

# IBM SPSS Forecasting 21



*Nota:* Prima di utilizzare queste informazioni e il relativo prodotto, leggere le informazioni generali disponibili in Note a pag. 113.

Questa versione si applica a IBM® SPSS® Statistics 21 e a tutte le successive versioni e modifiche fino a eventuali disposizioni contrarie indicate in nuove versioni.

Le schermate dei prodotti Adobe sono state ristampate su autorizzazione di Adobe Systems Incorporated.

Le schermate dei prodotti Microsoft sono state ristampate su autorizzazione di Microsoft Corporation.

Materiali concessi in licenza - Proprietà di IBM

© **Copyright IBM Corporation 1989, 2012.**

Tutti i diritti riservati.

---

# Prefazione

IBM® SPSS® Statistics è un sistema completo per l'analisi dei dati. Il modulo aggiuntivo opzionale Forecasting include le tecniche di analisi aggiuntive descritte nel presente manuale. Il modulo aggiuntivo Forecasting deve essere usato con il modulo Core SPSS Statistics in cui è completamente integrato.

## **Informazioni su Business Analytics di IBM**

Il software IBM Business Analytics fornisce informazioni complete, coerenti e accurate a cui i responsabili delle decisioni possono affidarsi per ottimizzare le prestazioni dell'azienda. Un ampio portafoglio di applicazioni di [business intelligence](#), [analisi predittiva](#), [gestione delle prestazioni e delle strategie finanziarie](#) e [analisi](#) offre una panoramica chiara, istantanea e interattiva delle prestazioni attuali e la possibilità di prevedere i risultati futuri. Utilizzato in combinazione con potenti soluzioni di settore, prassi consolidate e servizi professionali, questo software consente alle aziende di tutte le dimensioni di ottimizzare la produttività, automatizzare le decisioni senza problemi e fornire risultati migliori.

Come parte di questo portafoglio, il software IBM SPSS Predictive Analytics consente alle aziende di prevedere gli eventi futuri e di agire tempestivamente in modo da migliorare i risultati delle attività aziendali. Le aziende, gli enti governativi e le università di tutto il mondo si affidano alla tecnologia IBM SPSS perché rappresenta un vantaggio concorrenziale in termini di attrazione, retention e aumento dei clienti, riducendo al tempo stesso le frodi e limitando i rischi. Incorporando il software IBM SPSS nelle attività quotidiane, le aziende diventano imprese in grado di effettuare previsioni e di gestire e automatizzare le decisioni, per raggiungere gli obiettivi aziendali e vantaggi tangibili sulla concorrenza. Per ulteriori informazioni o per contattare un rappresentante, visitare il sito <http://www.ibm.com/spss>.

## **Supporto tecnico**

Ai clienti che richiedono la manutenzione, viene messo a disposizione un servizio di supporto tecnico. I clienti possono contattare il supporto tecnico per richiedere assistenza per l'utilizzo dei prodotti IBM Corp. o per l'installazione di uno degli ambienti hardware supportati. Per contattare il supporto tecnico, visitare il sito Web di IBM Corp. all'indirizzo <http://www.ibm.com/support>. Ricordare che durante la richiesta di assistenza sarà necessario fornire i dati di identificazione personali, i dati relativi alla propria società e il numero del contratto di manutenzione.

## **Supporto tecnico per studenti**

Gli studenti che utilizzano una versione accademica o grad pack di qualsiasi prodotto software IBM SPSS sono pregati di utilizzare le apposite pagine online per studenti [Solutions for Education](#) (<http://www.ibm.com/spss/rd/students/>). Gli studenti che utilizzano una copia del software IBM SPSS fornita dall'università, sono pregati di contattare il coordinatore del prodotto IBM SPSS presso l'università.

## ***Servizio clienti***

Per eventuali chiarimenti in merito alla spedizione o al proprio conto, rivolgersi alla sede locale. Tenere presente che sarà necessario fornire il numero di serie.

## ***Corsi di formazione***

IBM Corp. organizza corsi di formazione pubblici e onsite che includono esercitazioni pratiche. Tali corsi si terranno periodicamente nelle principali città. Per ulteriori informazioni su questi seminari, andare a <http://www.ibm.com/software/analytics/spss/training>.

---

# Contenuto

## **Parte I: Manuale dell'utente**

### **1 Introduzione alle serie storiche 1**

Dati di serie storica . . . . .	1
Trasformazioni di dati . . . . .	2
Periodi di stima e di convalida . . . . .	2
Creazione di modelli e previsioni . . . . .	2

### **2 Modelli serie storiche 3**

Per specificare opzioni per Expert Modeler . . . . .	7
Selezione del modello e specificazione dell'evento . . . . .	7
Gestione dei valori anomali con Expert Modeler . . . . .	9
Modelli di livellamento esponenziale personalizzati . . . . .	10
Modelli ARIMA personalizzati . . . . .	11
Specifica del modello per i modelli ARIMA personalizzati . . . . .	12
Funzioni di trasferimento nei modelli ARIMA personalizzati . . . . .	14
Valori anomali nei modelli ARIMA personalizzati . . . . .	16
Output . . . . .	17
Tabelle delle statistiche e delle previsioni . . . . .	18
Grafici . . . . .	20
Limitazione dell'output ai modelli migliori o peggiori . . . . .	22
Salvataggio dei valori stimati e delle specifiche del modello . . . . .	23
Opzioni . . . . .	24
Opzioni aggiuntive del comando TSMODEL . . . . .	26

### **3 Applica modelli di serie storiche 27**

Output . . . . .	30
Tabelle delle statistiche e delle previsioni . . . . .	30
Grafici . . . . .	32
Limitazione dell'output ai modelli migliori o peggiori . . . . .	34
Salvataggio dei valori stimati e delle specifiche del modello . . . . .	35
Opzioni . . . . .	36
Opzioni aggiuntive del comando TSAPPLY . . . . .	37

<b>4</b>	<b><i>Decomposizione stagionale</i></b>	<b>38</b>
	Decomposizione stagionale: Salva . . . . .	39
	Opzioni aggiuntive del comando SEASON . . . . .	40
<b>5</b>	<b><i>Grafici spettrali</i></b>	<b>41</b>
	Opzioni aggiuntive del comando SPECTRA . . . . .	43
 <b>Parte II: Esempi</b>		
<b>6</b>	<b><i>Uso di Expert Modeler per le previsioni di massa</i></b>	<b>46</b>
	Esame dei dati . . . . .	46
	Esecuzione dell'analisi . . . . .	48
	Grafici riassuntivi dei modelli . . . . .	54
	Previsioni dei modelli . . . . .	55
	Riepilogo . . . . .	56
<b>7</b>	<b><i>Uso dei modelli salvati per l'esecuzione di nuove previsioni di massa</i></b>	<b>57</b>
	Esecuzione dell'analisi . . . . .	57
	Statistiche dell'adattamento dei modelli . . . . .	60
	Previsioni dei modelli . . . . .	61
	Riepilogo . . . . .	61
<b>8</b>	<b><i>Uso di Expert Modeler per la determinazione di predittori significativi</i></b>	<b>62</b>
	Inserimento dei dati nel grafico. . . . .	62
	Esecuzione dell'analisi . . . . .	64
	Grafico di una serie storica. . . . .	70

Tabella descrittiva del modello . . . . .	70
Tabella delle statistiche del modello . . . . .	71
Tabella dei parametri del modello ARIMA . . . . .	71
Riepilogo . . . . .	72
<b>9 <i>Usa dei modelli salvati per l'esecuzione di prove</i></b>	<b>73</b>
Ampliamento della serie dei predittori. . . . .	73
Modifica dei valori dei predittori nel periodo di previsione . . . . .	77
Esecuzione dell'analisi . . . . .	79
<b>10 <i>Decomposizione stagionale</i></b>	<b>84</b>
Rimozione della stagionalità dai dati di vendita . . . . .	84
Determinazione e impostazione della periodicità . . . . .	84
Esecuzione dell'analisi. . . . .	89
Output . . . . .	90
Riepilogo . . . . .	91
Procedure correlate . . . . .	91
<b>11 <i>Grafici spettrali</i></b>	<b>92</b>
Utilizzo dei grafici spettrali per verificare le aspettative relative alla periodicità . . . . .	92
Esecuzione dell'analisi. . . . .	92
Periodogramma e densità spettrale . . . . .	94
Riepilogo . . . . .	96
Procedure correlate . . . . .	96

## ***Appendici***

<b><i>A</i></b>	<b><i>Misure della bontà di adattamento</i></b>	<b>97</b>
<b><i>B</i></b>	<b><i>Tipi di valori anomali</i></b>	<b>98</b>
<b><i>C</i></b>	<b><i>Guida ai grafici ACF/PACF</i></b>	<b>99</b>
<b><i>D</i></b>	<b><i>File di esempio</i></b>	<b>103</b>
<b><i>E</i></b>	<b><i>Note</i></b>	<b>113</b>
	<b><i>Bibliografia</i></b>	<b>116</b>
	<b><i>Indice</i></b>	<b>117</b>



***Parte I:***  
***Manuale dell'utente***



# Introduzione alle serie storiche

Una **serie storica** è un insieme di osservazioni ottenute misurando regolarmente una variabile per un periodo di tempo specifico. In una serie riferita a dati d'inventario, ad esempio, le osservazioni possono rappresentare i livelli di inventario giornalieri riferiti a più mesi. Una serie riferita alle quote di mercato di un prodotto può essere costituita dalle quote di mercato settimanali relative a più anni. Una serie riferita alle vendite totali può essere costituita da un'osservazione al mese relativa a più anni. L'elemento comune di tutti questi esempi è che almeno una variabile viene osservata su intervalli noti e regolari nell'arco di un periodo di tempo specifico. Quindi, la forma dei dati di una serie storica tipica è rappresentata da una sequenza o da un elenco di osservazioni che rappresentano le misurazioni effettuate ad intervalli regolari.

Tabella 1-1  
*Serie storica giornaliera dei dati d'inventario*

Ora	Settimana	Giorno	Livello scorte
t <sub>1</sub>	1	Lunedì	160
t <sub>2</sub>	1	Martedì	135
t <sub>3</sub>	1	Mercoledì	129
t <sub>4</sub>	1	Giovedì	122
t <sub>5</sub>	1	Venerdì	108
t <sub>6</sub>	2	Lunedì	150
		...	
t <sub>60</sub>	12	Venerdì	120

Le analisi delle serie storiche vengono principalmente effettuate per prevedere i valori futuri di una serie. Un modello di serie che fornisce indicazioni sui valori passati può essere utile anche per prevedere se e quanto aumenteranno o diminuiranno i valori successivi. Poter fare previsioni di questo tipo è molto importante a livello commerciale e scientifico.

## Dati di serie storica

Ai fini della definizione di dati per le serie storiche da utilizzare con il modulo aggiuntivo Forecasting, ciascuna serie corrisponde ad una singola variabile. Per definire, ad esempio, una serie storica nell'Editor dei dati, fare clic sulla scheda *Visualizzazione variabili* ed immettere il nome della variabile in una riga vuota. Ciascuna osservazione di una serie storica corrisponde ad un caso (ossia ad una riga dell'Editor dei dati).

Nei fogli di calcolo contenenti dati di serie storiche, le serie sono generalmente contenute all'interno delle colonne del foglio di calcolo. Se si dispone di un foglio di calcolo con le serie storiche già disposte su righe, è possibile aprire ed utilizzare l'opzione *Trasponi* del menu *Dati* per invertire le righe in colonne.

## **Trasformazioni di dati**

Per le analisi delle serie storiche possono essere utilizzate numerose procedure di trasformazione dei dati del modulo Core.

- La procedura **Definisci date** (nel menu **Dati**) genera variabili di dati che possono essere utilizzate per stabilire la periodicità e distinguere tra periodi storici, di convalida e previsione. Il modulo **Forecasting** è in grado di utilizzare le variabili create dalla procedura **Definisci date**.
- La procedura **Crea serie storica** (nel menu **Trasforma**) consente di creare nuove variabili per le serie storiche sotto forma di funzioni delle variabili delle serie storiche esistenti. La procedura comprende funzioni vicine per il livellamento, la media e la differenza.
- La procedura **Sostituisci valori mancanti** (nel menu **Trasforma**) sostituisce i valori mancanti specificati dal sistema e dall'utente con stime basate su più metodi. I dati mancanti all'inizio o alla fine 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.

Per informazioni dettagliate sulla trasformazione dei dati per le serie storiche, vedere il *Manuale dell'utente del sistema Core*.

## **Periodi di stima e di convalida**

Spesso può essere utile dividere la serie storica in un periodo *di stima* o *storico* e in un periodo di *convalida*. In questo caso è necessario sviluppare un modello sulla base delle osservazioni del periodo di stima (storico) e verificarlo per vedere se è adatto o meno al periodo di convalida. Forzando il modello in modo da fare previsioni per punti già conosciuti (i punti del periodo di convalida), si può avere un'idea della capacità di previsione del modello.

I casi del periodo di convalida sono generalmente riferiti a casi di controllo, ovvero a casi non compresi nel processo di creazione del modello. Il periodo di stima comprende i casi correntemente selezionati nel file di dati attivo. Tutti i casi restanti, che seguono l'ultimo caso selezionato, possono essere usati come casi di controllo. Dopo aver determinato che il modello è utile per le previsioni, è possibile ridefinire il periodo di stima in modo che includa i casi di controllo, quindi creare il modello finale.

## **Creazione di modelli e previsioni**

Il modulo aggiuntivo **Previsioni** fornisce due procedure che possono essere utilizzate per creare modelli e previsioni.

- La procedura **Modelli serie storiche** crea modelli per le serie storiche e genera previsioni. La procedura comprende anche **Expert Modeler** che stabilisce automaticamente il modello migliore per ciascuna serie storica. Gli analisti esperti che desiderano un maggior grado di controllo, possono usare questo strumento anche per creare modelli personalizzati.
- La procedura **Applica modelli di serie storiche** applica i modelli delle serie storiche esistenti, creati da **Modelli serie storiche**, all'insieme di dati attivo. Ciò consente di effettuare previsioni per le serie che contengono dati nuovi o rivisti senza dover ricreare il modello. Se si sospetta che il modello possa essere cambiato, è possibile ricrearlo con **Time Series Modeler**.

## ***Modelli serie storiche***

La procedura Modelli di serie storiche consente di stimare i modelli di livellamento esponenziale, i modelli ARIMA (Autoregressive Integrated Moving Average) univariati e multivariati (o modelli della funzione di trasferimento) e di generare previsioni. La procedura include un Expert Modeler che identifica e stima automaticamente il modello ARIMA o il modello di livellamento esponenziale più adatto per una o più serie di variabili dipendenti ed eliminando quindi la necessità di identificare un modello appropriato tramite prove ed errori. In alternativa, è possibile specificare un modello ARIMA o un modello di livellamento esponenziale personalizzati.

**Esempio.** Un manager di prodotto responsabile delle previsioni per il mese successivo relative alle vendite unitarie e ai ricavi di 100 prodotti differenti, con un'esperienza limitata o addirittura assente nell'ambito dei modelli di serie storiche. I dati storici delle vendite unitarie di tutti i 100 prodotti sono archiviati in un singolo foglio di lavoro Excel. Dopo aver aperto il foglio di lavoro in IBM® SPSS® Statistics, utilizzare Expert Modeler e richiedere le previsioni del prossimo mese. Expert Modeler individua il modello di vendite unitarie migliore per ogni prodotto e utilizza questi due modelli per generare le previsioni. Poiché Expert Modeler è in grado di gestire più serie di input, è sufficiente eseguire la procedura una sola volta per ottenere le previsioni di tutti i prodotti. Scegliendo di salvare le previsioni nell'insieme di dati attivo, si possono esportare facilmente i risultati di nuovo in Excel.

**Statistiche.** Misure della bontà di adattamento:  $R$ -quadrato stazionario,  $R$ -quadrato ( $R^2$ ), radice della media dei quadrati (RMSE), errore medio assoluto (MAE), errore percentuale medio assoluto (MAPE), errore massimo assoluto (MaxAE), errore percentuale massimo assoluto (MaxAPE), criterio di informazione bayesiano normalizzato (BIC). Residui: funzione di autocorrelazione, funzione di autocorrelazione parziale,  $Q$  di Ljung-Box. Per i modelli ARIMA: ordini ARIMA per le variabili dipendenti, ordini della funzione di trasferimento per le variabili indipendenti e stime dei valori anomali. Inoltre il parametro di livellamento consente eseguire le stime per i modelli di livellamento esponenziale.

**Grafici.** Grafici riassuntivi in tutti i modelli: istogrammi di  $R$ -quadrato stazionario,  $R$ -quadrato ( $R^2$ ), radice della media dei quadrati (RMSE), errore medio assoluto (MAE), errore percentuale medio assoluto (MAPE), errore massimo assoluto (MaxAE), errore percentuale massimo assoluto (MaxAPE), criterio di informazione bayesiano normalizzato (BIC). Risultati dei modelli individuali: valori di previsione, valori adattati, valori osservati, limiti di confidenza superiore e inferiore, autocorrelazioni residue e autocorrelazioni parziali.

### ***Considerazioni sui dati di Modelli serie storiche***

**Dati.** Le variabili dipendenti ed indipendenti devono essere numeriche.

**Assunzioni.** La variabile dipendente e qualsiasi variabile indipendente vengono considerate come serie storiche, vale a dire che ogni caso rappresenta un punto temporale con casi successivi separati da un intervallo di tempo costante.

- **Stazionarietà.** Per i modelli ARIMA personalizzati, le serie storiche da utilizzare per i modelli devono essere stazionarie. Il metodo più efficace per trasformare una serie non stazionaria in una stazionaria consiste nell'utilizzare la trasformazione della differenza tramite la finestra di dialogo Crea serie storica.
- **Previsioni.** Per effettuare delle previsioni utilizzando modelli con variabili indipendenti (predittore), l'insieme di dati attivo deve contenere i valori di tali variabili per tutti i casi del periodo di previsione. Le variabili indipendenti inoltre, nel periodo di stima non devono contenere alcun valore mancante.

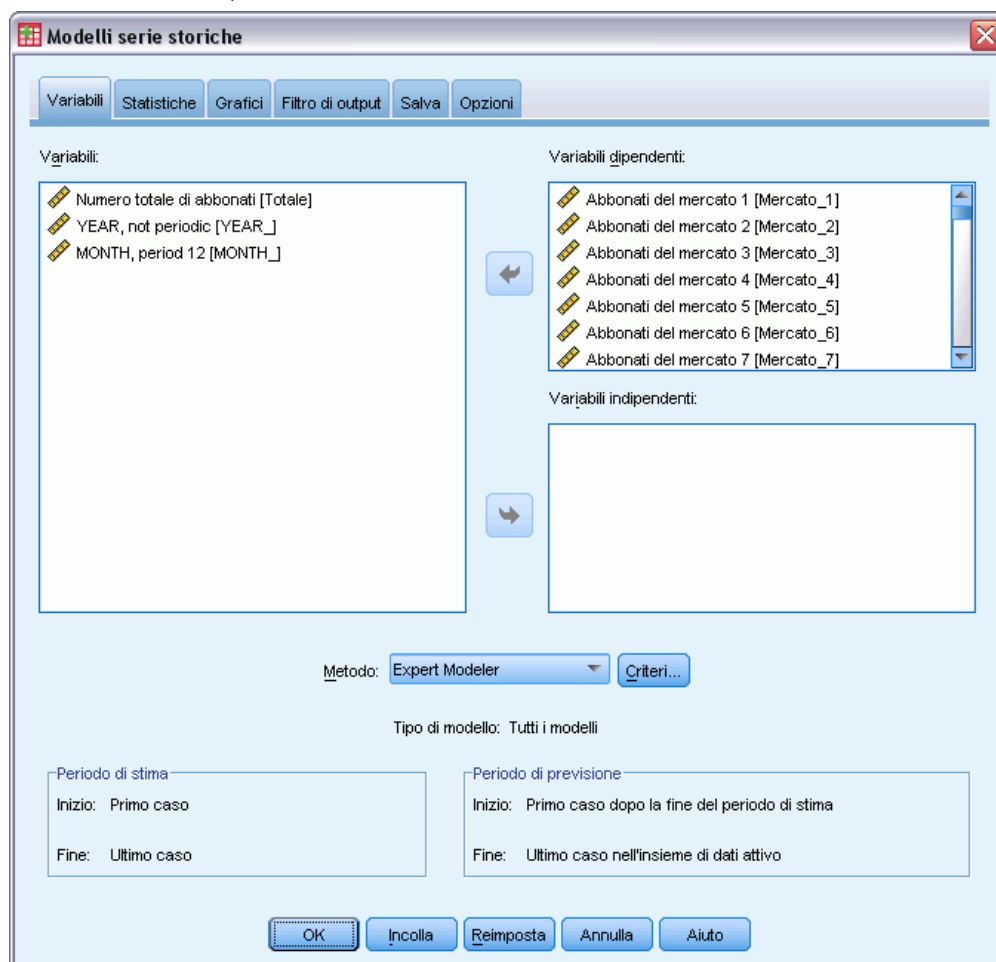
### ***Definizione delle date***

Sebbene non sia 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. Compiere tale operazione prima di utilizzare Modelli serie storiche e i risultati in un insieme di variabili che riportano una data associata a ogni caso. Tale operazione consente inoltre di impostare una periodicità presunta dei dati, ad esempio una periodicità di 12 se l'intervallo di tempo tra i casi successivi è un mese. La periodicità è necessaria se si desidera creare dei modelli stagionali. Se non si è interessati ai modelli stagionali e non si richiedono le etichette delle date nell'output, è possibile evitare la finestra di dialogo Definisci date. L'etichetta associata a ogni caso quindi, riporta semplicemente il numero di caso.

