

IBM SPSS Forecasting 30



Nota

Prima di utilizzare queste informazioni e il prodotto che supportano, leggere le informazioni in [“Informazioni particolari” a pagina 51](#).

Informazioni sul prodotto

Questa edizione si applica alla versione 30, release 0, modifica 0 di IBM® SPSS Statistiche e a tutte le release e modifiche successive se non diversamente indicato in nuove edizioni.

© **Copyright International Business Machines Corporation .**

Indice

Capitolo 1. Previsioni.....	1
Introduzione alla serie Time.....	1
Dati di serie temporali.....	1
Trasformazioni dati.....	2
Periodi di stima e validazione.....	2
Building Models e Previsioni di produzione.....	3
Modeler Serie Storiche.....	3
Specifiche delle Opzioni per il Modeler Expert.....	5
Modelli Personalizzati di livellamento esponenziale.....	6
Modelli ARIMA personalizzati.....	7
Output.....	10
Salvataggio Modello Previsioni e Modello Specifiche.....	12
Opzioni.....	12
TSMODEL Comando Funzioni Aggiuntive.....	13
Applica modelli di serie storiche.....	13
Output.....	15
Salvataggio Modello Previsioni e Modello Specifiche.....	18
Opzioni.....	18
TSAPPLICARE Comando Funzioni aggiuntive.....	19
Decomposizione stagionale.....	19
Decomposizione stagionale Salva.....	20
SEASON Comando Funzioni aggiuntive.....	20
Grafici spettrali.....	20
Funzioni Aggiuntive Comando SPECTRA.....	22
Modelli causali temporali.....	22
Come ottenere un modello causale temporale.....	23
Serie temporali al modello.....	24
Osservazioni	25
Intervallo di tempo per l'analisi	26
Aggregazione e distribuzione	27
Valori mancanti	27
Opzioni di dati generali	28
Opzioni di creazione generali	28
Serie da visualizzare	28
Opzioni di output	29
Periodo di stima	31
Previsione	32
Salva	32
Output interattivo	33
Applicazione Modelli Causali Temporali.....	34
Applicazione Modelli Causali Temporali.....	34
Previsione modello causale temporale.....	34
Scenari di modelli causali temporali.....	40
Misure di bontà - di - Fit.....	46
Tipi di valori anomali.....	46
Guida ai grafici ACF/PACF.....	47
Informazioni particolari.....	51
Marchi.....	52

Indice analitico.....	55
------------------------------	-----------

Capitolo 1. Previsioni

Le seguenti funzioni di previsione sono incluse in SPSS Statistiche Professional Edition o l'opzione Forecasting.

Introduzione alla serie Time

Una **serie temporale** è una serie di osservazioni ottenute misurando una singola variabile regolarmente in un periodo di tempo. In una serie di dati di inventario, ad esempio, le osservazioni potrebbero rappresentare livelli di inventario giornalieri per diversi mesi. Una serie che mostra la quota di mercato di un prodotto potrebbe consistere nella quota di mercato settimanale assunta nel corso di alcuni anni. Una serie di cifre totali di vendita potrebbe consistere in un'osservazione al mese per molti anni. Quello che ognuno di questi esempi ha in comune è che qualche variabile è stata osservata ad intervalli regolari, noti su un certo periodo di tempo. Così, la forma dei dati per una serie temporale tipica è una sequenza unica o un elenco di osservazioni che rappresentano misurazioni effettuate ad intervalli regolari.

Tabella 1. Serie temporali di inventario giornaliero

Ora	Settimana	Giorno	Livello di inventario
t ₁	1	Lunedì	160
t ₂	1	Martedì	135
t ₃	1	Mercoledì	129
t ₄	1	Giovedì	122
t ₅	1	Venerdì	108
t ₆	2	Lunedì	4
		...	
t ₆₀	12	Venerdì	120

Uno dei motivi più importanti per fare l'analisi delle serie storiche è quello di provare a prevedere i valori futuri della serie. Un modello della serie che ha spiegato i valori passati potrebbe anche prevedere se e quanto i prossimi valori aumenteranno o diminuiranno. La capacità di fare tali previsioni è ovviamente importante per qualsiasi campo aziendale o scientifico.

Dati di serie temporali

Dati basati su colonne

Ciascun campo di serie temporali contiene i dati di una singola serie temporale. Questa struttura è la struttura tradizionale dei dati delle serie temporali, così come utilizzato dalle procedure Modeler di serie temporali, Decomposizione stagionale e Grafici spettrali. Per definire, ad esempio, una serie temporale nell'Editor dei dati, fare clic sulla scheda **Vista variabile** ed immettere il nome della variabile in una riga vuota. Ciascuna osservazione di una serie temporale corrisponde ad un caso (ossia ad una riga dell'Editor dei dati).

Se si apre un foglio di calcolo che contiene dati di serie temporali, ciascuna serie deve essere disposta in una colonna nel foglio di calcolo. Se si dispone di un foglio di calcolo con le serie temporali già disposte su righe, è possibile aprire ed utilizzare l'opzione Trasponi del menu Dati per invertire le righe in colonne.

Dati multidimensionali

Per dati multidimensionali, ciascun campo delle serie temporali contiene i dati per più serie temporali. Serie temporali separate, all'interno di un particolare campo, vengono quindi identificate da un serie di valori di campi categoriali indicati come campi *dimensione*.

Ad esempio, i dati di vendita per regioni e marchi differenti potrebbero essere memorizzati in un singolo campo *vendite*, in modo che le dimensioni siano, in questo caso, *regione* e *marchio*. Ciascuna combinazione di *regione* e *marca* identifica una determinata serie temporale per *vendite*. Ad esempio, nella seguente tabella, i record che hanno 'north' per *regione* e 'brandX' per *marca* definiscono una singola serie temporale.

<i>Tabella 2. Dati multidimensionali</i>			
data	regione	marchio	vendite
01/01/2014	nord	brandX	82350
01/01/2014	nord	brandY	86380
01/01/2014	sud	brandX	91375
01/01/2014	sud	brandY	70320
01/02/2014	nord	brandX	83275
01/02/2014	nord	brandY	85260
01/02/2014	sud	brandX	94760
01/02/2014	sud	brandY	69870

Nota: I dati che vengono importati dai cubi OLAP, ad esempio da IBM Cognos TM1, sono rappresentati come dati multidimensionali.

Trasformazioni dati

Diverse procedure di trasformazione dei dati che vengono fornite nel sistema Core sono utili nell'analisi delle serie storiche. Queste trasformazioni si applicano solo ai dati a base di colonne, dove ogni campo della serie temporale contiene i dati per una singola serie temporale.

- La procedura Definire Date (sul menu Dati) genera variabili di data che vengono utilizzate per stabilire periodicità e per distinguere tra periodi storici, di validazione e previsione. La previsione è progettata per funzionare con le variabili create dalla procedura Definisci Date.
- La procedura Create Time Series (sul menu Transform) crea nuove variabili di serie storiche come funzioni delle variabili di serie storiche esistenti. Include funzioni che utilizzano le osservazioni vicine per il livellamento, la media e la differenziazione.
- La procedura Sostituisci valori mancanti (sul menu Transform) sostituisce i valori mancanti di sistema e utente con le stime basate su uno dei diversi metodi. I dati mancanti all'inizio o al termine della serie non pongono problemi, perché riducono semplicemente la lunghezza utile della serie. Intervalli vuoti al centro di una serie (dati mancanti *incorporati*) possono porre problemi molto più seri.

Consultare il manuale *Core System User's Guide* per informazioni dettagliate sulle trasformazioni dei dati per le serie storiche.

Periodi di stima e validazione

Spesso è utile dividere la propria serie temporale in una *stima storica*, periodo e un periodo *validazione*. Si sviluppa un modello sulla base delle osservazioni nel periodo di stima (storico) e poi lo prova a vedere come funziona bene nel periodo di validazione. Forzando il modello a fare previsioni per i punti già conosciuti (i punti nel periodo di validazione), si ottiene un'idea di quanto bene il modello faccia a previsione.

I casi nel periodo di validazione sono tipicamente indicati come casi holdout perché sono detenuti - indietro dal processo di costruzione del modello. Una volta soddisfatto che il modello faccia un adeguato lavoro di previsione, è possibile ridefinire il periodo di stima per includere i casi holdout e quindi costruire il tuo modello finale.

Building Models e Previsioni di produzione

Il Modulo aggiuntivo di previsione prevede le seguenti procedure per l'espletamento dei compiti di creazione di modelli e di produzione delle previsioni:

- La procedura “[Modeler Serie Storiche](#)” a [pagina 3](#) crea modelli per le serie storiche e produce previsioni. Include un Expert Modeler che determina automaticamente il modello migliore per ciascuna delle tue serie storiche. Per gli analisti esperti che desiderano un maggiore grado di controllo, fornisce anche strumenti per l'edificio di modello personalizzato.
- La procedura “[Applica modelli di serie storiche](#)” a [pagina 13](#) applica modelli di serie temporali esistenti -- creati da Time Series Modeler -- al dataset attivo. Questo ti permette di ottenere previsioni per serie per le quali sono disponibili nuovi o rivisti dati, senza ricostruire i tuoi modelli. Se c'è ragione di pensare che un modello sia cambiato, può essere ricostruito utilizzando il Modeler Time Series.
- La procedura “[Modelli causali temporali](#)” a [pagina 22](#) costruisce modelli di serie temporali autoregressive per ogni destinazione e determina automaticamente i migliori input che hanno una relazione causale con l'obiettivo. La procedura produce output interattivo che è possibile utilizzare per esplorare le relazioni causali. La procedura può anche generare previsioni, rilevare gli outliers e determinare la serie che più probabilmente causa un outlier.
- La procedura “[Previsione modello causale temporale](#)” a [pagina 34](#) applica un modello di causale temporale al dataset attivo. È possibile utilizzare questa procedura per ottenere previsioni per serie per le quali sono disponibili dati più attuali, senza ricostruire i propri modelli. È anche possibile utilizzarlo per determinare serie che più probabilmente causano degli outlier rilevati dalla procedura Temporal Causal Models.

Modeler Serie Storiche

La procedura di Time Series Modeler stima livellamento esponenziale, univariata Autoregressive Integrated Moving Average (ARIMA) e multivariate ARIMA (o modelli di funzione di trasferimento) modelli per serie storiche, e produce previsioni. La procedura include un Expert Modeler che tenta di identificare e stimare automaticamente il modello ARIMA o di livellamento espansivo per una o più serie variabili dipendenti, eliminando così la necessità di identificare un modello appropriato attraverso il processo e l'errore. In alternativa, è possibile specificare un modello personalizzato ARIMA o di livellamento esponenziale.

Esempio. Sei un product manager responsabile di previsione delle vendite e delle entrate unitari del mese prossimo per ciascuno di 100 prodotti separati, e ha poca o nessuna esperienza nella modellazione di serie storiche. I dati di vendita dell'unità storica per tutti i 100 prodotti sono memorizzati in un unico foglio di calcolo Excel. Dopo l'apertura del foglio di calcolo in IBM SPSS Statistiche, si utilizza l'Expert Modeler e si richiedono previsioni un mese sul futuro. L'Esperto Modeler trova il miglior modello di vendite unitari per ognuno dei tuoi prodotti e utilizza quei modelli per produrre le previsioni. Dal momento che Expert Modeler può gestire più serie di input, basta eseguire la procedura una volta per ottenere previsioni per tutti i vostri prodotti. Scegliendo di salvare le previsioni al dataset attivo, è possibile esportare facilmente i risultati in Excel.

Statistiche. Misure della bontà di adattamento: R -quadrato stazionario, R -quadrato (R^2), errore quadratico medio (RMSE), errore assoluto medio (MAE), errore percentuale assoluto medio (MAPE), errore assoluto massimo (MaxAE), errore percentuale assoluto massimo (MaxAPE), criterio di informazione bayesiano normalizzato (BIC). Residui: funzione autocorrelazione, funzione di autocorrelazione parziale, Ljung - Box Q . Per i modelli ARIMA: ordini ARIMA per variabili dipendenti, ordini di funzioni di trasferimento per variabili indipendenti e stime di outlier. Inoltre, livellare le stime dei parametri per i modelli di livellamento esponenziale.

Grafici. Grafici di riepilogo per tutti i modelli: istogrammi di R -quadrato stazionario, R -quadrato (R^2), errore quadratico medio (RMSE), errore assoluto medio (MAE), errore percentuale assoluto medio (MAPE), errore assoluto massimo (MaxAE), errore percentuale assoluto massimo (MaxAPE), criterio di informazione bayesiano normalizzato (BIC); grafici a scatole di autocorrelazioni residue e autocorrelazioni parziali. Risultati per singoli modelli: valori di previsione, valori di adattamento, valori osservati, limiti di confidenza superiore e inferiore, autocorrelazioni residue e autocorrelazioni parziali.

Considerazioni Dati Modeler di serie temporali

Dati. La variabile dipendente e le eventuali variabili indipendenti dovrebbero essere numeriche.

Ipotesi. La variabile dipendente e le eventuali variabili indipendenti sono trattate come serie temporali, il che significa che ogni caso rappresenta un punto temporale, con successivi casi separati da un intervallo di tempo costante.

- **Stazionarietà.** Per i modelli ARIMA personalizzati, la serie temporale da modellare dovrebbe essere stazionaria. Il modo più efficace per trasformare una serie non stazionaria in uno stazionario è attraverso una trasformazione di differenza - disponibile dalla finestra di dialogo Crea time Series .
- **Previsioni.** Per la produzione di previsioni utilizzando modelli con variabili indipendenti (predittore), il dataset attivo deve contenere valori di queste variabili per tutti i casi nel periodo di previsione. Inoltre, le variabili indipendenti non dovrebbero contenere valori mancanti nel periodo di stima.

Definizione Date

Sebbene non richiesto, si consiglia di utilizzare la finestra di dialogo Definisci Date per specificare la data associata al primo caso e l'intervallo di tempo tra i casi successivi. Ciò avviene prima di utilizzare il Modelli di Time Series e i risultati in una serie di variabili che etichettano la data associata ad ogni caso. Inoltre fissa una periodicità ipotizza dei dati - ad esempio una periodicità di 12 se l'intervallo di tempo tra i casi successivi è di un mese. Questa periodicità è richiesta se siete interessati a creare modelli stagionali. Se non sei interessato ai modelli stagionali e non richiami le etichette di data sul tuo output, puoi saltare la finestra di dialogo Definisci Date. L'etichetta associata ad ogni caso è quindi semplicemente il numero del caso.

Per utilizzare il Modeler di serie Time

1. Dai menu, scegliere:

Analizzare > Previsione > Creare Modelli tradizionali ...

2. Nella scheda Variabili, selezionare una o più variabili dipendenti da modellare.
3. Dalla casella di discesa Metodo, selezionare un metodo di modellazione. Per la modellazione automatica, lasciare il metodo predefinito di **Expert Modeler**. In questo modo verrà richiamato l'Expert Modeler per determinare il modello di adattamento migliore per ognuna delle variabili dipendenti.

Per produrre previsioni:

4. Fare clic sulla scheda **Opzioni**.
5. Specificare il periodo di previsione. Questo produrrà un grafico che include previsioni e valori osservati.

Facoltativamente, è possibile:

- Selezionare una o più variabili indipendenti. Le variabili indipendenti sono trattate molto come le variabili predittore nell'analisi di regressione ma sono facoltative. Possono essere inclusi nei modelli ARIMA ma non modelli di livellamento esponenziale. Se si specifica **Expert Modeler** come metodo di modellazione e includono variabili indipendenti, verranno considerati solo modelli ARIMA.
- Clicca su **Criteri** per specificare i dettagli di modellazione.
- Salva previsioni, intervalli di confidenza e residui di rumore.
- Salva i modelli stimati in formato XML. I modelli salvati possono essere applicati a dati nuovi o revisati per ottenere previsioni aggiornate senza modelli di ricostruzione.
- Ottenere statistiche di riepilogo su tutti i modelli stimati.
- Specificare le funzioni di trasferimento per le variabili indipendenti nei modelli ARIMA personalizzati.
- Abilita rilevamento automatico degli outliers.
- Modello di tempo specifico come outliers per i modelli ARIMA personalizzati.

Metodi Di Modellazione

I metodi di modellazione disponibili sono:

Expert Modeler. L'Esperto Modeler trova automaticamente il modello di adattamento migliore per ogni serie dipendente. Se si specificano le variabili indipendenti (predittore), Expert Modeler seleziona quelle che presentano una relazione statisticamente significativa con la serie dipendente, per includerle nei modelli ARIMA. Laddove è opportuno, le variabili del modello vengono trasformate tramite la differenziazione e/o una trasformazione a radice quadrata o logaritmica naturale. Per impostazione predefinita, Expert Modeler considera entrambi i modelli di livellamento esponenziale e ARIMA. È possibile, tuttavia, limitare il Modeler Expert per cercare solo modelli ARIMA o per cercare solo modelli di livellamento esponenziale. È inoltre possibile specificare il rilevamento automatico dei valori anomali.

Livellamento esponenziale. Utilizzare questa opzione per specificare un modello di livellamento esponenziale personalizzato. È possibile scegliere tra vari modelli di livellamento esponenziale che si differenziano tra loro per il trattamento della tendenza e della stagionalità.

ARIMA. Utilizzare questa opzione per specificare un modello ARIMA personalizzato. Questo implica esplicitamente specificare ordini di media autoregressivi e di spostamento, così come il grado di differenziazione. È possibile includere variabili indipendenti (predittore) e definire funzioni di trasferimento per qualsiasi o tutte. È inoltre possibile specificare l'individuazione automatica degli outliers o specificare un insieme esplicito di valori anomali.

Periodi di stima e previsione

Periodo di stima. Il periodo di stima definisce la serie di casi utilizzati per determinare il modello. Per impostazione predefinita, il periodo di stima include tutti i casi nel dataset attivo. Per impostare il periodo di stima, selezionare **In base all'intervallo di tempo o di caso** nella finestra di dialogo Seleziona casi . A seconda dei dati disponibili, il periodo di stima utilizzato dalla procedura può variare per variabile dipendente e quindi differire dal valore visualizzato. Per una determinata variabile dipendente, il periodo di stima vero è il periodo lasciato dopo l'eliminazione di eventuali valori mancanti contigui della variabile che si verificano all'inizio o alla fine del periodo di stima specificato.

Periodo di previsione. Il periodo di previsione inizia al primo caso dopo il periodo di stima e per impostazione predefinita passa all'ultimo caso nel dataset attivo. È possibile impostare la fine del periodo di previsione dalla scheda Opzioni .

Specifiche delle Opzioni per il Modeler Expert

L'Esperto Modeler fornisce opzioni per vincolare la serie di modelli candidati, specificando la gestione degli outliers, e incluse le variabili evento.

Selezione del modello e specifica evento

La scheda Modello consente di specificare i tipi di modelli considerati dal Modeler Expert e di specificare le variabili evento.

Tipo di modello. Sono disponibili le seguenti opzioni:

- **Tutti i modelli.** Expert Modeler considera entrambi i modelli ARIMA e di livellamento esponenziale.
- **Solo modelli di livellamento esponenziale.** Expert Modeler considera solo i modelli di livellamento esponenziale.
- **Solo modelli ARIMA.** Expert Modeler considera solo i modelli ARIMA.

Expert Modeler considera i modelli stagionali. Questa opzione viene attivata solo se è stata definita una periodicità per il file di dati attivo. Quando questa opzione viene selezionata (controllata), Expert Modeler considera sia modelli stagionali che non stagionali. Se l'opzione non è selezionata, Expert Modeler considera solo i modelli non stagionali.

Periodicità corrente. Indica la periodicità (se presente) attualmente definita per il dataset attivo. L'attuale periodicità è data come intero - ad esempio 12 per la periodicità annuale, con ogni caso che rappresenta un mese. Il valore *Nessuno* viene visualizzato se non è stata impostata alcuna periodicità. I modelli stagionali richiedono una periodicità. È possibile impostare la periodicità dalla finestra di dialogo Definisci Date.

Eventi. Selezionare eventuali variabili indipendenti che devono essere trattate come variabili evento. Per le variabili evento, i casi con un valore di 1 indicano i tempi in cui la serie dipendente dovrebbe essere influenzata dall'evento. Valori diversi da 1 indicano nessun effetto.

Gestione degli Outliers con il Modeler Expert

La scheda Outliers consente di scegliere il rilevamento automatico degli outliers nonché il tipo di outliers da rilevare.

Rileva automaticamente valori anomali. Per impostazione predefinita, il rilevamento automatico dei valori anomali non viene eseguito. Selezionare (controllare) questa opzione per eseguire il rilevamento automatico degli outliers, quindi selezionare uno o più dei seguenti tipi di outlier:

- Additivo
- Spostamento di livello
- Innovativo
- Transiente
- Additivo stagionale
- Tendenza locale
- Patch additivo

Modelli Personalizzati di livellamento esponenziale

Tipo di modello. Modelli exponential smoothing¹ sono classificati come stagionali o non stagionali. I modelli stagionali sono disponibili solo se è stata definita una periodicità per il dataset attivo (vedi "Periodicità corrente" di seguito).

- *Semplice.* Questo modello è appropriato per le serie in cui non vi è alcuna tendenza o stagionalità. Il suo unico parametro di livellamento è uniforme. Il livellamento esponenziale semplice è molto simile a un modello ARIMA con zero ordini di autoregressione, un ordine di differenziazione, un ordine di media mobile e nessuna costante.
- *Tendenza lineare di Holt.* Questo modello è appropriato per le serie in cui è presente una tendenza lineare e nessuna stagionalità. I parametri di livellamento sono livello e tendenza e non sono limitati dai rispettivi valori. Il modello di Holt è più generico del modello di Brown, ma allunga i tempi di calcolo per le serie molto grandi. Il livellamento esponenziale di Holt è molto simile al modello ARIMA con zero ordini di autoregressione, due ordini di differenziazione e due ordini di media mobile.
- *Tendenza lineare di Brown.* Questo modello è appropriato per le serie in cui è presente una tendenza lineare e nessuna stagionalità. I parametri di livellamento di questo modello sono uniformi, quindi la loro tendenza è uguale. Il modello di Brown rappresenta quindi un caso speciale del modello di Holt. Il livellamento esponenziale di Brown è molto simile al modello ARIMA con zero ordini di autoregressione, due ordini di differenziazione, due ordini di media mobile e un coefficiente per il secondo ordine di media mobile uguale al quadrato di metà del coefficiente per il primo ordine.
- *Tendenza smorzata.* Questo modello è appropriato per le serie con una tendenza lineare che si sta attenuando e senza stagionalità. I parametri di livellamento del modello sono rappresentati da una tendenza uniforme e smorzata. Il livellamento esponenziale smorzato è simile al modello ARIMA con 1 ordine di autoregressione, 1 ordine di differenze e 2 ordini di medie mobili.
- *Stagionale semplice.* Questo modello è appropriato per le serie senza tendenza e con un effetto stagionale costante nel tempo. I parametri di livellamento di questo modello sono uniformi e stagionali. Il livellamento esponenziale stagionale semplice è molto simile al modello ARIMA con zero ordini di autoregressione, un ordine di differenziazione, un ordine di differenziazione stagionale, e ordini 1, p e p + 1 di media mobile, dove p è il numero di periodi in un intervallo stagionale (per i dati mensili p = 12).
- *Additivo di Winters.* Questo modello è adatto per le serie con una tendenza lineare e un effetto stagionale che non dipende dal livello della serie. I parametri di livellamento di questo modello sono uniformi, in tendenza e stagionali. Il livellamento esponenziale additivo di Winters è molto simile

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

al modello ARIMA con zero ordini di autoregressione, un ordine di differenziazione, un ordine di differenziazione stagionale e $p + 1$ di ordini di media mobile, dove p è il numero di periodi in un intervallo stagionale (per i dati mensili, $p = 12$).

- **Moltiplicativo di Winters.** Questo modello è appropriato per le serie con una tendenza lineare e un effetto stagionale che dipende dal livello della serie. I parametri di livellamento di questo modello sono uniformi, in tendenza e stagionali. Il livellamento esponenziale moltiplicativo di Winters non è simile ad alcun altro modello ARIMA.

Periodicità corrente. Indica la periodicità (se presente) attualmente definita per il dataset attivo. L'attuale periodicità è data come intero - ad esempio 12 per la periodicità annuale, con ogni caso che rappresenta un mese. Il valore *Nessuno* viene visualizzato se non è stata impostata alcuna periodicità. I modelli stagionali richiedono una periodicità. È possibile impostare la periodicità dalla finestra di dialogo Definisci Date.

