuntos de valores para la métrica durante el próximo trimestre. La investigación se puede realizar fácilmente cargando el archivo del sistema de modelo en el procedimiento de escenarios de modelos causales temporales y especificando los conjuntos de valores para la métrica clave.

Ejecutar escenarios de modelos causales temporales

Esta función requiere la opción de predicción de estadísticas.

En los menús, seleccione:

Analizar > Previsiones > Aplicar modelos causales temporales...

1. Especifique la especificación de archivo para un archivo de sistema de modelo o pulse **Examinar** y seleccione un archivo de sistema de modelo. Los archivos de sistema de modelo se crean con el procedimiento de modelado causal temporal.
2. Pulse la opción para ejecutar escenarios.
3. Pulse **Continuar**.
4. En la pestaña Escenarios (en el diálogo Escenarios de modelos causales temporales), pulse **Definir periodo de escenario** y especifique el periodo de escenario.
5. Para datos basados en columna, pulse **Añadir escenario** para definir cada escenario. Para datos multidimensionales, pulse **Añadir escenario** para definir cada escenario individual y pulse **Añadir grupo de escenarios** para definir cada grupo de escenarios.
6. Pulse **Opciones** para especificar el contenido del resultado y para especificar la amplitud de las series que afectará el escenario.
7. Pulse en **Ejecutar** para ejecutar el procedimiento.

Definición del periodo de escenario

El periodo del escenario es el periodo durante el cual especifica los valores que se utilizan para ejecutar los escenarios. Puede empezar antes o después del final del periodo de estimación. De forma opcional, puede especificar predecir más allá del final del periodo del escenario. De forma predeterminada, las predicciones se generan hasta el final del periodo del escenario. Todos los escenarios utilizan el mismo periodo de escenario y especificaciones por lo que concierne a durante cuánto tiempo predecir.

Nota: las predicciones empiezan en el primer período de tiempo posterior al comienzo del período de escenario. Por ejemplo, si el periodo de escenario empieza en 2014-11-01 y el intervalo de tiempo es meses, la primera predicción será para 2014-12-01.

Especificar por hora de inicio, hora de finalización o tiempo de pronóstico

- Si las observaciones se definen mediante un campo de fecha/hora, especifique valores de inicio, final y tiempo de pronóstico con el mismo formato que se utiliza para el campo de fecha/hora. Los valores para los campos de fecha/hora se alinean al principio del intervalo de tiempo asociado. Por ejemplo, si el intervalo de tiempo del análisis es meses, el valor 10/10/2014 se ajusta a 10/01/2014, que es el inicio del mes.
- Para las observaciones que se han definido mediante periodos cíclicos, especifique un valor para cada uno de los campos de periodos cíclicos. Cada campo se visualiza en una columna separada.
- Si hay una especificación de fecha en vigor para el conjunto de datos activo, debe especificar un valor para cada componente (como Mes) de la especificación de fecha. Cada componente se visualiza en una columna separada.
- Cuando las observaciones se definen mediante el orden de registro, el inicio, el final y el tiempo de pronóstico se definen a través del número de fila (tal como se visualiza en el editor de datos) del caso relevante.

Especificar por intervalos de tiempo relativos al final del periodo de estimación

Define el inicio y el final en términos del número de intervalos de tiempo relativos al final del periodo de estimación, donde el intervalo de tiempo es el intervalo de tiempo del análisis. El final del periodo de estimación se define como el intervalo de tiempo 0. Los intervalos de tiempo antes del final del periodo de estimación tienen intervalos y valores negativos después de que el final del periodo de estimación tenga valores positivos. También puede especificar cuántos intervalos predecir más allá del final del periodo del escenario. El valor predeterminado es 0.

Por ejemplo, suponga que el intervalo de tiempo del análisis es meses y que especifica 1 para iniciar el intervalo y 3 para finalizarlo, y 1 para cuánto más ampliar la previsión. El periodo del escenario sería de 3 meses que siguen al final del periodo de estimación. Las predicciones se generan para el segundo y tercer mes del periodo de escenario y para 1 mes más después del final periodo de escenario.

Adición de escenarios y grupos de escenarios

La pestaña Escenarios especifica los escenarios que se van a ejecutar. Para definir escenarios, en primer lugar, debe definir el periodo del escenario pulsando **Definir periodo del escenario**. Los escenarios y los grupos de escenarios (solo se aplica a datos multidimensionales) se crean pulsando el botón asociado **Añadir escenario** o **Añadir grupo de escenarios**. Al seleccionar un escenario particular o un grupo de escenarios en la cuadrícula asociada, puede editarlo, hacer una copia o suprimirlo.

Datos basados en columna

La columna **Campo raíz** de la cuadrícula especifica el campo de serie temporal cuyos valores se sustituyen con los valores de escenario. La columna **Valores de escenario** muestra los valores de escenario especificados en el orden del primero al último. Si los valores del escenario se definen mediante una expresión, la columna muestra la expresión.

Datos multidimensionales

Escenarios individuales

Cada fila de la cuadrícula Escenarios individuales especifica una serie temporal cuyos valores se sustituyen por los valores de escenario especificados. La serie se define mediante la combinación del campo que se especifica en la columna **Métrica raíz** y el valor especificado para cada uno de los campos de dimensión. El contenido de la columna **Valores de escenario** es el mismo que para los datos basados en columnas.

Grupos de escenarios

Un *grupo de escenarios* define un conjunto de escenarios que se basan en un solo campo de métrica raíz y varios conjuntos de valores de dimensión. Cada conjunto de valores de dimensión (un valor por campo de dimensión), para el campo de métrica especificado, define una serie temporal. Un escenario individual se genera para cada una de dichas series temporales, cuyos valores se sustituyen después por los valores de escenario. Los valores de escenario para un grupo de escenarios se especifican mediante una expresión, que se aplica después a cada serie temporal del grupo.

La columna **Recuento de series** muestra el número de conjuntos de valores de dimensión que están asociados a un grupo de escenarios. El valor visualizado puede ser mayor que el número real de series temporales que están asociadas al grupo de escenarios (una serie por conjunto). Esta condición se produce cuando algunas de las combinaciones especificadas de valores de dimensión no corresponden a las series incluidas por la métrica raíz para el grupo.

Como ejemplo de un grupo de escenarios, considere un campo de métrica *publicidad* y dos campos de dimensión *región* y *marca*. Puede definir un grupo de escenarios que se basa en *publicidad* como la métrica raíz y que incluye todas las combinaciones de *región* y *marca*. A continuación, podría especificar " *advertising*1.2* como expresión para investigar el efecto de aumentar *la publicidad* en un 20% para cada una de las series temporales asociadas al campo de *publicidad*. Si hay 4 valores de *región* y 2 valores de *marca*, existen 8 series temporales de ese tipo y, por lo tanto, 8 escenarios definidos mediante el grupo.

Definición de escenario

Los valores para definir un escenario dependen de si los datos se basan en columnas o son multidimensionales.

Series raíz

Especifica la serie raíz para el escenario. Cada escenario se basa en un única serie raíz. Para los datos basados en columna, seleccione el campo que define la serie raíz. Para datos multidimensionales, especifique la serie raíz añadiendo una entrada a la cuadrícula para el campo de métrica que contiene la serie. A continuación, especifique los valores de los campos de dimensión que definen la serie raíz. Se aplica lo siguiente para especificar los valores de dimensión:

- Puede especificar el valor para cada campo de dimensión directamente en la cuadrícula o puede seleccionar en la lista de valores de dimensión disponibles. Para seleccionar en la lista de valores de dimensión disponibles, pulse el botón de puntos suspensivos en la casilla para la dimensión que desee. Esta acción abre el subdiálogo Seleccionar valor de dimensión.
- Puede buscar en la lista de valores de dimensión, en el subdiálogo Seleccionar valor de dimensión, pulsando el icono de los prismáticos y especificando un término de búsqueda. Los espacios se tratan como parte del término de búsqueda. Los asteriscos (*) en el término de búsqueda no indican caracteres comodín.