### ***Per utilizzare Modelli serie storiche***

- ▶ Dai menu, scegliere:  
Analizza > Previsioni > Crea modelli...

Figura 2-1  
Time Series Modeler, scheda Variabili



- ▶ Nella scheda Variabili, selezionare una o più variabili dipendenti da utilizzare per i modelli.
- ▶ Dall'elenco a discesa Metodo, selezionare un metodo di creazione dei modelli. Per la creazione automatica dei modelli, lasciare inalterato il metodo predefinito di Expert Modeler. In tal modo Expert Modeler si attiva per determinare il modello migliore per ogni variabile dipendente.

Per generare delle previsioni:

- ▶ Fare clic sulla scheda Opzioni.
- ▶ Specificare il periodo di previsione. Si genera un grafico che include le previsioni e i valori osservati.

Se lo si desidera, è possibile:

- Selezionare una o più variabili indipendenti. Nell'analisi di regressione le variabili indipendenti vengono trattate in modo molto simile alle variabili predittore, ma a differenza di queste ultime sono facoltative. È possibile includerle nei modelli ARIMA, ma non nei modelli

di livellamento esponenziale. Se si specifica Expert Modeler come metodo di creazione dei modelli e si includono le variabili indipendenti, vengono considerati solo i modelli ARIMA.

- Fare clic su Criteri per specificare i dettagli dei modelli.
- [Salvare le previsioni, gli intervalli di confidenza e i residui del rumore.](#)
- [Salvare i modelli stimati in formato XML.](#) È possibile applicare i modelli salvati a dei dati nuovi o rivisti per ottenere delle previsioni aggiornate senza ricreare i modelli. Per compiere tale operazione, utilizzare la procedura [Applica modelli di serie storiche.](#)
- [Ottenere statistiche riassuntive in tutti i modelli stimati.](#)
- [Nei modelli ARIMA personalizzati, specificare le funzioni di trasferimento delle variabili indipendenti.](#)
- [Attivare il rilevamento automatico dei valori anomali.](#)
- [Punti temporali specifici per modello come valori anomali per modelli ARIMA personalizzati.](#)

### ***Metodi di creazione dei modelli***

I metodi di creazione dei modelli disponibili sono:

**Expert Modeler.** Expert Modeler individua automaticamente il modello 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. Tuttavia, è possibile limitare la ricerca di Expert Modeler ai soli modelli ARIMA oppure ai soli 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 del trend e della stagionalità.

**ARIMA.** Utilizzare questa opzione per specificare un modello ARIMA personalizzato. Ciò comporta a specificare esplicitamente gli ordini autoregressivi e di media mobile, oltre che al grado di differenziazione. Si possono includere le variabili indipendenti (predittore) e definirne le funzioni di trasferimento. È inoltre possibile specificare il rilevamento automatico dei valori anomali o specificare un insieme di valori anomali esplicito.

### ***Periodi di stima e di previsione***

**Periodo di stima.** Il periodo di stima definisce l'insieme di casi utilizzato per determinare il modello. Per impostazione predefinita, il periodo di stima include tutti i casi dell'insieme di dati attivo. Per impostare il periodo di stima, selezionare Basato su intervallo di tempo o di casi nella finestra di dialogo Seleziona casi. In base ai dati disponibili, il periodo di stima utilizzato dalla procedura può variare a seconda della variabile dipendente, di conseguenza può differire dal valore visualizzato. Per una data variabile dipendente, il periodo di stima reale è il periodo rimasto dopo l'eliminazione di qualsiasi valore mancante consecutivo della variabile che si presenta all'inizio o alla fine del periodo di stima specificato.



**Periodo di previsione.** Il periodo di previsione ha inizio nel primo caso dopo il periodo di stima e per impostazione predefinita, si protrae fino all'ultimo caso dell'insieme di dati attivo. È possibile impostare la fine del periodo di previsione dalla scheda [Opzioni](#).

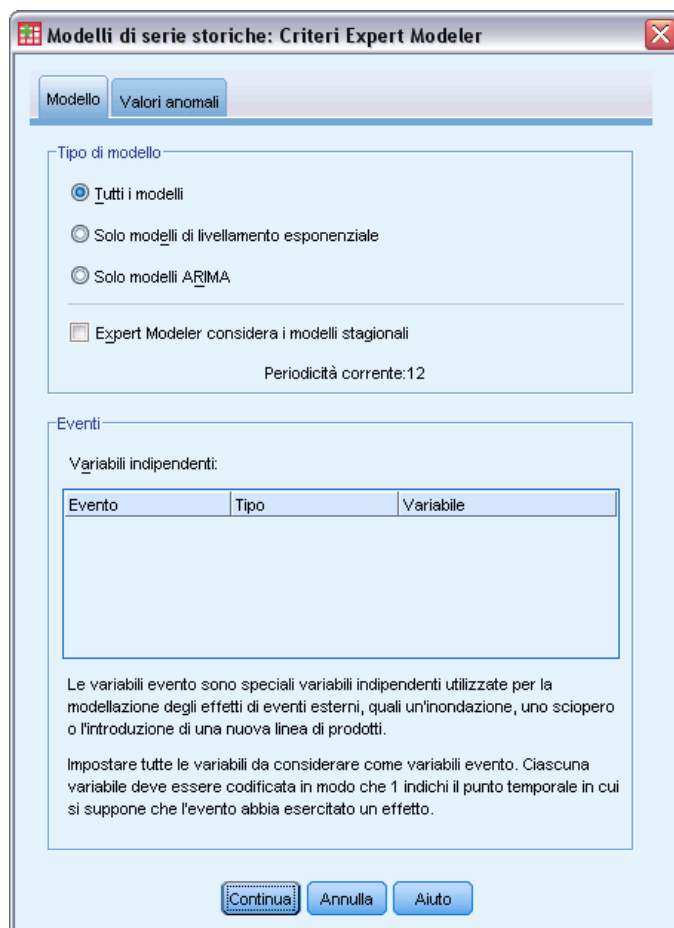
## Per specificare opzioni per Expert Modeler

Expert Modeler contiene le opzioni per imporre l'insieme di modelli candidati, specificare la gestione dei valori anomali e per includere le variabili evento.

## Selezione del modello e specificazione dell'evento

Figura 2-2

Finestra di dialogo Criteri di Expert Modeler, scheda Modello



La scheda Modello consente di specificare i tipi di modelli considerati da Expert Modeler 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 l'opzione è selezionata (contrassegnata), Expert Modeler considera entrambi i modelli stagionali e non stagionali. Se l'opzione non è selezionata, Expert Modeler considera solo i modelli non stagionali.

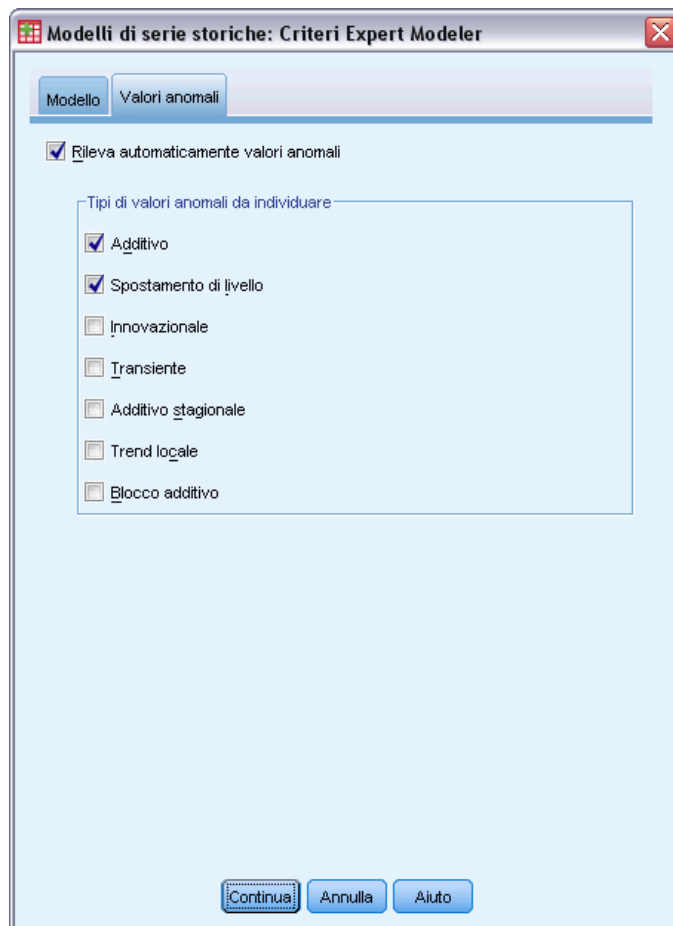
**Periodicità corrente.** Indica la periodicità corrente (se presente) definita per il file di dati attivo. La periodicità corrente viene data come numero intero, ad esempio 12 per la periodicità annuale in cui ogni caso rappresenta un mese. Se non è stata impostata alcuna periodicità, si visualizza il valore *Nessuna*. I modelli stagionali richiedono una periodicità. È possibile impostare la periodicità dalla finestra di dialogo Definisci date.

**Eventi.** Selezionare qualsiasi variabile indipendente da trattare come variabile evento. Per le variabili evento, i casi con un valore 1 indicano le volte in cui si prevede che le serie dipendenti vengano influenzate dall'evento. I valori diversi da 1 indicano nessun effetto.

## Gestione dei valori anomali con Expert Modeler

Figura 2-3

Finestra di dialogo Criteri di Expert Modeler, scheda Valori Anomali



La scheda Valori anomali consente di selezionare l'opzione di rilevamento automatico dei valori anomali e il tipo di valori anomali da rilevare.

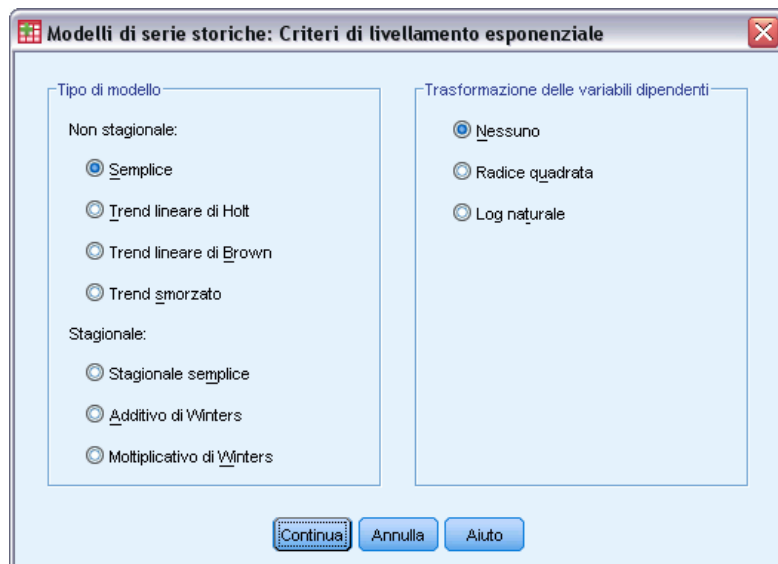
**Rileva automaticamente valori anomali.** Per impostazione predefinita, il rilevamento automatico dei valori anomali non viene eseguito. Selezionare (contrassegnare) l'opzione per eseguire il rilevamento automatico dei valori anomali, quindi selezionare uno o più dei seguenti tipi di valori anomali:

- Additivo
- Spostamento di livello
- Innovazionale
- Transiente
- Additivo stagionale
- Trend locale
- Blocco additivo

Per ulteriori informazioni, vedere l'argomento [Tipi di valori anomali](#) in l'appendice B a pag. 98.

## Modelli di livellamento esponenziale personalizzati

Figura 2-4  
Finestra di dialogo Criteri di livellamento esponenziale



**Tipo di modello.** I modelli di livellamento esponenziale (Gardner, 1985) vengono classificati come stagionali o non stagionali. I modelli stagionali sono disponibili solo se è stata definita una periodicità per l'insieme di dati dati attivo (vedere "Periodicità corrente" riportata sotto).

- **Semplice.** Questo modello è adatto per le serie in cui non sono presenti trend o stagionalità. Il suo unico parametro di livellamento è uniforme. Il livellamento esponenziale semplice è molto simile al modello ARIMA con nessun ordine di autoregressione, un ordine di differenze, un ordine di medie mobili e nessuna costante.
- **Modello di trend lineare di Holt.** Questo modello è adatto per le serie in cui è presente un trend lineare ma nessuna stagionalità. I parametri di livellamento sono uniformi e in trend, perché 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 differenze e due ordini di medie mobili.
- **Modello di trend lineare di Brown.** Questo modello è adatto per le serie in cui è presente un trend lineare ma nessuna stagionalità. I parametri di livellamento di questo modello sono uniformi, quindi il loro trend 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 perché non ha ordini di autoregressione, ha due ordini di differenze, due ordini di medie mobili e un coefficiente per il secondo ordine di medie mobili uguale al quadrato di metà del coefficiente del primo ordine.
- **Trend smorzato.** Questo modello è indicato per le serie con un trend lineare che tende a smorzarsi e che non presentano stagionalità. I parametri di livellamento del modello sono rappresentati da un trend uniforme e smorzato. 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 indicato per le serie senza trend ed effetti stagionali costanti 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 nessun ordine di autoregressione, un ordine di differenze, un ordine di differenze stagionali, e ordini 1,  $p$  e  $p + 1$  di medie mobile, dove  $p$  è il numero di periodi nell'intervallo stagionale (per i dati mensili  $p = 12$ ).
- **Additivo di Winters.** Questo modello è indicato per le serie con un trend lineare e un effetto stagionale che non dipende dal livello della serie. I parametri di livellamento di questo modello sono uniformi, in trend e stagionali. Il livellamento esponenziale dell'additivo di Winters è molto simile al modello ARIMA con nessun ordine di autoregressione, un ordine di differenze, un ordine di differenze stagionali, e ordini  $p + 1$  di medie mobile, dove  $p$  è il numero di periodi nell'intervallo stagionale (per i dati mensili  $p = 12$ ).
- **Moltiplicativo di Winters.** Questo modello è indicato per le serie con un trend lineare e un effetto stagionale che dipende dal livello della serie. I parametri di livellamento di questo modello sono uniformi, in trend e stagionali. Il livello esponenziale moltiplicativo di Winters non è simile a nessun modello ARIMA.

**Periodicità corrente.** Indica la periodicità corrente (se presente) definita per il file di dati attivo. La periodicità corrente viene data come numero intero, ad esempio 12 per la periodicità annuale in cui ogni caso rappresenta un mese. Se non è stata impostata alcuna periodicità, si visualizza il valore *Nessuna*. I modelli stagionali richiedono una periodicità. È possibile impostare la periodicità dalla finestra di dialogo Definisci date.

**Trasformazione delle variabili dipendenti.** È possibile specificare una trasformazione eseguita su ogni variabile dipendente prima utilizzarla per i modelli.

- **Nessuna.** Non viene eseguita nessuna trasformazione.
- **Radice quadrata.** Trasformazione a radice quadrata.
- **Log naturale.** Trasformazione logaritmica naturale.

## **Modelli ARIMA personalizzati**

La procedura Modelli di serie storiche consente di creare dei modelli ARIMA (Autoregressive Integrated Moving Average) non stagionali o stagionali, noti anche come modelli Box-Jenkins (Scatola, Jenkins, e Reinsel, 1994), con o senza un insieme stazionario di variabili predittore. È possibile definire le funzioni di trasferimento per qualsiasi variabile predittore e specificare il rilevamento automatico dei valori anomali o un insieme specifico di valori anomali.

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

## Specifica del modello per i modelli ARIMA personalizzati

Figura 2-5  
Finestra di dialogo Criteri ARIMA, scheda Modello

Modelli di serie storiche: Criteri ARIMA

Modello Valori anomali

Ordini ARIMA

Struttura:

	Non stagionale	Stagionale
Autoregressivo (p)	0	0
Differenza (d)	0	0
Media mobile (q)	0	0

Periodicità corrente: 12

Trasformazione

Nessuno  
 Radice quadrata  
 Log naturale

Includi costante nel modello

Continua Annulla Aiuto

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. Se per esempio si specifica 2, il modello include gli ordini 2 e 1. Le celle nella colonna Stagionale vengono attivate solo se è stata definita una periodicità per il file di dati attivo (Vedere "Periodicità corrente" riportata sotto).

- **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 dei trend (di norma, le serie che presentano dei trend 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 trend della serie, la differenziazione di primo grado tiene conto dei trend lineari, la differenziazione di secondo grado tiene conto dei trend quadratici 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 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 corrente (se presente) definita per il file di dati attivo. La periodicità corrente viene data come numero intero, ad esempio 12 per la periodicità annuale in cui ogni caso rappresenta un mese. Se non è stata impostata alcuna periodicità, si visualizza il valore *Nessuna*. I modelli stagionali richiedono una periodicità. È possibile impostare la periodicità dalla finestra di dialogo Definisci date.

**Trasformazione delle variabili dipendenti.** È possibile specificare una trasformazione eseguita su ogni variabile dipendente prima utilizzarla per i modelli.

- **Nessuna.** Non viene eseguita nessuna trasformazione.
- **Radice quadrata.** Trasformazione a radice quadrata.
- **Log naturale.** Trasformazione logaritmica naturale.

**Includi costante nel modello.** Il processo di inclusione di una costante è standard a meno che non si abbia la certezza che il valore generale medio della serie sia 0. Quando si applica la differenziazione, si consiglia di escludere la costante.

## Funzioni di trasferimento nei modelli ARIMA personalizzati

Figura 2-6  
Finestra di dialogo Criteri ARIMA, scheda Funzione di trasferimento

The screenshot shows a software dialog box titled "Modelli di serie storiche: Criteri ARIMA". It has three tabs: "Modello", "Funzione di trasferimento" (selected), and "Valori anomali".

Under "Variabili indipendenti", there is a list box containing "Numero cataloghi sped...".

The "Ordini della funzione di trasferimento" section contains a table:

Struttura:	Non stagionale	Stagionale
Numeratore	0	0
Denominatore	0	0
Differenza	0	0

Below the table, it says "Periodicità corrente: Nessuna".

There is a "Ritardo:" field with the value "0".

The "Trasformazione" section has three radio buttons: "Nessuno" (selected), "Radice quadrata", and "Log naturale".

At the bottom, there are three buttons: "Continua", "Annulla", and "Aiuto".

La scheda Funzione di trasferimento (presente solo se sono specificate le variabili indipendenti) consente di definire le funzioni di trasferimento delle variabili indipendenti specificate nella 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. Se per esempio si specifica 2 per il numeratore, il modello include gli ordini 2, 1 e 0. Se si specifica 3 per il denominatore, il modello include gli ordini 3, 2, e 1. Le celle nella colonna Stagionale vengono attivate solo se è stata definita una periodicità per l'insieme di dati attivo (vedere "Periodicità corrente" riportata sotto).

- Numeratore.** L'ordine numeratore della funzione di trasferimento. Specifica quali valori precedenti provenienti dalla serie indipendente (predittore) specificata 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 dei trend 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 corrente (se presente) definita per il file di dati attivo. La periodicità corrente viene data come numero intero, ad esempio 12 per la periodicità annuale in cui ogni caso rappresenta un mese. Se non è stata impostata alcuna periodicità, si visualizza il valore *Nessuna*. I modelli stagionali richiedono una periodicità. È possibile impostare la periodicità dalla finestra di dialogo Definisci date.