Trasformazione variabile dipendente. È possibile specificare una trasformazione eseguita su ciascuna variabile dipendente prima di essere modellata.

- **Messuno.** Non viene eseguita nessuna trasformazione.
- **Radice quadrata.** Trasformazione a radice quadrata.
- **Logaritmo naturale.** Trasformazione di un logaritmo naturale.

Modelli ARIMA personalizzati

Il Modeler di Time Series consente di costruire modelli custom non stagionali o stagionali ARIMA (Autoregressive Integrated Moving Average) - noto anche come Box-Jenkins² modelli -- con o senza serie fissa di variabili predittore. È possibile definire le funzioni di trasferimento per qualsiasi o tutte le variabili del predittore, e specificare il rilevamento automatico degli outliers o specificare un insieme esplicito di valori anomali.

- Tutte le variabili indipendenti (predittore) specificate sulla scheda Variabili sono esplicitamente incluse nel modello. Questo è in contrasto con l'utilizzo di Expert Modeler dove le variabili indipendenti sono incluse solo se hanno una relazione statisticamente significativa con la variabile dipendente.

Modello Specifica per i Modelli ARIMA personalizzati

La scheda Modello consente di specificare la struttura di un modello ARIMA personalizzato.

Ordini ARIMA. Immettere i valori per i vari componenti ARIMA del modello nelle celle corrispondenti della griglia struttura. Tutti i valori devono essere interi non negativi. Per i componenti autoregressivo e media mobile, il valore rappresenta l'ordine massimo. Nel modello vengono inclusi tutti gli ordini inferiori positivi. Ad esempio, se si specifica 2, il modello include gli ordini 2 e 1. Le celle nella colonna Stagionale sono abilitate solo se è stata definita una periodicità per il dataset attivo (vedi "Periodicità corrente" di seguito).

- **Autoregressivo (p).** Il numero di ordini autoregressivi nel modello. Gli ordini autoregressivi specificano quali valori precedenti della serie vengono utilizzati per prevedere i valori correnti. Per esempio, un ordine autoregressivo 2 specifica di utilizzare il valore dei due periodi precedenti della serie per prevedere il valore corrente.
- **Differenza (d).** Specifica l'ordine di differenziazione applicato alla serie prima di eseguire la stima dei modelli. La differenziazione è necessaria quando sono presenti delle tendenze (di norma, le serie che presentano delle tendenze sono non stazionarie e nei modelli ARIMA si presume che vi sia stazionarietà) e viene utilizzata per rimuoverne l'effetto. L'ordine di differenziazione corrisponde al grado di tendenza della serie, la differenziazione di primo grado tiene conto delle tendenze lineari, la differenziazione di secondo grado tiene conto delle tendenze quadratiche e così via.
- **Media mobile (q).** Il numero di ordini di media mobile nel modello. Gli ordini di media mobile specificano il modo in cui vengono utilizzate le deviazioni provenienti dalla media della serie per

² Casella, G. E. P., G. M. Jenkins, e G. C. Reinsel. 1994. *Analisi delle serie temporali: previsione e controllo*, 3rd ed. Scogliere di Englewood, N.J.: Sala Prentice.

prevedere i valori correnti. Per esempio, gli ordini di media mobile 1 e 2 specificano di considerare le deviazioni dalla media della serie degli ultimi due periodi precedenti per prevedere i valori correnti della serie.

Gradi stagionali. I componenti autoregressivo, media mobile e differenziazione stagionali hanno lo stesso ruolo delle corrispettive controparti non stagionali. Per gli ordini stagionali tuttavia, i valori di serie correnti vengono influenzati dai valori di serie precedenti separati da uno o più periodi stagionali. Ad esempio, per i dati mensili (periodo stagionale di 12), un ordine stagionale 1 è il valore della serie corrente è influenzato dal valore della serie che precede di 12 periodi quello corrente. Specificare un ordine stagionale 1, per i dati mensili, è quindi come specificare un ordine non stagionale 12.

Periodicità corrente. Indica la periodicità (se presente) attualmente definita per il dataset attivo. L'attuale periodicità è data come intero - ad esempio 12 per la periodicità annuale, con ogni caso che rappresenta un mese. Il valore *Nessuno* viene visualizzato se non è stata impostata alcuna periodicità. I modelli stagionali richiedono una periodicità. È possibile impostare la periodicità dalla finestra di dialogo Definisci Date.

Trasformazione variabile dipendente. È possibile specificare una trasformazione eseguita su ciascuna variabile dipendente prima di essere modellata.

- **Messuno.** Non viene eseguita nessuna trasformazione.
- **Radice quadrata.** Trasformazione a radice quadrata.
- **Logaritmo naturale.** Trasformazione di un logaritmo naturale.

Includi costante nel modello. L'inclusione di una costante è standard a meno che non si sia sicuri che il valore complessivo della serie media sia di 0. L'esclusione della costante è consigliata quando viene applicata la differenziazione.

Funzioni Transfer in Modelli ARIMA personalizzati

La scheda Funzione di trasferimento (presente solo se sono specificate variabili indipendenti) consente di definire le funzioni di trasferimento per qualsiasi o tutte le variabili indipendenti specificate sulla scheda Variabili. Le funzioni di trasferimento consentono di specificare il modo in cui i valori passati delle variabili indipendenti (predittore) vengono utilizzati per prevedere i valori futuri delle serie dipendenti.

Ordini della funzione di trasferimento. Immettere i valori per i vari componenti della funzione di trasferimento nelle celle corrispondenti della griglia struttura. Tutti i valori devono essere interi non negativi. Per i componenti numeratore e denominatore, il valore rappresenta l'ordine massimo. Nel modello vengono inclusi tutti gli ordini inferiori positivi. Inoltre, per i componenti numeratore l'ordine 0 viene incluso sempre. Ad esempio, se si specifica 2 per il numeratore, il modello include gli ordini 2, 1 e 0. Se si specifica 3 per denominatore, il modello include gli ordini 3, 2 e 1. Le celle nella colonna Stagionale sono abilitate solo se è stata definita una periodicità per il dataset attivo (vedi "Periodicità corrente" di seguito).

- **Numeratore.** L'ordine numerico della funzione di trasferimento. Specifica quali valori precedenti dalla serie indipendente (predittore) selezionate vengono utilizzati per prevedere i valori correnti della serie dipendente. Per esempio, un ordine numeratore 1 specifica che per prevedere il valore corrente di ogni serie dipendente viene utilizzato il valore di una serie indipendente di un periodo precedente, così come il valore corrente della serie indipendente.
- **Denominatore.** L'ordine denominatore della funzione di trasferimento. Specifica come vengono utilizzate le deviazioni provenienti dalla media della serie, per i valori precedenti della serie indipendente (predittore), per prevedere i valori correnti della serie dipendente. Per esempio, un ordine denominatore 1 specifica di considerare le deviazioni dal valore della media di una serie indipendente di un periodo precedente quando si prevede il valore corrente di ogni serie dipendente.
- **Differenza.** Specifica l'ordine di differenziazione applicato alla serie indipendente (predittore) selezionata prima di eseguire la stima dei modelli. La differenziazione è necessaria quando sono presenti delle tendenze e viene utilizzata per rimuoverne l'effetto.

Gradi stagionali. I componenti numeratore, denominatore e differenziazione stagionali, hanno lo stesso ruolo delle corrispettive controparti non stagionali. Per gli ordini stagionali tuttavia, i valori di serie correnti

vengono influenzati dai valori di serie precedenti separati da uno o più periodi stagionali. Ad esempio, per i dati mensili (periodo stagionale di 12), un ordine stagionale 1 è il valore della serie corrente è influenzato dal valore della serie che precede di 12 periodi quello corrente. Specificare un ordine stagionale 1, per i dati mensili, è quindi come specificare un ordine non stagionale 12.

Periodicità corrente. Indica la periodicità (se presente) attualmente definita per il dataset attivo. L'attuale periodicità è data come intero - ad esempio 12 per la periodicità annuale, con ogni caso che rappresenta un mese. Il valore *Nessuno* viene visualizzato se non è stata impostata alcuna periodicità. I modelli stagionali richiedono una periodicità. È possibile impostare la periodicità dalla finestra di dialogo Definisci Date.

Ritardo. L'impostazione di un ritardo induce a ritardare l'influenza della variabile indipendente per il numero di intervalli specificato. Ad esempio, se il ritardo è impostato su 5, il valore della variabile indipendente al tempo t non influiscono sulle previsioni fino a quando sono trascorsi cinque periodi ($t + 5$).

Trasformazione. Specifica di una funzione di trasferimento relativa a un insieme di variabili indipendenti che include anche una trasformazione facoltativa da eseguire su tali variabili.

- **Messuno.** Non viene eseguita nessuna trasformazione.
- **Radice quadrata.** Trasformazione a radice quadrata.
- **Logaritmo naturale.** Trasformazione di un logaritmo naturale.

Outliers in Modelli ARIMA personalizzati

La scheda Outliers fornisce le seguenti scelte per la gestione degli outliers³: individuarli automaticamente, specificare punti particolari come valori anomali o non rilevare o modellarli.

Non individuare valori anomali o modellarli. Per impostazione predefinita, i valori anomali non vengono né rilevati né modellati. Selezionare questa opzione per disattivare qualsiasi rilevamento o modello di valori anomali.

Rileva automaticamente valori anomali. Selezionare questa opzione per eseguire il rilevamento automatico dei valori anomali e selezionare uno o più dei seguenti tipi di valori anomali:

- Additivo
- Spostamento di livello
- Innovativo
- Transiente
- Additivo stagionale
- Tendenza locale
- Patch additivo

Modello di tempo specifico come outliers. Selezionare questa opzione per specificare i punti temporali particolari come valori anomali. Utilizzare una riga separata della griglia di definizione di Outlier per ogni outlier. Inserire valori per tutte le celle in una determinata riga.

- **Tipo.** Il tipo di outlier. I tipi supportati sono: additivo (default), shift di livello, innovational, transitorio, additivo stagionale e tendenza locale.

Nota 1: se non è stata definita alcuna specifica data per il dataset attivo, la griglia Definizione di Outlier mostra la singola colonna *Osservazione*. Per specificare un outlier, inserire il numero di riga (come visualizzato nell'editor dei dati) del relativo caso.

Nota 2: la colonna *Ciclo* (se presente) nella griglia Definizione di Outlier si riferisce al valore della variabile *CYCLE_* nel dataset attivo.

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

Output

L'output disponibile include i risultati per i singoli modelli nonché i risultati calcolati su tutti i modelli. I risultati per i singoli modelli possono essere limitati a una serie di modelli di adattamento più o più poveri basati su criteri definiti dall'utente.

Statistiche e tabelle di previsione

La scheda Statistiche fornisce opzioni per la visualizzazione delle tabelle dei risultati di modellazione.

Visualizza le misure di adattamento, la statistica Ljung - Box e il numero di outliers per modello.

Selezionare (controllare) questa opzione per visualizzare una tabella contenente le misure di adattamento selezionate, il valore Ljung - Box e il numero di outliers per ogni modello stimato.

Misure di adattamento. È possibile selezionare uno o più dei seguenti per l'inclusione nella tabella contenente le misure di adattamento per ogni modello stimato:

- Fermo R - quadrato
- R - quadrato
- Errore valore quadratico medio
- Errore percentuale assoluto medio
- Errore assoluto medio
- Errore percentuale assoluto massimo
- Errore assoluto massimo
- BIC normalizzato

Statistiche per i modelli di confronto. Questo gruppo di opzioni controlla la visualizzazione di tabelle contenenti statistiche calcolate su tutti i modelli stimati. Ogni opzione genera una tabella separata. È possibile selezionare una o più delle seguenti opzioni:

- **Bontà di adattamento.** Tabella delle statistiche di riepilogo e dei percentili per stazionari R - quadrato, R -quadrato, radice media errore quadrata, errore percentuale assoluta, errore assoluto medio, massimo errore percentuale assoluto, massimo errore assoluto e criterio di informazione Bayesiano normalizzato.
- **Funzione di autocorrelazione residua (ACF).** Tabella delle statistiche di riepilogo e dei percentili per le autocorrelazioni dei residui su tutti i modelli stimati.
- **Funzione di autocorrelazione parziale residua (PACF).** Tabella delle statistiche di riepilogo e dei percentili per autocorrelazioni parziali dei residui su tutti i modelli stimati.

Statistiche per i Modelli individuali. Questo gruppo di opzioni controlla la visualizzazione di tabelle contenenti informazioni dettagliate per ogni modello stimato. Ogni opzione genera una tabella separata. È possibile selezionare una o più delle seguenti opzioni:

- **Stime dei parametri.** Visualizza una tabella di stime dei parametri per ogni modello stimato. Le tabelle separate vengono visualizzate per i modelli di livellamento esponenziale e ARIMA. Se gli outliers esistono, le stime dei parametri per loro vengono visualizzate anche in una tabella separata.
- **Funzione di autocorrelazione residua (ACF).** Visualizza una tabella di autocorrelazioni residue per ritardo per ogni modello stimato. La tabella include gli intervalli di confidenza per le autocorrelazioni.
- **Funzione di autocorrelazione parziale residua (PACF).** Visualizza una tabella di autocorrelazioni parziali residue per lag per ogni modello stimato. La tabella include gli intervalli di confidenza per le autocorrelazioni parziali.

Visualizza le previsioni. Visualizza una tabella di previsioni del modello e intervalli di confidenza per ogni modello stimato. Il periodo di previsione è impostato dalla scheda Opzioni.

Grafici

La scheda Plotti fornisce opzioni per la visualizzazione dei tracciati dei risultati di modellazione.

Grafici per il confronto modelli

Questo gruppo di opzioni controlla la visualizzazione di grafici contenenti statistiche calcolate su tutti i modelli stimati. Ogni opzione genera una trama separata. È possibile selezionare una o più delle seguenti opzioni:

- Fermo R - quadrato
- R - quadrato
- Errore valore quadratico medio
- Errore percentuale assoluto medio
- Errore assoluto medio
- Errore percentuale assoluto massimo
- Errore assoluto massimo
- BIC normalizzato
- Funzione di autocorrelazione dei residui (ACF)
- Funzione di autocorrelazione parziale dei residui (PACF)

Grafici per singoli modelli

Serie. Selezionare (controllare) questa opzione per ottenere complotti dei valori previsti per ogni modello stimato. È possibile selezionare uno o più dei seguenti per l'inclusione nella trama:

- **Valori osservati.** I valori osservati della serie dipendente.
- **Previsioni.** Il modello ha previsto valori per il periodo di previsione.
- **Valori di adattamento.** Il modello ha previsto valori per il periodo di stima.
- **Intervalli di confidenza per le previsioni.** Gli intervalli di confidenza per il periodo di previsione.
- **Intervalli di confidenza per i valori di adattamento.** Gli intervalli di confidenza per il periodo di stima.

Funzione di autocorrelazione residua (ACF). Visualizza una trama di autocorrelazioni residue per ogni modello stimato.

Funzione di autocorrelazione parziale residua (PACF). Visualizza una trama di autocorrelazioni parziali residue per ogni modello stimato.

Limitare l'output ai modelli Best - or Peggiori - Fitting

La scheda Filtro di output fornisce opzioni per limitare l'output sia tabellare che grafico ad un sottoinsieme dei modelli stimati. Puoi scegliere di limitare l'output ai modelli di adattamento più convenienti e / o più poveri secondo i criteri di adattamento forniti. Per impostazione predefinita, tutti i modelli stimati sono inclusi nell'output.

Modelli di adattamento. Selezionare (controllare) questa opzione per includere i modelli di adattamento migliori nell'output. Selezionare una misura di bontà e specificare il numero di modelli da includere. La selezione di questa opzione non preclude anche la selezione dei modelli di adattamento più poveri. In quel caso, l'output sarà costituito dai modelli più poveri, oltre che dai migliori allestimento.

- **Numero fisso di modelli.** Specifica che i risultati vengono visualizzati per i modelli di best - fitting n . Se il numero supera il numero di modelli stimati, vengono visualizzati tutti i modelli.
- **Percentuale del numero totale di modelli.** Specifica che i risultati vengono visualizzati per i modelli con valori di bontà di adattamento nella parte superiore n in tutti i modelli stimati.

Modelli di adattamento più poveri. Selezionare (controllare) questa opzione per includere i modelli di adattamento più poveri nell'output. Selezionare una misura di bontà e specificare il numero di modelli da includere. La selezione di questa opzione non preclude anche la selezione dei modelli di adattamento migliori. In quel caso, l'output sarà costituito dai modelli di adattamento migliori così come quelli più poveri.

- **Numero fisso di modelli.** Specifica che i risultati vengono visualizzati per i modelli di adattamento più poveri n . Se il numero supera il numero di modelli stimati, vengono visualizzati tutti i modelli.

- **Percentuale del numero totale di modelli.** Specifica che i risultati vengono visualizzati per i modelli con valori di bontà di adattamento nella parte inferiore n in tutti i modelli stimati.

Bontà di Fit Misura. Selezionare la misura della bontà di adattamento da utilizzare per i modelli di filtraggio. Il default è fisso R quadrato.

Salvataggio Modello Previsioni e Modello Specifiche

La scheda Salva consente di salvare le previsioni del modello come nuove variabili nel dataset attivo e salvare le specifiche del modello in un file esterno in formato XML.

Salva Variabili. È possibile salvare le previsioni del modello, gli intervalli di confidenza e i residui come nuove variabili nel dataset attivo. Ogni serie dipendente dà origine alla propria serie di nuove variabili e ogni nuova variabile contiene valori sia per periodi di stima che di previsione. Si aggiungono nuovi casi se il periodo di previsione si estende oltre la lunghezza della serie variabile dipendente. Scegli di salvare nuove variabili selezionando la casella di controllo Salva associata per ciascuno. Per impostazione predefinita, non vengono salvate nuove variabili.

- **Valori previsti.** Il modello ha previsto valori.
- **Limiti di confidenza inferiore.** Abbassare i limiti di confidenza per i valori previsti.
- **Limiti di confidenza superiore.** Limiti di confidenza superiore per i valori previsti.
- **Residui di rumore.** I residui del modello. Quando vengono eseguite trasformazioni della variabile dipendente (ad esempio, log naturale), questi sono i residui per la serie trasformata.
- **Prefisso nome variabile.** Specificare prefissi da utilizzare per i nuovi nomi di variabili, oppure lasciare i prefissi predefiniti. I nomi delle variabili sono costituiti dal prefisso, dal nome della variabile dipendente associata e da un identificativo del modello. Il nome variabile viene esteso se necessario per evitare conflitti di denominazione variabili. Il prefisso deve essere conforme alle regole per i nomi variabili validi.

Export Modello file. Le specifiche del modello per tutti i modelli stimati vengono esportate nel file specificato in formato XML. I modelli salvati possono essere utilizzati per ottenere previsioni aggiornate.

- **File XML.** Le specifiche del modello vengono salvate in un file XML che può essere utilizzato con le applicazioni IBM SPSS .
- **File PMML.** Le specifiche del modello vengono salvate in un file XML compatibile con PMML che può essere utilizzato con applicazioni compatibili con PMML, tra cui le applicazioni IBM SPSS .

Opzioni

La scheda Opzioni consente di impostare il periodo di previsione, specificare la gestione dei valori mancanti, impostare la larghezza dell'intervallo di confidenza, specificare un prefisso personalizzato per gli identificativi del modello e impostare il numero di cartellini mostrati per le autocorrelazioni.

Periodo di previsione. Il periodo di previsione inizia sempre con il primo caso dopo la fine del periodo di stima (l'insieme dei casi utilizzati per determinare il modello) e passa attraverso l'ultimo caso nel dataset attivo o una data specificata dall'utente. Per impostazione predefinita, la fine del periodo di stima è l'ultimo caso nel dataset attivo, ma può essere modificato dalla finestra di dialogo Seleziona casi selezionando **In base all'intervallo di tempo o di caso**.

- **Primo caso dopo la fine del periodo di stima attraverso l'ultimo caso in dataset attivo.** Selezionare questa opzione quando la fine del periodo di stima è precedente all'ultimo caso nel dataset attivo e si desiderano le previsioni attraverso l'ultimo caso. Questa opzione è tipicamente utilizzata per produrre previsioni per un periodo holdout, consentendo il confronto delle previsioni del modello con un sottoinsieme dei valori effettivi.
- **Primo caso dopo la fine del periodo di stima attraverso una data specificata.** Selezionare questa opzione per specificare esplicitamente la fine del periodo di previsione. Questa opzione è tipicamente utilizzata per produrre previsioni oltre la fine della serie effettiva. Inserire valori per tutte le celle della griglia Data.

Se non è stata definita alcuna specifica data per il dataset attivo, la griglia Data mostra la singola colonna *Osservazione*. Per specificare la fine del periodo di previsione, inserire il numero di riga (come visualizzato nell'editor dei dati) del relativo caso.

La colonna *Ciclo* (se presente) nella griglia Data si riferisce al valore della variabile *CYCLE_* nel dataset attivo.

Valori mancanti definiti dall'utente. Queste opzioni controllano la gestione dei valori mancanti degli utenti.

- **Trattare come non valido.** I valori mancanti dell'utente sono trattati come valori mancanti di sistema.
- **Trattare come valido.** I valori mancanti dell'utente sono trattati come dati validi.

Politica di valore mancante. Le seguenti regole si applicano al trattamento dei valori mancanti (include valori mancanti di sistema e valori mancanti dell'utente trattati come non validi) durante la procedura di modellazione:

- I casi con valori mancanti di una variabile dipendente che si verificano all'interno del periodo di stima sono inclusi nel modello. La gestione specifica del valore mancante dipende dal metodo di stima.
- Viene emesso un avviso se una variabile indipendente ha valori mancanti all'interno del periodo di stima. Per il Modeler Expert, i modelli che coinvolgono la variabile indipendente sono stimati senza la variabile. Per l'ARIMA personalizzato, i modelli che coinvolgono la variabile indipendente non sono stimati.
- Se una qualsiasi variabile indipendente presenta valori mancanti entro il periodo di previsione, la procedura emette un avviso e le previsioni per quanto possibile.

Intervalli di confidenza (%). Intervalli di confidenza sono calcolati per le previsioni del modello e le autocorrelazioni residue. È possibile specificare qualsiasi valore positivo inferiore a 100. Per impostazione predefinita, viene utilizzato un intervallo di confidenza del 95%.

Prefisso per identificativi di modello in Output. Ogni variabile dipendente specificata sulla scheda Variabili dà origine ad un modello stimato separato. I modelli si distinguono con nomi univocamente costituiti da un prefisso personalizzabile insieme a un suffisso intero. È possibile inserire un prefisso o lasciare il default di *Modello*.

Numero massimo di Lags Shown in ACF e PACF Output. È possibile impostare il numero massimo di ritardi mostrato nelle tabelle e nei grafici delle autocorrelazioni e delle autocorrelazioni parziali.

TSMODEL Comando Funzioni Aggiuntive

È possibile personalizzare la modellazione della serie temporale se incollare le selezioni in una finestra di sintassi e modificare la sintassi del comando TSMODEL risultante. Il linguaggio di sintassi del comando consente di:

- Specificare il periodo stagionale dei dati (con la parola chiave *SEASONLENGTH* nel comando *AUXILIARY*). Questo sovrasta l'attuale periodicità (se presente) per il dataset attivo.
- Specificare i lags non consecutivi per i componenti di funzione ARIMA e di trasferimento personalizzati (con i sottocomandi *ARIMA* e *TRANSFERFUNCTION*). Ad esempio, è possibile specificare un modello ARIMA personalizzato con lags autoregressive degli ordini 1, 3 e 6; oppure una funzione di trasferimento con i cartellini numeratori degli ordini 2, 5 e 8.
- Fornire più di una serie di specifiche di modellazione (ad esempio, il metodo di modellazione, gli ordini ARIMA, le variabili indipendenti e così via) per una singola esecuzione della procedura di Time Series Modeler (con il comando *MODEL*).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

Applica modelli di serie storiche

I modelli La Procedura Applica Tempi Serie Tempi carica i modelli di serie temporali esistenti da un file esterno e li applica al dataset attivo. È possibile utilizzare questa procedura per ottenere previsioni per

serie per le quali sono disponibili nuovi o rivisti dati, senza ricostruire i propri modelli. vengono generati utilizzando la procedura Modeler Time Series .

Esempio. Sei un gestore di magazzino con un grande rivenditore, e responsabile di ciascuno di 5.000 prodotti. Hai usato il Modeler Expert per creare modelli che prevedono vendite per ogni prodotto tre mesi nel futuro. Il tuo data warehouse viene aggiornato ogni mese con dati di vendita reali che ti piacerebbe utilizzare per produrre previsioni aggiornate mensili. La Procedura Applicata Time Series Models consente di effettuare ciò utilizzando i modelli originali, e di ristimare semplicemente i parametri del modello per rendere conto dei nuovi dati.