Especificar objetivos afectados

Utilice esta opción cuando sepa los objetivos específicos afectados por la serie raíz y cuando desee investigar los efectos solo en estos objetivos. De forma predeterminada, los objetivos afectados por la serie raíz se determinan automáticamente. Puede especificar la amplitud de la serie afectada por el escenario con los valores de la pestaña Opciones.

Para datos basados en columnas, seleccione los objetivos que desea. Para datos multidimensionales, especifique series de objetivo añadiendo una entrada a la cuadrícula para el campo de métrica de objetivo que contiene la serie. Por defecto, se incluyen todas las series contenidas en el campo métrico especificado. Puede personalizar el conjunto de series incluidas personalizando los valores incluidos para uno o más de los campos de dimensión. Para personalizar los valores de dimensión que se incluyen, pulse el botón de los puntos suspensivos para la dimensión que desee. Esta acción abre el diálogo Seleccionar valores de dimensión.

La columna **Recuento de series** (para datos multidimensionales) muestra el número de conjuntos de valores de dimensión especificados actualmente para la métrica de objetivo asociado. El valor visualizado puede ser mayor que el número real de series de objetivos afectados (una serie por

conjunto). Esta condición se produce cuando algunas de las combinaciones especificadas de valores de dimensión no corresponden a las series incluidas en la métrica de objetivo asociado.

ID de escenario

Cada escenario debe tener un identificador exclusivo. El identificador se visualiza en el resultado asociado al escenario. No hay ninguna restricción, que no sea la exclusividad, en el valor del identificador.

Especificar valores del escenario para series raíz

Utilice esta opción para especificar valores explícitos para la serie raíz en el periodo del escenario. Debe especificar un valor numérico para cada intervalo de tiempo listado en la cuadrícula. Puede obtener los valores de la serie raíz (real o previsto) para cada intervalo del periodo del escenario pulsando **Leer**, **Previsión** o **Leer\Previsión**.

Especificar expresión de los valores del escenario para la serie raíz

Puede definir una expresión para calcular los valores de la serie raíz en el periodo de escenario. Puede especificar la expresión directamente o pulsar el botón de la calculadora y crear la expresión desde el Generador de expresiones de valores de escenario.

- La expresión puede contener cualquier destino o entrada en el sistema del modelo.
- Cuando el periodo del escenario se extiende más allá de los datos existentes, la expresión se aplica a valores previstos de los campos en la expresión.
- Para datos multidimensionales, cada campo de la expresión especifica una serie temporal definida mediante los valores del campo y de dimensión que se han especificado para la métrica raíz. Es una de esas series temporales que se utilizan para evaluar la expresión.

Por ejemplo, suponga que el campo raíz es *publicidad* y que la expresión es *advertising*1.2*. Los valores para *publicidad* que se utilizan en el escenario representan un aumento del 20 % con respecto a los valores existentes.

Nota: Se crean escenarios pulsando **Añadir escenario** en la pestaña Escenarios.

Generador de expresiones de valores de escenario

Utilice el Generador de expresiones de valores de escenario para crear una expresión que se utiliza para calcular valores de escenario para un solo escenario o un grupo de escenarios. Para generar una expresión, pegue componentes en el campo **Expresión** o escriba directamente en el campo **Expresión**.

- La expresión puede contener cualquier destino o entrada en el sistema del modelo.
- Puede pegar las funciones seleccionando un grupo de la lista Grupo de funciones y pulsando dos veces en la función de la lista de funciones (o seleccione la función y pulse en la flecha que se encuentra sobre la lista Grupo de funciones). Especifique los parámetros indicados por los signos de interrogación. El grupo de funciones con la etiqueta **Todo** contiene una lista de todas las funciones variables. En un área reservada del cuadro de diálogo se muestra una breve descripción de la función actualmente seleccionada.
- Las constantes de cadena deben ir entre comillas.
- Si los valores contienen decimales, debe utilizarse una coma(,) como indicador decimal.

Para acceder al Generador de expresiones de valores de escenario, pulse el botón de la calculadora en el cuadro de diálogo Definición de escenario o Definición del grupo de escenarios.

Seleccionar valores de dimensión

Para datos multidimensionales, puede personalizar los valores de dimensión que definen los objetivos afectados por un escenario o un grupo de escenarios. También puede personalizar los valores de dimensión que definen el conjunto de series raíz para un grupo de escenarios.

Todos los valores

Especifica que todos los valores del campo de dimensión actual se incluyen. Esta opción es la predeterminada.

Seleccionar valores

Utilice esta opción para especificar el conjunto de valores para el campo de dimensión actual. Puede filtrar el conjunto de valores entre los que elegir. Los valores que cumplen la condición de filtro aparecen en la pestaña **Coincidentes** y los valores que no cumplen la condición de filtro aparecen en la pestaña **No coincidentes** de la lista **Valores no seleccionados**. La pestaña **Todos** lista todos los valores no seleccionados, independientemente de cualquier condición de filtro.

- Puede utilizar asteriscos (*) para indicar caracteres comodín cuando especifique un filtro.
- Para borrar el filtro actual, especifique un valor vacío para el término de búsqueda en el diálogo Filtrar valores visualizados.

Para personalizar valores de dimensión para objetivos afectados:

1. En el diálogo Definición de escenario o Definición de grupo de escenarios, seleccione la métrica de objetivo para la cual desea personalizar valores de dimensión.
2. Pulse el botón de puntos suspensivos en la columna para la dimensión que desee personalizar.

Para personalizar valores de dimensión para las series raíz de un grupo de escenarios:

1. En el diálogo Definición de grupo de escenarios, pulse el botón de puntos suspensivos (en la cuadrícula de serie raíz) para la dimensión que desea personalizar.

Definición de grupo de escenario

Serie raíz

Especifica el conjunto de series raíz para el grupo de escenarios. Un escenario individual se genera para cada serie temporal del conjunto. Especifique la serie raíz añadiendo una entrada a la cuadrícula para el campo de métrica que contiene la serie que desea. A continuación, especifique los valores de los campos de dimensión que definen el conjunto. De forma predeterminada, todas las series incluidas en el campo de métrica raíz especificado se incluyen. Puede personalizar el conjunto de series incluidas personalizando los valores incluidos para uno o más de los campos de dimensión. Para personalizar los valores de dimensión que se incluyen, pulse el botón de los puntos suspensivos para una dimensión. Esta acción abre el diálogo Seleccionar valores de dimensión.

La columna **Recuento de series** muestra el número de conjuntos de valores de dimensión que están incluidos actualmente para la métrica raíz asociada. El valor visualizado puede ser mayor que el número real de series raíz para el grupo de escenarios (una serie por conjunto). Esta condición se produce cuando algunas de las combinaciones especificadas de valores de dimensión no corresponden a las series incluidas por la métrica raíz.

Especifique series objetivo afectadas

Utilice esta opción cuando sepa los objetivos específicos afectados por el conjunto de series raíz y desea investigar los efectos solo en estos objetivos. De forma predeterminada, los objetivos afectados por cada serie raíz se determinan automáticamente. Puede especificar el ancho de las series afectadas por cada escenario individual con la configuración en la pestaña Opciones.

Puede especificar series objetivo añadiendo una entrada a la cuadrícula para el campo de medida que contiene la serie. Por defecto, se incluyen todas las series contenidas en el campo métrico especificado. Puede personalizar el conjunto de series incluidas personalizando los valores incluidos para uno o más de los campos de dimensión. Para personalizar los valores de dimensión que se incluyen, pulse el botón de los puntos suspensivos para la dimensión que desee. Esta acción abre el diálogo Seleccionar valores de dimensión.

La columna **Recuento de series** muestra el número de conjuntos de valores de dimensión que están especificados actualmente para la métrica de objetivo asociada. El valor visualizado puede ser mayor que el número real de series de objetivos afectados (una serie por conjunto). Esta condición se produce cuando algunas de las combinaciones especificadas de valores de dimensión no corresponden a las series incluidas en la métrica de objetivo asociado.

Prefijo de ID de escenario

Cada grupo de escenarios debe tener un prefijo exclusivo. El prefijo se utiliza para construir un identificador que se visualiza en el resultado que está asociado a cada escenario individual e n

el grupo de escenarios. El identificador para un escenario individual es el prefijo, seguido por un subrayado, seguido por el valor de cada campo de dimensión que identifica la serie raíz. Los valores de dimensión se separan mediante subrayados. No hay ninguna restricción, que no sea la exclusividad, en el valor del prefijo.