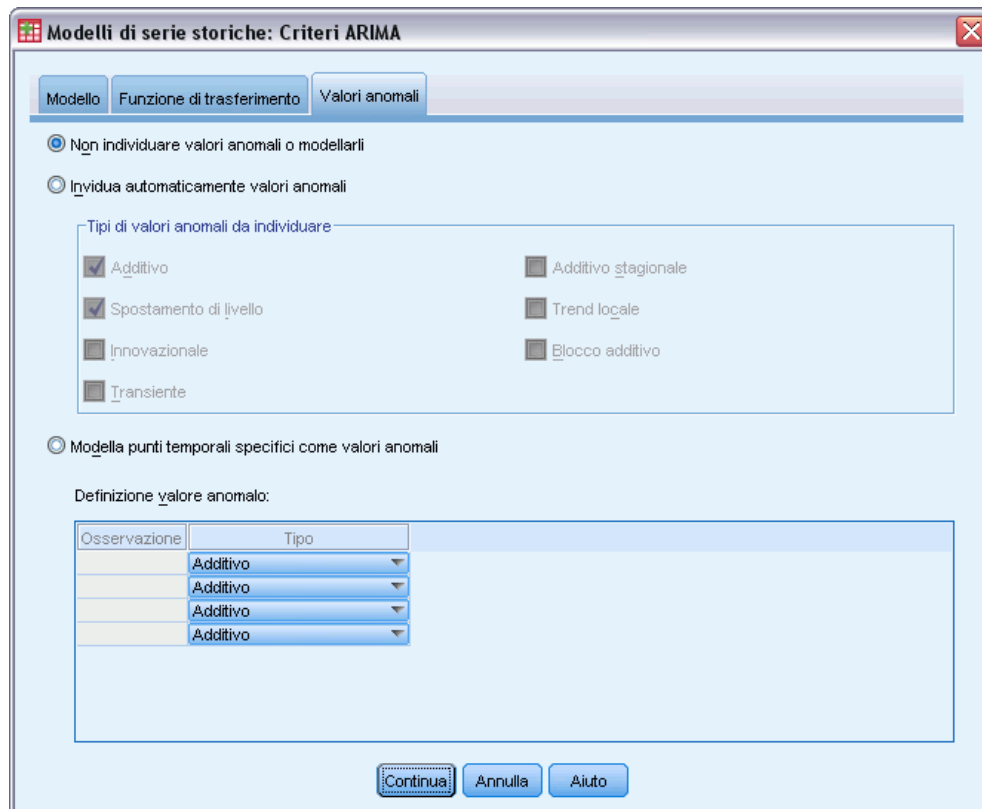
**Ritardo.** L'impostazione di un ritardo induce a ritardare l'influenza della variabile indipendente per il numero di intervalli specificato. Se per esempio il ritardo è impostato a 5, il valore della variabile indipendente all'ora  $t$  non incide sulle previsioni fino a quando non sono trascorsi 5 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.

- **Nessuna.** Non viene eseguita nessuna trasformazione.
- **Radice quadrata.** Trasformazione a radice quadrata.
- **Log naturale.** Trasformazione logaritmica naturale.

## Valori anomali nei modelli ARIMA personalizzati

Figura 2-7  
Finestra di dialogo Criteri ARIMA, scheda Valori anomali



La scheda Valori anomali contiene le seguenti opzioni per la gestione dei valori anomali (Pena, Tiao, e Tsay, 2001): rilevamento automatico, specificazione di punti particolari come valori anomali oppure nessun rilevamento o modello.

**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
- Innovazionale
- Transiente
- Additivo stagionale
- Trend locale
- Blocco additivo

Per ulteriori informazioni, vedere l'argomento [Tipi di valori anomali in l'appendice B a pag. 98](#).

**Modella punti temporali specifici come valori anomali.** Selezionare questa opzione per specificare dei punti temporali particolari come valori anomali. Per ogni valore anomalo utilizzare una riga diversa della griglia Definizione valore anomalo. Immettere i valori per tutte le celle di una data riga.

- **Tipo.** Il tipo di valore anomalo. I tipi supportati sono: additivo (default), spostamento di livello, innovazionale, transiente, additivo stagionale e trend locale.

*Nota 1:* se non è stata definita alcuna specificazione della data per l'insieme di dati attivo, la griglia Definizione valore anomalo mostra unicamente la colonna *Osservazione*. Per specificare un valore anomalo, immettere il numero di riga (come visualizzato nell'Editor dei dati) del caso relativo.

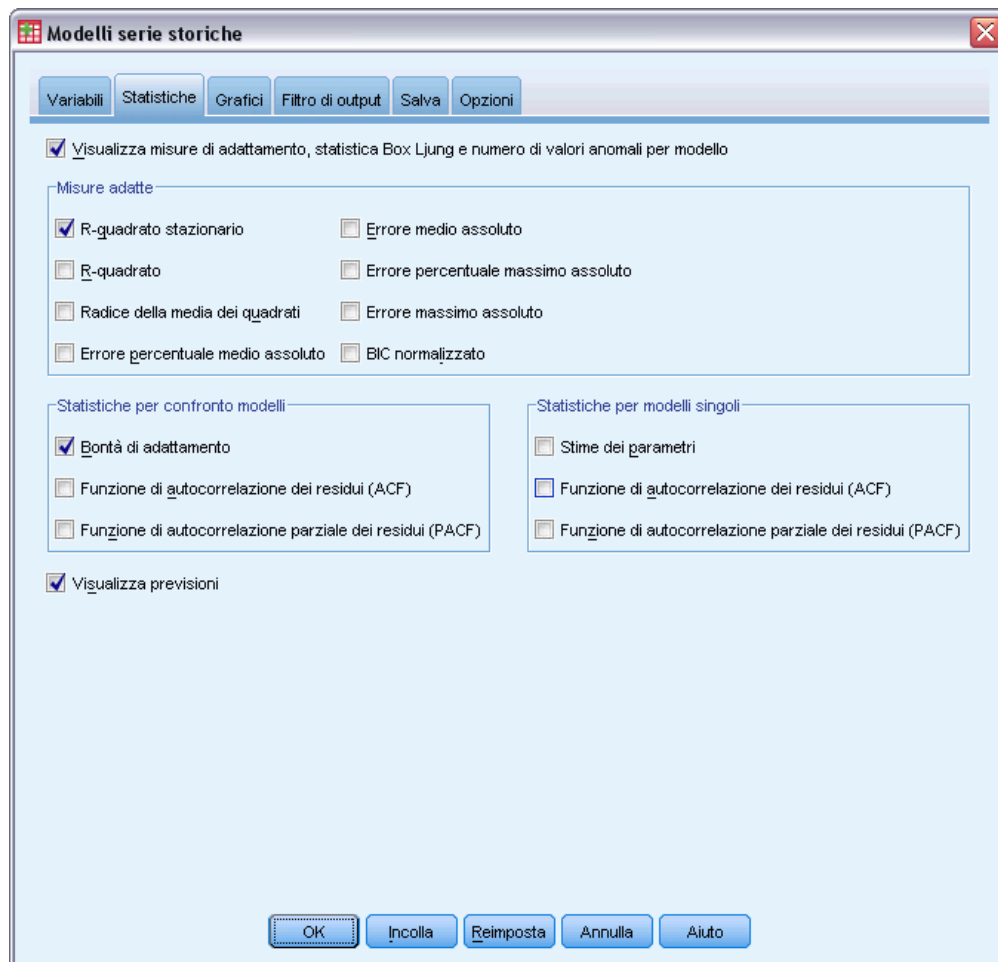
*Nota 2:* la colonna *Ciclo* (se presente) della griglia Definizione valore anomalo si riferisce al valore della variabile *CYCLE\_* dell'insieme di dati attivo.

## **Output**

Output disponibili include i risultati dei modelli individuali e quelli calcolati in tutti i modelli. È possibile limitare i risultati dei modelli individuali a un insieme di modelli migliori o peggiori, in base ai criteri specificati dall'utente.

## Tabelle delle statistiche e delle previsioni

Figura 2-8  
Time Series Modeler, scheda Statistiche



La scheda Statistiche contiene le opzioni per la visualizzazione delle tabelle dei risultati dei modelli.

### **Visualizza misure di adattamento, statistica Box Ljung e numero di valori anomali per modello.**

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

**Misure di adattamento.** È possibile selezionare una o più delle seguenti opzioni, per includerle nella tabella che contiene le misure di adattamento di ogni modello stimato:

- *R*-quadrato stazionario
- *R*-quadrato
- Radice della media dei quadrati
- Errore percentuale medio assoluto

- Errore medio assoluto
- Errore percentuale massimo assoluto
- Errore massimo assoluto
- BIC normalizzato

Per ulteriori informazioni, vedere l'argomento [Misure della bontà di adattamento in l'appendice A a pag. 97.](#)

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

- **Bontà di adattamento.** Tabella di statistiche riassuntive e di percentili per  $R$ -quadrato stazionario,  $R$ -quadrato, radice della media dei quadrati, errore percentuale medio assoluto, errore medio assoluto, errore percentuale massimo assoluto, errore massimo assoluto e Criterio di informazione bayesiano normalizzato.
- **Funzione di autocorrelazione dei residui (ACF).** Tabella di statistiche riassuntive e di percentili per le autocorrelazioni dei residui in tutti i modelli stimati.
- **Funzione di autocorrelazione parziale dei residui (PACF).** Tabella di statistiche riassuntive e di percentili per le autocorrelazioni parziali dei residui in tutti i modelli stimati.

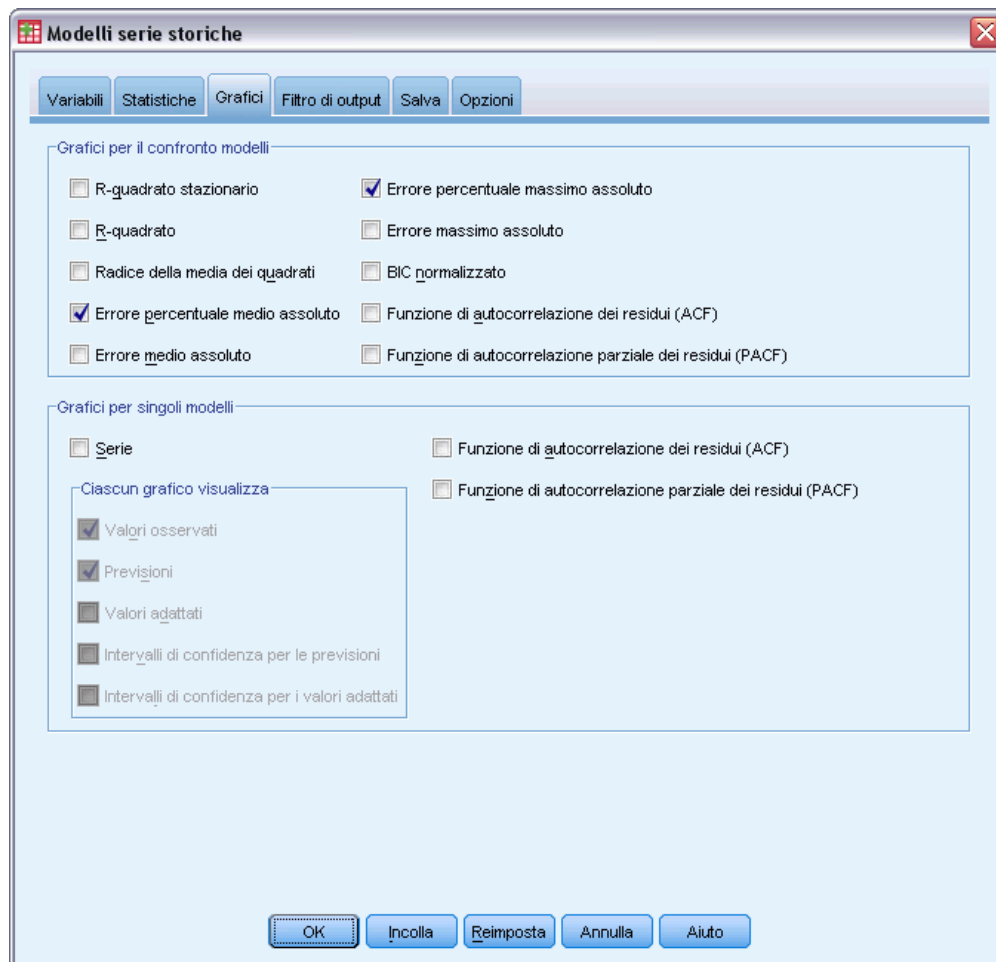
**Statistiche modelli singoli.** Questo gruppo di opzioni controlla la visualizzazione delle tabelle contenenti le informazioni dettagliate di ogni modello stimato. Ogni opzione genera una tabella distinta. È possibile selezionare una o più delle seguenti opzioni:

- **Stime dei parametri.** Visualizza una tabella di stime dei parametri di ogni modello stimato. Per i modelli di livellamento esponenziale e ARIMA vengono visualizzate delle tabelle distinte. Se sono presenti dei valori anomali, anche le stime dei parametri ad essi relative vengono visualizzate in una tabella distinta.
- **Funzione di autocorrelazione dei residui (ACF).** Visualizza una tabella di autocorrelazioni dei residui per ritardo di ogni modello stimato. La tabella include gli intervalli di confidenza delle autocorrelazioni.
- **Funzione di autocorrelazione parziale dei residui (PACF).** Visualizza una tabella di autocorrelazioni parziali dei residui per ritardo di ogni modello stimato. La tabella include gli intervalli di confidenza delle autocorrelazioni parziali.

**Visualizza previsioni.** Visualizza una tabella di previsioni dei modelli e di intervalli di confidenza di ogni modello stimato. Dalla scheda Opzioni è possibile impostare il periodo di previsione.

## Grafici

Figura 2-9  
Time Series Modeler, scheda Grafici



La scheda Grafici contiene le opzioni per la visualizzazione dei grafici dei risultati dei modelli.

### **Grafici per il confronto modelli**

Questo gruppo di opzioni controlla la visualizzazione dei grafici contenenti le statistiche calcolate in tutti i modelli stimati. Ogni opzione genera un grafico distinto. È possibile selezionare una o più delle seguenti opzioni:

- R-quadrato stazionario
- R-quadrato
- Radice della media dei quadrati
- Errore percentuale medio assoluto
- Errore medio assoluto
- Errore percentuale massimo assoluto

- Errore massimo assoluto
- BIC normalizzato
- Funzione di autocorrelazione dei residui (ACF)
- Funzione di autocorrelazione parziale dei residui (PACF)

Per ulteriori informazioni, vedere l'argomento [Misure della bontà di adattamento in l'appendice A a pag. 97.](#)

### ***Grafici per singoli modelli***

**Serie.** Selezionare (contrassegnare) questa opzione per ottenere i grafici dei valori attesi per ogni modello stimato. È possibile selezionare una o più delle opzioni seguenti da includere nel grafico:

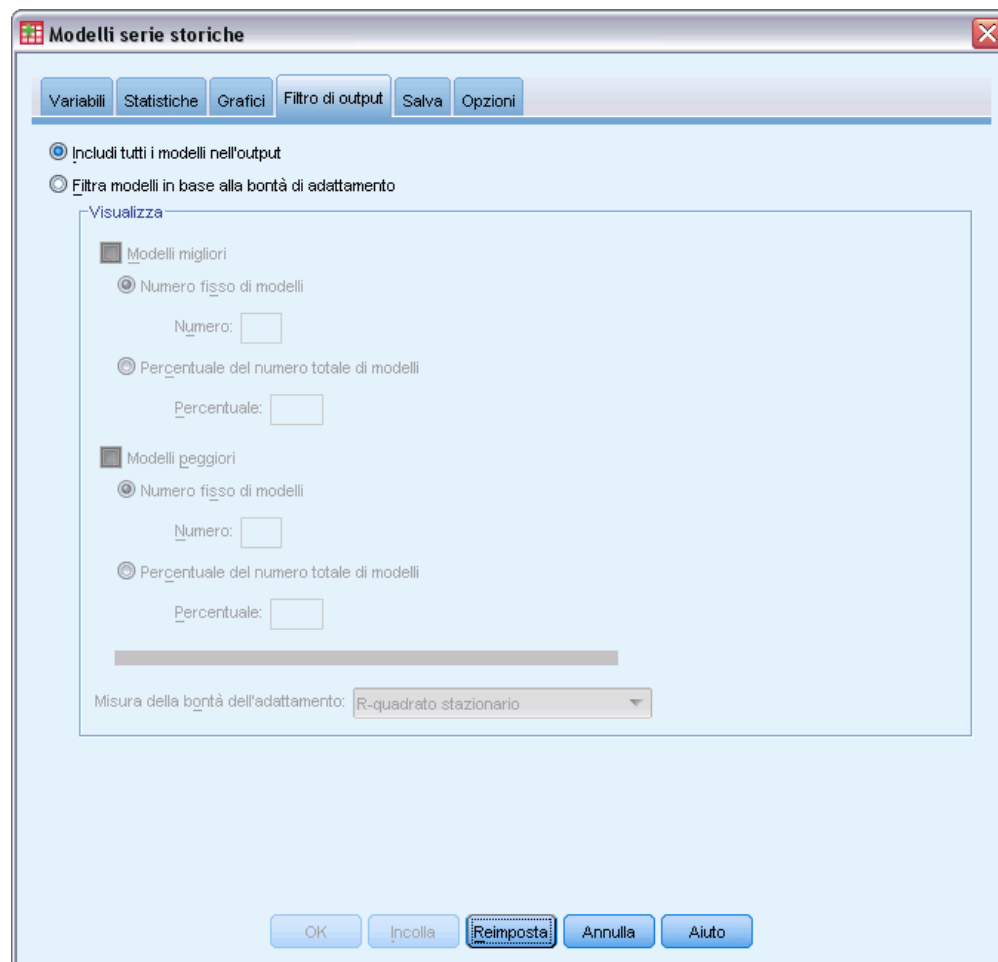
- **Valori osservati.** I valori osservati della serie dipendente.
- **Previsioni.** I valori attesi del modello per il periodo di previsione.
- **Valori adattati.** I valori attesi del modello per il periodo di stima.
- **Intervalli di confidenza per le previsioni.** Gli intervalli di confidenza per il periodo di previsione.
- **Intervalli di confidenza per i valori adattati.** Gli intervalli di confidenza per il periodo di stima.

**Funzione di autocorrelazione dei residui (ACF).** Visualizza un grafico di autocorrelazioni dei residui di ogni modello stimato.

**Funzione di autocorrelazione parziale dei residui (PACF).** Visualizza un grafico di autocorrelazioni parziali dei residui di ogni modello stimato.

## Limitazione dell'output ai modelli migliori o peggiori

Figura 2-10  
Modelli serie storiche, scheda Filtro di output



La scheda Filtro di output contiene le opzioni per limitare sia l'output della tabella, sia quello del grafico a un sottoinsieme di modelli stimati. È possibile scegliere di limitare l'output al modello migliore e/o peggiore, a seconda dei criteri di adattamento forniti. Per impostazione predefinita, tutti i modelli stimati vengono inclusi nell'output.

**Modelli migliori.** Selezionare (contrassegnare) questa opzione per includere nell'output i modelli migliori. Selezionare una misura della bontà di adattamento e specificare il numero di modelli da includere. La selezione dell'opzione non preclude anche la selezione dei modelli peggiori. In questo caso, l'output è costituito sia dai modelli peggiori, sia da quelli migliori.

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



**Modelli peggiori.** Selezionare (contrassegnare) questa opzione per includere nell'output i modelli peggiori. Selezionare una misura della bontà di adattamento e specificare il numero di modelli da includere. La selezione dell'opzione non preclude anche la selezione dei modelli migliori. In questo caso, l'output è costituito sia dai modelli migliori, sia da quelli peggiori.

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

**Misura della bontà di adattamento.** Selezionare la misura della bontà di adattamento da utilizzare per i modelli di filtro. Il default è  $R$ -quadrato stazionario.

## ***Salvataggio dei valori stimati e delle specifiche del modello***

La scheda Salva consente di salvare i valori stimati del modello come variabili nuove nell'insieme di dati attivo e di salvare le specifiche del modello in un file esterno in formato XML.

**Salva variabili.** Nell'insieme di dati attivo è possibile salvare i valori stimati del modello, gli intervalli di confidenza e i residui come variabili nuove. Ogni serie dipendente genera un insieme proprio di variabili nuove e ognuna di esse contiene i valori per i periodi di stima e di previsione. Se il periodo di previsione supera la serie di variabili dipendenti, si aggiungono i casi nuovi. Scegliere di salvare le variabili nuove selezionando per ognuna di esse la casella di controllo Salva associata. Per impostazione predefinita, nessuna variabile nuova è salvata.

- **Valori attesi.** I valori attesi del modello.
- **Limiti di confidenza inferiori.** I limiti di confidenza inferiori dei valori attesi.
- **Limiti di confidenza superiori.** I limiti di confidenza superiori dei valori attesi.
- **Residui del rumore.** I residui del modello. Quando si eseguono le trasformazioni della variabile dipendente (ad esempio, logaritmica naturale), questi sono i residui della serie trasformata.
- **Prefisso del nome variabile.** Consente di specificare i prefissi da utilizzare per i nomi delle nuove variabili oppure di mantenere i prefissi predefiniti. I nomi delle variabili sono costituiti dal prefisso, dal nome della variabile dipendente associata e da un identificatore di modello. Per evitare eventuali conflitti con i nomi delle variabili, è possibile estendere il nome della variabile. Il prefisso deve essere conforme alle regole stabilite per i nomi di variabile validi.

**Esporta file del modello.** Le specifiche del modello di tutti i modelli stimati vengono esportate nel file specificato in formato XML. Per ottenere delle previsioni aggiornate, basate su più dati correnti, si possono utilizzare i modelli salvati tramite la procedura [Applica modelli di serie storiche](#).

- **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 conforme a PMML che può essere utilizzato con applicazioni conformi a PMML, tra cui le applicazioni IBM SPSS.