Statistiche. Misure della bontà di adattamento: R -quadrato stazionario, R -quadrato (R^2), errore quadratico medio (RMSE), errore assoluto medio (MAE), errore percentuale assoluto medio (MAPE), errore assoluto massimo (MaxAE), errore percentuale assoluto massimo (MaxAPE), criterio di informazione bayesiano normalizzato (BIC). Residui: funzione autocorrelazione, funzione di autocorrelazione parziale, Ljung - Box Q .

Grafici. Grafici di riepilogo per tutti i modelli: istogrammi di R -quadrato stazionario, R -quadrato (R^2), errore quadratico medio (RMSE), errore assoluto medio (MAE), errore percentuale assoluto medio (MAPE), errore assoluto massimo (MaxAE), errore percentuale assoluto massimo (MaxAPE), criterio di informazione bayesiano normalizzato (BIC); grafici a scatole di autocorrelazioni residue e autocorrelazioni parziali. Risultati per singoli modelli: valori di previsione, valori di adattamento, valori osservati, limiti di confidenza superiore e inferiore, autocorrelazioni residue e autocorrelazioni parziali.

Applica Considerazioni Sui Modelli di serie temporali

Dati. Le variabili (dipendenti e indipendenti) a cui verranno applicate le modelli dovranno essere numeriche.

Ipotesi. I modelli vengono applicati alle variabili nel dataset attivo con gli stessi nomi delle variabili specificate nel modello. Tutte queste variabili sono trattate come serie temporali, il che significa che ogni caso rappresenta un punto temporale, con successivi casi separati da un intervallo di tempo costante.

- **Previsioni.** Per la produzione di previsioni utilizzando modelli con variabili indipendenti (predittore), il dataset attivo deve contenere valori di queste variabili per tutti i casi nel periodo di previsione. Se i parametri del modello sono rivalutati, allora le variabili indipendenti non dovrebbero contenere valori mancanti nel periodo di stima.

Definizione Date

La Procedura Applicata Tempi Serie richiede che la periodicità, se presente, del dataset attivo corrisponde alla periodicità dei modelli da applicare. Se si è semplicemente preclusi utilizzando lo stesso dataset (magari con dati nuovi o rivisti) come quello utilizzato per la build del modello, allora questa condizione sarà soddisfatta. Se non esiste alcuna periodicità per il dataset attivo, ti verrà data la possibilità di navigare nella finestra di dialogo Definisci Date per crearne una. Se però i modelli sono stati creati senza specificare una periodicità, allora anche il dataset attivo dovrebbe essere senza uno.

Per Applicare Modelli

1. Dai menu, scegliere:

Analizzare > Previsione > Applicare Modelli Tradizionali ...

2. Inserire la specifica del file per un file di modello o fare clic su **Sfoggia** e selezionare un file di modello (i file di modello vengono creati con la procedura Modeler di serie Time).

Facoltativamente, è possibile:

- Ristimare i parametri del modello utilizzando i dati nel dataset attivo. Le previsioni vengono create utilizzando i parametri ristimati.
- Salva previsioni, intervalli di confidenza e residui di rumore.
- Salva modelli rivalutati in formato XML.

Parametri dei modelli e misure della bontà di adattamento

Carica da file di modello. Le previsioni vengono prodotte utilizzando i parametri del modello dal file modello senza ristimare tali parametri. La bontà delle misure di adattamento visualizzate in output e

utilizzate per filtrare i modelli (best - or worst - fitting) vengono prese dal file modello e riflettono i dati utilizzati quando ogni modello è stato sviluppato (o ultimo aggiornamento). Con questa opzione, le previsioni non tengono conto dei dati storici - per variabili dipendenti o indipendenti - nel dataset attivo. È necessario scegliere **Ristima dai dati** se si desidera che i dati storici influenzino le previsioni. Inoltre, le previsioni non tengono conto dei valori della serie dipendente nel periodo di previsione - ma tengono conto dei valori di variabili indipendenti nel periodo di previsione. Se si hanno valori più attuali della serie dipendente e vogliono che siano inclusi nelle previsioni, è necessario rivalutare, regolando il periodo di stima per includere questi valori.

Ristima dai dati. I parametri del modello vengono ristimati utilizzando i dati nel dataset attivo. La rivalutazione dei parametri del modello non ha alcun effetto sulla struttura del modello. Ad esempio, un modello ARIMA (1,0, 1) rimarrà tale, ma verranno rivalutati i parametri autoregressivi e in movimento. La ristima non comporta la rilevazione di nuovi anomalie. Gli outliers, se presenti, sono sempre presi dal file modello.

- **Periodo di stima.** Il periodo di stima definisce la serie di casi utilizzati per ristimare i parametri del modello. Per impostazione predefinita, il periodo di stima include tutti i casi nel dataset attivo. Per impostare il periodo di stima, selezionare **In base all'intervallo di tempo o di caso** nella finestra di dialogo Seleziona casi . A seconda dei dati disponibili, il periodo di stima utilizzato dalla procedura può variare per modello e quindi differire dal valore visualizzato. Per un determinato modello, il periodo di stima vero è il periodo lasciato dopo l'eliminazione di eventuali valori mancanti contigui, dalla variabile dipendente del modello, verificandosi all'inizio o alla fine del periodo di stima specificato.

Periodo di previsione

Il periodo di previsione per ogni modello inizia sempre con il primo caso dopo la fine del periodo di stima e passa attraverso l'ultimo caso nel dataset attivo o una data specificata dall'utente. Se i parametri non vengono ristimati (questo è il default), allora il periodo di stima per ogni modello è la serie di casi utilizzati quando il modello è stato sviluppato (o ultimo aggiornamento).

- **Primo caso dopo la fine del periodo di stima attraverso l'ultimo caso in dataset attivo.** Selezionare questa opzione quando la fine del periodo di stima è precedente all'ultimo caso nel dataset attivo e si desiderano le previsioni attraverso l'ultimo caso.
- **Primo caso dopo la fine del periodo di stima attraverso una data specificata.** Selezionare questa opzione per specificare esplicitamente la fine del periodo di previsione. Inserire valori per tutte le celle della griglia Data.

Se non è stata definita alcuna specifica data per il dataset attivo, la griglia Data mostra la singola colonna *Osservazione*. Per specificare la fine del periodo di previsione, inserire il numero di riga (come visualizzato nell'editor dei dati) del relativo caso.

La colonna *Ciclo* (se presente) nella griglia Data si riferisce al valore della variabile *CYCLE_* nel dataset attivo.

Output

L'output disponibile include i risultati per i singoli modelli oltre che i risultati su tutti i modelli. I risultati per i singoli modelli possono essere limitati a una serie di modelli di adattamento più o più poveri basati su criteri definiti dall'utente.

Statistiche e tabelle di previsione

La scheda Statistiche fornisce opzioni per la visualizzazione delle tabelle delle statistiche di adattamento del modello, dei parametri del modello, delle funzioni di autocorrelazione e delle previsioni. A meno che i parametri del modello non vengano ristimati (**Ristima dai dati** nella scheda Modelli), i valori visualizzati delle misure di adattamento, i valori di Ljung - Box e i parametri del modello sono quelli dal file modello e riflettono i dati utilizzati quando ogni modello è stato sviluppato (o ultimo aggiornamento). Le informazioni dell'outlier vengono sempre prese dal file del modello.

Visualizza le misure di adattamento, la statistica Ljung - Box e il numero di outliers per modello.

Selezionare (controllare) questa opzione per visualizzare una tabella contenente le misure di adattamento selezionate, il valore Ljung - Box e il numero di outliers per ogni modello.

Misure di adattamento. È possibile selezionare uno o più dei seguenti per l'inclusione nella tabella contenente le misure di adattamento per ogni modello:

- Fermo R - quadrato
- R - quadrato
- Errore valore quadratico medio
- Errore percentuale assoluto medio
- Errore assoluto medio
- Errore percentuale assoluto massimo
- Errore assoluto massimo
- BIC normalizzato

Statistiche per i modelli di confronto. Questo gruppo di opzioni controlla la visualizzazione di tabelle contenenti statistiche su tutti i modelli. Ogni opzione genera una tabella separata. È possibile selezionare una o più delle seguenti opzioni:

- **Bontà di adattamento.** Tabella delle statistiche di riepilogo e dei percentili per stazionari R - quadrato, R -quadrato, radice media errore quadrata, errore percentuale assoluta, errore assoluto medio, massimo errore percentuale assoluto, massimo errore assoluto e criterio di informazione Bayesiano normalizzato.
- **Funzione di autocorrelazione residua (ACF).** Tabella delle statistiche di riepilogo e dei percentili per le autocorrelazioni dei residui su tutti i modelli stimati. Questa tabella è disponibile solo se i parametri del modello sono ristimati (**Ristima dai dati** sulla scheda Modelli).
- **Funzione di autocorrelazione parziale residua (PACF).** Tabella delle statistiche di riepilogo e dei percentili per autocorrelazioni parziali dei residui su tutti i modelli stimati. Questa tabella è disponibile solo se i parametri del modello sono ristimati (**Ristima dai dati** sulla scheda Modelli).

Statistiche per i Modelli individuali. Questo gruppo di opzioni controlla la visualizzazione di tabelle contenenti informazioni dettagliate per ogni modello. Ogni opzione genera una tabella separata. È possibile selezionare una o più delle seguenti opzioni:

- **Stime dei parametri.** Visualizza una tabella di stime dei parametri per ogni modello. Le tabelle separate vengono visualizzate per i modelli di livellamento esponenziale e ARIMA. Se gli outliers esistono, le stime dei parametri per loro vengono visualizzate anche in una tabella separata.
- **Funzione di autocorrelazione residua (ACF).** Visualizza una tabella di autocorrelazioni residue per ritardo per ogni modello stimato. La tabella include gli intervalli di confidenza per le autocorrelazioni. Questa tabella è disponibile solo se i parametri del modello sono ristimati (**Ristima dai dati** sulla scheda Modelli).
- **Funzione di autocorrelazione parziale residua (PACF).** Visualizza una tabella di autocorrelazioni parziali residue per lag per ogni modello stimato. La tabella include gli intervalli di confidenza per le autocorrelazioni parziali. Questa tabella è disponibile solo se i parametri del modello sono ristimati (**Ristima dai dati** sulla scheda Modelli).

Visualizza le previsioni. Visualizza una tabella di previsioni del modello e intervalli di confidenza per ogni modello.

Grafici

La scheda Plotti fornisce opzioni per la visualizzazione di tracciati di statistiche di adattamento del modello, funzioni di autocorrelazione e valori di serie (incluse le previsioni).

Grafici per il confronto modelli

Questo gruppo di opzioni controlla la visualizzazione di grafici contenenti statistiche su tutti i modelli. A meno che i parametri del modello non vengano ristimati (**Ristima dai dati** nella scheda Modelli), i valori visualizzati sono quelli dal file del modello e riflettono i dati utilizzati quando ogni modello è

stato sviluppato (o ultimo aggiornamento). Inoltre, i tracciati di autocorrelazione sono disponibili solo se vengono rivalutati i parametri del modello. Ogni opzione genera una trama separata. È possibile selezionare una o più delle seguenti opzioni:

- Fermo R - quadrato
- R - quadrato
- Errore valore quadratico medio
- Errore percentuale assoluto medio
- Errore assoluto medio
- Errore percentuale assoluto massimo
- Errore assoluto massimo
- BIC normalizzato
- Funzione di autocorrelazione dei residui (ACF)
- Funzione di autocorrelazione parziale dei residui (PACF)

Grafici per singoli modelli

Serie. Selezionare (controllare) questa opzione per ottenere complotti dei valori previsti per ogni modello. I valori osservati, i valori di adattamento, gli intervalli di confidenza per i valori di adattamento e le autocorrelazioni sono disponibili solo se vengono rivisti i parametri del modello (**Ristima dai dati** sulla scheda Modelli). È possibile selezionare uno o più dei seguenti per l'inclusione nella trama:

- **Valori osservati.** I valori osservati della serie dipendente.
- **Previsioni.** Il modello ha previsto valori per il periodo di previsione.
- **Valori di adattamento.** Il modello ha previsto valori per il periodo di stima.
- **Intervalli di confidenza per le previsioni.** Gli intervalli di confidenza per il periodo di previsione.
- **Intervalli di confidenza per i valori di adattamento.** Gli intervalli di confidenza per il periodo di stima.

Funzione di autocorrelazione residua (ACF). Visualizza una trama di autocorrelazioni residue per ogni modello stimato.

Funzione di autocorrelazione parziale residua (PACF). Visualizza una trama di autocorrelazioni parziali residue per ogni modello stimato.

Limitare l'output ai modelli Best - or Peggiori - Fitting

La scheda Filtro di output fornisce opzioni per limitare l'output sia tabellare che grafico ad un sottoinsieme di modelli. Puoi scegliere di limitare l'output ai modelli di adattamento più convenienti e / o più poveri secondo i criteri di adattamento forniti. Per impostazione predefinita, tutti i modelli sono inclusi nell'output. A meno che i parametri del modello non vengano ristimati (**Ristima dai dati** sulla scheda Modelli), i valori delle misure di adattamento utilizzati per i modelli di filtraggio sono quelli dal file modello e riflettono i dati utilizzati quando ogni modello è stato sviluppato (o ultimo aggiornamento).

Modelli di adattamento. Selezionare (controllare) questa opzione per includere i modelli di adattamento migliori nell'output. Selezionare una misura di bontà e specificare il numero di modelli da includere. La selezione di questa opzione non preclude anche la selezione dei modelli di adattamento più poveri. In quel caso, l'output sarà costituito dai modelli più poveri, oltre che dai migliori allestimento.

- **Numero fisso di modelli.** Specifica che i risultati vengono visualizzati per i modelli di best - fitting n . Se il numero supera il numero totale di modelli, vengono visualizzati tutti i modelli.
- **Percentuale del numero totale di modelli.** Specifica che i risultati vengono visualizzati per i modelli con valori di bontà di adattamento nella parte superiore n in tutti i modelli.

Modelli di adattamento più poveri. Selezionare (controllare) questa opzione per includere i modelli di adattamento più poveri nell'output. Selezionare una misura di bontà e specificare il numero di modelli da includere. La selezione di questa opzione non preclude anche la selezione dei modelli di adattamento migliori. In quel caso, l'output sarà costituito dai modelli di adattamento migliori così come quelli più poveri.

- **Numero fisso di modelli.** Specifica che i risultati vengono visualizzati per i modelli di adattamento più poveri n . Se il numero supera il numero totale di modelli, vengono visualizzati tutti i modelli.
- **Percentuale del numero totale di modelli.** Specifica che i risultati vengono visualizzati per i modelli con valori di bontà di adattamento nella parte inferiore n in tutti i modelli.

Bontà di Fit Misura. Selezionare la misura della bontà di adattamento da utilizzare per i modelli di filtraggio. Il default è fisso R - quadrato.

Salvataggio Modello Previsioni e Modello Specifiche

La scheda Salva consente di salvare le previsioni del modello come nuove variabili nel dataset attivo e salvare le specifiche del modello in un file esterno in formato XML.

Salva Variabili. È possibile salvare le previsioni del modello, gli intervalli di confidenza e i residui come nuove variabili nel dataset attivo. Ogni modello dà origine alla propria serie di nuove variabili. Si aggiungono nuovi casi se il periodo di previsione si estende oltre la lunghezza della serie variabile dipendente associata al modello. A meno che i parametri del modello non vengano ristimati (**Ristima dai dati** sulla scheda Modelli), i valori previsti e i limiti di confidenza sono creati solo per il periodo di previsione. Scegli di salvare nuove variabili selezionando la casella di controllo Salva associata per ciascuno. Per impostazione predefinita, non vengono salvate nuove variabili.

- **Valori previsti.** Il modello ha previsto valori.
- **Limiti di confidenza inferiore.** Abbassare i limiti di confidenza per i valori previsti.
- **Limiti di confidenza superiore.** Limiti di confidenza superiore per i valori previsti.
- **Residui di rumore.** I residui del modello. Quando vengono eseguite trasformazioni della variabile dipendente (ad esempio, log naturale), questi sono i residui per la serie trasformata. Questa scelta è disponibile solo se i parametri del modello sono ristimati (**Ristima dai dati** sulla scheda Modelli).
- **Prefisso nome variabile.** Specificare prefissi da utilizzare per i nuovi nomi variabili o lasciare i prefissi predefiniti. I nomi delle variabili sono costituiti dal prefisso, dal nome della variabile dipendente associata e da un identificativo del modello. Il nome variabile viene esteso se necessario per evitare conflitti di denominazione variabili. Il prefisso deve essere conforme alle regole per i nomi variabili validi.

Le specifiche del modello **Export Model File**, contenenti parametri ristimati e statistiche di adattamento, vengono esportate nel file specificato in formato XML. Questa opzione è disponibile solo se i parametri del modello sono ristimati (**Ristima dai dati** sulla scheda Modelli).

- **File XML.** Le specifiche del modello vengono salvate in un file XML che può essere utilizzato con le applicazioni IBM SPSS.
- **File PMML.** Le specifiche del modello vengono salvate in un file XML compatibile con PMML che può essere utilizzato con applicazioni compatibili con PMML, incluse le applicazioni IBM SPSS.

Opzioni

La scheda Opzioni consente di specificare la gestione dei valori mancanti, impostare la larghezza dell'intervallo di confidenza e impostare il numero di cartellini mostrati per le autocorrelazioni.

Valori mancanti definiti dall'utente. Queste opzioni controllano la gestione dei valori mancanti degli utenti.

- **Trattare come non valido.** I valori mancanti dell'utente sono trattati come valori mancanti di sistema.
- **Trattare come valido.** I valori mancanti dell'utente sono trattati come dati validi.

Politica di valore mancante. Le seguenti regole si applicano al trattamento dei valori mancanti (include valori mancanti di sistema e valori mancanti degli utenti trattati come non validi):

- I casi con valori mancanti di una variabile dipendente che si verificano all'interno del periodo di stima sono inclusi nel modello. La gestione specifica del valore mancante dipende dal metodo di stima.
- Per i modelli ARIMA viene emesso un avviso se un predittore ha valori mancanti all'interno del periodo di stima. Eventuali modelli che coinvolgono il predittore non sono ristimati.

- Se una qualsiasi variabile indipendente presenta valori mancanti entro il periodo di previsione, la procedura emette un avviso e le previsioni per quanto possibile.

Intervalli di confidenza (%). intervalli di confidenza sono calcolati per le previsioni del modello e le autocorrelazioni residue. È possibile specificare qualsiasi valore positivo inferiore a 100. Per impostazione predefinita, viene utilizzato un intervallo di confidenza del 95%.

Numero massimo di Lags Shown in ACF e PACF Output. È possibile impostare il numero massimo di ritardi mostrato nelle tabelle e nei grafici delle autocorrelazioni e delle autocorrelazioni parziali. Questa opzione è disponibile solo se i parametri del modello sono ristimati (**Ristima dai dati** sulla scheda Modelli).

TSAPPLICARE Comando Funzioni aggiuntive

Ulteriori funzioni sono disponibili se incolla le selezioni in una finestra di sintassi e modifica la sintassi del comando TSAPPLY risultante. Il linguaggio di sintassi del comando consente di:

- Specificare che solo un sottoinsieme dei modelli in un file di modello deve essere applicato al dataset attivo (con le parole chiave DROP e KEEP nel sottocomando MODEL).
- Applicare i modelli da due o più file di modello ai tuoi dati (con il comando MODEL). Ad esempio, un file di modello potrebbe contenere modelli per serie che rappresentano vendite unitari e un altro potrebbe contenere modelli per serie che rappresentano ricavi.

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

Decomposizione stagionale

La procedura Decomposizione stagionale decompone la serie in una componente stagionale, una tendenza combinata e una componente ciclo, ed in una componente "errore". La procedura è un'implementazione del Metodo I del Census Bureau, noto come metodo per il calcolo del rapporto della media mobile.

Esempio. Uno scienziato è interessato ad analizzare misurazioni mensili del livello di ozono in una particolare stazione meteorologica. L'obiettivo è determinare se c'è tendenza nei dati. Per scoprire qualsiasi tendenza reale, lo scienziato deve innanzitutto rendere conto della variazione delle letture a causa degli effetti stagionali. La Procedura di decomposizione stagionale può essere utilizzata per eliminare eventuali variazioni stagionali sistemiche. L'analisi di tendenza viene poi eseguita su una serie destagionalizzata.

Statistiche. L'insieme dei fattori stagionali.

Considerazioni Dati di decomposizione stagionale

Dati. Le variabili dovrebbero essere numeriche.

Ipotesi. Le variabili non dovrebbero contenere alcun dato mancante incorporato. È necessario definire almeno una componente data periodica.

Stima Fattori Stagionali

1. Dai menu, scegliere:

Analizzare > Previsione > Decomposizione stagionale ...

2. Selezionare una o più variabili dall'elenco disponibile e spostarle nella lista delle variabili. Da notare che l'elenco include solo variabili numeriche.

Tipo di modello. La Procedura Stagionale Decomposizione offre due approcci diversi per la modellazione dei fattori stagionali: moltiplicativo o additivo.

- *Moltiplicativo.* La componente stagionale è un fattore che moltiplicato per la serie adattata alla stagione restituisce la serie originale. Calcola componenti stagionali proporzionate al livello globale della serie. Le osservazioni senza la variazione stagionale hanno una componente stagionale pari a 1.

- **Additivo.** Alla serie adattata alla stagione, viene aggiunta una correzione stagionale per ottenere i valori osservati. Questa correzione tenta di rimuovere la componente stagionale da una serie storica per favorire l'osservazione di altre componenti che potrebbero essere "mascherate" dalla stagionalità. Si tratta di componenti stagionali che non dipendono dal livello globale della serie. Le osservazioni senza la variazione stagionale hanno una componente stagionale pari a 0.

Peso medio di spostamento. Le Opzioni di Moving Average Weight consentono di specificare come trattare la serie quando si calcolano le medie mobili. Queste opzioni sono disponibili solo se la periodicità della serie è pari. Se la periodicità è strana, tutti i punti sono ponderati ugualmente.

- **Tutti i punti uguali.** Le medie mobili vengono calcolate con un'estensione uguale alla periodicità e con tutti i punti ponderati in modo uguale. Questo metodo viene usato sempre se la periodicità è dispari.
- **Endpoint ponderati per .5.** Le medie mobili per le serie con periodicità pari sono calcolate con un'estensione pari alla periodicità più 1 e assegnando peso 0,5 agli endpoint dell'estensione.

Facoltativamente, è possibile:

- Clicca su **Salva** per specificare come devono essere salvate le nuove variabili.

Decomposizione stagionale Salva

Crea variabili. Consente di scegliere come trattare le nuove variabili.

- **Aggiungi al file.** Le nuove serie create dalla decomposizione stagionale vengono salvate come variabili regolari nel dataset attivo. I nomi di variabile sono formati da un prefisso di tre lettere, un carattere di sottolineatura e un numero.
- **Sostituisci esistenti.** Le nuove serie create dalla decomposizione stagionale vengono salvate come variabili temporanee nel dataset attivo. Allo stesso tempo, eventuali variabili temporanee esistenti create dalle procedure Previsioni vengono ignorate. I nomi delle variabili sono formati da un prefisso di tre lettere, da un segno di cancelletto (#) e da un numero.
- **Non creare.** Le nuove serie non vengono aggiunte al dataset attivo.

Nuovi Nomi Variabili

La Procedura Stagionale Decomposizione crea quattro nuove variabili (serie), con i seguenti prefissi a tre lettere, per ogni serie specificata:

Saf. *Fattori di regolazione stagionale.* Questi valori indicano l'effetto di ogni periodo sul livello della serie.

SAS. *Serie destagionalizzata.* Questi sono i valori ottenuti dopo la rimozione della variazione stagionale di una serie.

STC. *Componenti del ciclo di tendenza Smoothed.* Questi valori mostrano il trend e il comportamento ciclico presenti nella serie.

asimmetria *Valori o "di errore".* I valori che rimangono dopo i componenti stagionali, di tendenza e di ciclo sono stati rimossi dalla serie.

SEASON Comando Funzioni aggiuntive

Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare qualsiasi periodicità all'interno del comando SEASON piuttosto che selezionare una delle alternative offerte dalla procedura Definisci Date.