Expresión para valores de escenario para series raíz

Los valores de escenario para un grupo de escenarios se especifican mediante una expresión, que se utiliza después para calcular los valores para cada una de las series raíz del grupo. Puede especificar una expresión directamente o pulsar el botón de la calculadora y crear la expresión a partir del Generador de expresiones de valores de escenario.

- La expresión puede contener cualquier destino o entrada en el sistema del modelo.
- Cuando el periodo del escenario se extiende más allá de los datos existentes, la expresión se aplica a valores previstos de los campos en la expresión.
- Para cada serie raíz del grupo, los campos de la expresión especifican series temporales definidas mediante estos campos y los valores de dimensión que definen las series raíz. Es una de esas series temporales que se utilizan para evaluar la expresión. Por ejemplo, si `region='north'` y `brand='X'` definen una serie raíz, las series temporales que se utilizan en la expresión se definen mediante esos mismos valores de dimensión.

Como ejemplo, suponga que el campo de métrica raíz es *publicidad* y que hay dos campos de dimensión *región* y *marca*. Además, suponga que el grupo de escenarios incluye todas las combinaciones de los valores de campo de dimensión. A continuación, puede especificar `advertising*1.2` como expresión para investigar el efecto de aumentar *publicidad* en un 20 por ciento para cada una de las series temporales asociadas al campo *publicidad*.

Nota: Los grupos de escenarios se aplican a datos multidimensionales y se crean pulsando **Añadir grupo de escenarios** en la pestaña Escenarios.

Opciones

Máximo nivel de los objetivos afectados

Especifica el número máximo de niveles de objetivos afectados. Cada nivel sucesivo, hasta el máximo de 5, incluye objetivos a los que las series raíz afectan indirectamente. Específicamente, el primer nivel incluye objetivos que tienen las series raíz como entrada directa. Los objetivos del segundo nivel tienen objetivos del primer nivel como una entrada directa y, así, sucesivamente. Aumentar este valor aumenta la complejidad del cálculo y podría afectar al rendimiento.

Número máximo de objetivos detectados automáticamente

Especifica el número máximo de destinos afectados que se detectan automáticamente para cada serie raíz. Aumentar este valor aumenta la complejidad del cálculo y podría afectar al rendimiento.

Diagrama de impacto

Muestra una representación gráfica de las relaciones causales entre las series raíz para cada escenario y las series de objetivos a que afecta. Las tablas de valores de escenario y también las tablas de los valores pronosticados para los objetivos afectados se incluyen como parte del elemento de resultado. El gráfico incluye gráficos de los valores pronosticados de los objetos afectados. Si pulsa una vez en cualquier nodo del diagrama de impactos se abre un diagrama de secuencia detallada para la serie que está asociada al nodo. Se genera un diagrama de impactos separado para cada escenario.

Gráficos de series

Genera gráficos de series de los valores pronosticados para cada uno de los objetivos afectados en cada escenario.

Tablas de predicciones y escenarios

Tablas de valores pronosticados y valores de escenario para cada escenario. Estas tablas contienen la misma información que las tablas del diagrama de impactos. Estas tablas soportan todas las funciones estándar para tablas dinámicas y edición.

Incluir intervalos de confianza en gráficos y tablas

Especifica si los intervalos de confianza para las predicciones de escenario se incluyen en los resultados del gráfico y, también, de tabla.

Ancho de intervalo de confianza (%)

Este valor controla los intervalos de confianza para las predicciones de escenario. Puede especificar cualquier valor positivo menor que 100. De forma predeterminada, se utiliza un intervalo de confianza del 95%.

Modelos autorregresivos vectoriales

Un modelo autorregresivo vectorial (VAR) es una herramienta fundamental en el análisis de series temporales multivariantes. La idea consiste en ampliar el concepto de los modelos autorregresivos —que explican una variable a partir de sus propios valores históricos— al caso multivariante, en el que se modelan conjuntamente varias series temporales interrelacionadas. Un modelo VAR refleja las interdependencias lineales entre varias variables de series temporales. Es decir, un modelo VAR(p) expresa cada variable como una función lineal de sus propios valores pasados y de los valores pasados de todas las demás variables: cada variable del sistema se explica a partir de sus propios valores pasados (autorregresión) y de los valores pasados de todas las demás variables del sistema

Los pasos del análisis VAR incluyen la definición de la fecha y la hora, la comprobación de la estacionariedad, la estimación del modelo, la selección del orden de retraso, el diagnóstico del modelo y la previsión.

Estadísticos

- Para cada variable: identifica el número de casos válidos, calcula la media y la desviación estándar, y especifica los valores mínimo y máximo.
- Para cada modelo: evalúa las estimaciones de los parámetros (coeficientes), el error estándar de los coeficientes, la estadística t y el valor p (Sig.), intervalos de confianza inferior y superior, matriz de covarianza de los residuos, determinante de la matriz de covarianza de los residuos, matriz de correlación de los residuos, valores propios de la matriz de correlación de los residuos, suma de cuadrados de los residuos, R cuadrado, R cuadrado ajustado, MSE, RMSE, MAE, MAPE, prueba de Ljung-Box para la autocorrelación de los residuos y prueba de Jarque-Bera para la normalidad de los residuos.
- Para evaluar el ajuste del modelo: calcula los índices AIC, BIC y HQIC para distintos retrasos.
- Gráficos: series temporales, ACF, gráfico de residuos para cada variable, gráfico de residuos combinado, gráfico de ACF de residuos para cada variable, valores de ajuste y previsión.

Consideraciones sobre los datos

- Las variables endógenas deben ser escalares. Se requiere un mínimo de 2 variables.
- Las variables exógenas pueden ser de escala o categóricas. Esta información es opcional.

Obtención de modelos autorregresivos vectoriales

1. Los modelos VAR solo son aplicables a datos de series temporales. Haga clic en **«Definir fecha y hora»** para generar variables de fecha y hora.
2. La estacionariedad es fundamental para los modelos VAR. Las series no estacionarias suelen requerir un análisis de diferenciación o de cointegración.

Una serie temporal estacionaria es aquella cuyas propiedades estadísticas, como la media, la varianza y la autocorrelación, se mantienen constantes a lo largo del tiempo. Antes de proceder al ajuste de un modelo VAR, asegúrate de que la serie temporal sea estacionaria.

En el caso de series no estacionarias, calcula la primera diferencia de los datos reales y comprueba si son estacionarias. Vuelve a calcular la diferencia y continúa con el procedimiento hasta que los datos alcancen la estacionariedad.

La prueba de Dickey-Fuller aumentada (ADF) puede utilizarse para comprobar la estacionariedad. La prueba ADF forma parte del procedimiento de prórroga. `STATS TTESTS` Una vez instalado `STATS TTESTS`, ve a **«Analizar > pruebas > de estacionariedad y cointegración»**. Selecciona una variable y haz clic en **«Pruebas »**. Selecciona **«Calcular prueba (una variable) »** para la prueba de Dickey-Fuller aumentada. Ejecuta la prueba.

3. Tras la prueba de estacionariedad, haz clic en «**Analizar** > » > « > **Previsión**» > > «**Series temporales multivariantes**» > «**Modelos autorregresivos** vectoriales».

Nota: Los campos resaltados en rojo son obligatorios. Los botones «**Pegar**» y «**Aceptar**» se activan una vez que hayas introducido valores válidos en todos los campos obligatorios.

4. Seleccione las variables endógenas y exógenas; las variables exógenas son opcionales.
5. Para obtener mejores previsiones, es necesario elegir la longitud óptima del lag. Esto se puede llevar a cabo en función de la selección utilizando diversos criterios de ajuste o criterios de información: el AIC (Criterio de Información de Akaike), el BIC (Criterio de Información Bayesiano) o el HQIC (Criterio de Información de Hannan-Quin).
6. Seleccione las opciones de las pestañas «Estadísticas», «Gráficos», «Guardar» y «Opciones».
7. Pulse **Aceptar**.