## Opzioni

Figura 2-11  
Time Series Modeler, scheda Opzioni

Modelli serie storiche

Variabili Statistiche Grafici Filtro di output Salva Opzioni

Periodo di previsione

Primo caso tra la fine del periodo di stima e l'ultimo caso dell'insieme di dati attivo

Primo caso tra la fine del periodo di stima e la data specificata

Data:

Anno	Mese
2004	3

Valori mancanti definiti dall'utente

Considera non valido

Considera valido

Larghezza intervallo di confidenza (%): 95

Prefisso per identificatori di modello nell'output: Modello

Numero massimo di ritardi mostrati nell'output ACF e PACF: 24

OK Incolla Reimposta Annulla Aiuto

La scheda Opzioni consente di impostare il periodo di previsione, specificare la gestione dei valori mancanti, impostare l'ampiezza dell'intervallo di confidenza, specificare un prefisso personalizzato per gli identificatori di modelli e di impostare il numero di ritardi 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 di cause utilizzato per determinare il modello) e si protrae fino all'ultimo caso dell'insieme di dati attivo o fino a una data specificata dall'utente. Per impostazione predefinita, la fine del periodo di stima è l'ultimo caso dell'insieme di dati attivo, ma è possibile modificarlo dalla finestra di dialogo Seleziona casi selezionando Basato su intervallo di tempo o di casi.

- **Primo caso dopo la fine del periodo di stima fino all'ultimo caso dell'insieme di dati attivo.** Selezionare questa opzione quando la fine del periodo di stima è precedente all'ultimo caso dell'insieme di dati attivo e si desidera effettuare delle previsioni fino all'ultimo caso.

Di norma si utilizza questa opzione per creare le previsioni per un periodo di controllo, consentendo il confronto delle previsioni del modello con un sottoinsieme di valori effettivi.

- **Primo caso dopo la fine del periodo di stima fino a una data specificata.** Selezionare questa opzione per specificare esplicitamente la fine del periodo di previsione. Di norma si utilizza questa opzione per creare le previsioni che vanno oltre la fine della serie effettiva. Immettere i valori per tutte le celle della griglia Data.

Se non è stata definita alcuna specificazione della data per l'insieme di dati attivo, la griglia Data mostra unicamente la colonna *Osservazione*. Per specificare la fine del periodo di previsione, immettere il numero di riga (come visualizzato nell'Editor dei dati) del caso relativo.

La colonna *Ciclo* (se presente) della griglia Data si riferisce al valore della variabile *CYCLE\_* dell'insieme di dati attivo.

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

- **Considera come non valido.** I valori mancanti definiti dall'utente sono considerati come mancanti di sistema.
- **Considera come valido.** I valori mancanti definiti dall'utente vengono considerati come dati validi.

**Politica dei valori mancanti.** Le seguenti regole vengono applicate al trattamento dei valori mancanti (include i valori mancanti di sistema e i valori mancanti definiti dall'utente considerati non validi) durante la procedura di creazione dei modelli:

- I casi che presentano i valori mancanti di una variabile dipendente che si verificano nel periodo di stima vengono inclusi nel modello. La gestione specifica del valore mancante dipende dal metodo di stima.
- Se nel periodo di stima una variabile indipendente presenta dei valori mancanti, viene generato un avviso. Per Expert Modeler, i modelli relativi alla variabile indipendente vengono stimati senza variabile. Per ARIMA personalizzati, i modelli relativi alla variabile indipendente non vengono stimati.
- Se una variabile indipendente presenta dei valori mancanti nel periodo di previsione, la procedura genera un avviso ed effettua le previsioni per il periodo consentito.

**Ampiezza dell'intervallo di confidenza (%).** Gli intervalli di confidenza vengono calcolati per i valori stimati del modello e le autocorrelazioni dei residui. È possibile specificare qualsiasi valore positivo inferiore a 100. Per impostazione predefinita, si utilizza un intervallo di confidenza al 95%.

**Prefisso per identificatori modello nell'output.** Ogni variabile dipendente specificata nella scheda Variabili genera un modello stimato distinto. I modelli sono contraddistinti da nomi unici costituiti da un prefisso personalizzabile e da un suffisso rappresentato da un intero. È possibile immettere un prefisso o lasciare inalterato il default del *Modello*.

**Numero massimo di ritardi mostrato nell'output ACF e PACF.** È possibile impostare il numero massimo di ritardi mostrato nelle tabelle e nei grafici delle autocorrelazioni e delle autocorrelazioni parziali.

## ***Opzioni aggiuntive del comando TSMODEL***

Per personalizzare i modelli di serie storiche, è possibile incollare le impostazioni selezionate in una finestra di sintassi e modificare la sintassi del comando `TSMODEL` così ottenuta. Il linguaggio della sintassi dei comandi consente di:

- Specificare il periodo stagionale dei dati (con la parola chiave `SEASONLENGTH` nel sottocomando `AUXILIARY`). In tal modo si modifica la periodicità corrente (se presente) dell'insieme di dati attivo.
- Specificare i ritardi non consecutivi per i componenti `ARIMA` personalizzati e funzione di trasferimento (con i sottocomandi `ARIMA` e `TRANSFERFUNCTION`). Per esempio, è possibile specificare un modello `ARIMA` personalizzato con dei ritardi autoregressivi di ordine 1, 3 e 6; oppure una funzione di trasferimento con ritardi numeratore di ordine 2, 5 e 8.
- Fornire più di un insieme di specifiche di modelli (ad esempio, metodo di creazione dei modelli, ordini `ARIMA`, variabili indipendenti, e così via) per eseguire una sola volta la procedura Modelli serie storiche (con il sottocomando `MODEL`).

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

# *Applica modelli di serie storiche*

La procedura *Applica modelli delle serie storiche* carica i modelli delle serie storiche esistenti da un file esterno e li applica all'insieme di dati attivo. Questa procedura permette di ottenere previsioni per le serie con dati nuovi o rivisti, senza dover ricreare i modelli. Per generare i modelli si utilizza la procedura [Modelli serie storiche](#).

**Esempio.** Nel caso di un responsabile delle scorte di uno dei maggiori rivenditori e che gestisce 5.000 prodotti è stato utilizzato Expert Modeler per creare i modelli per effettuare le previsioni delle vendite di ogni prodotto per i tre mesi successivi. Il magazzino dati viene aggiornato ogni mese con i dati delle vendite effettive che si vorrebbero utilizzare per effettuare delle previsioni aggiornate mensilmente. La procedura *Applica modelli delle serie storiche* consente di compiere tale operazione utilizzando i modelli originali e ripetendo semplicemente la stima dei parametri dei modelli in modo da spiegare i dati nuovi.

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

**Grafici.** Grafici riassuntivi in tutti i modelli: istogrammi di  $R$ -quadrato stazionario,  $R$ -quadrato ( $R^2$ ), radice della media dei quadrati (RMSE), errore medio assoluto (MAE), errore percentuale medio assoluto (MAPE), errore massimo assoluto (MaxAE), errore percentuale massimo assoluto (MaxAPE), criterio di informazione bayesiano normalizzato (BIC). Risultati dei modelli individuali: valori di previsione, valori adattati, valori osservati, limiti di confidenza superiore e inferiore, autocorrelazioni residue e autocorrelazioni parziali.

## ***Considerazioni sui dati di Applica modelli delle serie storiche***

**Dati.** Le variabili (dipendenti e indipendenti) alle quali vengono applicati i modelli devono essere numeriche.

**Assunzioni.** I modelli vengono applicati alle variabili nell'insieme di dati attivo con gli stessi nomi delle variabili specificate nel modello. Tutte le variabili di questo tipo vengono considerate come serie storiche, vale a dire che ogni caso rappresenta un punto temporale con casi successivi separati da un intervallo di tempo costante.

- **Previsioni.** Per effettuare delle previsioni utilizzando modelli con variabili indipendenti (predittore), l'insieme di dati attivo deve contenere i valori di tali variabili per tutti i casi del periodo di previsione. Se si ripete la stima dei parametri del modello, nel periodo di stima le variabili indipendenti non devono contenere alcun valore mancante.

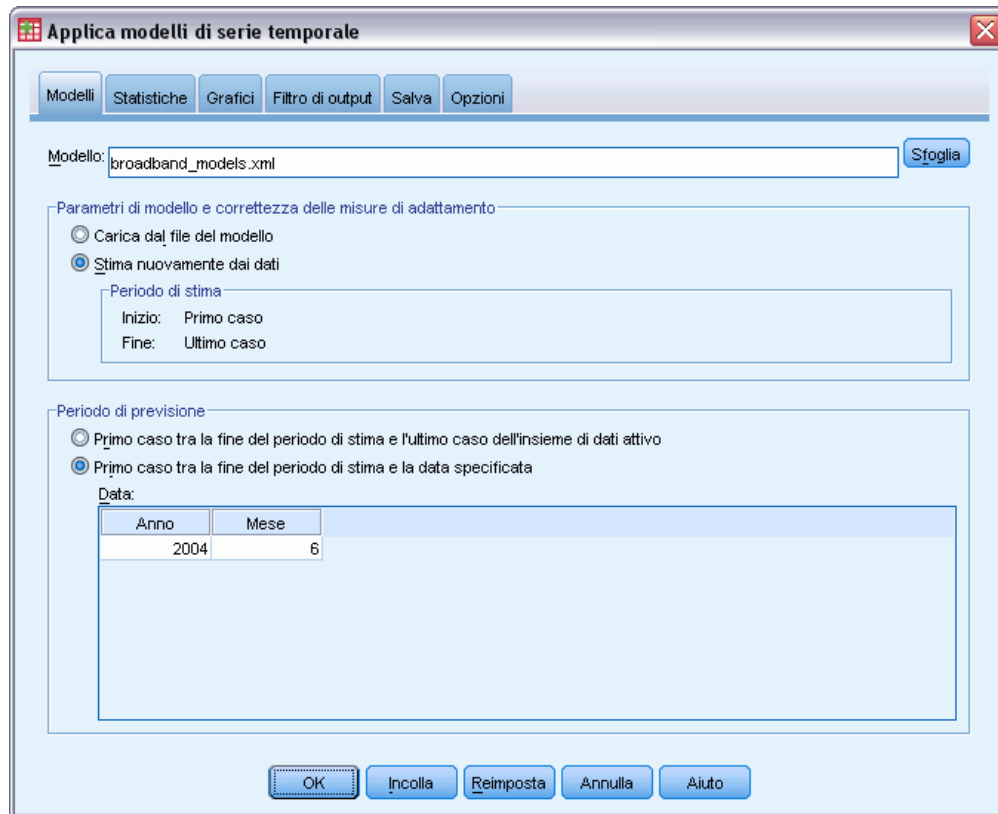
### Definizione delle date

La procedura Applica modelli delle serie storiche richiede che, se presente, la periodicità dell'insieme di dati attivo corrisponda a quella dei modelli da applicare. Se si sta semplicemente effettuando delle previsioni utilizzando lo stesso insieme di dati (magari con dei dati nuovi o aggiornati) utilizzato per creare il modello, tale condizione viene soddisfatta. Se non esiste alcuna periodicità per l'insieme di dati attivo, è possibile navigare fino alla finestra di dialogo Definisci date per crearne una. Se tuttavia i modelli sono stati creati senza specificare una periodicità, anche l'insieme di dati attivo deve esserne privo.

### Per applicare i modelli

- Dai menu, scegliere:  
Analizza > Forecasting > Applica modelli...

Figura 3-1  
Applica modelli delle serie storiche, scheda Modelli



- Immettere la specificazione del file relativa al file del modello o fare clic su Sfogli... e selezionare un file del modello (i file del modello vengono creati con la procedura [Modelli serie storiche](#)).

Se lo si desidera, è possibile:

- Ripetere la stima dei parametri del modello utilizzando i dati dell'insieme di dati attivo. Per creare le previsioni si utilizzano i parametri di cui si è ripetuta la stima.

- [Salvare le previsioni, gli intervalli di confidenza e i residui del rumore.](#)
- [Salvare i modelli di cui si è ripetuta la stima in formato XML.](#)

### **Parametri di modello e correttezza delle misure di adattamento**

**Carica da file del modello.** Le previsioni vengono generate utilizzando i parametri del modello dal file del modello senza ripetere la stima di tali parametri. [Le misure della bontà di adattamento](#) visualizzate nell'output e utilizzate per filtrare i modelli (migliori o peggiori), provengono dal file del modello e riflettono i dati utilizzati durante la creazione (o l'ultimo aggiornamento) di ogni modello. Con tale opzione, le previsioni non prendono in considerazione i dati storici, relativi alle variabili dipendenti o indipendenti, dell'insieme di dati attivo. Affinché i dati storici abbiano impatto sulle previsioni, selezionare Ripeti stima dai dati. Le previsioni inoltre, non prendono in considerazione i valori delle serie dipendenti nel periodo di previsione, bensì i valori delle variabili indipendenti nel periodo di previsione. Se si dispone di altri valori correnti delle serie dipendenti e si desidera includerli nelle previsioni, è necessario ripetere la stima regolando il periodo di stima affinché includa tali valori.

**Ripeti stima dai dati.** Per ripetere la stima dei parametri del modello si utilizzano i dati dell'insieme di dati attivo. La nuova stima dei parametri del modello non incide sulla struttura del modello. Per esempio, un modello ARIMA(1,0,1) rimane tale, ma il parametro autoregressivo e quello della media mobile vengono stimati di nuovo. La nuova stima non comporta il rilevamento di nuovi valori anomali. Se presenti, i valori anomali provengono sempre dal file del modello.

- **Periodo di stima.** Il periodo di stima definisce l'insieme di casi utilizzato per ripetere la stima dei parametri del modello. Per impostazione predefinita, il periodo di stima include tutti i casi dell'insieme di dati attivo. Per impostare il periodo di stima, selezionare Basato su intervallo di tempo o di casi nella finestra di dialogo Seleziona casi. A seconda dei dati disponibili, il periodo di stima utilizzato dalla procedura può variare da modello a modello, di conseguenza può differire dal valore visualizzato. Per un dato modello, il periodo di stima reale è il periodo rimasto dopo l'eliminazione di qualsiasi valore mancante consecutivo dalla variabile dipendente del modello, che si presenta all'inizio o alla fine del periodo di stima specificato.

### **Periodo di previsione**

Il periodo di previsione di ogni modello inizia sempre con il primo caso dopo la fine del periodo di stima e si protrae fino all'ultimo caso dell'insieme di dati attivo o fino a una data specificata dall'utente. Se non si ripete la stima dei parametri (come da impostazione predefinita), il periodo di stima di ogni modello è l'insieme di casi utilizzato durante la creazione (o l'ultimo aggiornamento) del modello.

- **Primo caso dopo la fine del periodo di stima fino all'ultimo caso dell'insieme di dati attivo.** Selezionare questa opzione quando la fine del periodo di stima è precedente all'ultimo caso dell'insieme di dati attivo e si desidera effettuare delle previsioni fino all'ultimo caso.
- **Primo caso dopo la fine del periodo di stima fino a una data specificata.** Selezionare questa opzione per specificare esplicitamente la fine del periodo di previsione. Immettere i valori per tutte le celle della griglia Data.

Se non è stata definita alcuna specificazione della data per l'insieme di dati attivo, la griglia Data mostra unicamente la colonna *Osservazione*. Per specificare la fine del periodo di previsione, immettere il numero di riga (come visualizzato nell'Editor dei dati) del caso relativo.

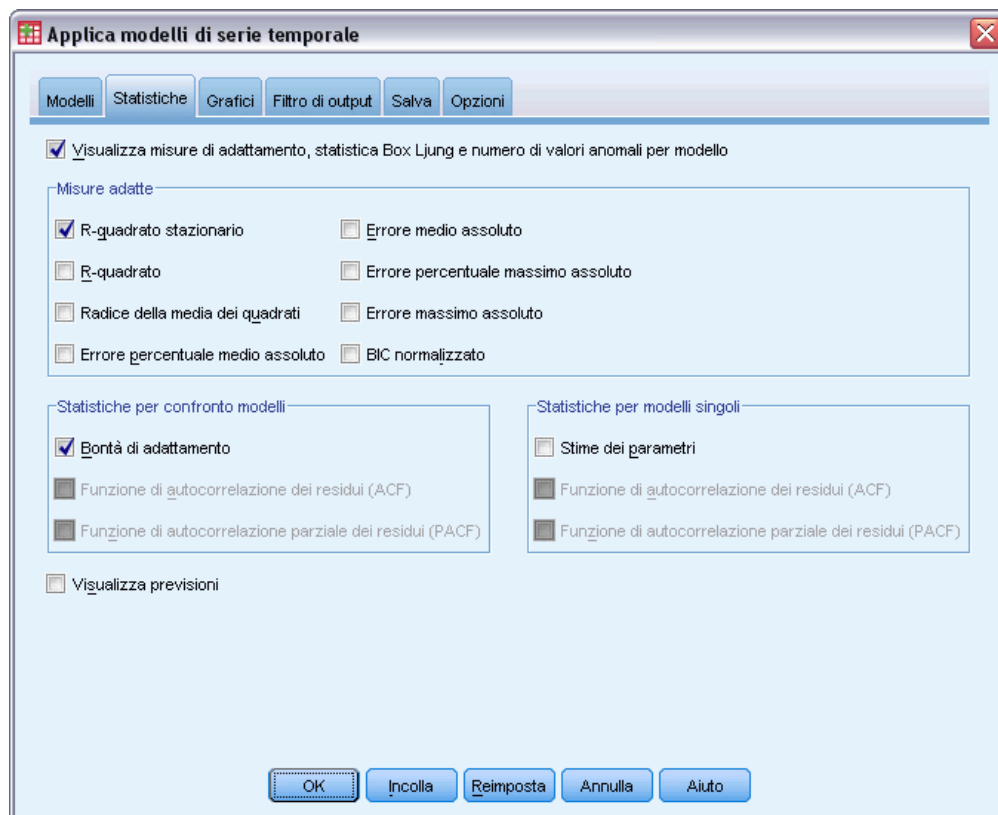
La colonna *Ciclo* (se presente) della griglia Data si riferisce al valore della variabile *CYCLE\_* dell'insieme di dati attivo.

## Output

Output disponibili include i risultati dei modelli individuali e quelli di tutti i modelli. È possibile limitare i risultati dei modelli individuali a un insieme di modelli migliori o peggiori, in base ai criteri specificati dall'utente.

### Tablelle delle statistiche e delle previsioni

Figura 3-2  
Applica modelli delle serie storiche, scheda Statistiche



La scheda Statistiche contiene le opzioni di visualizzazione delle tablelle delle statistiche di adattamento del modello, dei parametri del modello, delle funzioni di autocorrelazione e delle previsioni. A meno che non venga ripetuta la stima dei parametri del modello (Ripeti stima dei dati nella scheda Modelli), i valori delle misure di adattamento visualizzati, i valori Ljung-Box e i



parametri del modello sono i valori provenienti dal file del modello e riflettono i dati utilizzati durante la creazione (o l'aggiornamento) del modello. Le informazioni relative ai valori anomali provengono sempre dal file del modello.

**Visualizza misure di adattamento, statistica Box Ljung e numero di valori anomali per modello.**

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

**Misure di adattamento.** È possibile selezionare una o più delle seguenti opzioni, per includerle nella tabella che contiene le misure di adattamento di ogni modello:

- *R*-quadrato stazionario
- *R*-quadrato
- Radice della media dei quadrati
- Errore percentuale medio assoluto
- Errore medio assoluto
- Errore percentuale massimo assoluto
- Errore massimo assoluto
- BIC normalizzato

Per ulteriori informazioni, vedere l'argomento [Misure della bontà di adattamento in l'appendice A a pag. 97.](#)

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

- **Bontà di adattamento.** Tabella di statistiche riassuntive e di percentili per *R*-quadrato stazionario, *R*-quadrato, radice della media dei quadrati, errore percentuale medio assoluto, errore medio assoluto, errore percentuale massimo assoluto, errore massimo assoluto e Criterio di informazione bayesiano normalizzato.
- **Funzione di autocorrelazione dei residui (ACF).** Tabella di statistiche riassuntive e di percentili per le autocorrelazioni dei residui in tutti i modelli stimati. Questa tabella è disponibile solo se i parametri del modello vengono stimati di nuovo (Ripeti stima dei dati nella scheda Modelli).
- **Funzione di autocorrelazione parziale dei residui (PACF).** Tabella di statistiche riassuntive e di percentili per le autocorrelazioni parziali dei residui in tutti i modelli stimati. Questa tabella è disponibile solo se i parametri del modello vengono stimati di nuovo (Ripeti stima dei dati nella scheda Modelli).