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

Grafici spettrali

La procedura Spectral Plotti viene utilizzata per identificare il comportamento periodico nelle serie storiche. Anziché analizzare la variazione da un punto temporale al successivo, analizza infatti la variazione dell'intera serie in componenti periodiche di frequenza diversa. Le serie lisce hanno

componenti periodici più forti a basse frequenze; la variazione casuale ("rumore bianco") diffonde la forza del componente su tutte le frequenze.

Serie che includono i dati mancanti non possono essere analizzati con questa procedura.

Esempio. Il tasso in cui si costruiscono nuove case è un importante barometro dello stato dell'economia. I dati per le abitazioni iniziano tipicamente esibire una forte componente stagionale. Ma ci sono cicli più lunghi presenti nei dati di cui gli analisti devono essere consapevoli quando si valutano le cifre attuali?

Statistiche. Le trasformazioni del seno e della cosina, il valore periodogramma e la stima della densità spettrale per ogni componente di frequenza o periodo. Quando viene selezionata l'analisi bivariata: parti reali e immaginarie di cross - periodogram, densità cospettrale, spettro di quadratura, guadagno, coerenza quadrata e spettro di fase per ogni frequenza o componente di periodo.

Grafici. Per analisi univariate e bivariate: periodogramma e densità spettrale. Per le analisi bivariate: coerenza quadrata, spettro di quadratura, ampiezza incrociata, densità cospettrale, spettro di fase e guadagno.

Considerazioni Dati Grafici Spettrali

Dati. Le variabili dovrebbero essere numeriche.

Ipotesi. Le variabili non dovrebbero contenere alcun dato mancante incorporato. La serie temporale da analizzare dovrebbe essere stazionaria e qualsiasi media non zero dovrebbe essere sottratta alla serie.

- *Stazionario.* Una condizione che deve essere soddisfatta dalle serie temporali a cui si adatta un modello ARIMA. Semplici serie MA saranno comunque stazionarie, mentre serie AR e ARMA potrebbero non esserlo. Una serie stazionaria ha media e varianza costanti nel tempo.

Ottenimento di un'analisi spettrale

1. Dai menu, scegliere:

Analisi > Serie di tempo > Analisi spettrale ...

2. Selezionare una o più variabili dall'elenco disponibile e spostarle nella lista delle variabili. Da notare che l'elenco include solo variabili numeriche.

3. Seleziona una delle opzioni Finestra Finestra per scegliere come regolare il periodogramma per ottenere una stima di densità spettrale. Le opzioni di livellamento disponibili sono Tukey-Hamming, Tukey, Parzen, Bartlett, Daniell (Unità) e None.

- *Tukey-Hamming.* I pesi sono $W_k = .54D_p(2\pi f_k) + .23D_p(2\pi f_k + \pi/p) + .23D_p(2\pi f_k - \pi/p)$, per $k = 0, \dots, p$, dove p è la parte intera di mezzo span e D_p è il kernel di Dirichlet p .
- *Tukey.* I pesi sono $W_k = 0.5D_p(2\pi f_k) + 0.25D_p(2\pi f_k + \pi/p) + 0.25D_p(2\pi f_k - \pi/p)$, per $k = 0, \dots, p$, dove p è la parte intera di mezzo span e D_p è il kernel di Dirichlet p .
- *Parzen.* I pesi sono $W_k = 1/p(2 + \cos(2\pi f_k)) (F[p/2](2\pi f_k))^2$, per $k = 0, \dots, p$, dove p è la parte intera di mezzo span e $F[p/2]$ è il kernel di Fejer dell'ordine $p/2$.
- *Bartlett.* La forma di una finestra spettrale per la quale i pesi della metà superiore della finestra sono calcolati come $W_k = F_p(2\pi f_k)$, per $k = 0, \dots, p$, dove p è la parte intera di mezzo span e F_p è il kernel di Fejer. La metà inferiore è simmetrica con la metà superiore.
- *Daniell (Unità).* Una forma di finestra spettrale in cui i pesi sono tutti pari a 1.
- *Nessuno.* Nessuna lisciatura. Se viene scelta questa opzione, la stima della densità spettrale è la stessa del periodogramma.

Estensione. L'intervallo di valori consecutivi su cui viene eseguito il livellamento. Solitamente si usa un intero positivo dispari. Estensioni maggiori livellano il grafico di densità spettrale più delle estensioni di dimensioni minori.

Centra le variabili. Corregge le serie in modo che abbiano media zero, prima di calcolare lo spettro e di rimuovere il periodo esteso che può essere associato alla media della serie.

Analisi bivariata - prima variabile con ciascuna. Se avete selezionato due o più variabili, è possibile selezionare questa opzione per richiedere analisi spettrali bivariate.

- La prima variabile nell'elenco Variabili è trattata come variabile indipendente e tutte le variabili rimanenti sono trattate come variabili dipendenti.
- Ogni serie dopo il primo viene analizzata con la prima serie indipendentemente da altre serie denominate. Vengono eseguite anche analisi univariate di ogni serie.

Grafico. Periodogramma e densità spettrale sono disponibili sia per le analisi univariate che per le analisi bivariate. Tutte le altre scelte sono disponibili solo per le analisi bivariate.

- *Periodogramma.* Grafico non livellato dell'ampiezza spettrale (tracciato su una scala logaritmica) rispetto alla frequenza o al periodo. Le serie livellate sono caratterizzate da una bassa variazione della frequenza. Una variazione che si espande equamente su tutte le frequenze indica la presenza di "rumore bianco".
- *Coerenza quadrata.* Il prodotto dei guadagni delle due serie.
- *Spettro di quadratura.* La parte immaginaria del periodogramma incrociato, che è una misura della correlazione delle componenti di frequenza fuori fase di due serie temporali. Le componenti sono fuori fase di $\pi/2$ radianti.
- *Ampiezza incrociata.* La radice quadrata della somma del quadrato della densità cospettrale e del quadrato dello spettro di quadratura.
- *Densità spettrale.* Un periodogramma livellato per rimuovere la variazione irregolare.
- *Densità cospettrale.* La parte reale del periodogramma incrociato che è una misura della correlazione delle componenti di frequenza in fase di due serie temporali.
- *Densità spettrale di potenza.* La densità spettrale di potenza (o densità di potenza o spettro di densità di potenza o spettro di potenza) del segnale descrive la potenza presente nel segnale in funzione della frequenza, per unità di frequenza.
- *Coerenza.* La coerenza è una statistica che può essere utilizzata per esaminare la relazione tra due segnali o dataset nell'elaborazione del segnale. È comunemente usato per stimare il trasferimento di potenza tra l'input e l'output di un sistema lineare.
- *Spettro di fase.* Una misura di quanto ogni componente di frequenza di una serie precede o segue l'altro.
- *Guadagno.* Il quoziente di divisione dell'ampiezza della croce per la densità spettrale di una delle serie. Ciascuna delle due serie ha un proprio valore di guadagno.

Per frequenza. Tutti i grafici vengono prodotti per frequenza, partendo da 0 (la costante o il periodo medio) fino a 0,5 (il termine per un ciclo di due osservazioni).

Per periodo. Tutti i grafici vengono prodotti per periodo, che va da 2 (il termine per un ciclo di due osservazioni) a un periodo uguale al numero di osservazioni (la costante o il termine medio). Il periodo è visualizzato su una scala logaritmica.

Funzioni Aggiuntive Comando SPECTRA

Il linguaggio della sintassi dei comandi consente inoltre di:

- Salvare le variabili di analisi spettrale calcolate al dataset attivo per un utilizzo successivo.
- Specificare pesi personalizzati per la finestra spettrale.
- Produrre tracciati sia per frequenza che per periodo.
- Stampa una elencazione completa di ogni valore mostrato nella trama.

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

Modelli causali temporali

La modellazione causale temporale cerca di rilevare le relazioni causali di chiavi nei dati di serie temporali. Nella modellazione causale temporale, l'utente specifica un insieme di serie di destinazione e un insieme di input candidati per tali destinazioni. La procedura poi crea un modello di serie temporali autoregressivo per ogni destinazione e include solo gli input che hanno una relazione causale con la destinazione. Questo approccio differisce dalla modellazione di serie temporali tradizionale in quanto è

necessario specificare esplicitamente i predittori per una serie di destinazione. Poiché la modellazione causale temporale generalmente implica la creazione di modelli di più serie temporali correlate, il risultato è indicato come *sistema modello*.

In un contesto di modellazione causale temporale, il termine *causale* fa riferimento alla causalità di Granger. Una serie temporale X sottopone alla "causalità di Granger" un'altra serie temporale Y se la regressione per Y in termini di valori passati di X e Y comporta un modello migliore per Y rispetto alla sola regressione di valori passati di Y.

Esempi

I responsabili delle decisioni di business possono utilizzare la modellazione causale temporale per scoprire relazioni causali nell'ambito di un elevato insieme di metriche basate sul tempo che descrivono il business. L'analisi potrebbe rivelare alcune immissioni controllabili che hanno il maggiore impatto sugli indicatori di prestazioni chiave (KPI, Key Performance Indicator).

I gestori di sistemi informatici di grandi dimensioni possono utilizzare la modellazione causale temporale per rilevare anomalie in un'ampia serie di metriche operative interconnesse. Il modello causale consente quindi di andare oltre il rilevamento delle anomalie e di scoprire le cause più probabili delle anomalie.

Requisiti dei campi

Deve essere presente almeno un obiettivo. Per impostazione predefinita, i campi con un ruolo predefinito Nessuno non sono utilizzati.

Struttura dei dati

La modellazione causale temporale supporta due tipi di strutture dati.

Dati basati su colonne

Per dati basati su colonne, ciascun campo delle serie temporali contiene i dati per una singola serie temporale. Questa è la struttura tradizionale dei dati delle serie temporali utilizzata da Modeler di serie temporali.

Dati multidimensionali

Per dati multidimensionali, ciascun campo delle serie temporali contiene i dati per più serie temporali. Serie temporali separate, all'interno di un particolare campo, vengono quindi identificate da un serie di valori di campi categoriali indicati come campi *dimensione*. Ad esempio, i dati di vendita per due diversi canali di vendita (al dettaglio e web) possono essere memorizzati in un singolo campo *vendita*. Un campo dimensione denominato *canale*, con valori 'retail' e 'web', identifica i record associati a ciascuno dei due canali di vendita.

Come ottenere un modello causale temporale

Questa funzione richiede l'utilizzo dell'opzione Previsione statistiche.

Dai menu, scegliere:

Analizzare > Previsione > Creare Modelli Causali ...

1. Se le osservazioni sono definite da un campo data/ora, specificare il campo.
2. Se i dati sono multidimensionali, specificare i campi dimensione che identificano le serie temporali.
 - L'ordine specificato dei campi dimensione definisce l'ordine in cui essi vengono visualizzati in tutte le finestre successive o nell'output. Utilizzare i pulsanti freccia verso l'alto e verso il basso per riordinare i campi dimensione.
 - È possibile specificare sottoinsiemi di valori della dimensione che limitano l'analisi a particolari valori dei campi dimensione. Ad esempio, se si dispone di dimensioni per la regione ed il marchio, è possibile limitare l'analisi ad una regione specifica. I sottoinsiemi dimensione si applicano a tutti i campi di metriche utilizzati nell'analisi.

- È anche possibile personalizzare l'analisi specificando dei valori della dimensione in base al ruolo a livello del campo di metriche. Ad esempio, se si dispone di una dimensione per il canale di vendita (con valori 'retail' e 'web') e di metriche per la vendita e la pubblicità in tali canali, è possibile specificare la pubblicità web come input per le vendite retail e web. Per impostazione predefinita, questo tipo di personalizzazione è abilitato e limitato alla selezione da un numero specificato di valori distinti (250, per impostazione predefinita) di ciascun campo della dimensione.

3. Fare clic su **Continua**.

Nota: I passi 1, 2 e 3 non vengono applicati se il dataset attivo ha una specifica della data. Le specifiche della data vengono create dalla finestra di dialogo Definisci Date o dal comando DATE .

4. Fare clic su **Campi** per specificare le serie temporali da includere nel modello e specificare il modo in cui sono definite le osservazioni. È necessario specificare almeno un campo come obiettivo o come obiettivo ed input.
5. Fare clic su **Specifiche dati** per specificare impostazioni facoltative che includono l'intervallo di tempo per l'analisi, le impostazioni di aggregazione e distribuzione e la gestione dei valori mancanti.
6. Fare clic su **Opzioni di creazione** per definire il periodo di stima, specificare il contenuto dell'output e specificare le impostazioni di creazione, come il numero massimo di input per obiettivo.
7. Fare clic su **Opzioni modello** per richiedere previsioni, salvare previsioni ed esportare il sistema di modello in un file esterno.
8. Fare clic su **Esegui** per eseguire la procedura.

Serie temporali al modello

Nella scheda Campi, utilizzare le impostazioni **Serie temporali** per specificare le serie da includere nel sistema di modelli.

Per dati basati su colonne, il termine *serie* ha lo stesso significato del termine *campo*. Per dati multidimensionali, i campi che contengono serie temporali vengono indicati come campi di *metriche*. Una serie temporale, per dati multidimensionali, è definita da un campo di metriche e un valore per ognuno dei campi dimensione. Le seguenti considerazioni si applicano ad entrambi i dati basati su colonne e multidimensionali.

- Le serie specificate come input candidato o come obiettivo e input vengono considerate per l'inclusione nel modello per ciascuna destinazione. Il modello per ciascuna destinazione include sempre valori ritardati della destinazione stessa.
- Le serie specificate come input forzati vengono sempre incluse nel modello per ciascuna destinazione.
- È necessario specificare almeno una serie come obiettivo o come obiettivo e input.
- Quando è selezionata l'opzione **Utilizza ruoli predefiniti**, i campi con il ruolo Input vengono impostati come input candidati. Nessun ruolo predefinito viene associato ad un input forzato.

Dati multidimensionali

Per dati multidimensionali, i campi di metriche ed i ruoli associati vengono specificati in una griglia, in cui ogni riga nella griglia specifica uno singolo ruolo e metrica. Per impostazione predefinita, il sistema di modelli include le serie per tutte le combinazioni dei campi dimensione di ciascuna riga nella griglia. Ad esempio, se vi sono dimensioni per *region* e *brand*, per impostazione predefinita, specificando la metrica *sales* come obiettivo, si indica che esiste una serie obiettivo di vendita separata per ciascuna combinazione di *region* e *brand*.

Per ciascuna riga nella griglia, è possibile personalizzare l'insieme di valori per qualsiasi campo dimensione facendo clic sul pulsante con i puntini di sospensione per una dimensione. Questa azione apre la finestra Seleziona valori dimensione. È anche possibile aggiungere, eliminare o copiare le righe della griglia.

La colonna **Numero di serie** visualizza il numero delle serie di valori dimensione che sono attualmente specificate per la metrica associata. Il valore visualizzato può essere più grande del numero effettivo di

serie (una serie per insieme). Questa condizione si verifica quando alcune delle combinazioni specificate di valori dimensione non corrispondono alle serie contenute dalla metrica associata.

Seleziona valori dimensione

Per dati multidimensionali, è possibile personalizzare l'analisi specificando quali valori dimensione si applicano a un determinato campo di metriche con un ruolo specifico. Ad esempio, se *sales* è un campo di metriche e *channel* è una dimensione con i valori 'retail' e 'web', è possibile specificare che la vendita 'web' sia un input e la vendita 'retail' sia un obiettivo. È anche possibile specificare dei sottoinsiemi dimensione che si applicano a tutti i campi di metriche utilizzati nell'analisi. Ad esempio, se *region* è un campo dimensione che indica una regione geografica, è possibile limitare l'analisi a particolari regioni.

Tutti i valori

Specifica che sono inclusi tutti i valori del campo dimensione corrente. Questa opzione è il valore predefinito.

Seleziona valori da includere o escludere

Utilizzare questa opzione per specificare la serie di valori per il campo dimensione corrente. Quando si seleziona **Includi** per la **Modalità**, solo i valori specificati nell'elenco **Valori selezionati** vengono inclusi. Quando si seleziona **Escludi** per la **Modalità**, tutti i valori diversi dai valori specificati nell'elenco **Valori selezionati** vengono inclusi.

È possibile filtrare l'insieme di valori da cui effettuare la scelta. I valori che soddisfano la condizione filtro vengono visualizzati nella scheda **Corrispondenti**, mentre i valori che non soddisfano la condizione filtro vengono visualizzati nella scheda **Non corrispondenti** dell'elenco **Valori non selezionati**. La scheda **Tutti** elenca tutti i valori non selezionati, indipendentemente da qualsiasi condizione filtro.

- È possibile utilizzare gli asterischi (*) per indicare i caratteri jolly quando si specifica un filtro.
- Per deselezionare il filtro corrente, specificare un valore vuoto per il termine di ricerca nella finestra Filtra valori visualizzati.

Osservazioni

Nella scheda Campi, utilizzare le impostazioni **Osservazioni** per specificare i campi che definiscono le osservazioni.

Nota: Se il dataset attivo dispone di una specifica di data, le osservazioni vengono definite dalla specifica di data e non possono essere modificate nella procedura di modellazione causale temporale. Le specifiche della data vengono create dalla finestra di dialogo Definisci Date o dal comando DATE .

Osservazioni che vengono definite dalla data/ora

È possibile specificare che le osservazioni siano definite da un campo con un formato data, ora o data/ora o da un campo stringa che rappresenta una data/ora. I campi stringa possono rappresentare una data in formato AAAA-MM-GG, un'ora in formato HH:MM:SS, o una data/ora in formato AAAA-MM-GG HH:MM:SS. Gli zeri iniziali possono essere omessi nella rappresentazione stringa. Ad esempio, la stringa 2014-9-01 è equivalente a 2014-09-01.

Oltre al campo che definisce le osservazioni, selezionare l'intervallo di tempo appropriato che descrive le osservazioni. A seconda dell'intervallo di tempo specificato, è possibile anche specificare altre impostazioni, quali l'intervallo tra le osservazioni (incremento) o il numero di giorni per settimana. Le seguenti considerazioni si applicano all'intervallo di tempo:

- Utilizzare il valore **Irregolare** quando le osservazioni non sono regolarmente distanziate nel tempo, come ad esempio l'ora in cui un ordine di vendita viene elaborato. Quando viene selezionato il valore **Irregolare**, è necessario specificare l'intervallo di tempo utilizzato per l'analisi dalle impostazioni **Intervallo di tempo** nella scheda Specifiche dati.
- Quando le osservazioni rappresentano una data e ora, e l'intervallo di tempo è espresso in ore, minuti o secondi, utilizzare le opzioni **Ore al giorno**, **Minuti al giorno** o **Secondi al giorno**. Quando le osservazioni rappresentano un tempo (durata) senza riferimento ad una data, mentre l'intervallo

di tempo è espresso in ore, minuti o secondi, utilizzare le opzioni **Ore (non periodiche)**, **Minuti (non periodici)** o **Secondi (non periodici)**.

- In base all'intervallo di tempo selezionato, la procedura può rilevare le osservazioni mancanti. Il rilevamento delle osservazioni mancanti è necessario poiché la procedura presuppone che tutte le osservazioni siano equamente distanziate nel tempo e che non vi siano osservazioni mancanti. Ad esempio, se l'intervallo di tempo è Giorni e la data 2014-10-27 è seguita da 2014-10-29, vi è un'osservazione mancante per il 2014-10-28. I valori sono assegnati per qualsiasi osservazione mancante. Le impostazioni per la gestione di valori mancanti possono essere specificate dalla scheda Specifiche dei dati.
- L'intervallo di tempo specificato fa sì che la procedura rilevi più osservazioni nello stesso intervallo di tempo da sottoporre ad aggregazione e di allinearle rispetto a un limite di intervallo, come il primo giorno del mese, al fine di assicurare che le osservazioni siano equamente distanziate. Ad esempio, se l'intervallo di tempo è Mesi, più date nello stesso mese vengono aggregate insieme. Questo tipo di aggregazione viene indicato come *raggruppamento*. Per impostazione predefinita, le osservazioni vengono sommate quando raggruppate. È possibile specificare un metodo diverso per il raggruppamento, ad esempio la media delle osservazioni, dalle impostazioni **Aggregazione e distribuzione** nella scheda Specifiche dei dati.
- Per alcuni intervalli di tempo, le impostazioni aggiuntive possono definire le interruzioni nei normali intervalli equamente distanziati. Ad esempio, se l'intervallo di tempo è Giorni, ma sono validi solo i giorni feriali, è possibile specificare che vi sono cinque giorni in una settimana, e che la settimana inizia il lunedì.

Osservazioni che vengono definite da periodi o periodi ciclici

Le osservazioni possono essere definite da uno o più campi di tipo intero che rappresentano periodi o cicli ripetitivi di periodi, fino ad un numero arbitrario di livelli di ciclo. Con questa struttura, è possibile descrivere le serie di osservazioni che non si adattano ad uno degli intervalli di tempo standard. Ad esempio, un anno fiscale con solo 10 mesi può essere descritto con un campo ciclo che rappresenta gli anni e un campo periodo che rappresenta i mesi, in cui la lunghezza di un ciclo è pari a 10.

I campi che specificano i periodi ciclici definiscono una gerarchia di livelli periodici, in cui il livello più basso è definito dal campo **Periodo**. Il successivo livello più alto è specificato da un campo ciclo il cui livello è 1, seguito da un campo ciclo il cui livello è 2 e così via. I valori campo per ciascun livello, ad eccezione del più alto, devono essere periodici rispetto al livello più alto successivo. I valori per il livello più alto non possono essere periodici. Ad esempio, nel caso di anno fiscale di 10 mesi, i mesi sono periodici all'interno degli anni e gli anni non sono periodici.

- La lunghezza di un ciclo a un particolare livello è la periodicità del successivo livello più basso. Per l'esempio dell'anno fiscale, vi è un solo livello di ciclo e la relativa lunghezza è pari a 10 poiché il successivo livello più basso rappresenta i mesi e vi sono 10 mesi nell'anno fiscale specificato.
- Specificare il valore di partenza per qualsiasi campo periodico che non inizi da 1. Questa impostazione è necessaria per rilevare i valori mancanti. Ad esempio, se un campo periodico inizia da 2 ma il valore iniziale viene specificato come 1, la procedura presuppone che vi sia un valore mancante per il primo periodo in ogni ciclo di tale campo.

Osservazioni che vengono definite dall'ordine di record

Per dati basati su colonne, è possibile specificare che le osservazioni siano definite dall'ordine di record in modo che il primo record rappresenti la prima osservazione, il secondo record rappresenti la seconda osservazione, e così via. Quindi, si presuppone che i record rappresentino le osservazioni equamente distanziate nel tempo.

Intervallo di tempo per l'analisi

L'intervallo di tempo utilizzato per l'analisi può differire dall'intervallo di tempo utilizzato per le osservazioni. Ad esempio, se l'intervallo di tempo delle osservazioni è Giorni, è possibile scegliere Mesi per l'intervallo di tempo per l'analisi. I dati vengono quindi aggregati dai dati quotidiani a quelli mensili prima che il modello venga creato. È possibile inoltre scegliere di distribuire i dati da un intervallo di tempo più lungo ad uno più breve. Ad esempio, se le osservazioni sono trimestrali, è possibile distribuire i dati da un intervallo trimestrale ad uno mensile.

Le scelte disponibili per l'intervallo di tempo in cui l'analisi viene eseguita dipende dal modo in cui le osservazioni vengono definite e dall'intervallo di tempo di tali osservazioni. In particolare, quando le osservazioni vengono definite dai periodi ciclici o quando una specifica di data viene definita per il dataset attivo, solo l'aggregazione è supportata. In quel caso, l'intervallo di tempo dell'analisi deve essere maggiore o uguale all'intervallo di tempo delle osservazioni.

L'intervallo di tempo per l'analisi è specificato dalle impostazioni **Intervallo di tempo** nella scheda Specifiche dati. Il metodo con cui i dati vengono aggregati o distribuiti è specificato dalle impostazioni **Aggregazione e distribuzione** nella scheda Specifiche dei dati.

Aggregazione e distribuzione

Funzioni di aggregazione

Quando l'intervallo di tempo che viene utilizzato per l'analisi è superiore all'intervallo di tempo delle osservazioni, i dati di input vengono aggregati. Ad esempio, l'aggregazione viene eseguita quando l'intervallo di tempo delle osservazioni è Giorni e l'intervallo di tempo per l'analisi è Mesi. Sono disponibili le seguenti funzioni di aggregazione: valore medio, somma, moda, minimo o massimo.