Variables

Variables

La lista de variables sirve como fuente de la que puedes seleccionar tanto variables endógenas como variables exógenas.

- Selecciona dos o más variables endógenas. Las variables endógenas son aquellas que se determinan dentro del modelo o del sistema que se está estudiando. Están condicionadas por otras variables del modelo, tanto exógenas como endógenas.
- El campo «**Variables exógenas**» es opcional. Las variables exógenas son factores independientes que afectan al sistema desde el exterior. Se determinan fuera del modelo o del sistema que se está estudiando. Estas variables no se ven influidas por otras variables del modelo, pero pueden influir en las variables endógenas.

Selección de desfase

Para obtener mejores previsiones, es necesario seleccionar la longitud óptima del lag. Puedes seleccionar automáticamente el mejor orden de retardo introduciendo un valor para el retardo máximo. Por defecto, el retraso mínimo está establecido en 1. Esto se hace utilizando cualquiera de los tres criterios de ajuste disponibles en SPSS Statistics:

AIC (Criterio de información de Akaike), **BIC** (Criterio de información bayesiano) o **HQIC** (Criterio de información de Hannan-Quin)

- **El retraso máximo** debe ser un número entero. En IBM SPSS Estadística, el retraso máximo permitido es 24.

Método

El método se refiere al método de estimación del modelo VAR. El método de mínimos cuadrados ordinarios (OLS) es el método predeterminado.

Período de estimación

El período de estimación define el conjunto de casos utilizados para determinar el modelo. De forma predeterminada, el período de estimación incluye todos los casos del conjunto de datos activo. Para establecer el período de estimación, seleccione «Según el tiempo» o «Según el intervalo de casos» en el cuadro de diálogo «Seleccionar casos». En función de los datos disponibles, el período de estimación utilizado por el procedimiento puede variar según la variable endógena y, por lo tanto, diferir del valor mostrado. Para una variable endógena dada, el período de estimación real es el período que queda tras eliminar cualquier valor perdido contiguo de la variable que se produzca al principio o al final del período de estimación especificado.

Período de predicción

El período de predicción empieza en el primer caso después del período de estimación y, de forma predeterminada, llega hasta el último caso del conjunto de datos activo. Puede establecer el final del período de predicción en la pestaña **Opciones**.

Estadísticos

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

Medidas de ajuste y previsiones

Indica los distintos índices de ajuste de los modelos VAR desarrollados. Puede seleccionar una o varias de las siguientes opciones para incluirlas en la tabla que contiene las medidas de ajuste del modelo estimado:

- **Estadística descriptiva** : Determina las medidas descriptivas, como la media, la desviación estándar y los valores mínimo y máximo de cada una de las variables seleccionadas en «Variables endógenas».
- **Ajuste del modelo** : Determina los valores de AIC, BIC y HQIC para el mejor orden de retardo correspondiente a las variables seleccionadas.
- **Estimaciones de los parámetros** : Determina las estimaciones (coeficientes), el error estándar, la estadística t, el valor p y los intervalos de confianza inferior y superior mediante el método OLS para cada variable.

Nota: Los modelos VAR requieren estimaciones de los parámetros. Por lo tanto, la opción «Estimaciones de parámetros» está seleccionada de forma predeterminada como opción integrada y no se puede modificar.

- **Suma de cuadrados de los residuos** : Determina la suma de cuadrados de los residuos (o errores), donde los residuos se obtienen a partir de la diferencia entre los valores reales y los valores previstos (ajustados).
- **Función de autocorrelación de los residuos (ACF)** : Determina la función de autocorrelación de los residuos.
- **Error cuadrático medio** : calcula el error cuadrático medio, que es la media de los cuadrados de los residuos.
- **Error cuadrático medio** : Calcula el error cuadrático medio, que es la raíz cuadrada del error cuadrático.
- **Error absoluto medio** : Calcula el error absoluto medio, que es la media de los valores absolutos de los residuos.
- **Error porcentual absoluto medio** : determina el error porcentual absoluto medio.
- **R cuadrado** : Determina el valor de R cuadrado o el coeficiente de determinación.
- **R cuadrado ajustado** : determina el valor del R cuadrado ajustado.
- **Matriz de correlación de los residuos** : calcula la matriz de correlación de los residuos para todas las variables. En los resultados también se incluyen la matriz de covarianza de los residuos y el valor de su determinante.
- **Valores propios de la matriz de correlación residual** : calcula los valores propios de la matriz de correlación residual.
- **Prueba de normalidad de los residuos** : comprueba si los residuos siguen una distribución normal o no. Para ello se podría utilizar la prueba de Jarque-Bera.
- **Prueba de autocorrelación** : comprueba si un conjunto de valores de autocorrelación de una serie (o de los residuos de un modelo) difiere significativamente de cero, o bien determina si los datos son ruido blanco (aleatorios) o presentan patrones subyacentes (correlación serial). Para ello se podría utilizar la prueba de Ljung.
- **Previsiones** : muestra los valores previstos de las variables.

Gráficos

En la pestaña **Gráficos**, especifique las opciones para visualizar los gráficos de los resultados del modelo. Puedes generar visualizaciones para comprender mejor las relaciones entre los valores observados, los valores previstos, los residuos y otras características de las variables. La pestaña incluye diversos gráficos, como gráficos de series temporales correspondientes a cada variable, el gráfico de la función

de autocorrelación (ACF), gráficos de residuos para cada variable, un gráfico combinado de residuos, la función de autocorrelación de los residuos para cada variable, los valores de ajuste y la previsión.

Cada visualización ofrece una perspectiva única del modelo VAR y se puede personalizar según las preferencias del usuario.

Parcelas disponibles

- **Serie temporales** : Un gráfico de series temporales es una representación gráfica de puntos de datos registrados en orden cronológico, en la que el tiempo se muestra en el eje horizontal y la variable medida en el eje vertical. Puedes visualizar el gráfico de series temporales de cada variable seleccionada en las secciones «Variables endógenas» y «Variables exógenas».
- **ACF** : Un gráfico de la función de autocorrelación (ACF) es una herramienta gráfica que muestra la correlación entre una serie temporal y sus propios valores retrasados para diferentes intervalos de tiempo. El eje horizontal representa el número de retrasos, y el eje vertical representa el coeficiente de autocorrelación en cada retraso.
- **Residuales (cada variable)** : Un gráfico de residuales es un gráfico de diagnóstico que muestra los residuales (valores observados menos valores estimados, o valores estimados menos valores observados) de un modelo estadístico (VAR), normalmente representados en función del tiempo. Este gráfico se obtiene representando el tiempo en el eje horizontal y los residuos en el eje vertical. El gráfico de residuos es aplicable a todas las variables endógenas y exógenas seleccionadas.
- **Residual combinado** : se trata de una forma combinada del **modelo residual (cada variable)**. En lugar de mostrar los gráficos de residuos de cada variable en gráficos separados, los residuos de todas las variables se muestran en un mismo gráfico, con el tiempo representado en el eje X.
- **ACF residual (por variable)** : un gráfico de ACF residual muestra la correlación entre los errores (o residuos) de un modelo y sus valores anteriores (retrosos) para detectar posibles patrones residuales, lo que indica si el modelo ha captado adecuadamente la estructura de los datos. En teoría, los picos significativos fuera de los intervalos de confianza indican autocorrelación. En este gráfico se representa el desfase (la diferencia temporal entre observaciones) en el eje horizontal y el valor de autocorrelación en el eje vertical. Los intervalos de confianza (por defecto, del 95 %) están marcados con líneas discontinuas azules.
- **Valores de ajuste** : Este gráfico muestra los valores previstos correspondientes a cada variable en distintos intervalos de tiempo.
- **Previsión** : Este gráfico muestra la previsión para distintos periodos de tiempo.

Las opciones «Series temporales» y «Valores de ajuste» están seleccionadas de forma predeterminada.

Guardar

En la pestaña «**Guardar**», puedes seleccionar opciones para guardar los valores previstos (ajustados) y los residuos como columnas adicionales en un nuevo conjunto de datos.