**Statistiche modelli singoli.** Questo gruppo di opzioni controlla la visualizzazione delle tabelle contenenti le informazioni dettagliate di ogni modello. Ogni opzione genera una tabella distinta. È possibile selezionare una o più delle seguenti opzioni:

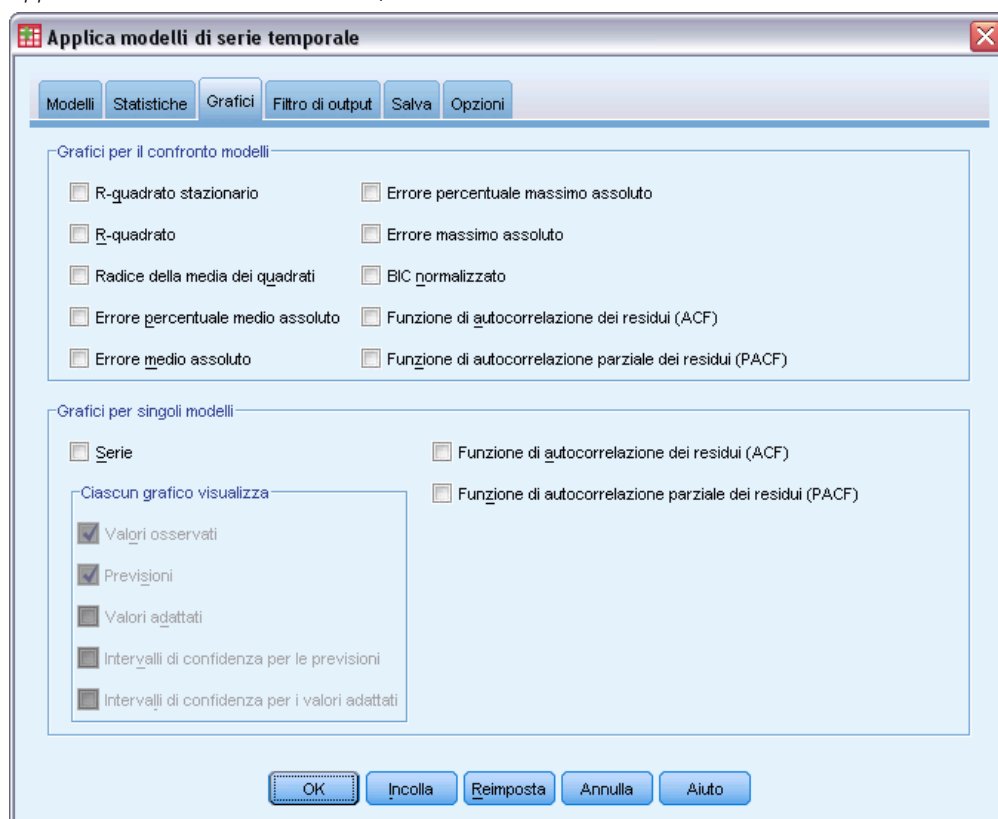
- **Stime dei parametri.** Visualizza una tabella di stime dei parametri di ogni modello. Per i modelli di livellamento esponenziale e ARIMA vengono visualizzate delle tabelle distinte. Se sono presenti dei valori anomali, anche le stime dei parametri ad essi relative vengono visualizzate in una tabella distinta.

- **Funzione di autocorrelazione dei residui (ACF).** Visualizza una tabella di autocorrelazioni dei residui per ritardo di ogni modello stimato. La tabella include gli intervalli di confidenza delle autocorrelazioni. Questa tabella è disponibile solo se i parametri del modello vengono stimati di nuovo (Ripeti stima dei dati nella scheda Modelli).
- **Funzione di autocorrelazione parziale dei residui (PACF).** Visualizza una tabella di autocorrelazioni parziali dei residui per ritardo di ogni modello stimato. La tabella include gli intervalli di confidenza delle autocorrelazioni parziali. Questa tabella è disponibile solo se i parametri del modello vengono stimati di nuovo (Ripeti stima dei dati nella scheda Modelli).

**Visualizza previsioni.** Visualizza una tabella di previsioni dei modelli e di intervalli di confidenza di ogni modello.

## Grafici

Figura 3-3  
Applica modelli delle serie storiche, scheda Grafici



La scheda Grafici contiene le opzioni per la visualizzazione dei grafici delle statistiche di adattamento del modello, delle funzioni di autocorrelazione e dei valori di serie, tra cui le previsioni.

### **Grafici per il confronto modelli**

Questo gruppo di opzioni controlla la visualizzazione dei grafici contenenti le statistiche di tutti i modelli. A meno che non venga ripetuta la stima dei parametri del modello (Ripeti stima dei dati nella scheda Modelli), i valori visualizzati sono i valori provenienti dal file del modello e riflettono i dati utilizzati durante la creazione (o l'aggiornamento) di ogni modello. Inoltre, i grafici di autocorrelazione sono disponibili solo se viene ripetuta la stima dei parametri del modello. Ogni opzione genera un grafico distinto. È possibile selezionare una o più delle seguenti opzioni:

- *R*-quadrato stazionario
- *R*-quadrato
- Radice della media dei quadrati
- Errore percentuale medio assoluto
- Errore medio assoluto
- Errore percentuale massimo assoluto
- Errore massimo assoluto
- BIC normalizzato
- Funzione di autocorrelazione dei residui (ACF)
- Funzione di autocorrelazione parziale dei residui (PACF)

Per ulteriori informazioni, vedere l'argomento [Misure della bontà di adattamento in l'appendice A a pag. 97](#).

### **Grafici per singoli modelli**

**Serie.** Selezionare (contrassegnare) questa opzione per ottenere i grafici dei valori attesi per ogni modello. I valori osservati, i valori adattati, gli intervalli di confidenza per i valori adattati e le autocorrelazioni sono disponibili solo se viene ripetuta la stima dei parametri del modello (Ripeti stima dei dati nella scheda Modelli). È possibile selezionare una o più delle opzioni seguenti da includere nel grafico:

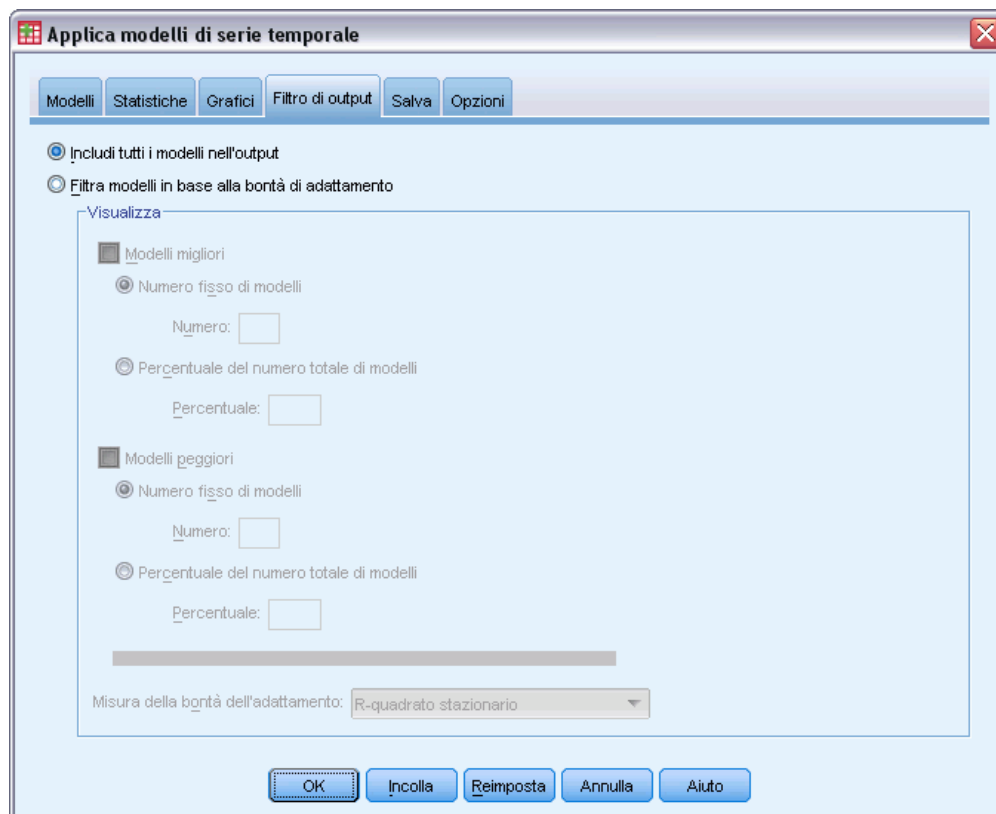
- **Valori osservati.** I valori osservati della serie dipendente.
- **Previsioni.** I valori attesi del modello per il periodo di previsione.
- **Valori adattati.** I valori attesi del modello per il periodo di stima.
- **Intervalli di confidenza per le previsioni.** Gli intervalli di confidenza per il periodo di previsione.
- **Intervalli di confidenza per i valori adattati.** Gli intervalli di confidenza per il periodo di stima.

**Funzione di autocorrelazione dei residui (ACF).** Visualizza un grafico di autocorrelazioni dei residui di ogni modello stimato.

**Funzione di autocorrelazione parziale dei residui (PACF).** Visualizza un grafico di autocorrelazioni parziali dei residui di ogni modello stimato.

## Limitazione dell'output ai modelli migliori o peggiori

Figura 3-4  
Applica modelli delle serie storiche, scheda Filtro di output



La scheda Filtro di output contiene le opzioni per limitare sia l'output della tabella, sia quello del grafico a un sottoinsieme di modelli. È possibile scegliere di limitare l'output al modello migliore e/o peggiore, a seconda dei criteri di adattamento forniti. Per impostazione predefinita, tutti i modelli vengono inclusi nell'output. A meno che non venga ripetuta la stima dei parametri del modello (Ripeti stima dei dati nella scheda Modelli), i valori delle misure di adattamento utilizzati per i modelli di filtro sono i valori provenienti dal file del modello e riflettono i dati utilizzati durante la creazione (o l'aggiornamento) di ogni modello.

**Modelli migliori.** Selezionare (contrassegnare) questa opzione per includere nell'output i modelli migliori. Selezionare una misura della bontà di adattamento e specificare il numero di modelli da includere. La selezione dell'opzione non preclude anche la selezione dei modelli peggiori. In questo caso, l'output è costituito sia dai modelli peggiori, sia da quelli migliori.

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

**Modelli peggiori.** Selezionare (contrassegnare) questa opzione per includere nell'output i modelli peggiori. Selezionare una misura della bontà di adattamento e specificare il numero di modelli da includere. La selezione dell'opzione non preclude anche la selezione dei modelli migliori. In questo caso, l'output è costituito sia dai modelli migliori, sia da quelli peggiori.

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

**Misura della bontà di adattamento.** Selezionare la misura della bontà di adattamento da utilizzare per i modelli di filtro. Il default è  $R$ -quadrato stazionario.

## ***Salvataggio dei valori stimati e delle specifiche del modello***

La scheda Salva consente di salvare i valori stimati del modello come variabili nuove nell'insieme di dati attivo e di salvare le specifiche del modello in un file esterno in formato XML.

**Salva variabili.** Nell'insieme di dati attivo è possibile salvare i valori stimati del modello, gli intervalli di confidenza e i residui come variabili nuove. Ogni modello genera un insieme proprio di variabili nuove. Se il periodo di previsione supera la serie di variabili dipendenti associata al modello, si aggiungono i casi nuovi. A meno che non venga ripetuta la stima dei parametri del modello (Ripeti stima dei dati nella scheda Modelli), i valori attesi e i limiti di confidenza vengono creati solo per il periodo di previsione. Scegliere di salvare le variabili nuove selezionando per ognuna di esse la casella di controllo Salva associata. Per impostazione predefinita, nessuna variabile nuova è salvata.

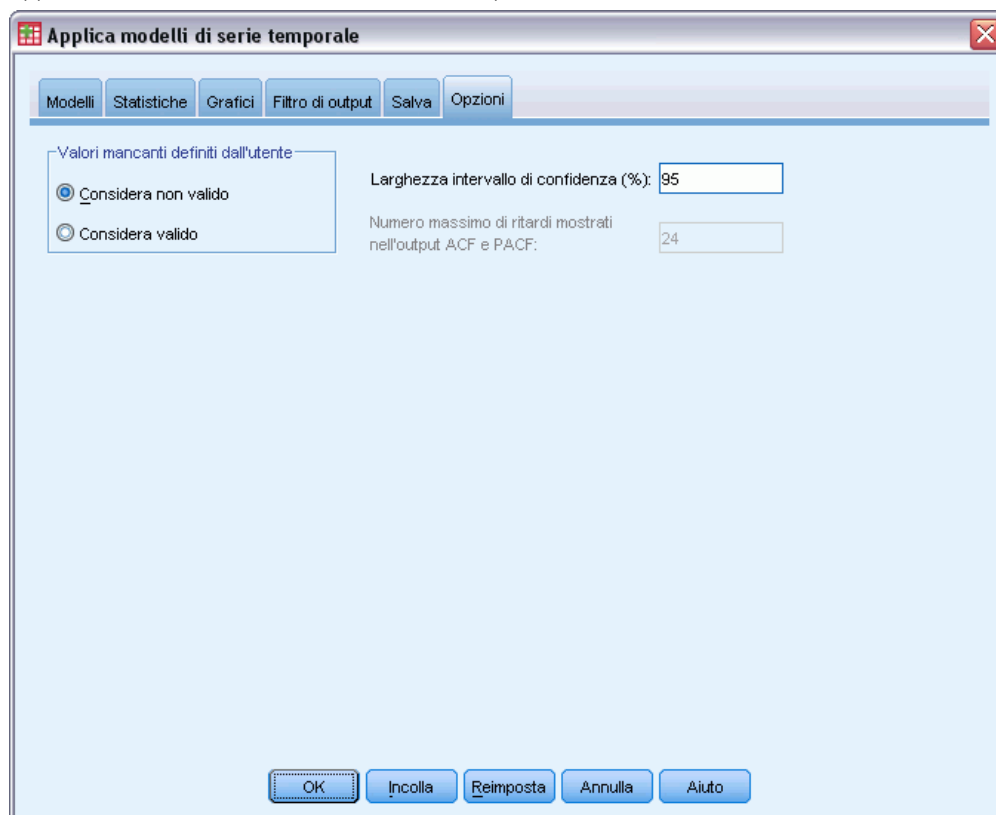
- **Valori attesi.** I valori attesi del modello.
- **Limiti di confidenza inferiori.** I limiti di confidenza inferiori dei valori attesi.
- **Limiti di confidenza superiori.** I limiti di confidenza superiori dei valori attesi.
- **Residui del rumore.** I residui del modello. Quando si eseguono le trasformazioni della variabile dipendente (ad esempio, logaritmica naturale), questi sono i residui della serie trasformata. Questa opzione è disponibile solo se i parametri del modello vengono stimati di nuovo (Ripeti stima dei dati nella scheda Modelli).
- **Prefisso del nome variabile.** Consente di specificare i prefissi da utilizzare per i nomi delle nuove variabili oppure di mantenere i prefissi predefiniti. I nomi delle variabili sono costituiti dal prefisso, dal nome della variabile dipendente associata e da un identificatore di modello. Per evitare eventuali conflitti con i nomi delle variabili, è possibile estendere il nome della variabile. Il prefisso deve essere conforme alle regole stabilite per i nomi di variabile validi.

**Esporta file del modello** Le specifiche del modello che contengono i parametri stimati nuovamente e le statistiche dell'adattamento vengono esportate nel file specificato in formato XML. Questa opzione è disponibile solo se i parametri del modello vengono stimati di nuovo (Ripeti stima dei dati nella 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 conforme a PMML che può essere utilizzato con applicazioni conformi a PMML, tra cui le applicazioni IBM SPSS.

## Opzioni

Figura 3-5  
Applica modelli delle serie storiche, scheda Opzioni



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

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

- **Considera come non valido.** I valori mancanti definiti dall'utente sono considerati come mancanti di sistema.
- **Considera come valido.** I valori mancanti definiti dall'utente vengono considerati come dati validi.

**Politica dei valori mancanti.** Le seguenti regole vengono applicate al trattamento dei valori mancanti (include i valori mancanti di sistema e i valori mancanti definiti dall'utente considerati non validi):

- I casi che presentano i valori mancanti di una variabile dipendente che si verificano nel periodo di stima vengono inclusi nel modello. La gestione specifica del valore mancante dipende dal metodo di stima.
- Per i modelli ARIMA, se una variabile predittore riporta dei valori mancanti nel periodo di stima viene generato un avviso. Per tutti i modelli che prevedono un predittore non viene ripetuta la stima.
- Se una variabile indipendente presenta dei valori mancanti nel periodo di previsione, la procedura genera un avviso ed effettua le previsioni per il periodo consentito.

**Ampiezza dell'intervallo di confidenza (%).** Gli intervalli di confidenza vengono calcolati per i valori stimati del modello e le autocorrelazioni dei residui. È possibile specificare qualsiasi valore positivo inferiore a 100. Per impostazione predefinita, si utilizza un intervallo di confidenza al 95%.

**Numero massimo di ritardi mostrato nell'output ACF e PACF.** 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 vengono stimati di nuovo (Ripeti stima dei dati nella scheda Modelli).

## ***Opzioni aggiuntive del comando TSAPPLY***

Per rendere disponibili le opzioni aggiuntive, incollare le selezioni in una finestra della sintassi e modificare la sintassi del comando `TSAPPLY` che ne risulta. Il linguaggio della sintassi dei comandi consente di:

- Specificare che solo un sottoinsieme di modelli di un file del modello deve essere applicato all'insieme di dati attivo (con le parole chiave `DROP` e `KEEP` nel sottocomando `MODEL`).
- Applicare i modelli provenienti da due o più file del modello ai dati (con il sottocomando `MODEL`). Per esempio, un file del modello può contenere i modelli delle serie che rappresentano le vendite unitarie e un altro può contenere i modelli delle serie che rappresentano gli introiti.

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

## ***Decomposizione stagionale***

La procedura Decomposizione stagionale decompone la serie in una componente stagionale, un trend combinato 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 desidera analizzare le misurazioni mensili del livello di ozono di una stazione meteorologica specifica. L’obiettivo è quello di stabilire se è possibile rilevare dei trend nei dati. Per poter rilevare i trend, lo scienziato deve prima tener conto delle variazioni nelle misurazioni causate dagli effetti stagionali. La procedura Decomposizione stagionale permette di rimuovere tutte le variazioni stagionali sistematiche. L’analisi del trend viene quindi eseguita sulle serie con le correzioni stagionali.

**Statistiche.** Insieme di fattori stagionali.

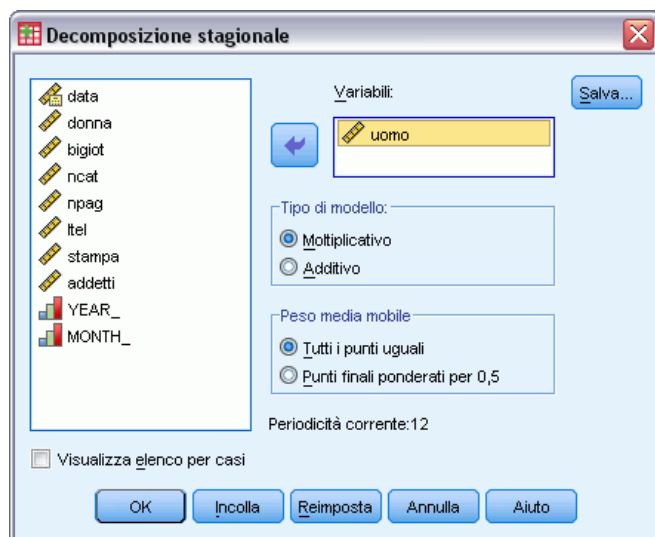
**Dati.** Le variabili devono essere numeriche.

**Assunzioni.** Le variabili non devono contenere dati mancanti incorporati. È necessario definire almeno una componente data periodica.

### ***Per stimare i fattori stagionali***

- Dai menu, scegliere:  
Analizza > Previsioni > Decomposizione stagionale...

Figura 4-1  
*Finestra di dialogo Decomposizione stagionale*



- Selezionare una o più variabili dall’elenco e sporarle nell’elenco Variabili. È utile ricordare che l’elenco comprende solo variabili numeriche.



**Tipo di modello.** La procedura Decomposizione stagionale offre due diverse opzioni per la creazione di fattori stagionali: moltiplicativa o additiva.

- **Moltiplicativo.** La componente stagionale è un fattore per il quale si moltiplica la serie corretta stagionalmente, per riprodurre la serie originale. Calcola componenti stagionali proporzionate al livello della serie. Le osservazioni prive di variazione stagionale hanno una componente stagionale pari a 1.
- **Additivo.** Al valore non corretto della serie, viene aggiunta una correzione stagionale additiva. 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 della serie. Le osservazioni prive di variazione stagionale hanno una componente stagionale pari a 0.

**Peso della media mobile.** Le opzioni del gruppo Peso della media mobile consentono di specificare la modalità di elaborazione delle serie durante il calcolo delle medie mobili. Le opzioni sono disponibili solo se la periodicità della serie è pari. Se la periodicità è dispari, tutti i punti vengono ponderati equamente.