Funzioni di distribuzione

Quando l'intervallo di tempo che viene utilizzato per l'analisi è inferiore all'intervallo di tempo delle osservazioni, i dati di input vengono distribuiti. Ad esempio, la distribuzione viene eseguita quando l'intervallo di tempo delle osservazioni è Trimestri e l'intervallo di tempo per l'analisi è Mesi. Sono disponibili le seguenti funzioni di distribuzione: mean o sum.

Funzioni di raggruppamento

Il raggruppamento viene applicato quando le osservazioni sono definite dalle date/ore e quando si verificano più osservazioni nello stesso intervallo di tempo. Ad esempio, se l'intervallo di tempo delle osservazioni è Mesi, più date nello stesso mese vengono raggruppate e associate al mese in cui si verificano. Sono disponibili le seguenti funzioni di raggruppamento: mean, sum, mode, min o max. Il raggruppamento viene sempre eseguito quando le osservazioni vengono definite dalla data/ora e l'intervallo di tempo delle osservazioni è specificato come Irregolare.

Nota: Sebbene il raggruppamento sia una forma di aggregazione, esso viene eseguito prima di qualsiasi gestione di valori mancanti, mentre l'aggregazione formale viene eseguita dopo la gestione di eventuali valori mancanti. Quando l'intervallo di tempo delle osservazioni è specificato come irregolare, l'aggregazione viene eseguita solo con la funzione di raggruppamento.

Aggrega le osservazioni di giorno al giorno precedente

Specifica se le osservazioni a cavallo tra due giorni vengono aggregate ai valori del giorno precedente. Ad esempio, per le osservazioni orarie con un giorno di otto ore che inizia alle 20.00, questa impostazione specifica se le osservazioni comprese tra le 00:00 e le 04:00 sono incluse nei risultati aggregati del giorno precedente. Questa impostazione si applica solo se l'intervallo di tempo delle osservazioni è Ore al giorno, Minuti al giorno o Secondi al giorno e l'intervallo di tempo per l'analisi è Giorni.

Impostazioni personalizzate per campi specificati

È possibile specificare le funzioni di aggregazione, distribuzione e raggruppamento su un campo in base al campo. Tali impostazioni sostituiscono le impostazioni predefinite per le funzioni di aggregazione, distribuzione e raggruppamento.

Valori mancanti

I valori mancanti nei dati di input vengono sostituiti con un valore assegnato. Sono disponibili i seguenti metodi di sostituzione:

Interpolazione lineare

Sostituisce i valori mancanti utilizzando un'interpolazione lineare. Per l'interpolazione verranno utilizzati l'ultimo valore valido prima del valore mancante e il primo valore valido dopo il valore mancante. Se la prima o ultima osservazione nella serie ha un valore mancante, vengono utilizzati i due valori non mancanti più prossimi all'inizio o alla fine della serie.

Media della serie

Consente di sostituire i valori mancanti con la media dell'intera serie.

Media di punti vicini

Consente di sostituire i valori mancanti con la media dei valori validi circostanti. L'intervallo di punti vicini è il numero di valori validi prima e dopo il valore mancante utilizzati per calcolare la media.

Mediana di punti vicini

Consente di sostituire i valori mancanti con la mediana dei valori validi circostanti. L'intervallo di punti vicini è il numero di valori validi prima e dopo il valore mancante utilizzati per calcolare la mediana.

Tendenza lineare

Questa opzione utilizza tutte le osservazioni non mancanti nella serie di adattamento di un modello di regressione lineare semplice, che viene quindi utilizzato per assegnare i valori mancanti.

Altre impostazioni:

Massima percentuale di valori mancanti (%)

Specifica la percentuale massima di valori mancanti consentita per ogni serie. Le serie con più valori mancanti del valore massimo specificato sono escluse dall'analisi.

Valori mancanti definiti dall'utente

Questa opzione specifica se i valori mancanti definiti dall'utente vengono trattati come dati validi e pertanto inclusi nella serie. Per impostazione predefinita, i valori mancanti definiti dall'utente vengono esclusi e trattati come valori mancanti di sistema, che vengono poi assegnati.

Opzioni di dati generali

Numero massimo di valori distinti per campo dimensione

Questa impostazione si applica ai dati multidimensionali e specifica il numero massimo di valori distinti consentiti per qualsiasi campo dimensione. Per impostazione predefinita, questo limite è impostato su 10000 ma può essere aumentato a un numero arbitrariamente più grande.

Opzioni di creazione generali

Larghezza intervallo di confidenza (%)

Questa impostazione controlla gli intervalli di confidenza sia per le previsioni che per i parametri del modello. È possibile specificare qualsiasi valore positivo inferiore a 100. Per impostazione predefinita, viene utilizzato un intervallo di confidenza del 95%.

Numero massimo di input per ciascuna destinazione

Questa impostazione specifica il numero massimo di input consentiti nel modello per ciascuna destinazione. È possibile specificare un numero intero compreso nell'intervallo da 1 a 20. Il modello di ciascuna destinazione include sempre valori ritardati di se stesso, pertanto l'impostazione di questo valore su 1 specifica che l'unico input è la destinazione stessa.

Tolleranza del modello

Questa impostazione controlla il processo iterativo che viene utilizzato per determinare la migliore serie di input per ciascuna destinazione. È possibile specificare qualsiasi valore maggiore di zero. Il valore predefinito è 0.001.

Soglia valore anomalo (%)

Un'osservazione è contrassegnata come valore anomalo se la probabilità, così come calcolata dal modello, che è un valore anomalo supera questa soglia. È possibile specificare un valore nell'intervallo compreso tra 50 e 100.

Numero di ritardi per ciascun input

Questa impostazione specifica il numero di ritardi per ciascun input nel modello per ciascuna destinazione. Per impostazione predefinita, il numero di ritardi viene determinato automaticamente dall'intervallo di tempo utilizzato per l'analisi. Ad esempio, se l'intervallo di tempo è mesi (con un incremento di un mese), il numero di ritardi è 12. Come opzione, è possibile specificare in modo esplicito il numero di ritardi. Il valore specificato deve essere un numero intero compreso tra 1 e 20.

Serie da visualizzare

Queste opzioni specificano le serie (destinazioni o input) per le quali viene visualizzato l'output. Il contenuto dell'output per le serie specificate è determinato dalle impostazioni **Opzioni di output**.

Visualizza destinazioni associate con modelli più adatti

Per impostazione predefinita, l'output viene visualizzato per le destinazioni associate ai 10 modelli più adatti, come determinato dal valore R quadrato. È possibile specificare un numero fisso differente di modelli più adatti oppure è possibile specificare una percentuale di modelli più adatti. È anche possibile scegliere tra le seguenti misure della bontà di adattamento:

R-quadrato

Misura della bontà dell'adattamento di un modello lineare, detto anche coefficiente di determinazione. È la proporzione di variabilità della variabile obiettivo spiegata dal modello. Varia in valore dal 0 al 1. I piccoli valori indicano che il modello non si adatta bene ai dati.

Errore percentuale valore quadratico medio

Una misura di quanto i valori previsti dal modello differiscono dai valori osservati della serie. È indipendente dalle unità utilizzate e può pertanto essere utilizzata per confrontare delle serie con unità differenti.

Errore valore quadratico medio

La radice quadrata dell'errore quadratico medio. Indica di quanto una serie dipendente varia rispetto al livello stimato del modello. Questa misura viene espressa nelle stesse unità della serie dipendente.

BIC

(BIC) normalizzato. Una misura per selezionare e confrontare modelli basata sulla verosimiglianza log ridotta -2. I valori più bassi indicano i modelli migliori. Anche BIC "penalizza" i modelli sovraparametrizzati (modelli complessi con un elevato numero di input, ad esempio) ma in modo più rigoroso rispetto ad AIC.

AIC

Criterio di informazione di Akaike. Una misura per selezionare e confrontare modelli basata sulla verosimiglianza log ridotta -2. I valori più bassi indicano i modelli migliori. AIC "penalizza" i modelli sovraparametrizzati (modelli complessi con un elevato numero di input, ad esempio).

Specifica serie individuali

È possibile specificare singole serie per le quali si desidera un output.

- Per dati basati su colonne, si specificano i campi che contengono le serie desiderate. L'ordine dei campi specificati definisce l'ordine in cui vengono visualizzati nell'output.
- Per dati multidimensionali, si specifica una particolare serie aggiungendo una voce alla griglia per il campo metrico che contiene la serie. Quindi, è possibile specificare i valori dei campi dimensione che definiscono le serie.
 - È possibile immettere il valore per ciascun campo dimensione direttamente nella griglia oppure effettuare una scelta dall'elenco di valori della dimensione disponibili. Per effettuare una selezione dall'elenco di valori dimensione disponibili, fare clic sul pulsante con i puntini di sospensione nella cella per la dimensione desiderata. Questa azione apre la finestra secondaria Seleziona valore dimensione.
 - È possibile ricercare l'elenco di valori della dimensione, nella finestra di dialogo secondaria Seleziona valore dimensione, facendo clic sull'icona a forma di binocolo e specificando un termine di ricerca. Gli spazi vengono considerati come parte del termine di ricerca. Gli asterischi (*) nel termine di ricerca non indicano caratteri jolly.
 - L'ordine delle serie nella griglia definisce l'ordine in cui vengono visualizzate nell'output.

Per entrambi i dati basati su colonne ed i dati multidimensionali, l'output è limitato a 30 serie. Questo limite include le serie individuali (input o obiettivi) specificate e gli obiettivi associati ai modelli più adatti. Le serie individuali specificate hanno la precedenza sugli obiettivi associati ai modelli più adatti.

Opzioni di output

Queste opzioni specificano il contenuto dell'output. Le opzioni nel gruppo **Output per destinazioni** generano l'output per le destinazioni associate ai modelli più adatti nelle impostazioni **Serie da**

visualizzare. Le opzioni nel gruppo **Output per serie** generano l'output per le serie individuali specificate nelle impostazioni **Serie da visualizzare.**

Sistema del modello globale

Visualizza una rappresentazione grafica delle relazioni causali tra le serie nel sistema di modelli. Le tabelle sia di valori anomali che di statistiche di adattamento di modelli per le destinazioni visualizzate sono incluse come parte dell'elemento di output. Quando questa opzione viene selezionata nel gruppo **Output per serie**, viene creato un elemento di output separato per ciascuna serie individuale specificata nelle impostazioni **Serie da visualizzare.**

Le relazioni causali tra le serie hanno un livello di significatività associato, in cui un livello di significatività inferiore indica una connessione più significativa. È possibile scegliere di nascondere le relazioni con un livello di significatività superiore al valore specificato.

Valori anomali e statistiche di adattamento modello

Tabelle di valori anomali e statistiche di adattamento modello per le serie di destinazione selezionate per la visualizzazione. Queste tabelle contengono le stesse informazioni delle tabelle nella visualizzazione Sistema complessivo di modelli. Queste tabelle supportano tutte le funzioni standard per la riorganizzazione e la modifica delle tabelle.

Effetti e parametri di modelli

Le tabelle di parametri e di test degli effetti di modelli per le serie di destinazione selezionate per la visualizzazione. I test degli effetti di modelli includono le statistiche F e il valore di significatività associato per ciascun input incluso nel modello.

Diagramma di impatto

Visualizza una rappresentazione grafica delle relazioni casuali tra una serie di interesse e altre serie che possono avere effetto l'una sulle altre. Le serie che influiscono sulla serie di interesse sono indicate *cause*. Selezionando **Effetti** viene generato un diagramma di impatto inizializzato per visualizzare gli effetti. Selezionando **Cause** viene generato un diagramma di impatto inizializzato per visualizzare le cause. Selezionando **Entrambe le cause e gli effetti** vengono generati due diagrammi di impatto separati, uno inizializzato per le cause ed uno inizializzato per gli effetti. È possibile passare in modo interattivo dalle cause agli effetti e viceversa nell'elemento di output che visualizza il diagramma di impatto.

È possibile specificare il numero di livelli delle cause o degli effetti da visualizzare, dove il primo livello è solo la serie di interesse. Ogni ulteriore livello mostra ulteriori cause o effetti indiretti della serie di interesse. Ad esempio, il terzo livello nella visualizzazione degli effetti è composto dalle serie che contengono serie nel secondo livello come input diretto. Le serie nel terzo livello sono quindi indirettamente interessate dalla serie di interesse, poiché la serie di interesse è un input diretto per le serie nel secondo livello.

Grafico serie

Grafici dei valori osservati e previsti per le serie di destinazione selezionate per la visualizzazione. Quando le previsioni sono richieste, il grafico mostra anche i valori previsti e gli intervalli di confidenza per le previsioni.

Grafico residui

Grafici dei residui di modelli per le serie di destinazione selezionate per la visualizzazione.

Input migliori

Grafici di ogni destinazione visualizzata, nel tempo, insieme ai primi 3 input per la destinazione. Gli input principali sono gli input con il valore di significatività più basso. Per accogliere diverse scale di input e di destinazione, l'asse y rappresenta il punteggio z per ciascuna serie.

Tabella delle previsioni

Tabelle di valori previsti e di intervalli di confidenza di quelle previsioni per le serie di destinazione selezionate per la visualizzazione.

Analisi causa principale valore anomalo

Determina quali serie sono più probabilmente la causa di ciascun valore anomalo in una serie di interesse. L'analisi della causa principale del valore anomalo viene eseguita per ciascuna serie di destinazione inclusa nell'elenco di serie individuali delle impostazioni **Serie da visualizzare.**

Output

Tabella e grafico valori anomali interattivi

Tabella e grafico di valori anomali e di cause principali di tali valori anomali per ciascuna serie di interesse. La tabella contiene una singola riga per ciascun valore anomalo. Il grafico è un diagramma di impatto. La selezione di una riga nella tabella evidenzia il percorso, nel diagramma di impatto, dalla serie di interesse alla serie che più probabilmente è la causa del valore anomalo associato.

Tabella pivot di valori anomali

Tabella di valori anomali e di cause principali di tali valori anomali per ciascuna serie di interesse. Questa tabella contiene le stesse informazioni della tabella nella visualizzazione interattiva. Questa tabella supporta tutte le funzioni standard per la riorganizzazione e la modifica delle tabelle.

Livelli causali

È possibile specificare il numero di livelli da includere nella ricerca per le cause principali. Il concetto dei livelli qui utilizzato è lo stesso descritto per i diagrammi di impatto.

Adattamento modello tra tutti i modelli

Istogramma dell'adattamento modello per tutti i modelli e per le statistiche di adattamento selezionate. Sono disponibili le seguenti statistiche di adattamento:

R-quadrato

Misura della bontà dell'adattamento di un modello lineare, detto anche coefficiente di determinazione. È la proporzione di variabilità della variabile obiettivo spiegata dal modello. Varia in valore dal 0 al 1. I piccoli valori indicano che il modello non si adatta bene ai dati.

Errore percentuale valore quadratico medio

Una misura di quanto i valori previsti dal modello differiscono dai valori osservati della serie. È indipendente dalle unità utilizzate e può pertanto essere utilizzata per confrontare delle serie con unità differenti.

Errore valore quadratico medio

La radice quadrata dell'errore quadratico medio. Indica di quanto una serie dipendente varia rispetto al livello stimato del modello. Questa misura viene espressa nelle stesse unità della serie dipendente.

BIC

(BIC) normalizzato. Una misura per selezionare e confrontare modelli basata sulla verosimiglianza log ridotta -2. I valori più bassi indicano i modelli migliori. Anche BIC "penalizza" i modelli sovraparametrizzati (modelli complessi con un elevato numero di input, ad esempio) ma in modo più rigoroso rispetto ad AIC.

AIC

Criterio di informazione di Akaike. Una misura per selezionare e confrontare modelli basata sulla verosimiglianza log ridotta -2. I valori più bassi indicano i modelli migliori. AIC "penalizza" i modelli sovraparametrizzati (modelli complessi con un elevato numero di input, ad esempio).

Valori anomali nel tempo

Grafico a barre del numero di valori anomali, su tutte le destinazioni, per ciascun intervallo di tempo nel periodo di stima.

Trasformazioni delle serie

Tabella delle eventuali trasformazioni che sono state applicate alla serie nel sistema di modelli. Le possibili trasformazioni sono distribuzione, aggregazione e assegnazione di valori mancanti.

Periodo di stima

Per impostazione predefinita, il periodo di stima inizia nell'ora della prima osservazione e termina nell'ora dell'ultima osservazione in tutte le serie.

In base alle ore di inizio e di fine

È possibile specificare sia l'inizio che la fine del periodo di stima oppure è possibile specificare solo l'inizio o solo la fine. Se si omette l'inizio o la fine del periodo di stima, viene utilizzato il valore predefinito.

- Se le osservazioni sono definite da un campo data/ora, immettere i valori per l'inizio e la fine nello stesso formato utilizzato per il campo data/ora.
- Per le osservazioni definite in base a periodi ciclici, specificare un valore per ognuno dei campi di periodi ciclici. Ogni campo viene visualizzato in una colonna separata.
- Se esiste una specifica data in vigore per il dataset attivo, è necessario specificare un valore per ogni componente (ad esempio Mese) della specifica data. Ogni componente viene visualizzato in una colonna separata.
- Quando le osservazioni sono definite dall'ordine dei record, l'inizio e la fine del periodo di stima vengono definiti dal numero di riga (come visualizzato nell'Editor dei dati) del caso pertinente.

In base agli intervalli di tempo più recenti o meno recenti

Definisce il periodo di stima come un numero specificato di intervalli di tempo che iniziano nell'intervallo di tempo meno recente e terminano nell'intervallo di tempo più recente nei dati con un offset facoltativo. In questo contesto, l'intervallo temporale indica l'intervallo di tempo dell'analisi. Ad esempio, si supponga che le osservazioni siano mensili ma che l'intervallo di tempo dell'analisi sia trimestrale. Specificando **Più recente** e il valore 24 per **Numero di intervalli di tempo** si indicano i 24 dati trimestrali.

Facoltativamente, è possibile escludere un numero specificato di intervalli di tempo. Ad esempio, specificando gli ultimi 24 intervalli di tempo e 1 per il numero da escludere, si indica che il periodo di stima è costituito da 24 intervalli che precedono l'ultimo.

Previsione

L'opzione **Estensione record nel futuro** imposta il numero di intervalli di tempo per cui si effettua una previsione oltre la fine del periodo di stima. In questo caso, l'intervallo di tempo è l'intervallo dell'analisi, che viene specificato nella scheda Specifiche dati. Quando le previsioni sono richieste, i modelli autoregressivi vengono automaticamente creati per tutte le serie di input che non sono anche obiettivi. Tali modelli vengono quindi utilizzati per generare i valori per queste serie di input nel periodo di previsione.

Salva

Opzioni di destinazione

È possibile salvare sia le trasformazioni dei dati (come l'aggregazione o l'imputazione dei valori mancanti) e le nuove variabili (specificate sulle impostazioni di **Save Targets**) ad un file di dati IBM SPSS Statistiche o un nuovo dataset nella sessione corrente. I valori di data / ora nei dati salvati sono allineati all'inizio di ogni intervallo di tempo, come il primo del mese, e rappresentano l'intervallo di tempo dell'analisi per il sistema modello. È possibile salvare eventuali nuove variabili nel dataset attivo solo se le osservazioni sono definite da una specifica data, o da un ordine record e i dati non sono aggregati.

Salva destinazioni

È possibile salvare le previsioni del modello, gli intervalli di confidenza e i residui come nuove variabili. Ogni obiettivo specificato per salvare genera la propria serie di nuove variabili e ogni nuova variabile contiene valori sia per periodi di stima che di previsione. Per i valori previsti, gli intervalli di confidenza e i residui di rumore è possibile specificare il nome radice da utilizzare come prefisso per le nuove variabili. Il nome variabile completo è la concatenazione del nome radice e il nome del campo che contiene la serie target. Il nome radice deve essere conforme alle regole per i nomi variabili validi. Il nome variabile viene esteso se necessario per evitare conflitti di denominazione variabili.

Indica casi contenenti previsioni

Crea una variabile che indica se un record contiene dati di previsione. È possibile specificare il nome variabile. Il valore predefinito è *ForecastIndicator*.

Obiettivi da salvare

Specifica se vengono create nuove variabili per tutte le serie di destinazione nel sistema modello o solo la serie di destinazione che vengono specificate sulle impostazioni **Series to Display**.

Sistema del modello di esportazione

Salva il sistema modello in un archivio file compresso (file .zip). Il file system del modello può essere utilizzato dalla procedura Temporal Causal Model Forecasting per ottenere previsioni aggiornate o per generare qualsiasi output disponibile. Può essere utilizzato anche dalla procedura Temporale Causal Model Scenarios per eseguire l'analisi di scenario.

Output interattivo

L'output della modellazione causale temporale include un numero di oggetti di output interattivi. Le funzioni interattive sono disponibili attivando (facendo doppio clic) l'oggetto nel Visualizzatore output.

Sistema del modello globale

Visualizza le relazioni causali tra le serie nel sistema di modelli. Tutte le linee che collegano un particolare obiettivo ai propri input hanno lo stesso colore. Lo spessore della linea indica il significato della connessione causale, dove le linee più spesse rappresentano una connessione più significativa. Gli input che non sono anche obiettivi sono indicati con un quadrato nero.

- È possibile visualizzare le relazioni dei modelli più importanti, di una serie specificata, di tutte le serie, o dei modelli senza input. I modelli più importanti sono quelli che soddisfano i criteri specificati per i modelli più adatti nelle impostazioni **Serie da visualizzare**.
- È possibile generare i diagrammi di impatto per una o più serie selezionando i nomi delle serie nel grafico, facendo clic con il tasto destro del mouse e selezionando **Crea diagramma di impatto** dal menu di contesto.
- È possibile scegliere di nascondere le relazioni causali con un livello di significatività superiore al valore specificato. Livelli più bassi di significatività indicano una relazione causale più significativa.
- È possibile visualizzare le relazioni per una particolare serie selezionando il nome della serie nel grafico, facendo clic con il tasto destro del mouse e selezionando **Evidenzia relazioni per la serie** dal menu di contesto.

Diagramma di impatto

Visualizza una rappresentazione grafica delle relazioni casuali tra una serie di interesse e altre serie che possono avere effetto l'una sulle altre. Le serie che influiscono sulla serie di interesse sono indicate *cause*.

- È possibile modificare la serie di interessi specificando il nome della serie che si desidera. Facendo doppio clic su qualsiasi nodo nel diagramma di impatto si modificherà la serie di interessi con la serie associata a tale nodo.
- È possibile alternare la visualizzazione tra le cause e gli effetti ed è possibile modificare il numero di livelli di cause o effetti da visualizzare.
- Facendo clic una sola volta su qualsiasi nodo si aprirà un diagramma di sequenza dettagliato della serie associata al nodo.

Analisi causa principale valore anomalo

Determina quali serie sono più probabilmente la causa di ciascun valore anomalo in una serie di interesse.

- È possibile visualizzare la causa principale per qualsiasi valore anomalo selezionando la riga ad esso relativa nella tabella Valori anomali. È anche possibile visualizzare la causa principale facendo clic sull'icona del valore anomalo in un diagramma di sequenza.
- Facendo clic una sola volta su qualsiasi nodo si aprirà un diagramma di sequenza dettagliato della serie associata al nodo.

Qualità generale del modello

L'istogramma di adattamento del modello di tutti i modelli, di una determinata statistica di adattamento. Facendo clic su una barra nel grafico a barre si filtrerà il grafico a punti in modo da

visualizzare solo i modelli associati con la barra selezionata. È possibile trovare il modello di una determinata serie di obiettivi nel grafico a punti specificando il nome della serie.

Distribuzione dei valori anomali

Grafico a barre del numero di valori anomali, su tutte le destinazioni, per ciascun intervallo di tempo nel periodo di stima. Facendo clic su una barra nel grafico a barre si filtrerà il grafico a punti in modo da visualizzare solo i valori anomali associati con la barra selezionata.

Nota:

- Se si salva un documento di output che contiene output interattivo dalla modellazione causale temporale e si desidera conservare le funzioni interattive, verificare che l'opzione **Memorizza le informazioni dei modelli richiesti assieme al documento di output** sia selezionata nella finestra Salva output con nome.
- Alcune funzioni interattive richiedono che il dataset attivo sia quello utilizzato per creare il sistema di modelli causale temporale.

Applicazione Modelli Causali Temporal

Applicazione Modelli Causali Temporal

Sono disponibili due procedure per l'applicazione di modelli che sono stati creati con la procedura Temporale Causal Modeling. Entrambe le procedure richiedono il file system del modello, che può essere salvato come parte della procedura di Modelli di Causale Temporale.