Valores pronosticados

Guardar los valores previstos

Residuos de ruido

Guarda los residuos de ruido.

Elegir conjunto de datos

Puedes guardar el conjunto de datos actual, junto con los valores predictivos y los residuos, en un nuevo .sav archivo. En el campo «**Nuevo conjunto de datos** », haz clic en el botón de puntos suspensivos («...»). Selecciona la ubicación del archivo, introduce un nombre y haz clic en «**Guardar** ».

También puedes guardar el nuevo conjunto de datos en IBM SPSS Repositorio de servicios de colaboración y despliegue. Consulte .

Opciones

La pestaña «Opciones» te permite establecer el periodo de previsión, especificar cómo se deben tratar los valores perdidos y fijar la amplitud del intervalo de confianza.

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. De forma predeterminada, el final del periodo de estimación es el último caso del conjunto de datos activos, 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 previsiones 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.

Valores perdidos del usuario

Estas opciones controlan el tratamiento de los valores perdidos del usuario.

- En el caso de «**Tratar como no válidos**», los valores perdidos se tratan como valores perdidos del sistema.
- En el caso de «**Tratar como válidos**», los valores que faltan se consideran datos válidos.

Ancho de intervalo de confianza (%)

Se calculan los intervalos de confianza para las predicciones del modelo. Puede especificar cualquier valor positivo menor que 100. De forma predeterminada, se utiliza un intervalo de confianza del 95%.

Medidas de bondad de ajuste

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

- *R cuadrado estacionaria*. Medida que compara la parte estacionaria del modelo con un modelo de media simple. Esta medida es preferible al R-cuadrado ordinario cuando existe tendencia o patrón estacional. El valor de R cuadrado estacionario puede ser negativo con un rango de infinidad negativa 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*. Estimación de la proporción de la variación total de la serie explicada por el modelo. Esta medida es más útil cuando la serie es estacionaria. R cuadrado puede ser negativo con un rango de infinito negativo 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 de la media cuadrática de los errores. 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 de porcentaje absoluto medio. 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 promedio absoluto. 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 de porcentaje absoluto máximo. Mayor error predicho, expresado como porcentaje. Esta medida es útil para imaginar el peor escenario de un caso en las predicciones.
- *MaxAE*. Error máximo absoluto. Mayor error predicho, 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 las predicciones. El error absoluto máximo y el error porcentual absoluto máximo 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 porcentual absoluto máximo 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 la media cuadrática de los errores 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 por una constante, empezando por un punto determinado de la serie. Un cambio de nivel puede ser el resultado de un cambio de política.
- *De Innovación*. Un valor atípico que actúa como una adición al término de ruido en un punto determinado 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.
- *transient*. Un valor atípico cuyo impacto decae exponencialmente hacia cero.
- *Aditivo estacional*. Un valor atípico que afecta a una observación determinada y a todas las observaciones posteriores separadas de ella por uno o varios periodos 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 valores atípicos individuales además de sus grupos.

Guía de gráficos ACF/PACF

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

- Las series no estacionarias tienen una ACF que sigue siendo significativa para media docena o más retardos, en lugar de disminuir rápidamente a 0. Debe diferenciar una serie de este tipo hasta que esté estacionaria antes de poder identificar el proceso.
- Los procesos autorregresivos tienen una FAS que disminuye exponencialmente y líneas de unión en el primer retardo o en más de la FAP. El número de picos indica el orden de la autorregresión.
- Los procesos de medias móviles tienen líneas de unión en el primer retardo o en más de la FAS y una FAP que disminuye exponencialmente. El número de picos indica el orden de la media móvil.
- Los procesos mixtos (ARMA) suelen mostrar declinaciones exponenciales tanto en el ACF como en el PACF.

En la fase de identificación, no es necesario preocuparse por el signo de la ACF o PACF, o por la velocidad con la que una ACF o PACF en declive exponencial se acerca a 0. Estos dependen del signo y del valor real de los coeficientes AR y MA. En algunos casos, una ACF que disminuye exponencialmente se 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. Tenga en cuenta que:

- Los procesos estacionales muestran estos patrones en los retardos estacionales (los múltiplos del periodo estacional).
- Tiene derecho a tratar los valores no significativos como 0. Es decir, puede ignorar los valores que se encuentran dentro de los intervalos de confianza en los gráficos. Sin embargo, no tiene que ignorarlos, sobre todo si continúan 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 alto, y si no se produce en un retardo estacional.

Consulte cualquier texto sobre el análisis ARIMA para una discusión más completa de las parcelas ACF y PACF.

Tabla 3. ARIMA (0,0,1), $q > 0$

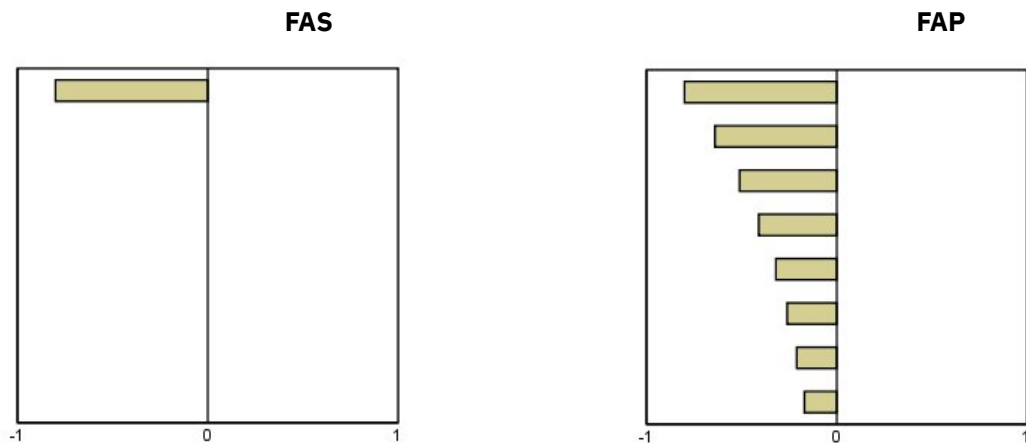
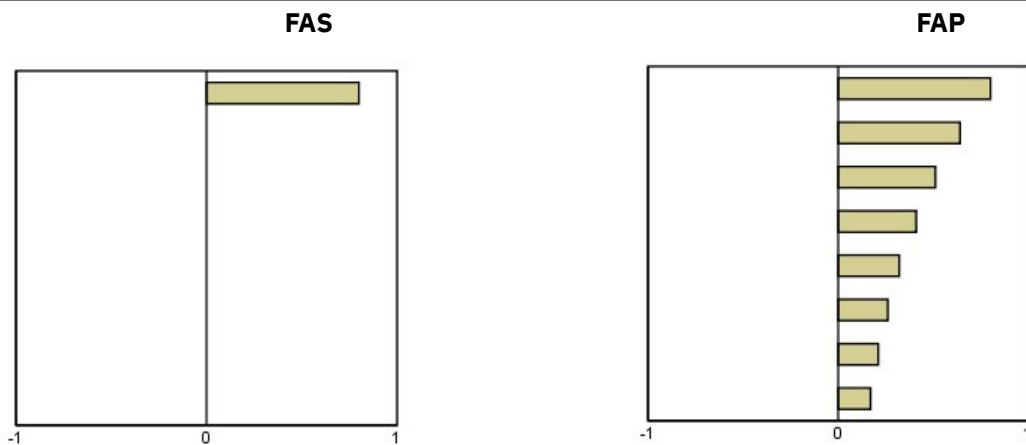