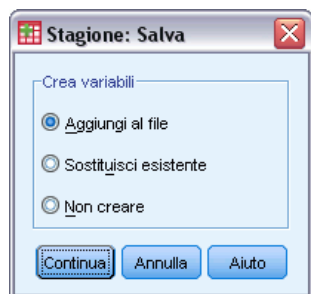
- **Tutti i punti uguali.** Le medie mobili vengono calcolate con un'estensione pari alla periodicità e tutti i punti pesati in modo uniforme. Questo metodo viene sempre utilizzato se la periodicità è valore specificato è dispari.
- **Punti finali ponderati per 0,5.** Le medie mobili per le serie con periodicità pari sono calcolate con un'estensione pari alla periodicità più 1 e assegnando peso 0,5 ai punti finali dell'estensione.

Se lo si desidera, è possibile:

- Fare clic su Salva per specificare come salvare le nuove variabili.

## Decomposizione stagionale: Salva

Figura 4-2  
Finestra di dialogo Salva stagione



**Crea variabili.** Permette di scegliere come considerare le nuove variabili.

- **Aggiungi al file.** Le nuove serie create da Decomposizione stagionale vengono salvate come variabili regolari nell'insieme di dati attivo. I nomi delle variabili sono formati da un prefisso di tre lettere, da un carattere di sottolineatura e da un numero.

- **Sostituisci esistenti.** Le nuove serie create da Decomposizione stagionale vengono salvate come variabili temporanee nell'insieme di dati 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 all'insieme di dati attivo.

### ***Nomi nuove variabili***

La procedura Decomposizione stagionale crea quattro variabili nuove (serie) con i seguenti prefissi da tre lettere per ogni serie specificata:

**SAF.** *Fattori di correzione per stagionalità.* Questi valori indicano l'effetto di ogni periodo sul livello della serie.

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

**STC.** *Componenti di trend-ciclo livellate.* Questi valori mostrano il comportamento trend e ciclico presente nella serie.

**ERR.** *Valori residui o di "errore".* I valori che rimangono dopo aver rimosso dalla serie le componenti stagionali, di trend e di ciclo.

## ***Opzioni aggiuntive del comando SEASON***

Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare una delle periodicità del comando SEASON anziché selezionare una delle opzioni della procedura Definizione di date.

Per informazioni dettagliate sulla sintassi, vedere *Command Syntax Reference*.

## ***Grafici spettrali***

La procedura dei Grafici spettrali è utile per identificare un comportamento periodico all'interno di 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 componenti periodiche delle serie di livellamento sono più significative alle frequenze basse, la variazione casuale diffonde la rilevanza delle componenti in tutte le frequenze.

Con questa procedura non è possibile analizzare le serie che includono dati mancanti.

**Esempio.** La velocità con cui vengono costruite nuove case è un importante barometro dello stato dell'economia. I dati relativi alla costruzione di nuove case sono generalmente caratterizzati da una forte componente stagionale. Ma esistono anche cicli più lunghi nei dati di cui gli analisti devono tenere conto ai fini della valutazione dei dati correnti?

**Statistiche.** Trasformazioni di seno e coseno, valore del periodogramma e stima della densità spettrale per ciascuna frequenza o componente del periodo. Se si seleziona l'analisi bivariata: parti reali o immaginarie del periodogramma incrociato, densità cospettrale, spettro di quadratura, guadagno, coerenza quadrata e spettro di fase per ciascuna frequenza o componente del periodo.

**Grafici.** Per le analisi univariate e bivariate: periodogramma e densità spettrale. Per le analisi bivariate: coerenza quadrata, spettro di quadratura, entità dell'incrocio, densità cospettrale, spettro di fase e guadagno.

**Dati.** Le variabili devono essere numeriche.

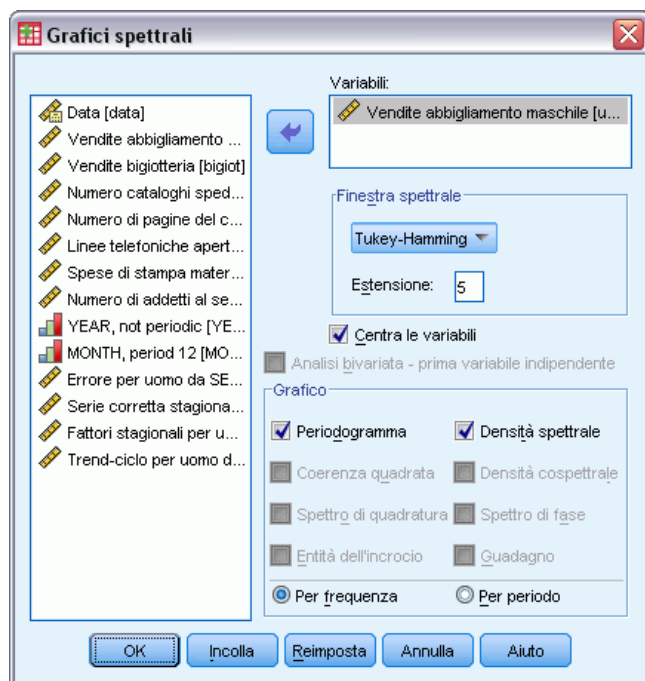
**Assunzioni.** Le variabili non devono contenere dati mancanti incorporati. Le serie storiche da analizzare devono essere stazionarie e tutte le medie diverse da zero devono essere sottratte dalla serie.

- **Stazionario.** Una condizione che deve essere rispettata da serie storiche alle quali viene adattato 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.

### ***Esecuzione di un'analisi spettrale***

- Dai menu, scegliere:  
Analisi > Serie storiche > Analisi spettrale...

Figura 5-1  
Finestra di dialogo Grafici spettrali



- ▶ Selezionare una o più variabili dall'elenco e spostarle nell'elenco Variabili. È utile ricordare che l'elenco comprende solo variabili numeriche.
- ▶ Selezionare una delle opzioni di Finestra spettrale per scegliere come livellare il periodogramma per ottenere la stima della densità spettrale. Le opzioni di livellamento disponibili sono: Tukey-Hamming, Tukey, Parzen, Bartlett, Daniell (Unitaria) e Nessuna.
  - **Tukey-Hamming.** I pesi sono calcolati come  $W_k = ,54Dp(2 \pi f_k) + ,23Dp(2 \pi f_k + \pi/p) + ,23Dp(2 \pi f_k - \pi/p)$ , per  $k = 0, \dots, p$  dove  $p$  è la parte intera dell'estensione diviso 2 e  $Dp$  è il kernel di Dirichlet di ordine  $p$ .
  - **Tukey.** I pesi vengono calcolati come:  $W_k = 0,5Dp(2 \pi f_k) + 0,25Dp(2 \pi f_k + \pi/p) + 0,25Dp(2 \pi f_k - \pi/p)$ , per  $k = 0, \dots, p$ , dove  $p$  è la parte intera di estensione diviso 2 e  $Dp$  è il kernel di Dirichlet di ordine  $p$ .
  - **Parzen.** I pesi sono  $P_k = 1/p(2 + \cos(2 \pi f_k)) (F[p/2] (2 \pi f_k))^{*2}$ , per  $k=0, \dots, p$  dove  $p$  è la parte intera di metà dell'estensione e  $F[p/2]$  è il kernel di Fejer di ordine  $p/2$ .
  - **Bartlett.** La forma di una finestra spettrale per la quale i pesi della metà superiore della finestra sono valutati in questo modo:  $W_k = F_p(2 * \pi * f_k)$ , per  $k = 0, \dots, p$ , dove  $p$  equivale alla parte intera dell'intervallo diviso 2 e  $F_p$  è il nucleo di Fejer di ordine  $p$ . La metà inferiore della finestra è simmetrica alla metà superiore.
  - **Daniell(unità).** Una forma di finestra spettrale in cui i pesi sono tutti pari a 1.
  - **Nessuno.** Livellamento assente. Se viene scelta questa opzione, la stima della densità spettrale è la stessa del periodogramma.

**Estensione.** In un processo di livellamento è il numero di osservazioni consecutive su cui viene effettuato il livellamento. Solitamente si usa un intero positivo dispari. Maggiore è l'estensione, maggiore è l'effetto di livellamento nei grafici di densità spettrale.

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

**Analisi bivariata — la prima variabile di ciascuna.** Se si selezionano due o più variabili, è possibile selezionare questa opzione per eseguire analisi spettrali bivariate.

- La prima variabile dell'elenco Variabili è considerata variabile indipendente, mentre tutte le altre variabili sono considerate variabili dipendenti.
- Le serie dopo la prima vengono analizzate con la prima, indipendentemente dalle altre serie denominate. Vengono inoltre eseguite analisi univariate di ciascuna serie.

**Grafico.** Il periodogramma e la densità spettrale possono essere utilizzati sia per le analisi univariate che per quelle bivariate. Tutte le altre opzioni possono essere selezionate solo per le analisi bivariate.

- **Periodogramma.** Grafico non livellato dell'ampiezza spettrale (su 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 storiche. I 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 storiche.
- **Spettro di fase.** Una misura di fino a che punto ogni componente di frequenza di una serie precede o segue l'altra.
- **Guadagno.** Il quoziente di divisione dell'ampiezza incrociata per la densità spettrale per 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 termine media) fino 0,5 (il termine per un ciclo di due osservazioni).

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

## ***Opzioni aggiuntive del comando SPECTRA***

Il linguaggio della sintassi dei comandi consente inoltre di:

- Salvare le variabili dell'analisi spettrale calcolate nell'insieme di dati attivo per poterle utilizzare successivamente.

- Specificare pesi personalizzati per la finestra spettrale.
- Creare grafici per frequenza e per periodo.
- Stampare un elenco completo di tutti i valori visualizzati nel grafico.

Per informazioni dettagliate sulla sintassi, vedere *Command Syntax Reference*.

# ***Parte II: Esempi***

## ***Uso di Expert Modeler per le previsioni di massa***

Un analista di un provider nazionale di servizi a banda larga deve creare delle previsioni sugli abbonamenti in modo da poter stimare l'uso dei servizi a banda larga. Le previsioni devono essere effettuate per ciascuno degli 85 mercati locali che costituiscono la base degli abbonati nazionali. I dati storici mensili sono stati raccolti nel file *broadband\_1.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice D in IBM SPSS Forecasting 21.](#)

Questo esempio mostra come usare Expert Modeler per creare previsioni per i tre mesi successivi per ciascuno degli 85 mercati locali nonché come salvare i modelli creati in un file XML esterno. Al termine, potrebbe essere utile utilizzare l'esempio che segue, [Uso dei modelli salvati per l'esecuzione di nuove previsioni di massa in il capitolo 7 a pag. 57](#), che applica i modelli salvati a un file di dati aggiornato, in modo che sia possibile estendere le previsioni di altri tre mesi senza dover ricreare i modelli.

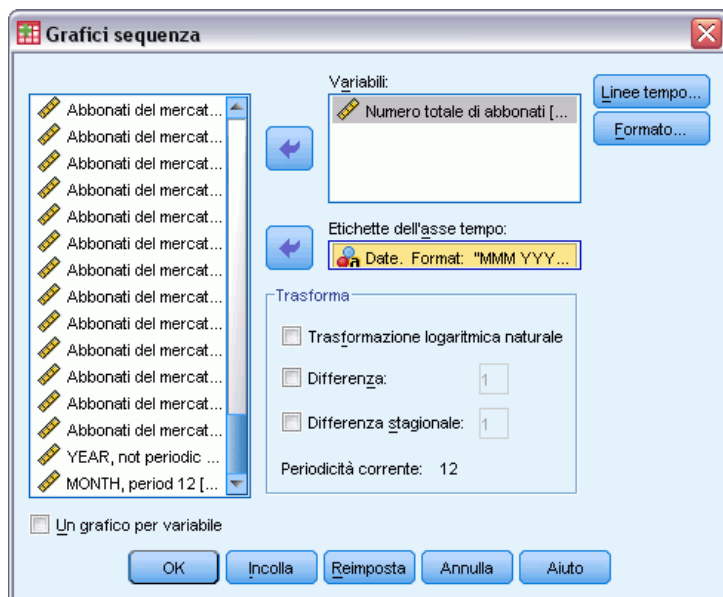
### ***Esame dei dati***

Prima di creare un modello, è sempre utile esaminare dettagliatamente le caratteristiche dei dati usati. I dati presentano variazioni stagionali? Benché Expert Modeler ricerchi automaticamente il miglior modello stagionale e non per ciascuna serie, per rendere più veloce la procedura è spesso possibile limitare la ricerca a modelli non stagionali se si ritiene che i dati non presentino caratteristiche di stagionalità. Per avere un quadro di insieme senza esaminare i dati, è sufficiente riportare su un grafico il numero totale di abbonati di tutti i mercati.

- Dai menu, scegliere:  
Analizza > Previsioni > Grafici sequenza...

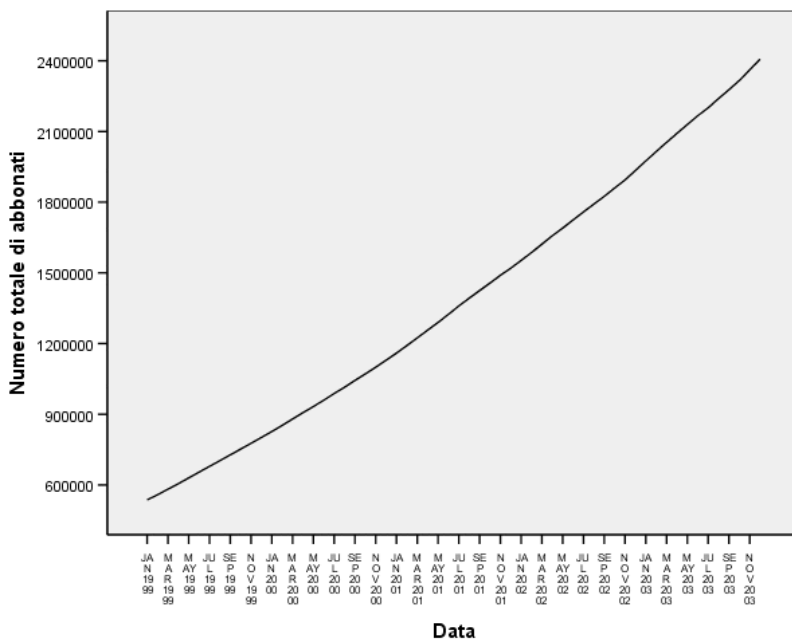


Figura 6-1  
Finestra di dialogo Grafici sequenza



- ▶ Selezionare *Numero totale di abbonati* e passare all'elenco Variabili.
- ▶ Selezionare *Data* e passare all'elenco Etichette dell'asse tempo.
- ▶ Fare clic su OK.

Figura 6-2  
*Numero totale di abbonati a servizi a banda larga in tutti i mercati*



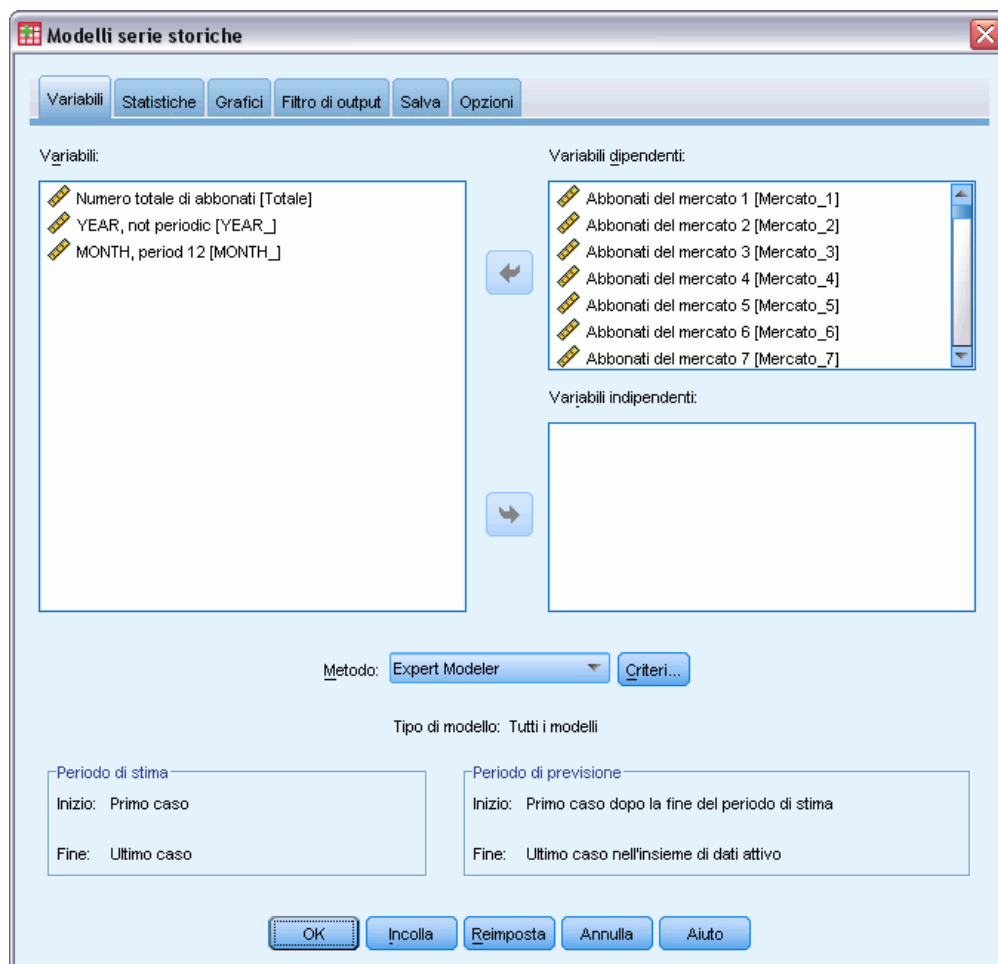
La serie mostra un trend crescente uniforme che non presenta tracce di variazioni stagionali. Anche se non è possibile escludere la presenza di serie con elementi stagionali, la stagionalità non è generalmente una caratteristica significativa dei dati. Naturalmente, prima di escludere i modelli stagionali è necessario esaminare ciascuna serie, in modo da poterle separare e creare modelli specifici. Nel caso dell'esempio, nessuna delle 85 serie presenta elementi di stagionalità.

## Esecuzione dell'analisi

Per usare Expert Modeler:

- Dai menu, scegliere:  
Analizza > Previsioni > Crea modelli...

Figura 6-3  
finestra di dialogo Serie storiche



- Selezionare le variabili dipendenti da *Abbonati del mercato 1* a *Abbonati del mercato 85*.
- Verificare di aver selezionato Expert Modeler nell'elenco a discesa Metodo. Expert Modeler ricerca automaticamente il miglior modello per la serie di variabili dipendenti.

L'insieme di casi usati per stimare il modello viene chiamato **periodo di stima**. Per impostazione predefinita comprende tutti i casi del file di dati attivo. Per impostare il periodo di stima, selezionare Basato su intervallo di tempo o casi nella finestra di dialogo Seleziona casi.

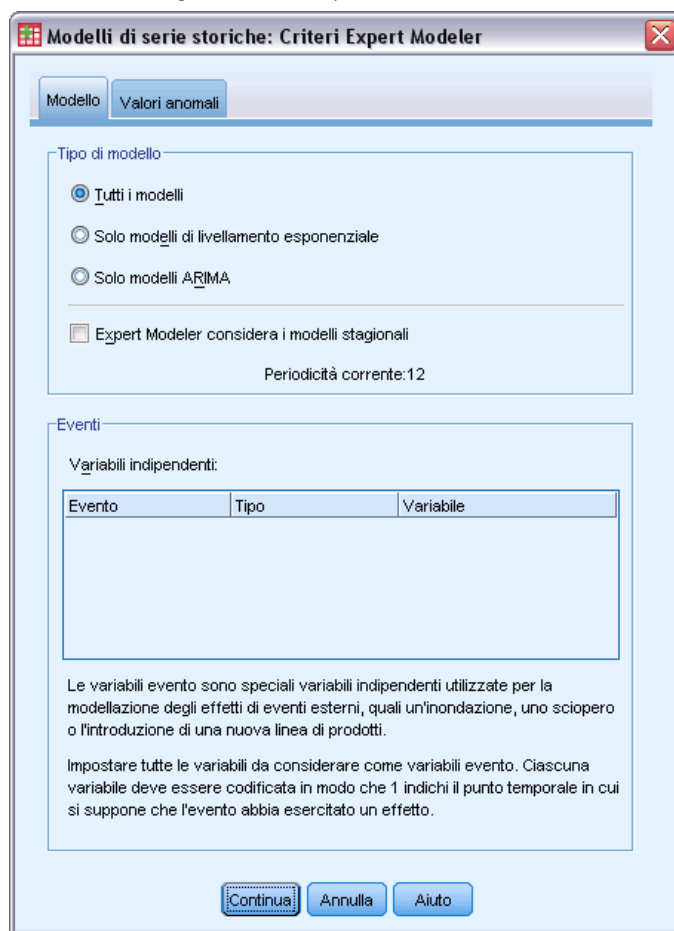
Nell'esempio verranno usati i dati predefiniti.

Vale la pena notare anche che il periodo di previsione predefinito inizia alla fine del periodo di stima e si estende fino all'ultimo caso del file di dati attivo. Per effettuare previsioni che si estendano oltre questo caso, è necessario estendere il periodo di previsione. Questa operazione può essere eseguita tramite la scheda Opzioni, come mostra più oltre l'esempio.