Previsione modello causale temporale

Si può utilizzare questa procedura per ottenere previsioni per serie per le quali sono disponibili dati più attuali, senza ricostruire i propri modelli. È inoltre possibile generare qualsiasi output disponibile con la procedura Temporale Causal Modeling.

Scenari di modelli causali temporali

Utilizzare questa procedura per indagare come valori specifici di una determinata serie temporale, in un sistema modello, influenzano i valori previsti delle serie storiche che ne sono causalmente correlate.

Previsione modello causale temporale

La Procedura di previsione del modello di causale temporale carica un file di sistema modello creato dalla procedura Temporale Causale Modeling e applica i modelli al dataset attivo. Si può utilizzare questa procedura per ottenere previsioni per serie per le quali sono disponibili dati più attuali, senza ricostruire i propri modelli. È inoltre possibile generare qualsiasi output disponibile con la procedura Temporale Causal Modeling.

Ipotesi

- La struttura dei dati nel dataset attivo, a base di colonne o multidimensionali, deve essere la stessa struttura utilizzata quando è stato costruito il sistema modello. Per i dati multidimensionali, i campi di quota devono essere gli stessi utilizzati per costruire il sistema modello. Inoltre, i valori di dimensione utilizzati per costruire il sistema modello devono esistere nel dataset attivo.
- I modelli vengono applicati ai campi nel dataset attivo con gli stessi nomi dei campi specificati nel sistema modello.
- Il campo, o i campi, che hanno definito le osservazioni quando il sistema modello è stato costruito devono esistere nel dataset attivo. L'intervallo di tempo tra le osservazioni si ipotizza come quando i modelli sono stati costruiti. Se le osservazioni sono state definite da una specifica data, allora la stessa specifica di data deve esistere nel dataset attivo. Le specifiche della data vengono create dalla finestra di dialogo Definisci Date o dal comando DATE .
- L'intervallo di tempo dell'analisi e le eventuali impostazioni per l'aggregazione, la distribuzione e i valori mancanti sono gli stessi di quando i modelli sono stati costruiti.

Per Utilizzare Temporal Causale Modello Previsione

Questa funzione richiede l'utilizzo dell'opzione Previsione statistiche.

Dai menu, scegliere:

Analizzare > Previsione > Applicare Modelli Causali ...

1. Inserire la specifica file per un file di sistema di modelli oppure fare clic su **Sfoggia** e selezionare un file di sistema di modelli. I file di sistema di modelli vengono creati con la procedura Modellazione causale temporale.
2. Clicca sull'opzione per ristimare i modelli, creare previsioni e generare output.
3. Fare clic su **Continua**.
4. Specificare se si desidera utilizzare i parametri del modello esistenti o ristimare i parametri del modello dai dati presenti nel dataset attivo.
5. Specificare fino a che punto in futuro previsione, o specificare di non prevedere.
6. Fare clic su **Opzioni** per specificare il contenuto dell'output.
7. Fare clic su **Salva** per salvare le previsioni e esportare il sistema modello aggiornato in un file esterno quando i parametri del modello vengono ristimati.
8. Fare clic su **Esegui** per eseguire la procedura.

Parametri Modello e Previsioni

Carica dal file del modello

Le previsioni sono prodotte utilizzando i parametri del modello dal file system del modello, e i dati del dataset attivo, senza rivalutare i parametri del modello. La bontà delle misure di adattamento che vengono visualizzate in output e utilizzate per selezionare i modelli di adattamento migliori sono prese dal file system del modello. Le misure di adattamento riflettono poi i dati che sono stati utilizzati quando ogni modello è stato sviluppato (o ultimo aggiornamento). Questa opzione è appropriata per generare previsioni e output dai dati utilizzati per costruire il sistema modello.

Stima nuovamente dai dati

I parametri del modello vengono ristimati utilizzando i dati nel dataset attivo. La rivalutazione dei parametri del modello non influenza quali input sono inclusi nel modello per ogni destinazione. Questa opzione è appropriata quando si hanno nuovi dati oltre il periodo di stima originale e si desidera generare previsioni o altro output con i dati aggiornati.

Tutte le osservazioni

Specifica che il periodo di stima inizia al momento della prima osservazione e termina al momento dell'ultima osservazione in tutta la serie.

In base alle ore di inizio e di fine

È possibile specificare sia l'inizio che la fine del periodo di stima oppure è possibile specificare solo l'inizio o solo la fine. Se si omette l'inizio o la fine del periodo di stima, viene utilizzato il valore predefinito.

- Se le osservazioni sono definite da un campo data/ora, immettere i valori per l'inizio e la fine nello stesso formato utilizzato per il campo data/ora.
- Per le osservazioni definite in base a periodi ciclici, specificare un valore per ognuno dei campi di periodi ciclici. Ogni campo viene visualizzato in una colonna separata.
- Se esiste una specifica data in vigore per il dataset attivo, è necessario specificare un valore per ogni componente (ad esempio Mese) della specifica data. Ogni componente viene visualizzato in una colonna separata.
- Quando le osservazioni sono definite dall'ordine dei record, l'inizio e la fine del periodo di stima vengono definiti dal numero di riga (come visualizzato nell'Editor dei dati) del caso pertinente.

In base agli intervalli di tempo più recenti o meno recenti

Definisce il periodo di stima come un numero specificato di intervalli di tempo che iniziano nell'intervallo di tempo meno recente e terminano nell'intervallo di tempo più recente nei dati con un offset facoltativo. In questo contesto, l'intervallo temporale indica l'intervallo di tempo dell'analisi. Ad esempio, si supponga che le osservazioni siano mensili ma che l'intervallo di tempo dell'analisi sia trimestrale. Specificando **Più recente** e il valore 24 per **Numero di intervalli di tempo** si indicano i 24 dati trimestrali.

Facoltativamente, è possibile escludere un numero specificato di intervalli di tempo. Ad esempio, specificando gli ultimi 24 intervalli di tempo e 1 per il numero da escludere, si indica che il periodo di stima è costituito da 24 intervalli che precedono l'ultimo.

Estendi record nel futuro

Imposta il numero di intervalli di tempo per cui si effettua una previsione oltre la fine del periodo di stima. L'intervallo di tempo in questo caso è l'intervallo di tempo dell'analisi. Quando le previsioni sono richieste, i modelli autoregressivi vengono automaticamente creati per tutte le serie di input che non sono anche obiettivi. Questi modelli vengono poi utilizzati per generare valori per quelle serie di input nel periodo di previsione per ottenere le previsioni per gli obiettivi di quegli input.

Opzioni generali

Larghezza intervallo di confidenza (%)

Questa impostazione controlla gli intervalli di confidenza sia per le previsioni che per i parametri del modello. È possibile specificare qualsiasi valore positivo inferiore a 100. Per impostazione predefinita, viene utilizzato un intervallo di confidenza del 95%.

Soglia valore anomalo (%)

Un'osservazione è contrassegnata come valore anomalo se la probabilità, così come calcolata dal modello, che è un valore anomalo supera questa soglia. È possibile specificare un valore nell'intervallo compreso tra 50 e 100.

Serie da visualizzare

Queste opzioni specificano le serie (destinazioni o input) per le quali viene visualizzato l'output. Il contenuto dell'output per le serie specificate è determinato dalle impostazioni **Opzioni di output**.

Visualizza destinazioni associate con modelli più adatti

Per impostazione predefinita, l'output viene visualizzato per le destinazioni associate ai 10 modelli più adatti, come determinato dal valore R quadrato. È possibile specificare un numero fisso differente di modelli più adatti oppure è possibile specificare una percentuale di modelli più adatti. È anche possibile scegliere tra le seguenti misure della bontà di adattamento:

R-quadrato

Misura della bontà dell'adattamento di un modello lineare, detto anche coefficiente di determinazione. È la proporzione di variabilità della variabile obiettivo spiegata dal modello. Varia in valore dal 0 al 1. I piccoli valori indicano che il modello non si adatta bene ai dati.

Errore percentuale valore quadratico medio

Una misura di quanto i valori previsti dal modello differiscono dai valori osservati della serie. È indipendente dalle unità utilizzate e può pertanto essere utilizzata per confrontare delle serie con unità differenti.

Errore valore quadratico medio

La radice quadrata dell'errore quadratico medio. Indica di quanto una serie dipendente varia rispetto al livello stimato del modello. Questa misura viene espressa nelle stesse unità della serie dipendente.

BIC

(BIC) normalizzato. Una misura per selezionare e confrontare modelli basata sulla verosimiglianza log ridotta -2. I valori più bassi indicano i modelli migliori. Anche BIC "penalizza" i modelli sovraparametrizzati (modelli complessi con un elevato numero di input, ad esempio) ma in modo più rigoroso rispetto ad AIC.

AIC

Criterio di informazione di Akaike. Una misura per selezionare e confrontare modelli basata sulla verosimiglianza log ridotta -2. I valori più bassi indicano i modelli migliori. AIC "penalizza" i modelli sovraparametrizzati (modelli complessi con un elevato numero di input, ad esempio).

Specifiche serie individuali

È possibile specificare singole serie per le quali si desidera un output.

- Per dati basati su colonne, si specificano i campi che contengono le serie desiderate. L'ordine dei campi specificati definisce l'ordine in cui vengono visualizzati nell'output.
- Per dati multidimensionali, si specifica una particolare serie aggiungendo una voce alla griglia per il campo metrico che contiene la serie. Quindi, è possibile specificare i valori dei campi dimensione che definiscono le serie.
 - È possibile immettere il valore per ciascun campo dimensione direttamente nella griglia oppure effettuare una scelta dall'elenco di valori della dimensione disponibili. Per effettuare una selezione dall'elenco di valori dimensione disponibili, fare clic sul pulsante con i puntini di sospensione nella cella per la dimensione desiderata. Questa azione apre la finestra secondaria *Seleziona valore dimensione*.
 - È possibile ricercare l'elenco di valori della dimensione, nella finestra di dialogo secondaria *Seleziona valore dimensione*, facendo clic sull'icona a forma di binocolo e specificando un termine di ricerca. Gli spazi vengono considerati come parte del termine di ricerca. Gli asterischi (*) nel termine di ricerca non indicano caratteri jolly.
 - L'ordine delle serie nella griglia definisce l'ordine in cui vengono visualizzate nell'output.

Per entrambi i dati basati su colonne ed i dati multidimensionali, l'output è limitato a 30 serie. Questo limite include le serie individuali (input o obiettivi) specificate e gli obiettivi associati ai modelli più adatti. Le serie individuali specificate hanno la precedenza sugli obiettivi associati ai modelli più adatti.

Opzioni di output

Queste opzioni specificano il contenuto dell'output. Le opzioni nel gruppo **Output per destinazioni** generano l'output per le destinazioni associate ai modelli più adatti nelle impostazioni **Serie da visualizzare**. Le opzioni nel gruppo **Output per serie** generano l'output per le serie individuali specificate nelle impostazioni **Serie da visualizzare**.

Sistema del modello globale

Visualizza una rappresentazione grafica delle relazioni causali tra le serie nel sistema di modelli. Le tabelle sia di valori anomali che di statistiche di adattamento di modelli per le destinazioni visualizzate sono incluse come parte dell'elemento di output. Quando questa opzione viene selezionata nel gruppo **Output per serie**, viene creato un elemento di output separato per ciascuna serie individuale specificata nelle impostazioni **Serie da visualizzare**.

Le relazioni causali tra le serie hanno un livello di significatività associato, in cui un livello di significatività inferiore indica una connessione più significativa. È possibile scegliere di nascondere le relazioni con un livello di significatività superiore al valore specificato.

Valori anomali e statistiche di adattamento modello

Tabelle di valori anomali e statistiche di adattamento modello per le serie di destinazione selezionate per la visualizzazione. Queste tabelle contengono le stesse informazioni delle tabelle nella visualizzazione Sistema complessivo di modelli. Queste tabelle supportano tutte le funzioni standard per la riorganizzazione e la modifica delle tabelle.

Effetti e parametri di modelli

Le tabelle di parametri e di test degli effetti di modelli per le serie di destinazione selezionate per la visualizzazione. I test degli effetti di modelli includono le statistiche F e il valore di significatività associato per ciascun input incluso nel modello.

Diagramma di impatto

Visualizza una rappresentazione grafica delle relazioni casuali tra una serie di interesse e altre serie che possono avere effetto l'una sulle altre. Le serie che influiscono sulla serie di interesse sono indicate *cause*. Selezionando **Effetti** viene generato un diagramma di impatto inizializzato per visualizzare gli effetti. Selezionando **Cause** viene generato un diagramma di impatto inizializzato per visualizzare le cause. Selezionando **Entrambe le cause e gli effetti** vengono generati due diagrammi di impatto separati, uno inizializzato per le cause ed uno inizializzato per gli effetti. È possibile passare in modo interattivo dalle cause agli effetti e viceversa nell'elemento di output che visualizza il diagramma di impatto.

È possibile specificare il numero di livelli delle cause o degli effetti da visualizzare, dove il primo livello è solo la serie di interesse. Ogni ulteriore livello mostra ulteriori cause o effetti indiretti della serie di interesse. Ad esempio, il terzo livello nella visualizzazione degli effetti è composto dalle serie che contengono serie nel secondo livello come input diretto. Le serie nel terzo livello sono quindi indirettamente interessate dalla serie di interesse, poiché la serie di interesse è un input diretto per le serie nel secondo livello.

Grafico serie

Grafici dei valori osservati e previsti per le serie di destinazione selezionate per la visualizzazione. Quando le previsioni sono richieste, il grafico mostra anche i valori previsti e gli intervalli di confidenza per le previsioni.

Grafico residui

Grafici dei residui di modelli per le serie di destinazione selezionate per la visualizzazione.

Input migliori

Grafici di ogni destinazione visualizzata, nel tempo, insieme ai primi 3 input per la destinazione. Gli input principali sono gli input con il valore di significatività più basso. Per accogliere diverse scale di input e di destinazione, l'asse y rappresenta il punteggio z per ciascuna serie.

Tabella delle previsioni

Tabella di valori previsti e di intervalli di confidenza di quelle previsioni per le serie di destinazione selezionate per la visualizzazione.

Analisi causa principale valore anomalo

Determina quali serie sono più probabilmente la causa di ciascun valore anomalo in una serie di interesse. L'analisi della causa principale del valore anomalo viene eseguita per ciascuna serie di destinazione inclusa nell'elenco di serie individuali delle impostazioni **Serie da visualizzare**.

Output

Tabella e grafico valori anomali interattivi

Tabella e grafico di valori anomali e di cause principali di tali valori anomali per ciascuna serie di interesse. La tabella contiene una singola riga per ciascun valore anomalo. Il grafico è un diagramma di impatto. La selezione di una riga nella tabella evidenzia il percorso, nel diagramma di impatto, dalla serie di interesse alla serie che più probabilmente è la causa del valore anomalo associato.

Tabella pivot di valori anomali

Tabella di valori anomali e di cause principali di tali valori anomali per ciascuna serie di interesse. Questa tabella contiene le stesse informazioni della tabella nella visualizzazione interattiva. Questa tabella supporta tutte le funzioni standard per la riorganizzazione e la modifica delle tabelle.

Livelli causali

È possibile specificare il numero di livelli da includere nella ricerca per le cause principali. Il concetto dei livelli qui utilizzato è lo stesso descritto per i diagrammi di impatto.

Adattamento modello tra tutti i modelli

Istogramma dell'adattamento modello per tutti i modelli e per le statistiche di adattamento selezionate. Sono disponibili le seguenti statistiche di adattamento:

R-quadro

Misura della bontà dell'adattamento di un modello lineare, detto anche coefficiente di determinazione. È la proporzione di variabilità della variabile obiettivo spiegata dal modello. Varia in valore dal 0 al 1. I piccoli valori indicano che il modello non si adatta bene ai dati.

Errore percentuale valore quadratico medio

Una misura di quanto i valori previsti dal modello differiscono dai valori osservati della serie. È indipendente dalle unità utilizzate e può pertanto essere utilizzata per confrontare delle serie con unità differenti.

Errore valore quadratico medio

La radice quadrata dell'errore quadratico medio. Indica di quanto una serie dipendente varia rispetto al livello stimato del modello. Questa misura viene espressa nelle stesse unità della serie dipendente.

BIC

(BIC) normalizzato. Una misura per selezionare e confrontare modelli basata sulla verosimiglianza log ridotta -2. I valori più bassi indicano i modelli migliori. Anche BIC "penalizza" i modelli sovraparametrizzati (modelli complessi con un elevato numero di input, ad esempio) ma in modo più rigoroso rispetto ad AIC.

AIC

Criterio di informazione di Akaike. Una misura per selezionare e confrontare modelli basata sulla verosimiglianza log ridotta -2. I valori più bassi indicano i modelli migliori. AIC "penalizza" i modelli sovraparametrizzati (modelli complessi con un elevato numero di input, ad esempio).

Valori anomali nel tempo

Grafico a barre del numero di valori anomali, su tutte le destinazioni, per ciascun intervallo di tempo nel periodo di stima.

Trasformazioni delle serie

Tabella delle eventuali trasformazioni che sono state applicate alla serie nel sistema di modelli. Le possibili trasformazioni sono distribuzione, aggregazione e assegnazione di valori mancanti.

Salva

Salva destinazioni

È possibile salvare le previsioni del modello, gli intervalli di confidenza e i residui come nuove variabili. Ogni obiettivo specificato per salvare genera la propria serie di nuove variabili e ogni nuova variabile contiene valori sia per periodi di stima che di previsione. Per i valori previsti, gli intervalli di confidenza e i residui di rumore è possibile specificare il nome radice da utilizzare come prefisso per le nuove variabili. Il nome variabile completo è la concatenazione del nome radice e il nome del campo che contiene la serie target. Il nome radice deve essere conforme alle regole per i nomi variabili validi. Il nome variabile viene esteso se necessario per evitare conflitti di denominazione variabili.

Indica casi contenenti previsioni

Crea una variabile che indica se un record contiene dati di previsione. È possibile specificare il nome variabile. Il valore predefinito è *ForecastIndicator*.

Obiettivi da salvare

Specifica se vengono create nuove variabili per tutte le serie di destinazione nel sistema modello o solo la serie di destinazione che vengono specificate sulle impostazioni **Series to Display**.

Opzioni di destinazione

È possibile salvare sia le trasformazioni dei dati (come l'aggregazione o l'imputazione dei valori mancanti) e le nuove variabili (specificate sulle impostazioni di **Save Targets**) ad un file di dati IBM SPSS Statistiche o un nuovo dataset nella sessione corrente. I valori di data / ora nei dati salvati sono allineati all'inizio di ogni intervallo di tempo, come il primo del mese, e rappresentano l'intervallo di tempo dell'analisi per il sistema modello. È possibile salvare eventuali nuove variabili nel dataset attivo solo se le osservazioni sono definite da una specifica data, o da un ordine record e i dati non sono aggregati.

Sistema del modello di esportazione

Salva il sistema modello in un archivio file compresso (file .zip). Il file di sistema modello può essere riutilizzato da questa procedura. Può essere utilizzato anche dalla procedura Temporal Causal Model Scenarios per eseguire l'analisi di scenario. Questa opzione è disponibile solo quando i parametri del modello vengono ristimati.

Scenari di modelli causali temporali

La procedura Scenari del modello causale temporale esegue gli scenari definiti dall'utente per un sistema di modelli causali temporali, con i dati dal dataset attivo. Uno *scenario* è definito da una serie temporale, a cui si fa riferimento come *serie root*, e da una serie di valori definiti dall'utente per tale serie in un intervallo di tempo specificato. I valori specificati vengono quindi utilizzati per generare le previsioni per serie temporali interessate dalla serie root. La procedura richiede un file di sistema di modelli che è stato creato dalla procedura Modellazione causale temporale. Si presuppone che il dataset attivo sia lo stesso dataset che è stato utilizzato per creare il file di sistema di modelli.

Esempio

Utilizzando la procedura Modellazione causale temporale, un responsabile dei processi decisionali aziendali ha rilevato una metrica chiave che influisce su diversi importanti indicatori di prestazioni. La metrica è controllabile, quindi il responsabile dei processi decisionali desidera esaminare l'effetto delle diverse serie di valori per la metrica nel trimestre successivo. La verifica è facilmente eseguita caricando il file di sistema di modelli nella procedura Scenari del modello causale temporale e specificando le serie di valori per la metrica chiave.

Esecuzione di scenari del modello causale temporale

Questa funzione richiede l'utilizzo dell'opzione Previsione statistiche.

Dai menu, scegliere:

Analizzare > Previsione > Applicare Modelli Causali ...

1. Inserire la specifica file per un file di sistema di modelli oppure fare clic su **Sfoggia** e selezionare un file di sistema di modelli. I file di sistema di modelli vengono creati con la procedura Modellazione causale temporale.
2. Fare clic sull'opzione per eseguire gli scenari.
3. Fare clic su **Continua**.
4. Nella scheda Scenari (nella finestra Scenari di modelli causali temporali), fare clic su **Definisci periodo dello scenario** e specificare il periodo dello scenario.
5. Per i dati basati su colonne, fare clic su **Aggiungi scenario** per definire ciascun scenario. Per dati multidimensionali, fare clic su **Aggiungi scenario** per definire ogni singolo scenario e su **Aggiungi gruppo di scenari** per definire ciascun gruppo di scenari.
6. Fare clic su **Opzioni** per specificare il contenuto dell'output e l'ampiezza della serie che sarà interessata dallo scenario.
7. Fare clic su **Esegui** per eseguire la procedura.

Definizione del periodo dello scenario

Il periodo dello scenario è il periodo in cui è possibile specificare i valori utilizzati per eseguire gli scenari. Può iniziare prima o dopo la fine del periodo di stima. È possibile facoltativamente specificare di effettuare una previsione oltre la fine del periodo dello scenario. Per impostazione predefinita, le previsioni vengono generate alla fine del periodo dello scenario. Tutti gli scenari utilizzano lo stesso periodo dello scenario e le specifiche sulla distanza di previsione.

Nota: Le previsioni iniziano dal primo periodo di tempo dopo l'inizio del periodo dello scenario. Ad esempio, se il periodo dello scenario inizia il giorno 1 novembre 2014 e l'intervallo di tempo è Mesi, la prima previsione è per il giorno 1 dicembre 2014.

Specifica in base a Ora di inizio, Ora di fine e Prevedi fino a

- Se le osservazioni sono definite da un campo data/ora, immettere i valori in base a Ora di inizio, Ora di fine e Prevedi fino a, nello stesso formato utilizzato per il campo data/ora. I valori dei campi data/ora vengono allineati all'inizio dell'intervallo di tempo associato. Ad esempio, se l'intervallo di tempo dell'analisi è Mesi, il valore 10/10/2014 viene modificato in 10/01/2014, per indicare l'inizio del mese.
- Per le osservazioni definite in base a periodi ciclici, specificare un valore per ognuno dei campi di periodi ciclici. Ogni campo viene visualizzato in una colonna separata.
- Se esiste una specifica data in vigore per il dataset attivo, è necessario specificare un valore per ogni componente (ad esempio Mese) della specifica data. Ogni componente viene visualizzato in una colonna separata.
- Quando le osservazioni sono definite dall'ordine dei record, i valori Ora di inizio, Ora di fine e Prevedi fino a vengono definiti dal numero di riga (come visualizzato nell'editor dei dati) del caso pertinente.

Specificare gli intervalli in base al tempo relativi alla fine del periodo di stima

Definisce l'inizio e la fine in termini di numero di intervalli di tempo relativi alla fine del periodo di stima, dove l'intervallo di tempo è l'intervallo dell'analisi. La fine del periodo di stima è definita come intervallo di tempo 0. Gli intervalli di tempo prima della fine del periodo di stima hanno valori e intervalli negativi dopo la fine del periodo di stima hanno valori positivi. È possibile inoltre specificare il numero di intervalli di previsione che superano la fine del periodo dello scenario. Il valore predefinito è 0.

Ad esempio, si supponga che l'intervallo di tempo dell'analisi sia Mesi e che venga specificato 1 per l'inizio dell'intervallo, 3 per la fine dell'intervallo e 1 per la distanza di previsione oltre tale periodo. Il periodo dello scenario è quindi rappresentato dai 3 mesi che seguono la fine del periodo di stima. Le previsioni vengono generate per il secondo ed il terzo mese del periodo dello scenario e per 1 ulteriore mese dopo la fine del periodo dello scenario.

Aggiunta di scenari e gruppi di scenari

La scheda Scenari specifica gli scenari da eseguire. Per definire gli scenari, è necessario prima definire il periodo dello scenario facendo clic su **Definisci periodo dello scenario**. Gli scenari ed i gruppi di scenari (solo per i dati multidimensionali) vengono creati facendo clic sul pulsante **Aggiungi scenario** o **Aggiungi gruppo di scenari** associato. Selezionando un particolare scenario o gruppo di scenari nella griglia associata, è possibile modificarlo, eseguirne una copia o eliminarlo.