Tabla 4. ARIMA (0,0,1), $q < 0$



ARIMA (0,0,2), $\alpha_1 > 0$

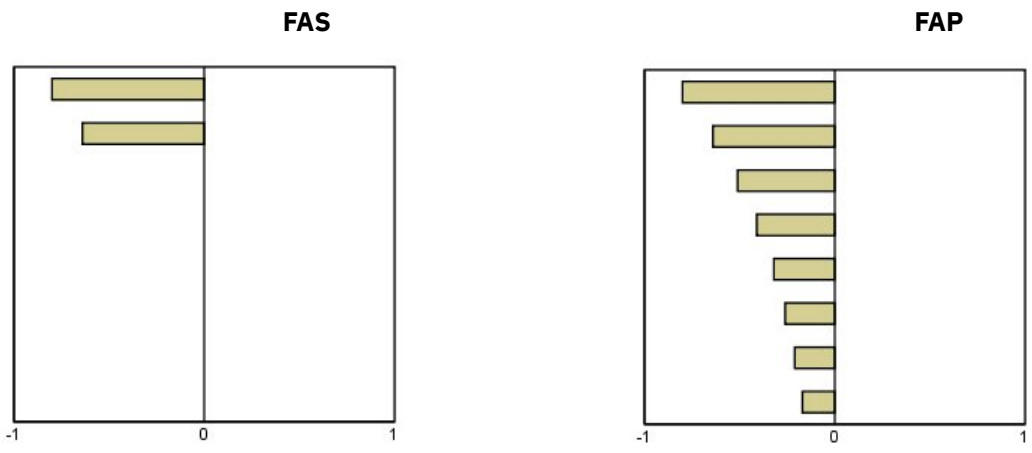


Tabla 5. ARIMA (1,0, 0), $f > 0$

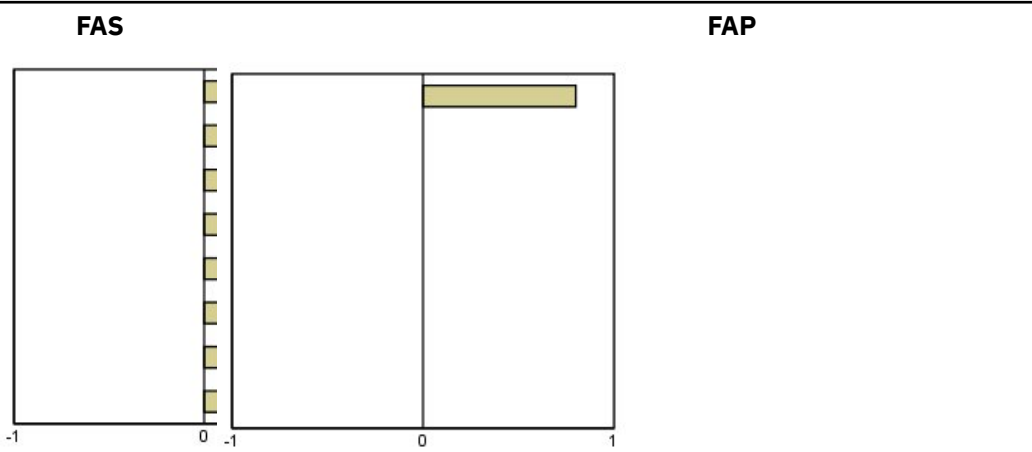
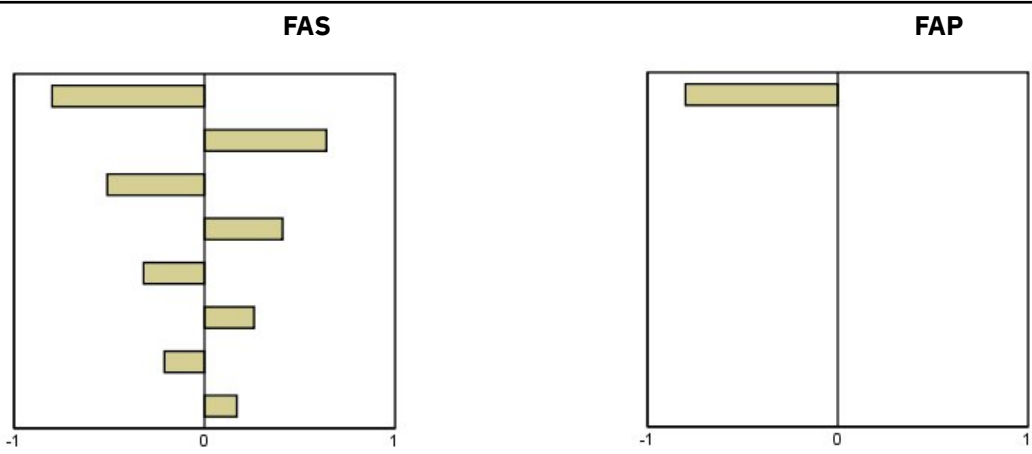
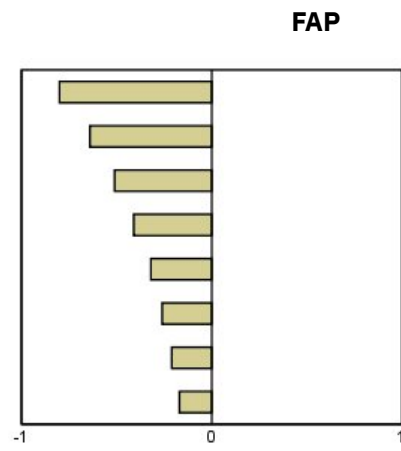
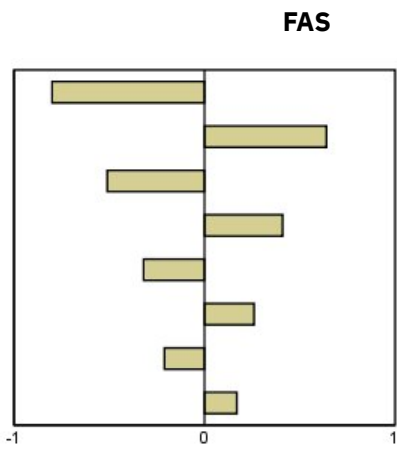


Tabla 6. ARIMA (1,0, 0), $f < 0$



ARIMA (1,0, 1), $< 0, > 0$



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

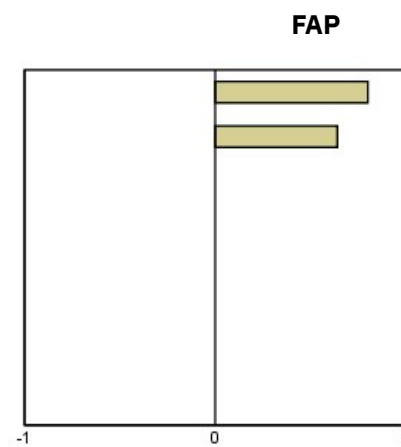
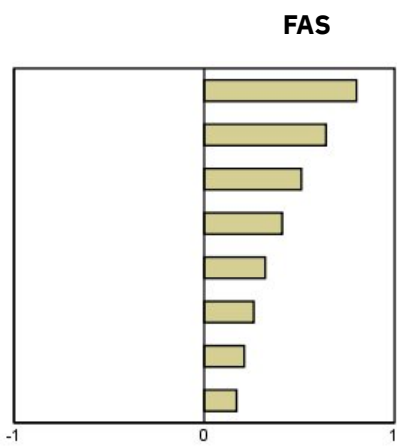
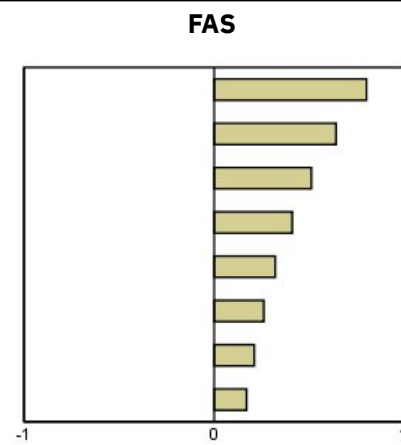


Tabla 7. ARIMA (0,1,0) (serie integrada)



Avisos

Esta información se ha desarrollado para productos y servicios ofrecidos en Estados Unidos. Este material puede estar disponible en otros idiomas en IBM. Sin embargo, deberá poseer una copia del producto o de la versión del producto en ese idioma para poder acceder a él.

Es posible que IBM no ofrezca los productos, servicios o características que se tratan en este documento en otros países. Consulte al representante local de IBM para obtener información sobre los productos y servicios que actualmente pueden adquirirse en su zona. Las referencias a programas, productos o servicios de IBM no pretenden establecer ni implicar que sólo puedan utilizarse dichos productos, programas o servicios de IBM. En su lugar, se puede utilizar cualquier producto, programa o servicio funcionalmente equivalente que no infrinja ninguno de los derechos de propiedad intelectual de IBM. Sin embargo, es responsabilidad del usuario evaluar y comprobar el funcionamiento de todo producto, programa o servicio que no sea de IBM.