- Fare clic su Criteri.

Figura 6-4

Finestra di dialogo Criteri di Expert Modeler, scheda Modello



- Deselezionare Expert Modeler usa i modelli stagionali nel gruppo Tipo di modello.

Sebbene i dati siano mensili e la periodicità corrente sia 12, i dati non presentano elementi stagionali, quindi l'analisi non richiede l'uso di modelli stagionali. Ciò riduce lo spazio dei modelli ricercati da Expert Modeler e può rendere il calcolo molto più veloce.

- Fare clic su Continua.
- Fare clic sulla scheda Opzioni nella finestra di dialogo Time Series Modeler.

Figura 6-5  
Time Series Modeler, scheda Opzioni

Anno	Mese
2004	3

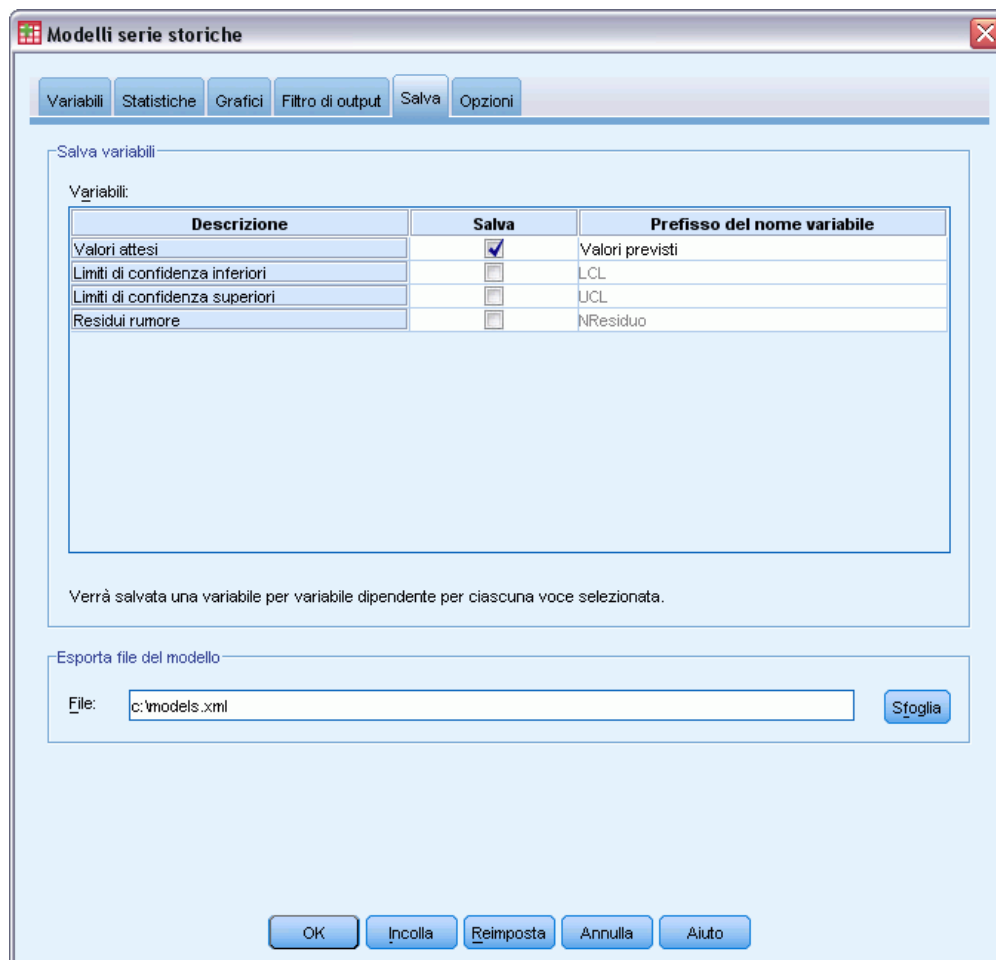
- Nel gruppo Periodo di previsione selezionare Primo caso tra la fine del periodo di previsione e una data specifica.

- Nella griglia Data immettere 2004 per l'anno e 3 per il mese.

L'insieme dei dati contiene tutti i dati riferiti al periodo compreso tra il gennaio 1999 e il dicembre 2003. Con le impostazioni correnti, il periodo di previsione sarà compreso tra il gennaio 2004 e il marzo dello stesso anno.

- Fare clic sulla scheda Salva.

Figura 6-6  
Time Series Modeler, scheda Salva



- Selezionare la voce per Valori attesi nella colonna *Salva* e lasciare il valore predefinito *Previsto* come prefisso del nome della variabile.

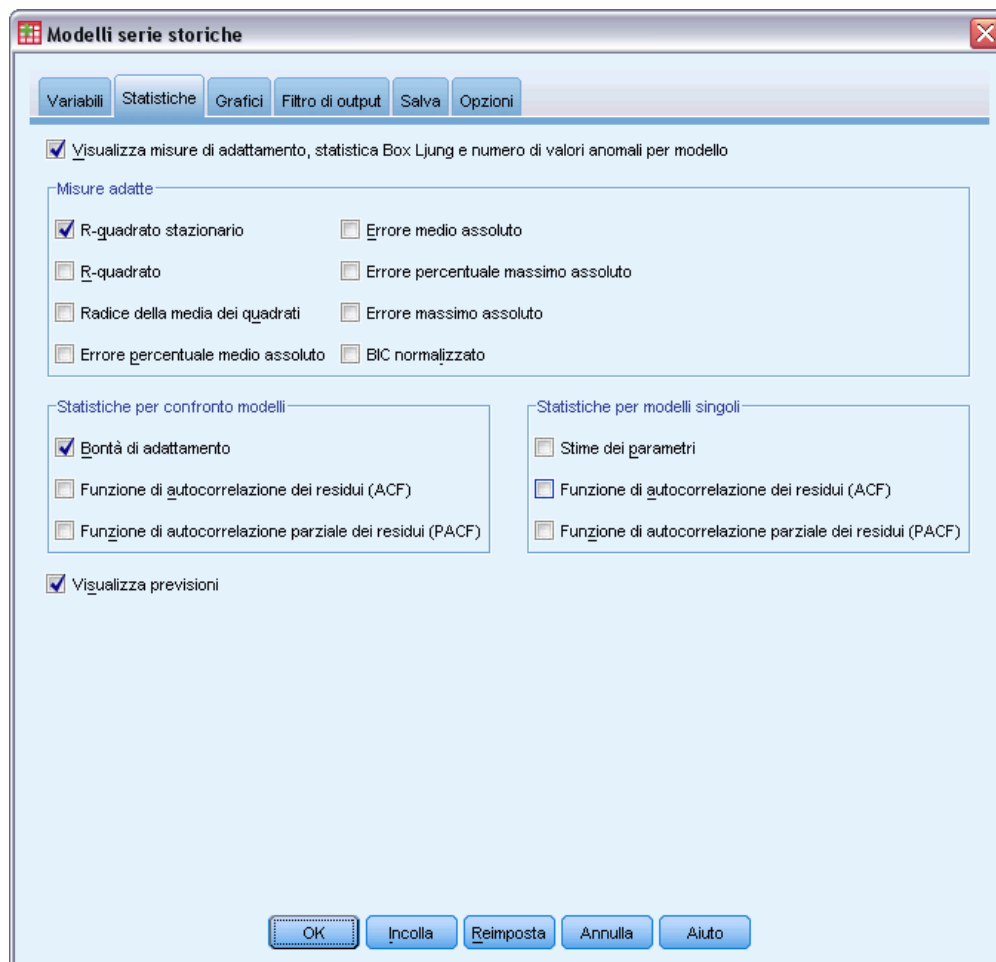
Le previsioni del modello vengono salvate come nuove variabili nel file di dati attivo, con il prefisso *Previsto* impostato come nome variabile. Le specifiche di ciascun modello possono essere salvate anche in un file XML esterno. Ciò può essere utile se si desidera riutilizzare i modelli per estendere le previsioni man mano che si rendono disponibili nuovi dati.

- Fare clic sul pulsante *Sfoggia* nella scheda *Salva*.

Viene aperta la finestra di dialogo standard in cui è possibile salvare il file.

- Individuare la cartella in cui salvare il file del modello XML, immettere il nome del file, quindi fare clic su *Salva*.
- Fare clic sulla scheda *Statistiche*.

Figura 6-7  
Time Series Modeler, scheda Statistiche



- Selezionare Visualizza previsioni.

Questa opzione consente di creare una tabella dei valori stimati per ciascuna serie di variabili dipendenti e fornisce un'alternativa per ottenere questi valori senza salvare le previsioni come nuove variabili.

La selezione predefinita di Bontà dell'adattamento (nel gruppo Statistiche per il confronto modelli) crea una tabella con statistiche di adattamento - —come  $R$  quadrato, media assoluta, errore percentuale assoluto e BIC— normalizzato - calcolate per tutti i modelli. Inoltre, fornisce un'indicazione generale del grado di adattamento del modello ai dati.

- Fare clic sulla scheda Grafici.

Figura 6-8  
Time Series Modeler, scheda Grafici



- Deselezionare Serie nel gruppo Grafici per modelli singoli.

Questa selezione evita che vengano generati grafici per le serie di ciascun modello. Tuttavia, lo scopo di questo esempio non è quello di salvare le previsioni come nuove variabili, ma di generare grafici per le previsioni.

Il gruppo Grafici per il confronto modelli fornisce numerosi grafici (sotto forma di istogrammi) delle statistiche di adattamento calcolate per tutti i modelli.

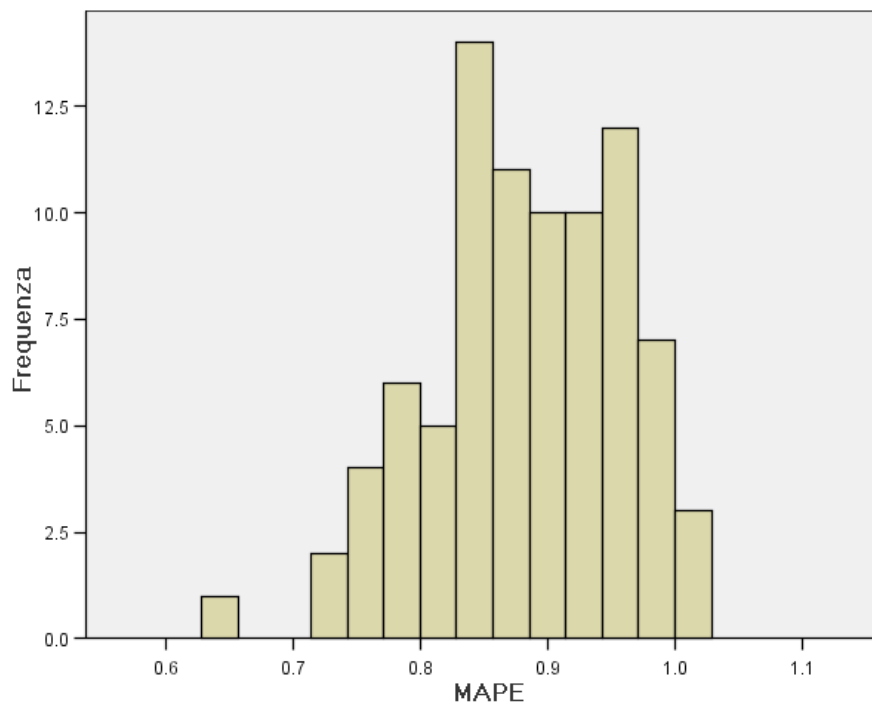
- Selezionare Errore percentuale assoluto medio e Errore percentuale assoluto massimo nel gruppo Grafici per il confronto modelli.

L'errore percentuale assoluto indica di quanto varia la serie dipendente rispetto al livello predittivo del modello. L'analisi degli errori medi e massimi dei modelli permette di avere un'indicazione del grado di incertezza delle previsioni. È generalmente preferibile esaminare i grafici riassuntivi degli errori percentuali anziché quelli relativi agli errori assoluti, poiché in questo caso la serie dipendente rappresenta il numero di abbonati di mercati di varie dimensioni.

- Fare clic su OK nella finestra di dialogo Time Series Modeler.

## Grafici riassuntivi dei modelli

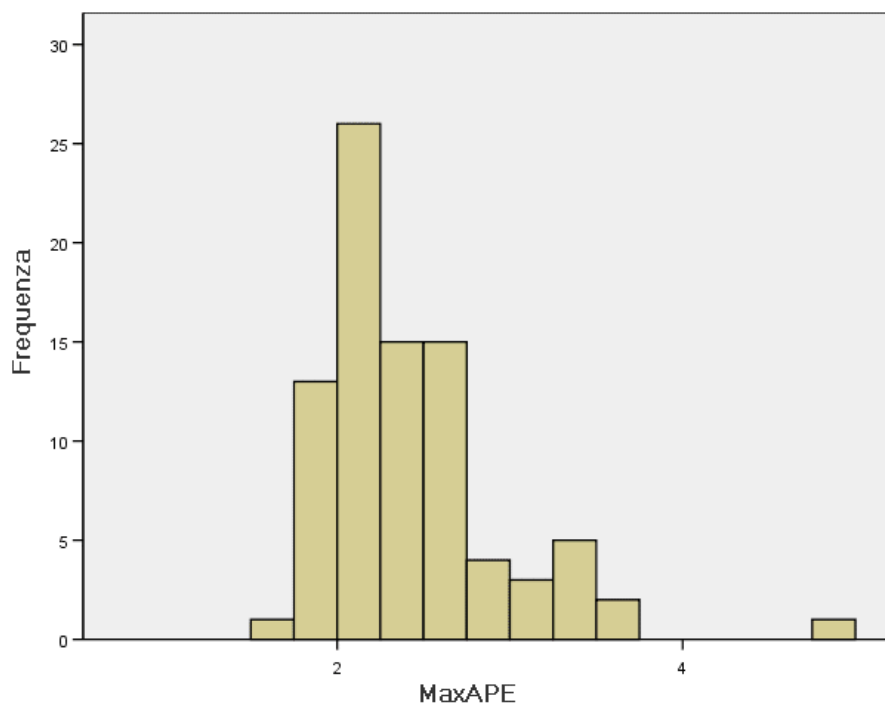
Figura 6-9  
Istogramma dell'errore percentuale assoluto medio



Questo istogramma visualizza l'errore percentuale assoluto medio (MAPE) dei modelli. In questo caso mostra che tutti i modelli presentano un grado di incertezza approssimativamente pari all'1%.



Figura 6-10  
Istogramma dell'errore percentuale assoluto massimo



Questo istogramma visualizza l'errore percentuale assoluto massimo (MaxAPE) dei modelli ed è utile per stimare gli scenari peggiori per le previsioni. In questo caso mostra che l'errore percentuale maggiore di ciascun modello è compreso nell'intervallo da 1 a 5%. Questi valori rappresentano un grado di incertezza accettabile? Questo è il momento in cui entra in azione il senso per gli affari, poiché il concetto di rischio accettabile cambia a seconda delle situazioni.

## Previsioni dei modelli

Figura 6-11  
Nuove variabili contenenti le previsioni dei modelli

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

L'Editor dei dati visualizza le nuove variabili che contengono le previsioni dei modelli. Benché l'esempio non mostri solo due variabili, ci sono 85 nuove variabili, ovvero una per ciascuna delle 85 serie dipendenti. I nomi delle variabili contengono il prefisso predefinito *Previsto*, seguito dal nome associato alla variabile dipendente (ad esempio *Mercato\_1*) e dall'identificatore del modello (ad esempio *Modello\_1*).

All'insieme dei dati sono stati aggiunti tre nuovi casi che contengono le previsioni per il periodo compreso tra il gennaio e il marzo 2004, oltre alle etichette delle date generate automaticamente. Ciascuna nuova variabile contiene le previsioni dei modelli per il periodo di stima (da gennaio 1999 a dicembre 2003) e permette quindi di stabilire se il modello è adatto o meno ai valori noti.

**Figura 6-12**  
*Tabella delle previsioni*

Modello		JAN 2004	FEB 2004	MAR 2004
Market_1-Modello_1	Previsione	11503	11447	11390
	UCL	11686	11767	11870
	LCL	11321	11126	10910
Market_2-Modello_2	Previsione	54893	55856	56704
	UCL	55632	57195	58575
	LCL	54154	54518	54832
Market_3-Modello_3	Previsione	59656	59305	58954
	UCL	60457	60753	61158
	LCL	58856	57857	56750
Market_4-Modello_4	Previsione	18235	18424	18628
	UCL	18413	18731	19121
	LCL	18058	18116	18136

È possibile anche scegliere di creare una tabella con i valori stimati. La tabella comprenderà i valori attesi per il periodo di previsione ma, — a differenza delle nuove variabili che contengono le previsioni del modello—, non comprende i valori stimati del periodo di stima. I risultati sono disposti per modello e identificati con il nome del modello, che comprende il nome (o l'etichetta) della variabile dipendente associata seguito dall'identificatore del modello, al pari dei nomi delle nuove variabili che contengono le previsioni del modello. La tabella comprende anche i limiti di confidenza superiori e i limiti di confidenza inferiori per i valori delle previsioni (pari al 95% per impostazione predefinita).

In questo esempio sono stati illustrati due approcci diversi per ottenere valori di previsione: il salvataggio delle previsioni come nuove variabili nel file di dati attivo e la creazione di una tabella delle previsioni. Entrambi questi approcci consentono di ottenere varie opzioni per esportare le previsioni (ad esempio in un foglio di calcolo di Excel).

## Riepilogo

In questo esempio è stato mostrato come usare Expert Modeler per creare previsioni per più serie e come salvare i modelli risultanti in un file XML esterno. Nell'esempio successivo, vedremo come estendere le previsioni quando si rendono disponibili nuovi dati, senza ricreare i modelli, grazie alla procedura Applica modelli delle serie storiche.

## ***Uso dei modelli salvati per l'esecuzione di nuove previsioni di massa***

Si supponga di aver usato Time Series Modeler per creare modelli per i dati delle serie storiche e per produrre previsioni iniziali basate sui dati disponibili. Per poter riutilizzare i modelli ed estendere le previsioni man mano che si rendono disponibili nuovi dati, si è deciso di salvare i modelli in un file esterno. A questo punto è possibile applicare i modelli salvati.

Questo esempio può essere considerato la continuazione naturale dell'esempio precedente, [Uso di Expert Modeler per le previsioni di massa in il capitolo 6 a pag. 46](#), ma può essere usato anche in modo indipendente. In questo scenario, si supponga di essere un analista di un fornitore nazionale di servizi a banda larga e di dover produrre delle previsioni mensili relative agli abbonati per ciascuno degli 85 mercati regionali. Si supponga di aver già usato Expert Modeler per creare modelli e previsioni per i successivi tre mesi, di aver aggiornato il proprio database con dati più recenti rispetto al periodo di previsione originale e di voler usare tali dati per ampliare l'orizzonte della previsione di altri tre mesi.

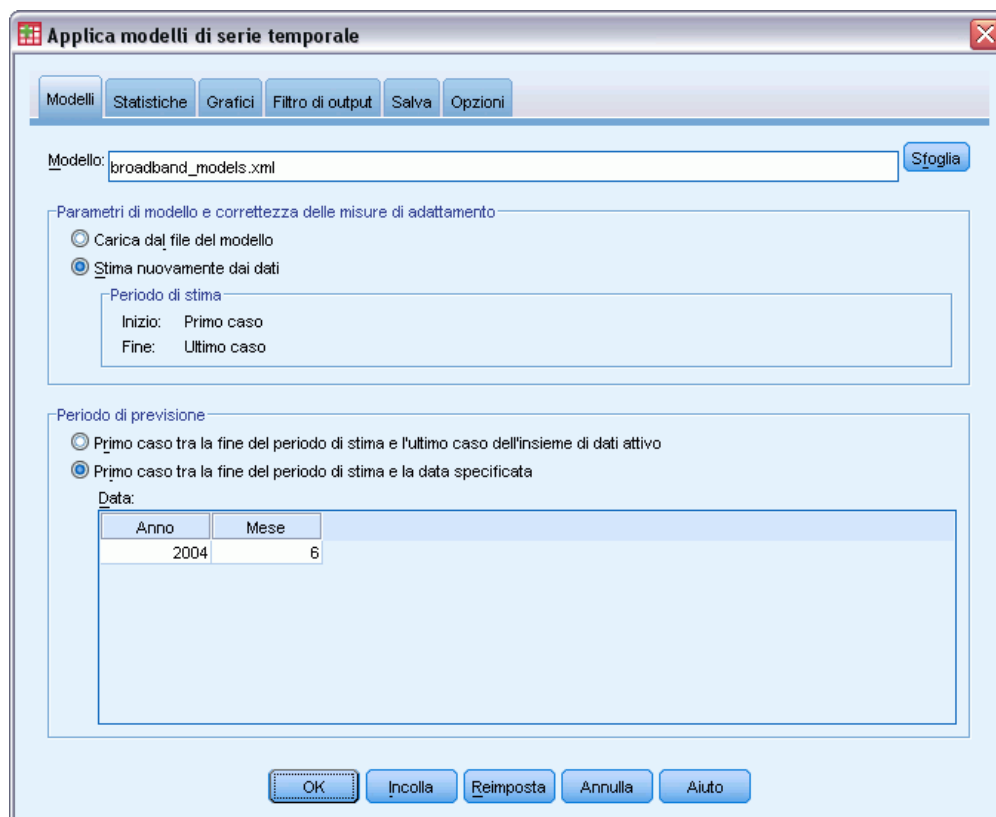
I dati storici mensili aggiornati sono contenuti nel file *broadband\_2.sav*, mentre i modelli salvati sono contenuti nel file *broadband\_models.xml*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice D in IBM SPSS Forecasting 21](#). Naturalmente se si utilizza l'esempio descritto in precedenza e si salva un file del modello personalizzato, è possibile usare il file salvato anziché *broadband\_models.xml*.