Dati basati su colonne

La colonna **Campo root** nella griglia specifica il campo delle serie temporali i cui valori sono sostituiti con i valori dello scenario. La colonna **Valori dello scenario** visualizza i valori dello scenario specificati nell'ordine dal meno recente al più recente. Se i valori dello scenario sono definiti da un'espressione, la colonna visualizza l'espressione.

Dati multidimensionali

Scenari individuali

Ciascuna riga nella griglia Scenari individuali specifica una serie temporale i cui valori sono sostituiti dai valori dello scenario specificati. La serie è definita dalla combinazione del campo specificato nella colonna **Metrica root** e dal valore specificato per ciascuno dei campi dimensione. Il contenuto della colonna **Valori dello scenario** è lo stesso utilizzato per i dati basati su colonne.

Gruppi di scenari

Un *gruppo di scenari* definisce una serie di scenari che si basano su un singolo campo di metriche root e su più serie di valori dimensione. Ciascuna serie di valori della dimensione (un valore per campo dimensione), per il campo di metriche specificato, definisce una serie temporale. Un singolo scenario viene quindi generato per ogni serie temporale di questo tipo, i cui valori vengono sostituiti dai valori di scenario. I valori di scenario di un gruppo di scenari vengono specificati da un'espressione, che è quindi applicata ad ogni serie temporale nel gruppo.

La colonna **Numero di serie** visualizza il numero delle serie di valori dimensione che sono associate a un gruppo di scenari. Il valore visualizzato può essere più grande del numero effettivo di serie temporali associate al gruppo di scenari (una serie per insieme). Questa condizione si verifica quando alcune delle combinazioni specificate di valori dimensione non corrispondono alle serie contenute dalla metrica root per il gruppo.

Come esempio di un gruppo di scenari, considerare un campo di metriche *advertising* e due campi dimensione *region* e *brand*. È possibile definire un gruppo di scenari che si basa su *advertising* come metrica root e che include tutte le combinazioni di *region* e *brand*. È quindi possibile specificare *advertising*1.2* come l'espressione che esamina l'effetto che si ottiene aumentando del 20 per cento *advertising* per ognuna delle serie temporali associate al campo *advertising*. Se sono presenti 4 valori di *region* e 2 valori di *brand*, esistono 8 serie temporali di questo tipo e pertanto 8 scenari definiti dal gruppo.

Definizione dello scenario

Le impostazioni per la definizione di uno scenario dipendono dal fatto che i dati siano basati su colonne o multidimensionali.

Serie root

Specifica la serie root dello scenario. Ogni scenario si basa su una serie root singola. Per dati basati su colonne, è possibile selezionare il campo che definisce la serie root. Per dati multidimensionali, si specifica la serie root aggiungendo una voce alla griglia per il campo metrico che contiene la serie. Quindi, è possibile specificare i valori dei campi dimensione che definiscono le serie root. Le seguenti opzioni si applicano per specificare i valori della dimensione:

- È possibile immettere il valore per ciascun campo dimensione direttamente nella griglia oppure effettuare una scelta dall'elenco di valori della dimensione disponibili. Per effettuare una selezione dall'elenco di valori dimensione disponibili, fare clic sul pulsante con i puntini di sospensione nella cella per la dimensione desiderata. Questa azione apre la finestra secondaria *Seleziona valore dimensione*.
- È possibile ricercare l'elenco di valori della dimensione, nella finestra di dialogo secondaria *Seleziona valore dimensione*, facendo clic sull'icona a forma di binocolo e specificando un termine di ricerca. Gli spazi vengono considerati come parte del termine di ricerca. Gli asterischi (*) nel termine di ricerca non indicano caratteri jolly.

Specifica destinazioni interessate

Utilizzare questa opzione quando si conoscono destinazioni specifiche interessate dalle serie root e si desidera esaminare gli effetti solo su quelle destinazioni. Per impostazione predefinita, le destinazioni che sono interessate dalle serie root sono determinate automaticamente. È possibile specificare l'ampiezza delle serie interessate dallo scenario con le impostazioni nella scheda *Opzioni*.

Per dati basati su colonne, selezionare le destinazioni desiderate. Per dati multidimensionali, si specifica la serie di destinazione aggiungendo una voce alla griglia per il campo metrico di destinazione che contiene la serie. Per impostazione predefinita, vengono incluse tutte le serie contenute nel campo di metrica specificato. È possibile personalizzare l'insieme di serie incluse personalizzando i valori inclusi per uno o più campi dimensione. Per personalizzare i valori della dimensione inclusi, fare clic sul pulsante con i puntini di sospensione della dimensione desiderata. Questa azione apre la finestra *Selezione valori dimensione*.

La colonna **Numero di serie** (per dati multidimensionali) visualizza il numero delle serie di valori dimensione che sono attualmente specificate per la metrica di destinazione associata. Il valore visualizzato può essere più grande del numero effettivo di serie di destinazione interessate (una serie per insieme). Questa condizione si verifica quando alcune delle combinazioni specificate di valori dimensione non corrispondono alle serie contenute dalla metrica di destinazione associata.

ID scenario

Ogni scenario deve avere un identificativo univoco. L'identificativo viene visualizzato nell'output che è associato allo scenario. Non vi sono limitazioni, oltre all'univocità, sul valore dell'identificativo.

Specifica valori scenario per serie radice

Utilizzare questa opzione per specificare valori espliciti per la serie root nel periodo dello scenario. È necessario specificare un valore numerico per ogni intervallo di tempo che viene elencato nella griglia. È possibile ottenere i valori delle serie root (effettivi o previsti) per ciascun intervallo nel periodo dello scenario facendo clic su **Lettura, Previsione** o **Lettura\Previsione**.

Specifica espressione per valori scenario per serie radice

È possibile definire un'espressione per calcolare i valori della serie root nel periodo dello scenario. È possibile immettere l'espressione direttamente o facendo clic sul pulsante del calcolatore e creando l'espressione dal Builder di espressioni dei valori di scenario.

- L'espressione può contenere qualsiasi obiettivo o input nel sistema del modello.
- Quando il periodo dello scenario si estende oltre i dati esistenti, l'espressione viene applicata ai valori previsti dei campi nell'espressione.
- Per dati multidimensionali, ciascun campo nell'espressione specifica una serie temporale definita dal campo e dai valori dimensione che sono stati specificati per la metrica root. Si tratta di quelle serie temporali utilizzate per valutare l'espressione.

Come esempio, si supponga che il campo root sia *advertising* e che l'espressione sia *advertising*1.2*. I valori per *advertising* utilizzati nello scenario rappresentano un incremento del 20 per cento sui valori esistenti.

Nota: Gli scenari vengono creati facendo clic su **Aggiungi scenario** nella scheda Scenari.

Builder di espressioni dei valori di scenario

Utilizzare il Builder di espressioni dei valori di scenario per creare un'espressione utilizzata per calcolare i valori di scenario di un singolo scenario o di un gruppo di scenari. Per creare un'espressione, incollare i componenti nel campo **Espressione** oppure digitare direttamente nel campo **Espressione**.

- L'espressione può contenere qualsiasi obiettivo o input nel sistema del modello.
- È possibile incollare le funzioni selezionando un gruppo nell'elenco Gruppo di funzioni e facendo doppio clic sulla funzione nell'elenco Funzioni (o selezionare la funzione e fare clic sulla freccia accanto all'elenco Gruppo di funzioni). Immettere i parametri indicati dai punti interrogativi. Il gruppo di funzioni etichettato **Tutti** offre un elenco di tutte le funzioni disponibili. Una breve descrizione della funzione attualmente selezionata viene visualizzata in un'area riservata della finestra di dialogo.
- Le costanti stringa devono essere incluse tra virgolette.
- Se i valori contengono numeri decimali, è necessario utilizzare un punto (.) come indicatore decimale.

Per accedere al Builder di espressioni dei valori di scenario, fare clic sul pulsante del calcolatore nella finestra di dialogo Definizione dello scenario o Definizione del gruppo di scenari.

Selezione valori dimensione

Per dati dimensionali, è possibile personalizzare i valori della dimensione che definiscono le destinazioni interessate da uno scenario o gruppo di scenari. È inoltre possibile personalizzare i valori della dimensione che definiscono l'insieme delle serie root per un gruppo di scenari.

Tutti i valori

Specifica che sono inclusi tutti i valori del campo dimensione corrente. Questa opzione è il valore predefinito.

Selezione valori

Utilizzare questa opzione per specificare la serie di valori per il campo dimensione corrente. È possibile filtrare l'insieme di valori da cui effettuare la scelta. I valori che soddisfano la condizione filtro vengono visualizzati nella scheda **Corrispondenti**, mentre i valori che non soddisfano la condizione filtro vengono visualizzati nella scheda **Non corrispondenti** dell'elenco **Valori non selezionati**. La scheda **Tutti** elenca tutti i valori non selezionati, indipendentemente da qualsiasi condizione filtro.

- È possibile utilizzare gli asterischi (*) per indicare i caratteri jolly quando si specifica un filtro.

- Per deselezionare il filtro corrente, specificare un valore vuoto per il termine di ricerca nella finestra Filtra valori visualizzati.

Per personalizzare i valori della dimensione per le destinazioni interessate:

1. Dalla finestra Definizione dello scenario o Definizione del gruppo di scenari, selezionare la metrica di destinazione per cui si desidera personalizzare i valori della dimensione.
2. Fare clic sul pulsante con i puntini di sospensione nella colonna per la dimensione che si desidera personalizzare.

Per personalizzare i valori della dimensione per le serie root di un gruppo di scenari:

1. Dalla finestra Definizione del gruppo di scenari, fare clic sul pulsante con i puntini di sospensione (nella griglia della serie root) per la dimensione che si desidera personalizzare.

Definizione del gruppo di scenari

Serie root

Specifica l'insieme di serie root per il gruppo di scenari. Un singolo scenario viene generato per ciascuna serie temporale nell'insieme. La serie root viene specificata aggiungendo una voce alla griglia per il campo di metriche che contiene la serie desiderata. Quindi, è possibile specificare i valori dei campi dimensione che definiscono l'insieme. Per impostazione predefinita, tutte le serie contenute nel campo di metriche root specificato sono incluse. È possibile personalizzare l'insieme di serie incluse personalizzando i valori inclusi per uno o più campi dimensione. Per personalizzare i valori della dimensione inclusi, fare clic sul pulsante con i puntini di sospensione di una dimensione. Questa azione apre la finestra Selezione valori dimensione.

La colonna **Numero di serie** visualizza il numero delle serie di valori dimensione che sono attualmente incluse per la metrica root associata. Il valore visualizzato può essere più grande del numero effettivo di serie root per il gruppo di scenari (una serie per insieme). Questa condizione si verifica quando alcune delle combinazioni specificate di valori dimensione non corrispondono alle serie contenute dalla metrica root.

Specifica serie di obiettivi interessati

Utilizzare questa opzione quando si conoscono destinazioni specifiche interessate dall'insieme di serie root e si desidera esaminare gli effetti solo su quelle destinazioni. Per impostazione predefinita, le destinazioni che sono interessate da ciascuna serie root sono determinate automaticamente. È possibile specificare l'ampiezza delle serie interessate da ciascun scenario individuale con le impostazioni nella scheda Opzioni.

La serie di destinazione viene specificata aggiungendo una voce alla griglia per il campo di metriche che contiene la serie. Per impostazione predefinita, vengono incluse tutte le serie contenute nel campo di metrica specificato. È possibile personalizzare l'insieme di serie incluse personalizzando i valori inclusi per uno o più campi dimensione. Per personalizzare i valori della dimensione inclusi, fare clic sul pulsante con i puntini di sospensione della dimensione desiderata. Questa azione apre la finestra Selezione valori dimensione.

La colonna **Numero di serie** visualizza il numero delle serie di valori dimensione che sono attualmente specificate per la metrica di destinazione associata. Il valore visualizzato può essere più grande del numero effettivo di serie di destinazione interessate (una serie per insieme). Questa condizione si verifica quando alcune delle combinazioni specificate di valori dimensione non corrispondono alle serie contenute dalla metrica di destinazione associata.

Prefisso ID scenario

Ogni gruppo di scenari deve avere un prefisso univoco. Il prefisso viene utilizzato per creare un identificativo visualizzato nell'output che è associato ad ogni singolo scenario nel gruppo di scenari. L'identificativo di un singolo scenario è il prefisso, seguito da un carattere di sottolineatura, e seguito dal valore di ciascun campo dimensione che identifica la serie root. I valori della dimensione sono separati dai caratteri di sottolineatura. Non vi sono limitazioni, oltre all'univocità, sul valore del prefisso.

Espressione dei valori dello scenario per serie root

I valori di scenario di un gruppo di scenari sono specificati da un'espressione, che viene quindi utilizzata per calcolare i valori di ciascuna delle serie root nel gruppo. È possibile immettere un'espressione direttamente o facendo clic sul pulsante del calcolatore e creando l'espressione dal Builder di espressioni dei valori di scenario.

- L'espressione può contenere qualsiasi obiettivo o input nel sistema del modello.
- Quando il periodo dello scenario si estende oltre i dati esistenti, l'espressione viene applicata ai valori previsti dei campi nell'espressione.
- Per ciascuna serie root nel gruppo, i campi nell'espressione specificano le serie temporali definite da questi campi e i valori dimensione che definiscono la serie root. Si tratta di quelle serie temporali utilizzate per valutare l'espressione. Ad esempio, se una serie root è definita da `region= ' north '` e `brand= ' X '`, le serie temporali utilizzate nell'espressione sono definite da quegli stessi valori dimensione.

Come esempio, si supponga che il campo della metrica root sia *advertising* e che vi siano due campi dimensione *region* e *brand*. Inoltre, si supponga che il gruppo di scenari includa tutte le combinazioni di valori del campo dimensione. È quindi possibile specificare `advertising*1.2` come l'espressione che esamina l'effetto che si ottiene aumentando del 20 percento *advertising* per ognuna delle serie temporali associate al campo *advertising*.

Nota: I gruppi di scenari si applicano solo ai dati multidimensionali e vengono creati facendo clic su **Aggiungi gruppo di scenari** nella scheda Scenari.

Opzioni

Massimo livello delle destinazioni interessate

Specifica il numero massimo di livelli delle destinazioni interessate. Ogni livello successivo, fino ad un massimo di 5, include le destinazioni che sono più indirettamente interessate dalle serie root. In particolare, il primo livello include le destinazioni che hanno la serie root come input diretto. Le destinazioni nel secondo livello hanno destinazioni nel primo livello come input diretto e così via. L'aumento del valore di questa impostazione aumenta la complessità del calcolo e potrebbe influire sulle prestazioni.

Numero massimo di destinazioni rilevate automaticamente

Specifica il numero massimo di destinazioni interessate che sono rilevate automaticamente per ciascuna serie root. L'aumento del valore di questa impostazione aumenta la complessità del calcolo e potrebbe influire sulle prestazioni.

Diagramma di impatto

Visualizza una rappresentazione grafica delle relazioni causali tra la serie root di ciascuno scenario e la serie di destinazione da essa interessata. Le tabelle dei valori dello scenario e dei valori previsti per le destinazioni interessate sono inclusi come parte dell'elemento di output. Il grafico include i grafici dei valori previsti delle destinazioni interessate. Facendo clic una volta su qualsiasi nodo nel diagramma di impatto, viene visualizzato un diagramma di sequenza dettagliata per la serie associata al nodo. Per ogni scenario verrà creato un diagramma di impatto separato.

Grafici serie

Genera i grafici delle serie di valori previsti per ciascuno degli obiettivi interessati in ogni scenario.

Tabelle scenari e previsioni

Tabelle di valori previsti e di valori di scenario per ogni scenario. Queste tabelle contengono le stesse informazioni delle tabelle nel diagramma di impatto. Queste tabelle supportano tutte le funzioni standard per la riorganizzazione e la modifica delle tabelle.

Includi intervalli di confidenza in grafici e tabelle

Specifica se gli intervalli di confidenza per le previsioni di scenario sono inclusi nell'output del grafico e della tabella.

Larghezza intervallo di confidenza (%)

Questa impostazione controlla gli intervalli di confidenza per le previsioni di scenario. È possibile specificare qualsiasi valore positivo inferiore a 100. Per impostazione predefinita, viene utilizzato un intervallo di confidenza del 95%.

Misure di bontà - di - Fit

Questa sezione fornisce le definizioni delle misure di bontà di adattamento utilizzate nella modellazione della serie temporale.

- *R-quadrato stazionario*. Una misura che confronta la parte stazionaria del modello con un semplice modello di media. Questa misura è preferibile all'*R-quadrato semplice* se è presente una tendenza o un modello stagionale. *R-quadrato stazionario* può essere negativo con un intervallo negativo infinito tendente a 1. I valori negativi indicano che il modello esaminato è peggiore del modello baseline. I valori positivi indicano che il modello esaminato è migliore del modello baseline.
- *R-quadrato*. Una stima della proporzione della variazione totale della serie spiegata dal modello. Questa misura è particolarmente utile quando la serie è fissa. *R - quadrato* può essere negativo con una gamma di infinito negativo a 1. I valori negativi indicano che il modello esaminato è peggiore del modello baseline. I valori positivi indicano che il modello esaminato è migliore del modello baseline.
- *RMSE*. Errore della media dei quadrati. La radice quadrata dell'errore quadratico medio. Indica di quanto una serie dipendente varia rispetto al livello previsto del modello; questa misura viene espressa nelle stesse unità della serie dipendente.
- *MAPE*. Errore percentuale assoluta media. Indica di quanto la serie dipendente varia rispetto al livello predittivo del modello. È indipendente dalle unità usate e può quindi essere usato per comparare le serie con unità diverse.
- *MAE*. Errore assoluto medio. Misura di quanto la serie varia rispetto al livello predittivo del modello. *MAE* è notificato in unità originali delle serie.
- *MaxAPE*. Errore percentuale assoluta massima. Errore stimato più alto, espresso in percentuale. Questa misura è utile per immaginare uno scenario peggiore per le previsioni.
- *MaxAE*. Errore assoluto massimo. L'errore stimato più alto, espresso nelle stesse unità della serie dipendente. Al pari di *MaxAPE*, è utile per stimare lo scenario peggiore per le previsioni. L'errore assoluto massimo e l'errore percentuale assoluto massimo possono verificarsi in punti diversi della serie; ad esempio se l'errore assoluto di una serie grande è leggermente maggiore dell'errore assoluto del valore minore della serie. In questo caso l'errore assoluto massimo si verifica nel valore maggiore della serie mentre l'errore percentuale assoluto massimo si verifica nel valore minore della serie.
- *BIC normalizzato*. Criterio di informazioni di Bayes normalizzato. Misura generale dell'adattamento generale di un modello che tenta di tenere conto della complessità del modello. Questa misura è un punteggio basato sull'errore quadratico medio e comprende una penalità per il numero di parametri del modello e una lunghezza per la serie. La penalità rimuove il vantaggio dei modelli con più parametri, consentendo di comparare più facilmente la statistica su modelli differenti per la stessa serie.

Tipi di valori anomali

Questa sezione fornisce le definizioni dei tipi di outlier utilizzati nella modellazione delle serie storiche.

- *Additivo*. Un valore anomalo che influenza una singola osservazione. Ad esempio, un errore di codifica dati può essere identificato come un valore anomalo additivo.
- *Spostamento di livello*. Un valore anomalo che sposta tutte le osservazioni di una costante, a partire da un determinato punto della serie. Uno spostamento di livello può essere provocato da una variazione nelle politiche.
- *Innovativo*. Un valore anomalo che agisce come aggiunta al termine rumore in un determinato punto della serie. Nelle serie stazionarie, il valore anomalo innovativo influisce su svariate osservazioni. Per le serie non stazionarie, può influenzare tutte le osservazioni che iniziano da un punto della serie specifico.
- *Transiente*. Un valore anomalo il cui impatto decade esponenzialmente a 0.
- *Additivo stagionale*. Un valore anomalo che influenza una particolare osservazione e tutte le osservazioni successive separate da essa da uno o più periodi stagionali. Il valore anomalo influisce nello stesso modo su tutte le osservazioni. Un valore anomalo additivo stagionale può verificarsi se, a iniziare da un certo anno, le vendite sono maggiori ogni gennaio.

- *Tendenza locale*. Un valore anomalo che dà inizio a una tendenza locale in un punto particolare della serie.
- *Patch additivo*. Un gruppo di due o più valori anomali additivi consecutivi. La selezione di questo tipo di valore anomalo determina il rilevamento dei singoli valori anomali additivi, in aggiunta a gruppi di essi.

Guida ai grafici ACF/PACF

I tracciati qui riportati sono quelli di processi ARIMA puro o teorico. Ecco alcune linee guida generali per l'identificazione del processo:

- Le serie non stazionarie hanno un ACF che rimane significativo per una mezza dozzina o più lags, piuttosto che rapidamente in calo a 0. Devi fare la differenza di una serie fino a quando non è fermo prima di poter identificare il processo.
- I processi autoregressivi hanno un ACF e le spiagge in calo esponenziale nel primo uno o più lags del PACF. Il numero di spiagge indica l'ordine dell'autoregressione.
- I processi medi di spostamento hanno delle spiagge nel primo uno o più lags dell'ACF e un PACF in calo esponenziale. Il numero di spiedini indica l'ordine della media mobile.
- I processi misti (ARMA) tipicamente mostrano declinazioni esponenziali sia in ACF che in PACF.

In fase di identificazione non è necessario preoccuparsi del segno dell'ACF o del PACF, o della velocità con cui un ACF o PACF in calo esponenziale si avvicina al 0. Questi dipendono dal segno e dal valore effettivo dei coefficienti AR e MA. In alcune istanze, un ACF in calo esponenziale si alterna tra valori positivi e negativi.

I grafici ACF e PACF da dati reali non sono mai così puliti come i tracciati qui riportati. Devi imparare a scegliere ciò che è essenziale in ogni trama data. Controllare sempre ACF e PACF dei residui, nel caso in cui la vostra identificazione sia errata. Tenere presente che:

- I processi stagionali mostrano questi modelli ai ritardi stagionali (i multipli del periodo stagionale).
- Si ha diritto a trattare valori non significativi come 0. Ovvero, è possibile ignorare i valori che risiedono all'interno degli intervalli di confidenza sui tracciati. Non bisogna ignorarli, però, in particolare se continuano lo schema dei valori statisticamente significativi.
- Un'autocorrelazione occasionale sarà statisticamente significativa per caso da sola. È possibile ignorare un'autocorrelazione statisticamente significativa se è isolata, preferibilmente ad un ritardo elevato, e se non si verifica ad un ritardo stagionale.

Consulta qualsiasi testo su analisi ARIMA per una discussione più completa dei tracciati ACF e PACF.