IBM puede tener patentes o solicitudes de patente en tramitación que cubran la materia descrita en este documento. La posesión de este documento no le otorga ninguna licencia sobre dichas patentes. Puede enviar preguntas acerca de las licencias, por escrito, a:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
EE.UU.*

Para consultas sobre licencias relacionadas con información de doble byte (DBCS), póngase en contacto con el departamento de propiedad intelectual de IBM de su país o envíe sus consultas, por escrito, a:

*Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokio 103-8510, Japón*

INTERNATIONAL BUSINESS MACHINES CORPORATION PROPORCIONA ESTA PUBLICACIÓN "TAL CUAL" SIN GARANTÍAS DE NINGUNA CLASE, NI EXPLÍCITAS NI IMPLÍCITAS, QUE INCLUYEN, PERO NO SE LIMITAN A, LAS GARANTÍAS IMPLÍCITAS DE NO VULNERACIÓN, MERCANTIBILIDAD O ADECUACIÓN A UN FIN DETERMINADO. Algunas jurisdicciones no permiten la renuncia a las garantías explícitas o implícitas en ciertas transacciones; por lo tanto, es posible que esta declaración no sea aplicable en su caso.

Esta información puede incluir imprecisiones técnicas o errores tipográficos. Periódicamente se efectúan cambios en la información aquí contenida; estos cambios se incorporarán en nuevas ediciones de la publicación. IBM puede realizar mejoras y/o cambios en los productos y/o programas descritos en esta publicación en cualquier momento sin previo aviso.

Las referencias hechas en esta publicación a sitios web que no son de IBM se proporcionan sólo para la comodidad del usuario y no constituyen de modo alguno un aval de esos sitios web. Los materiales de estos sitios web no forman parte de los materiales para este producto IBM, por lo que la utilización de dichos sitios web es a cuenta y riesgo del usuario.

IBM puede utilizar o distribuir la información que se le proporcione del modo que considere adecuado sin incurrir por ello en ninguna obligación con el remitente.

Los tenedores de licencias de este programa que deseen obtener información acerca de éste con el fin de permitir: (i) el intercambio de información entre programas creados independientemente y otros programas (incluido el presente) y (ii) la utilización mutua de la información que se ha intercambiado, deben ponerse en contacto con:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
EE.UU.*

Dicha información puede estar disponible, sujeta a los términos y condiciones correspondientes, incluyendo, en algunos casos, el pago de una tarifa.

IBM proporciona el programa bajo licencia descrito en este documento y todo el material bajo licencia disponible para el mismo según los términos del acuerdo de cliente de IBM, el acuerdo internacional de licencia de programa de IBM o cualquier otro acuerdo equivalente entre las partes.

Los ejemplos de datos de rendimiento y de clientes citados se presentan solamente a efectos ilustrativos. Los resultados reales de rendimiento pueden variar en función de las configuraciones específicas y condiciones de operación.

La información acerca de productos que no son de IBM se ha obtenido de los proveedores de dichos productos, sus publicaciones anunciadas u otras fuentes disponibles públicamente. IBM no ha probado esos productos y no puede confirmar la exactitud del rendimiento, la compatibilidad ni ninguna otra afirmación relacionada con productos no IBM. Las preguntas relacionadas con las prestaciones de los productos que no son de IBM deben dirigirse a los proveedores de esos productos.

Las declaraciones sobre el futuro rumbo o intención de IBM están sujetas a cambio o retirada sin previo aviso y representan únicamente metas y objetivos.

Esta información contiene ejemplos de datos e informes utilizados en operaciones empresariales cotidianas. Para ilustrarlos de la forma más completa posible, los ejemplos incluyen nombres de personas, empresas, marcas y productos. Todos estos nombres son ficticios y cualquier parecido con los nombres de personas o empresas reales es pura coincidencia.

LICENCIA DE DERECHOS DE AUTOR:

Esta información contiene programas de aplicación de muestra en lenguaje fuente, que ilustran técnicas de programación en distintas plataformas operativas. Puede copiar, modificar y distribuir estos programas de ejemplo en cualquier formato sin abonar ninguna cantidad a IBM, con el fin de desarrollar, utilizar, comercializar o distribuir programas de aplicaciones de acuerdo con la interfaz de programación de aplicaciones correspondiente a la plataforma operativa para la que se han escrito los programas de ejemplo. Estos ejemplos no se han probado exhaustivamente bajo todas las condiciones. Por lo tanto, IBM no puede garantizar ni dar por supuesta la fiabilidad, la capacidad de servicio ni la funcionalidad de estos programas. Los programas de ejemplo se proporcionan "TAL CUAL" sin garantías de ningún tipo. IBM no será responsable de ningún daño resultante del uso de los programas de ejemplo.

Cada copia, parcial o completa, de estos programas de ejemplo, o cualquier trabajo obtenido a partir de los mismos, debe incluir el siguiente aviso de copyright:

© Copyright IBM Corp. 2021. Las partes de este código se derivan de IBM Corp. Programas de ejemplo

© Copyright IBM Corp. 1989 - 2021. Reservados todos los derechos.

Marcas comerciales

IBM, el logotipo de IBM e ibm.com son marcas comerciales o marcas registradas de International Business Machines Corp., registradas en numerosas jurisdicciones de todo el mundo. Otros nombres de productos y servicios pueden ser marcas registradas de IBM o de otras empresas. En la web hay disponible una lista actualizada de las marcas registradas de IBM en "Copyright and trademark information" en www.ibm.com/legal/copytrade.shtml.

Adobe, el logotipo Adobe, PostScript y el logotipo PostScript son marcas registradas o marcas comerciales de Adobe Systems Incorporated en Estados Unidos y/o otros países.

Intel, el logotipo de Intel, Intel Inside, el logotipo de Intel Inside, Intel Centrino, el logotipo de Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium y Pentium son marcas comerciales o marcas registradas de Intel Corporation o sus filiales en Estados Unidos y otros países.

Linux es una marca registrada de Linus Torvalds en los Estados Unidos o en otros países.

Microsoft, Windows, Windows NT, y el logotipo de Windows son marcas comerciales de Microsoft Corporation en Estados Unidos, otros países o ambos.

UNIX es una marca registrada de The Open Group en Estados Unidos y otros países.

Java y todas las marcas registradas y logotipos basados en Java son marcas registradas de Oracle o sus filiales.

Índice

A

- análisis armónico [22](#)
- análisis de series temporales
 - modelos causales temporales [24](#)
- Aplicar modelos de series temporales
 - estadístico de Box-Ljung [16](#)
 - estadísticos de bondad de ajuste [16](#), [17](#)
 - estadísticos en todos los modelos [16](#), [17](#)
 - función de autocorrelación parcial (FAP) de residuo [16](#), [17](#)
 - función de autocorrelación simple (FAS) de residuo [16](#), [17](#)
 - guardar modelos estimados de nuevo en XML [19](#)
 - guardar predicciones [19](#)
 - intervalos de confianza [17](#), [20](#)
 - modelos de mejor o peor ajuste [18](#)
 - nombres de variables nuevas [19](#)
 - parámetros de modelo [16](#)
 - período de estimación [14](#)
 - período de predicción [14](#)
 - previsiones [16](#), [17](#)
 - valores ajustados [17](#)
 - valores perdidos [20](#)
 - volver a estimar parámetros de modelo [14](#)

B

- BIC normalizado (criterio de información bayesiano)
 - en Aplicar modelos de series temporales [16](#), [17](#)
 - en modelizador de series temporales [10](#), [11](#)
- bondad de ajuste
 - definiciones [52](#)
 - en Aplicar modelos de series temporales [16](#), [17](#)
 - en modelizador de series temporales [10](#), [11](#)

C

- casos reservados [2](#)

D

- datos históricos
 - en Aplicar modelos de series temporales [17](#)
 - en modelizador de series temporales [11](#)
- Descomposición estacional
 - cálculo de las medias móviles [20](#)
 - crear variables [21](#)
 - guardar variables nuevas [21](#)
 - modelos [20](#)
 - supuestos [20](#)
- Diagramas espectrales
 - análisis espectral bivariado [22](#)
 - centrado de transformación [22](#)
 - supuestos [22](#)
 - ventanas espectrales [22](#)

E

- error absoluto máximo
 - en Aplicar modelos de series temporales [16](#), [17](#)
 - en modelizador de series temporales [10](#), [11](#)
- error absoluto porcentual promedio
 - en Aplicar modelos de series temporales [16](#), [17](#)
 - en modelizador de series temporales [10](#), [11](#)
- error porcentual absoluto máximo
 - en Aplicar modelos de series temporales [16](#), [17](#)
 - en modelizador de series temporales [10](#), [11](#)
- escenarios de modelos causales temporales [42–47](#)
- estacionaria R2
 - en Aplicar modelos de series temporales [16](#), [17](#)
 - en modelizador de series temporales [10](#), [11](#)
- estadístico de Box-Ljung
 - en Aplicar modelos de series temporales [16](#)
 - en modelizador de series temporales [10](#)
- eventos
 - en modelizador de series temporales [6](#)

F

- FAP
 - en Aplicar modelos de series temporales [16](#), [17](#)
 - en modelizador de series temporales [10](#), [11](#)
 - gráficos para procesos ARIMA puros [53](#)
- FAS
 - en Aplicar modelos de series temporales [16](#), [17](#)
 - en modelizador de series temporales [10](#), [11](#)
 - gráficos para procesos ARIMA puros [53](#)
- función de autocorrelación
 - en Aplicar modelos de series temporales [16](#), [17](#)
 - en modelizador de series temporales [10](#), [11](#)
 - gráficos para procesos ARIMA puros [53](#)
- función de autocorrelación parcial
 - en Aplicar modelos de series temporales [16](#), [17](#)
 - en modelizador de series temporales [10](#), [11](#)
 - gráficos para procesos ARIMA puros [53](#)
- funciones de transferencia
 - órdenes de denominador [9](#)
 - órdenes de diferencia [9](#)
 - órdenes de numerador [9](#)
 - órdenes estacionales [9](#)
 - retardo [9](#)

G

- guardar
 - especificaciones de modelo en XML [12](#)
 - modelos estimados de nuevo en XML [19](#)
 - nombres de variables nuevas [12](#), [19](#)
 - predicciones de modelo [12](#), [19](#)

I

intervalos de confianza
en Aplicar modelos de series temporales [17](#), [20](#)
en modelizador de series temporales [11](#), [13](#)

M

MAE
en Aplicar modelos de series temporales [16](#), [17](#)
en modelizador de series temporales [10](#), [11](#)

MAPE
en Aplicar modelos de series temporales [16](#), [17](#)
en modelizador de series temporales [10](#), [11](#)

MaxAE
en Aplicar modelos de series temporales [16](#), [17](#)
en modelizador de series temporales [10](#), [11](#)

MaxAPE
en Aplicar modelos de series temporales [16](#), [17](#)
en modelizador de series temporales [10](#), [11](#)

media del error absoluto
en Aplicar modelos de series temporales [16](#), [17](#)
en modelizador de series temporales [10](#), [11](#)

Modelizador de series temporales
ARIMA [3](#), [8](#)
estadístico de Box-Ljung [10](#)
estadísticos de bondad de ajuste [10](#), [11](#)
estadísticos en todos los modelos [10](#), [11](#)
eventos [6](#)
función de autocorrelación parcial (FAP) de residuo [10](#),
[11](#)
función de autocorrelación simple (FAS) de residuo [10](#),
[11](#)
funciones de transferencia [9](#)
guardar especificaciones de modelo en XML [12](#)
guardar predicciones [12](#)
intervalos de confianza [11](#), [13](#)
Modelizador experto [3](#)
modelos de mejor o peor ajuste [12](#)
nombres de modelo [13](#)
nombres de variables nuevas [12](#)
parámetros de modelo [10](#)
periodicidad [6](#), [8](#), [9](#)
período de estimación [3](#)
período de predicción [3](#), [13](#)
previsiones [10](#), [11](#)
suavizado exponencial [3](#), [6](#)
transformación de serie [6](#), [8](#), [9](#)
valores ajustados [11](#)
valores atípicos [6](#), [10](#)
valores perdidos [13](#)

Modelizador experto
limitar espacio de modelo [6](#)
valores atípicos [6](#)

modelo de suavizado exponencial amortiguado [6](#)
modelo de suavizado exponencial de Brown [6](#)
modelo de suavizado exponencial de Holt [6](#)
modelo de suavizado exponencial estacional simple [6](#)
modelo de suavizado exponencial simple [6](#)
modelo de suavizado exponencial simple de Winters
aditivo [6](#)
multiplicativo [6](#)

modelos
ARIMA [3](#)

modelos (*continuación*)
Modelizador experto [3](#)
suavizado exponencial [3](#), [6](#)

modelos ARIMA
funciones de transferencia [9](#)
valores atípicos [10](#)

modelos causales temporales [24](#), [25](#), [27–31](#), [33–35](#)

modelos de suavizado exponencial [3](#), [6](#)

N

nombres de modelo
en modelizador de series temporales [13](#)

nombres de variables
en Aplicar modelos de series temporales [19](#)
en modelizador de series temporales [12](#)

P

parámetros de modelo
en Aplicar modelos de series temporales [16](#)
en modelizador de series temporales [10](#)

periodicidad
en modelizador de series temporales [6](#), [8](#), [9](#)

período de estimación
en Aplicar modelos de series temporales [14](#)
en modelizador de series temporales [3](#)

período de predicción
en Aplicar modelos de series temporales [14](#)
en modelizador de series temporales [3](#), [13](#)

período de validación [2](#)

período histórico [2](#)

predicción de modelos causales temporales [36–39](#), [41](#)

previsiones
en Aplicar modelos de series temporales [16](#), [17](#)
en modelizador de series temporales [10](#), [11](#)

R

R2
en Aplicar modelos de series temporales [16](#), [17](#)
en modelizador de series temporales [10](#), [11](#)

raíz de la media cuadrática de los errores
en Aplicar modelos de series temporales [16](#), [17](#)
en modelizador de series temporales [10](#), [11](#)

residuos
en Aplicar modelos de series temporales [16](#), [17](#)
en modelizador de series temporales [10](#), [11](#)

RMSE
en Aplicar modelos de series temporales [16](#), [17](#)
en modelizador de series temporales [10](#), [11](#)

T

transformación de raíz cuadrada
en modelizador de series temporales [6](#), [8](#), [9](#)

transformación logarítmica
en modelizador de series temporales [6](#), [8](#), [9](#)

transformación logarítmica natural
en modelizador de series temporales [6](#), [8](#), [9](#)

V

- valor atípico aditivo
 - en modelizador de series temporales [6](#), [10](#)
- valor atípico aditivo estacional
 - en modelizador de series temporales [6](#), [10](#)
- valor atípico de cambio de nivel
 - en modelizador de series temporales [6](#), [10](#)
- valor atípico de parche aditivo
 - en modelizador de series temporales [6](#), [10](#)
- valor atípico de tendencia local
 - en modelizador de series temporales [6](#), [10](#)
- valor atípico innovador
 - en modelizador de series temporales [6](#), [10](#)
- valor atípico transitorio
 - en modelizador de series temporales [6](#), [10](#)
- valores ajustados
 - en Aplicar modelos de series temporales [17](#)
 - en modelizador de series temporales [11](#)
- valores atípicos
 - definiciones [53](#)
 - Modelizador experto [6](#)
 - modelos ARIMA [10](#)
- valores perdidos
 - en Aplicar modelos de series temporales [20](#)
 - en modelizador de series temporales [13](#)
- volver a estimar parámetros de modelo
 - en Aplicar modelos de series temporales [14](#)

X

- XML
 - guardar modelos de series temporales en XML [12](#)
 - guardar modelos estimados de nuevo en XML [19](#)