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

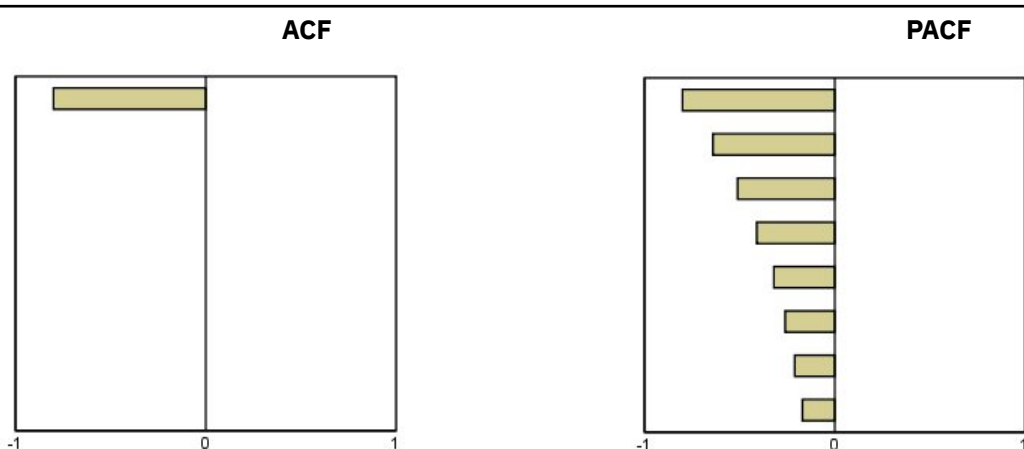
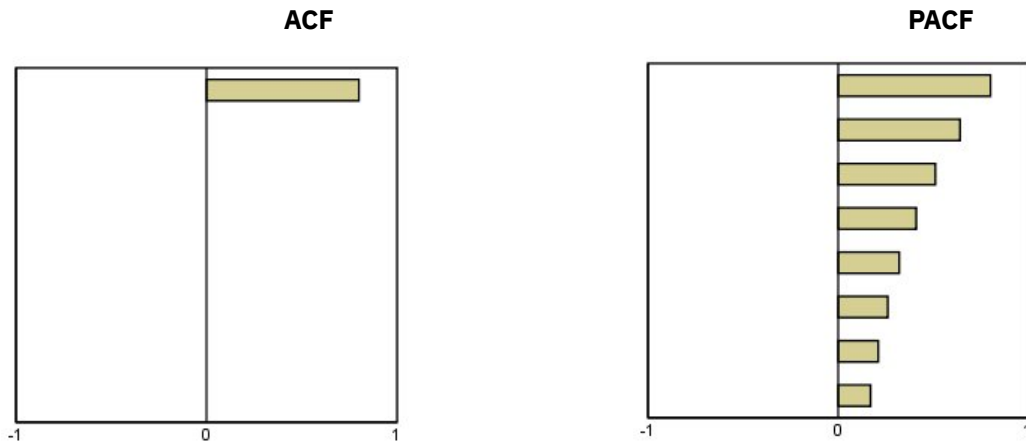


Tabella 4. ARIMA (0,0, 1), $q < 0$



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

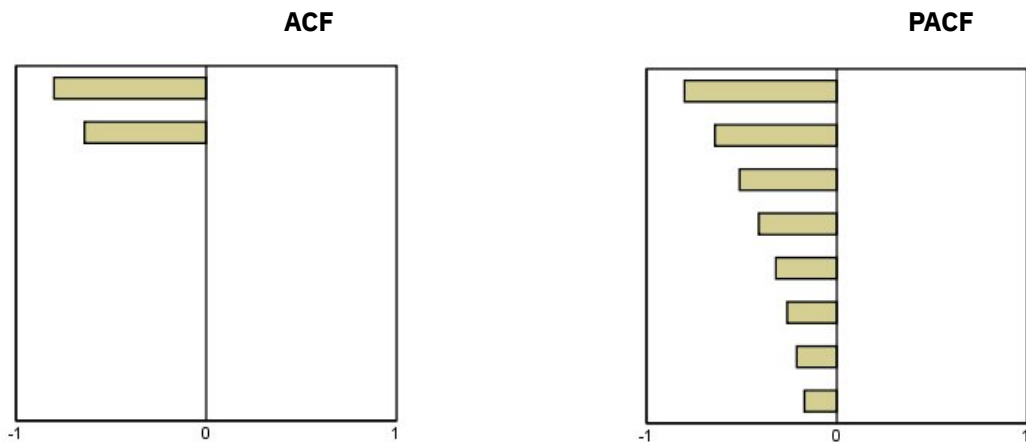


Tabella 5. ARIMA (1,0, 0), $f > 0$

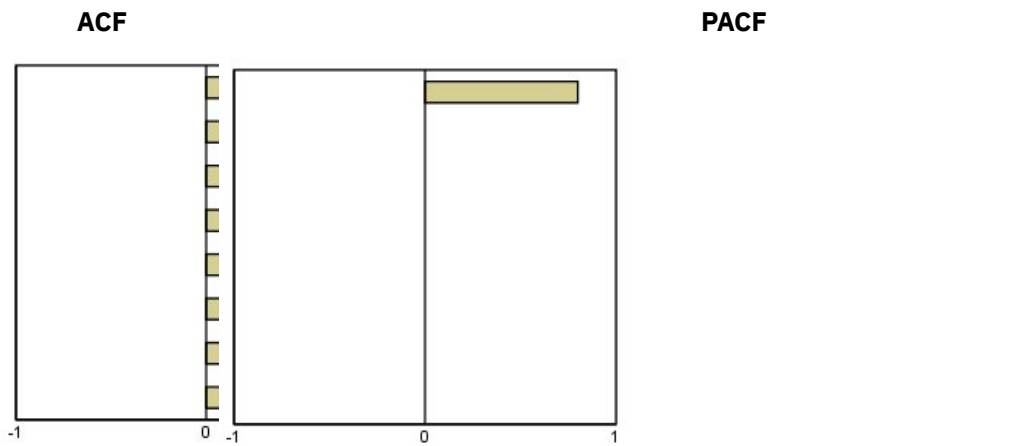
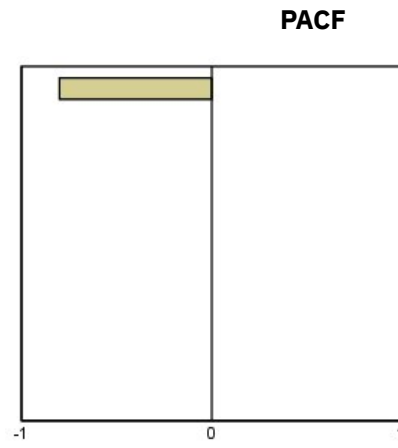
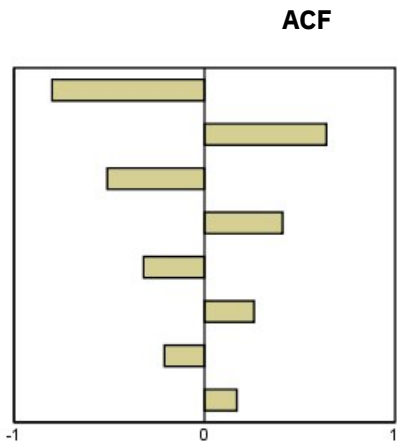
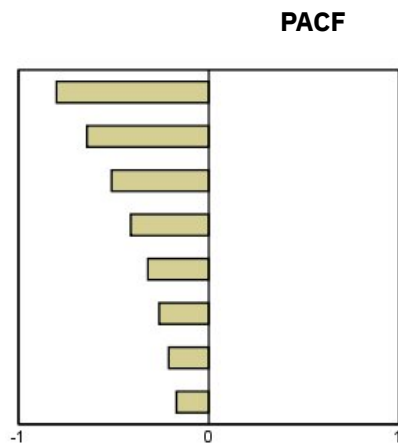
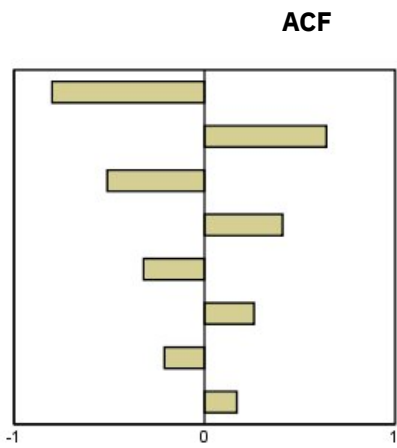


Tabella 6. ARIMA (1,0,0), $f < 0$



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



ARIMA (2,0,0), $1_2 > 0$

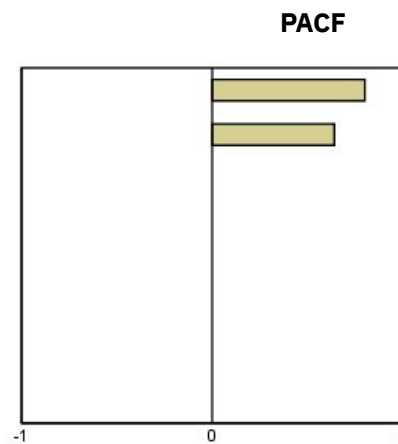
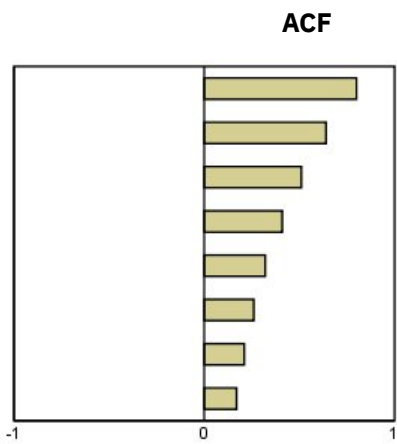
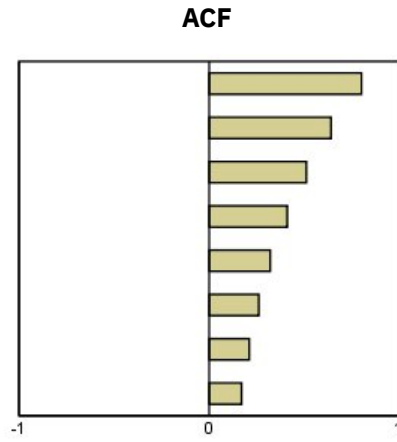


Tabella 7. ARIMA (0,1, 0) (serie integrata)



Informazioni particolari

Queste informazioni sono state sviluppate per prodotti e servizi offerti negli Stati Uniti. Questo materiale potrebbe essere disponibile da IBM in altre lingue. Tuttavia, può essere necessario possedere una copia del prodotto o versione del prodotto in tale lingua per potervi accedere.

IBM può non offrire i prodotti, i servizi o le funzioni presentati in questo documento in altri paesi. Consultare il proprio rappresentante locale IBM per informazioni sui prodotti ed i servizi attualmente disponibili nella propria zona. Qualsiasi riferimento ad un prodotto, programma o servizio IBM non implica o intende dichiarare che solo quel prodotto, programma o servizio IBM può essere utilizzato. Qualsiasi prodotto funzionalmente equivalente al prodotto, programma o servizio che non violi alcun diritto di proprietà intellettuale IBM può essere utilizzato. È comunque responsabilità dell'utente valutare e verificare la possibilità di utilizzare altri prodotti, programmi o servizi non IBM.

IBM può avere applicazioni di brevetti o brevetti in corso relativi all'argomento descritto in questo documento. La fornitura del presente documento non concede alcuna licenza a tali brevetti. È possibile inviare per iscritto richieste di licenze a:

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

Per richieste di licenze relative ad informazioni double-byte (DBCS), contattare il Dipartimento di Proprietà Intellettuale IBM nel proprio paese o inviare richieste per iscritto a:

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

IBM (INTERNATIONAL BUSINESS MACHINES CORPORATION) FORNISCE LA PRESENTE PUBBLICAZIONE "NELLO STATO IN CUI SI TROVA" SENZA GARANZIE DI ALCUN TIPO, ESPRESSE O IMPLICITE, IVI INCLUSE, A TITOLO DI ESEMPIO, GARANZIE IMPLICITE DI NON VIOLAZIONE, DI COMMERCIALIZZABILITÀ E DI IDONEITÀ PER UNO SCOPO PARTICOLARE. Alcune giurisdizioni non consentono la rinuncia ad alcune garanzie espresse o implicite in determinate transazioni, pertanto, la presente dichiarazione può non essere applicabile.

Queste informazioni potrebbero contenere imprecisioni tecniche o errori tipografici. Le modifiche alle presenti informazioni vengono effettuate periodicamente; tali modifiche saranno incorporate nelle nuove pubblicazioni della pubblicazione. IBM si riserva il diritto di apportare miglioramenti e/o modifiche al prodotto o al programma descritto nel manuale in qualsiasi momento e senza preavviso.

I riferimenti in queste informazioni a siti Web non IBM vengono forniti solo per comodità e non implicano in alcun modo l'approvazione di tali siti web. I materiali presenti in tali siti web non sono parte dei materiali per questo prodotto IBM e l'utilizzo di tali siti è a proprio rischio.

IBM può utilizzare o distribuire qualsiasi informazione fornita in qualsiasi modo ritenga appropriato senza incorrere in alcun obbligo verso l'utente.

Coloro che detengano la licenza su questo programma e desiderano avere informazioni su di esso allo scopo di consentire: (i) uno scambio di informazioni tra programmi indipendenti ed altri (compreso questo) e (ii) l'utilizzo reciproco di tali informazioni, dovrebbe rivolgersi a:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119*

Armonk, NY 10504-1785

US

Tali informazioni possono essere disponibili, in base ad appropriate clausole e condizioni, includendo in alcuni casi, il pagamento di una tassa.

Il programma concesso in licenza descritto nel presente documento e tutto il materiale concesso in licenza disponibile sono forniti da IBM in base alle clausole dell'Accordo per Clienti IBM (IBM Customer Agreement), dell'IBM IPLA (IBM International Program License Agreement) o qualsiasi altro accordo equivalente tra le parti.

Gli esempi client e dati delle prestazioni citati sono presentati esclusivamente a scopo illustrativo. Gli effettivi risultati delle prestazioni possono variare in base alle configurazioni e alle condizioni operative specifiche.

Le informazioni relative a prodotti non IBM sono ottenute dai fornitori di quei prodotti, dagli annunci pubblicati o da altre fonti disponibili al pubblico. IBM non ha testato quei prodotti e non può confermarne la precisione della prestazione, la compatibilità o qualsiasi altro reclamo relativo ai prodotti non IBM. Le domande sulle funzionalità dei prodotti non IBM devono essere indirizzate ai fornitori di tali prodotti.

Le dichiarazioni relative all'orientamento o alle intenzioni future di IBM sono soggette a modifica o a ritiro senza preavviso e rappresentano solo mete e obiettivi.

Questa pubblicazione contiene esempi di dati e prospetti utilizzati quotidianamente nelle operazioni aziendali. Pertanto, per maggiore completezza, gli esempi includono nomi di persone, società, marchi e prodotti. Tutti i nomi contenuti nel manuale sono fittizi e ogni riferimento a persone o aziende reali è puramente casuale.

LICENZA DI COPYRIGHT:

Queste informazioni contengono programmi applicativi di esempio in linguaggio sorgente, che illustrano tecniche di programmazione su varie piattaforme operative. È possibile copiare, modificare e distribuire questi programmi di esempio sotto qualsiasi forma senza alcun pagamento a IBM, allo scopo di sviluppare, utilizzare, commercializzare o distribuire i programmi applicativi in conformità alle API (application programming interface) a seconda della piattaforma operativa per cui i programmi di esempio sono stati scritti. Questi esempi non sono stati testati approfonditamente tenendo conto di tutte le condizioni possibili. IBM, quindi, non può garantire o sottintendere l'affidabilità, l'utilità o il funzionamento di questi programmi. I programmi di esempio sono forniti "NELLO STATO IN CUI SI TROVANO", senza alcun tipo di garanzia. IBM non intende essere responsabile per alcun danno derivante dall'uso dei programmi di esempio.

Ogni copia o qualsiasi parte di questi programmi di esempio o qualsiasi lavoro derivato, devono contenere le seguenti informazioni relative alle leggi sul diritto d'autore:

© Copyright IBM Corp. 2021. Le porzioni di questo codice derivano da IBM Corp. Programmi di esempio.

© Copyright IBM Corp. 1989 - 2021. Tutti i diritti riservati.

Marchi

IBM, il logo IBM e ibm.com sono marchi o marchi registrati di International Business Machines Corp., registrati in molte giurisdizioni in tutto il mondo. Altri nomi di prodotti e servizi possono essere marchi di IBM o di altre società. Un elenco corrente dei marchi IBM è disponibile sul web in "Copyright and trademark information" all'indirizzo www.ibm.com/legal/copytrade.shtml.

Adobe, il logo Adobe, PostScript e il logo PostScript sono marchi o marchi registrati di Adobe Systems Incorporated negli Stati Uniti e/o in altri paesi.

Intel, il logo Intel, Intel Inside, il logo Intel Inside, Intel Centrino, il logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium sono marchi o marchi registrati di Intel Corporation o delle sue consociate negli Stati Uniti e in altri paesi.

Linux è un marchio registrato di Linus Torvalds negli Stati Uniti e/o negli altri paesi.

Microsoft, Windows, Windows NT e il logo Windows sono marchi di Microsoft Corporation negli Stati Uniti e/o negli altri paesi.

UNIX è un marchio della The Open Group negli Stati Uniti e/o negli altri paesi.

Java e tutti i marchi e i logo basati su Java sono marchi o marchi registrati di Oracle e/o associate.

Indice analitico

A

- ACF
 - complotti per processi ARIMA puro [47](#)
 - in Applica serie temporali [15, 16](#)
 - in Time Series Modeler [10](#)
- analisi armonica [20](#)
- analisi serie storiche
 - modelli causali temporali [22](#)
- Applica modelli di serie storiche
 - funzione di autocorrelazione parziale residua [15, 16](#)
 - funzione di autocorrelazione residua [15, 16](#)
 - intervalli di confidenza [16, 18](#)
 - modelli di adattamento migliori e più poveri [17](#)
 - nuovi nomi variabili [18](#)
 - parametri del modello [15](#)
 - periodo di previsione [13](#)
 - periodo di stima [13](#)
 - previsioni [15, 16](#)
 - previsioni di risparmio [18](#)
 - ristima dei parametri del modello [13](#)
 - salvataggio dei modelli ristimati in XML [18](#)
 - Statistica di Box-Ljung [15](#)
 - statistiche su tutti i modelli [15, 16](#)
 - statistiche sulla bontà dell'adattamento. [15, 16](#)
 - Valori di adattamento [16](#)
 - valori mancanti [18](#)

B

- bontà di adattamento
 - definizioni [46](#)
 - in Applica serie temporali [15, 16](#)
 - in Time Series Modeler [10](#)

C

- Casi di holdout [2](#)

D

- dati cronologici
 - in Applica serie temporali [16](#)
 - in Time Series Modeler [10](#)
- Decomposizione stagionale
 - calcolo delle medie mobili [19](#)
 - Crea variabili [20](#)
 - Ipotesi [19](#)
 - modelli [19](#)
 - salvataggio di nuove variabili [20](#)

E

- Errore assoluto massimo
 - in Applica serie temporali [15, 16](#)
 - in Time Series Modeler [10](#)

- errore medio assoluto
 - in Applica serie temporali [15, 16](#)
 - in Time Series Modeler [10](#)
- Errore percentuale assoluto massimo
 - in Applica serie temporali [15, 16](#)
 - in Time Series Modeler [10](#)
- Errore percentuale assoluto medio
 - in Applica serie temporali [15, 16](#)
 - in Time Series Modeler [10](#)
- eventi
 - in Time Series Modeler [5](#)
- Expert Modeler
 - limitare lo spazio del modello [5](#)
 - valori anomali [6](#)

F

- fermo R2
 - in Applica serie temporali [15, 16](#)
 - in Time Series Modeler [10](#)
- funzione di autocorrelazione
 - complotti per processi ARIMA puro [47](#)
 - in Applica serie temporali [15, 16](#)
 - in Time Series Modeler [10](#)
- funzione di autocorrelazione parziale
 - complotti per processi ARIMA puro [47](#)
 - in Applica serie temporali [15, 16](#)
 - in Time Series Modeler [10](#)
- funzioni di trasferimento
 - ordini della differenza [8](#)
 - ordini denominatore [8](#)
 - ordini numeratore [8](#)
 - ordini stagionali [8](#)
 - ritardo [8](#)

G

- Grafici spettrali
 - analisi spettrale bivariata [20](#)
 - finestre spettrali [20](#)
 - Ipotesi [20](#)
 - trasformazione di centro [20](#)

I

- intervalli di confidenza
 - in Applica serie temporali [16, 18](#)
 - in Time Series Modeler [10, 12](#)

M

- MAE
 - in Applica serie temporali [15, 16](#)
 - in Time Series Modeler [10](#)
- MAPE
 - in Applica serie temporali [15, 16](#)

MAPE (*Continua*)
 in Time Series Modeler [10](#)

MaxAE
 in Applica serie temporali [15](#), [16](#)
 in Time Series Modeler [10](#)

MaxAPE
 in Applica serie temporali [15](#), [16](#)
 in Time Series Modeler [10](#)

Modeler Serie Storiche
 ARIMA [3](#), [7](#)
 eventi [5](#)
 Expert Modeler [3](#)
 funzione di autocorrelazione parziale residua [10](#)
 funzione di autocorrelazione residua [10](#)
 funzioni di trasferimento [8](#)
 intervalli di confidenza [10](#), [12](#)
 livellamento esponenziale [3](#), [6](#)
 modelli di adattamento migliori e più poveri [11](#)
 Nomi modello [12](#)
 nuovi nomi variabili [12](#)
 parametri del modello [10](#)
 periodicità [5–8](#)
 periodo di previsione [3](#), [12](#)
 periodo di stima [3](#)
 previsioni [10](#)
 previsioni di risparmio [12](#)
 salvataggio delle specifiche del modello in XML [12](#)
 Statistica di Box-Ljung [10](#)
 statistiche su tutti i modelli [10](#)
 statistiche sulla bontà dell'adattamento. [10](#)
 trasformazione della serie [6–8](#)
 valori anomali [6](#), [9](#)
 Valori di adattamento [10](#)
 valori mancanti [12](#)

modelli
 ARIMA [3](#)
 Expert Modeler [3](#)
 livellamento esponenziale [3](#), [6](#)

modelli ARIMA
 funzioni di trasferimento [8](#)
 valori anomali [9](#)

modelli causali temporali [22](#), [24–29](#), [31–33](#)

Modelli exponential smoothing [3](#), [6](#)

Modello di livellamento esponenziale degli inverni
 additivo [6](#)
 Moltiplicativo [6](#)

Modello di livellamento esponenziale di Brown [6](#)

Modello di livellamento esponenziale di Holt [6](#)

Modello di livellamento esponenziale semplice [6](#)

modello di livellamento esponenziale smorzato [6](#)

modello di livellamento esponenziale stagionale semplice [6](#)

N

nomi di variabili
 in Applica serie temporali [18](#)
 in Time Series Modeler [12](#)

Nomi modello
 in Time Series Modeler [12](#)

normalizzato BIC (criterio di informazione Bayesiana)
 in Applica serie temporali [15](#), [16](#)
 in Time Series Modeler [10](#)

O

outlier additivo
 in Time Series Modeler [6](#), [9](#)

outlier additivo stagionale
 in Time Series Modeler [6](#), [9](#)

outlier di tendenza locale
 in Time Series Modeler [6](#), [9](#)

outlier patch additivo
 in Time Series Modeler [6](#), [9](#)

outlier shift di livello
 in Time Series Modeler [6](#), [9](#)

P

PACF
 complotti per processi ARIMA puro [47](#)
 in Applica serie temporali [15](#), [16](#)
 in Time Series Modeler [10](#)

parametri del modello
 in Applica serie temporali [15](#)
 in Time Series Modeler [10](#)

periodicità
 in Time Series Modeler [5–8](#)

periodo di convalida [2](#)

periodo di previsione
 in Applica serie temporali [13](#)
 in Time Series Modeler [3](#), [12](#)

periodo di stima
 in Applica serie temporali [13](#)
 in Time Series Modeler [3](#)

periodo storico [2](#)

Previsione modello causale temporale [34–37](#), [39](#)

previsioni
 in Applica serie temporali [15](#), [16](#)
 in Time Series Modeler [10](#)

R

R2
 in Applica serie temporali [15](#), [16](#)
 in Time Series Modeler [10](#)

Radice errore quadratico medio
 in Applica serie temporali [15](#), [16](#)
 in Time Series Modeler [10](#)

residui
 in Applica serie temporali [15](#), [16](#)
 in Time Series Modeler [10](#)

ristima dei parametri del modello
 in Applica serie temporali [13](#)

RMSE
 in Applica serie temporali [15](#), [16](#)
 in Time Series Modeler [10](#)

S

salvare
 modelli rivalutati in XML [18](#)
 nuovi nomi variabili [12](#), [18](#)
 previsioni del modello [12](#), [18](#)
 specifiche del modello in XML [12](#)

scenari del modello causale temporale [40–45](#)

Statistica di Box-Ljung

Statistica di Box-Ljung (*Continua*)
in *Applica serie temporali* [15](#)
in *Time Series Modeler* [10](#)

T

trasformazione a radice quadrata
in *Time Series Modeler* [6–8](#)
trasformazione logaritmica
in *Time Series Modeler* [6–8](#)
trasformazione logaritmica naturale
in *Time Series Modeler* [6–8](#)

V

valore anomalo innovativo
in *Time Series Modeler* [6, 9](#)
valore anomalo transiente
in *Time Series Modeler* [6, 9](#)
valori anomali
definizioni [46](#)
Expert Modeler [6](#)
modelli ARIMA [9](#)
Valori di adattamento
in *Applica serie temporali* [16](#)
in *Time Series Modeler* [10](#)
valori mancanti
in *Applica serie temporali* [18](#)
in *Time Series Modeler* [12](#)

X

XML
salvataggio dei modelli ristimati in XML [18](#)
salvataggio modelli di serie temporali in XML [12](#)