### ***Esecuzione dell'analisi***

Per applicare i modelli:

- ▶ Dai menu, scegliere:  
Analizza > Previsioni > Applica modelli...

Figura 7-1  
Finestra di dialogo *Applica modelli delle serie storiche*



- Fare clic su *Sfoggia*, accedere e selezionare *broadband\_models.xml* oppure scegliere il file personale per l'esempio descritto in precedenza. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice D in IBM SPSS Forecasting 21.](#)

- Selezionare *Calcola nuova stima per i dati*.

Per inserire i nuovi valori della serie storica nelle previsioni, la procedura *Applica modelli delle serie storiche* deve effettuare una nuova stima dei parametri del modello. La struttura dei modelli rimane la stessa, quindi il calcolo della nuova stima è molto più veloce rispetto al tempo richiesto per la creazione iniziale dei modelli.

L'insieme dei casi usati per la nuova stima deve comprendere i nuovi dati. Per ottenere questo risultato, usare il periodo di stima predefinito compreso tra il primo e l'ultimo caso. Per impostare un periodo di stima diverso da quello predefinito, selezionare *Basato su intervallo di tempo o di casi* nella finestra di dialogo *Seleziona casi*.

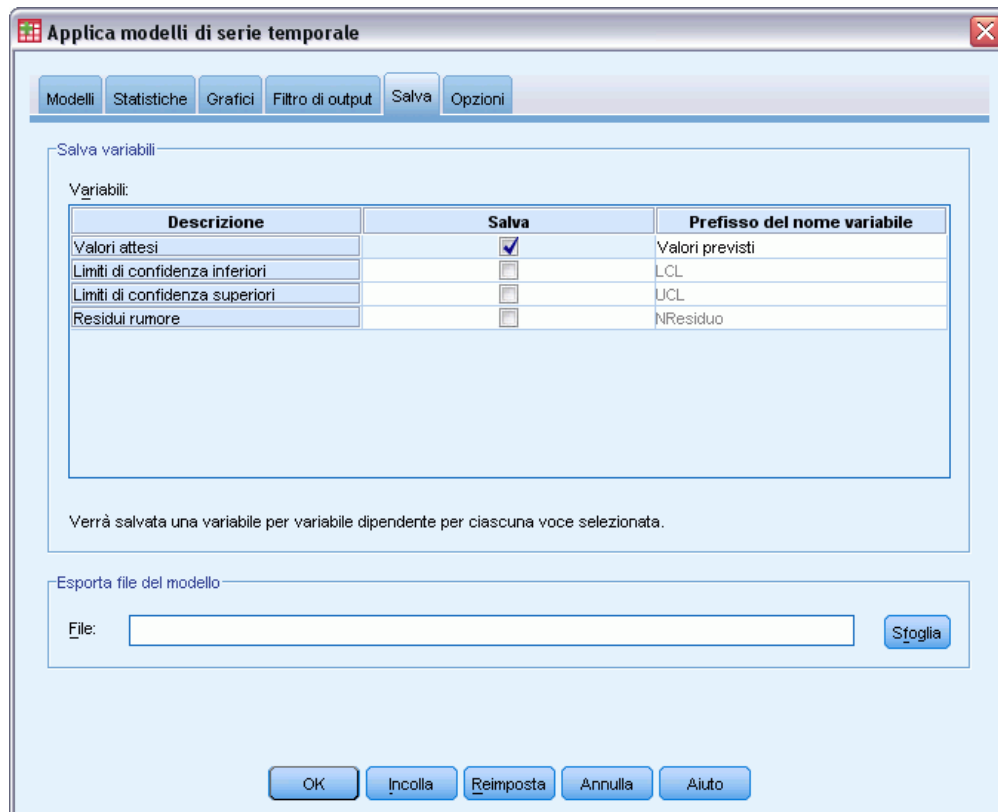
- Nel gruppo *Periodo di previsione* selezionare *Primo caso tra la fine del periodo di previsione e una data specifica*.
- Nella griglia *Data* immettere 2004 per l'anno e 6 per il mese.

L'insieme dei dati contiene tutti i dati riferiti al periodo compreso tra il gennaio 1999 e il marzo 2004. Con le impostazioni correnti, il periodo di previsione sarà compreso tra l'aprile 2004 e il giugno dello stesso anno.

- Fare clic sulla scheda Salva.

Figura 7-2

Applica modelli delle serie storiche, scheda Salva

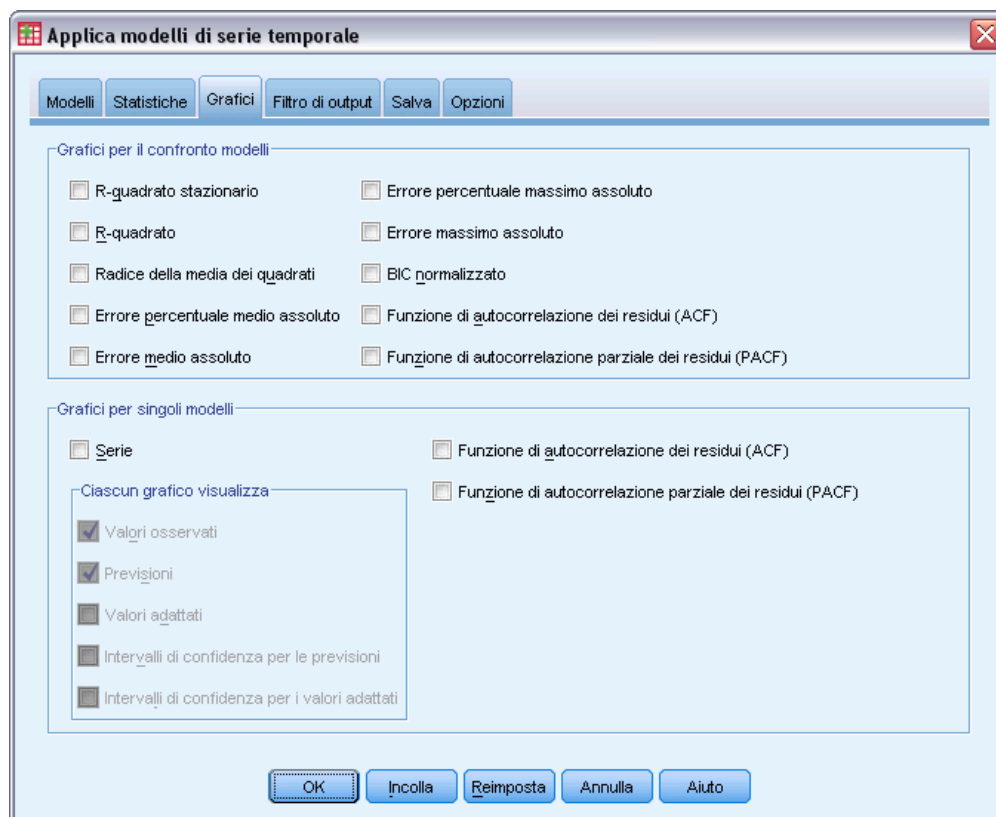


- Selezionare la voce per Valori attesi nella colonna Salva e lasciare il valore predefinito *Previsto* come prefisso del nome della variabile.

Le previsioni del modello vengono salvate come nuove variabili nel file di dati attivo, con il prefisso *Previsto* impostato come nome variabile.

- Fare clic sulla scheda Grafici.

Figura 7-3  
Applica modelli delle serie storiche, scheda Grafici



- Deselezionare Serie nel gruppo Grafici per modelli singoli.

Questa selezione evita che vengano generati grafici per le serie di ciascun modello. Tuttavia, lo scopo di questo esempio non è quello di salvare le previsioni come nuove variabili, ma di generare grafici per le previsioni.

- Fare clic su OK nella finestra di dialogo Applica modelli delle serie storiche.

## Statistiche dell'adattamento dei modelli

Figura 7-4  
Tabella Adattamento del modello

Statistiche di adattamento	Media	SE	Minimo	Massimo	Percentile						
					5	10	25	50	75	90	95
R-quadrato stazionario	,181	,141	-4,34E-016	,609	-9,59E-016	1,22E-016	,070	,193	,252	,359	,453
R-quadrato	,999	,000	,998	1,000	,999	,999	,999	,999	,999	,999	1,000
RMSE	187,766	140,959	45,048	764,698	52,558	59,848	93,155	138,913	205,742	433,883	493,904
MAPE	,896	,074	,669	1,026	,754	,792	,832	,894	,948	,976	,996
MaxAPE	2,446	,512	1,742	4,373	1,840	1,952	2,103	2,321	2,663	3,313	3,676
MAE	146,373	113,749	35,573	612,989	41,209	48,742	73,553	111,599	162,306	331,207	364,582
MaxAE	499,714	434,112	116,966	2143,993	131,261	144,918	226,960	346,221	574,071	1165,207	1587,793
BIC normalizzato	10,086	1,343	7,749	13,412	8,049	8,285	9,193	9,988	10,717	12,210	12,494

La tabella Adattamento del modello fornisce statistiche sull'adattamento del modello riferite a tutti i modelli. Questa tabella permette di verificare quanto i modelli, con i nuovi parametri stimati, si adattano ai dati. La tabella fornisce, per ciascuna statistica, la media, l'errore standard e i valori minimo e massimo di tutti i modelli. La tabella contiene anche il percentile che fornisce informazioni sulla distribuzione delle statistiche nei modelli. Per ciascun percentile, una certa percentuale di modelli ha un valore statistico di adattamento inferiore al valore specificato. Ad esempio il 95% dei modelli ha un valore MaxAPE (errore percentuale massimo assoluto) inferiore a 3,676.

Benché siano disponibili più statistiche, in questo caso ne esamineremo solo due: MAPE (errore percentuale medio assoluto) e MaxAPE (errore percentuale massimo assoluto). L'errore percentuale assoluto indica quanto una serie dipendente varia rispetto al livello stimato sul modello e fornisce un'indicazione dell'incertezza delle previsioni. L'errore percentuale medio assoluto varia da un minimo di 0,669% a un massimo di 1,026% per tutti i modelli. L'errore percentuale massimo assoluto varia da 1,742% a 4,373% per tutti i modelli. Quindi, l'incertezza media di ciascuna previsione del modello è pari a circa l'1%, mentre l'incertezza massima è intorno al 2,5% (ovvero al valore medio di MaxAPE), con uno scenario peggiore pari a circa il 4%. L'accettabilità dell'incertezza risultante dal valore dipende dal livello di rischio che si desidera assumere.

## Previsioni dei modelli

Figura 7-5  
Nuove variabili contenenti le previsioni dei modelli

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

L'Editor dei dati visualizza le nuove variabili che contengono le previsioni dei modelli. Benché l'esempio non mostri solo due variabili, ci sono 85 nuove variabili, ovvero una per ciascuna delle 85 serie dipendenti. I nomi delle variabili contengono il prefisso predefinito *Previsto*, seguito dal nome associato alla variabile dipendente (ad esempio *Mercato\_1*) e dall'identificatore del modello (ad esempio *Modello\_1*).

All'insieme dei dati sono stati aggiunti tre nuovi casi che contengono le previsioni per il periodo compreso tra l'aprile e il giugno 2004, oltre alle etichette delle date generate automaticamente.

## Riepilogo

In questo esempio si è appreso come applicare i modelli salvati per ampliare le previsioni esistenti senza ricreare i modelli, se si rendono disponibili nuovi dati. Ovviamente se si sospetta che il modello possa essere cambiato, è necessario ricrearlo utilizzando la procedura Time Series Modeler.

## ***Uso di Expert Modeler per la determinazione di predittori significativi***

Un'azienda di vendite per corrispondenza, interessata nello sviluppo di un modello di previsione, ha raccolto dati sulle vendite mensili di abbigliamento maschile oltre a numerose serie utilizzabili per spiegare alcune delle variazioni nelle vendite. I predittori possibili includono il numero di cataloghi inviati e il numero di pagine relativo, il numero di linee telefoniche attive per la ricezione degli ordini, l'importo speso per la pubblicità su carta stampata e il numero degli operatori dell'assistenza clienti. Uno o più di tali predittori sono utili per le previsioni?

Nell'esempio verrà usato Expert Modeler con tutti i predittori candidati per trovare il modello migliore. Poiché Expert Modeler seleziona solo i predittori che hanno solo relazioni statisticamente significative con la serie dipendente, consentirà anche di individuare i predittori utili e di creare il modello più idoneo per stimarli. Al termine, potrebbe essere utile utilizzare l'esempio che segue, [Uso dei modelli salvati per l'esecuzione di prove in il capitolo 9 a pag. 73](#), che esamina l'effetto dell'uso di scenari con predittori diversi sulle vendite utilizzando il modello creato nell'esempio.

I dati dell'esempio corrente sono contenuti nel file *catalog\_seasfac.sav*. Per ulteriori informazioni, vedere l'argomento [File di esempio in l'appendice D in IBM SPSS Forecasting 21](#).

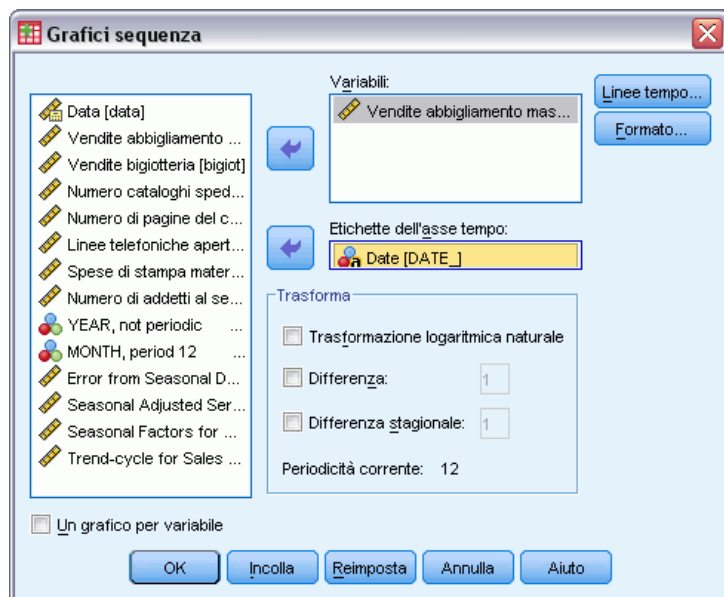
### ***Inserimento dei dati nel grafico***

È sempre consigliabile inserire i dati in un grafico, soprattutto se si lavora con un'unica serie:

- ▶ Dai menu, scegliere:  
Analizza > Previsioni > Grafici sequenza...

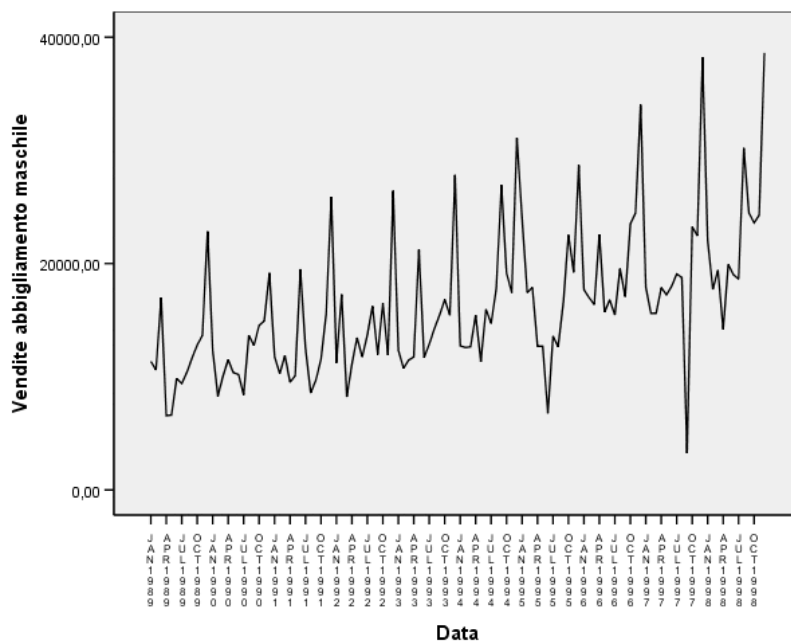


Figura 8-1  
Finestra di dialogo Grafici sequenza



- ▶ Selezionare *Vendite abbigliamento maschile* e spostare i dati nell'elenco Variabili.
- ▶ Selezionare *Data* e passare all'elenco Etichette dell'asse tempo.
- ▶ Fare clic su OK.

Figura 8-2  
Vendite di abbigliamento maschile (in Dollari USA)



La serie presenta numerosi picchi, molti dei quali equamente distanziati, nonché un chiaro trend di crescita. La presenza di picchi equamente distanziati suggerisce la presenza di un componente periodico nella serie storica. Data la natura stagionale delle vendite, con picchi che si verificano normalmente durante le vacanze natalizie, non è sorprendente rilevare un componente stagionale annuale nei dati.

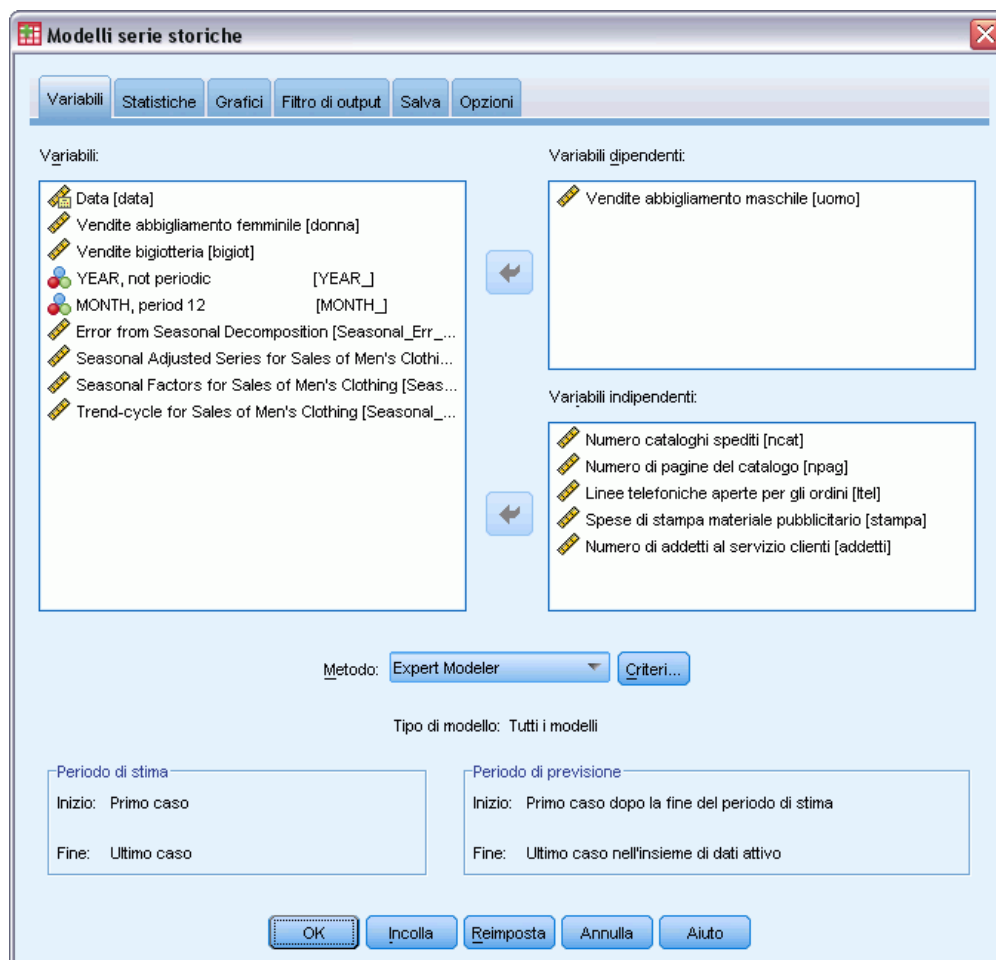
Sono presenti anche picchi non correlati all'andamento stagionale, che rappresentano deviazioni significative rispetto ai punti dati vicini. Questi punti possono essere valori anomali che possono e devono essere gestiti tramite Expert Modeler.

## ***Esecuzione dell'analisi***

Per usare Expert Modeler:

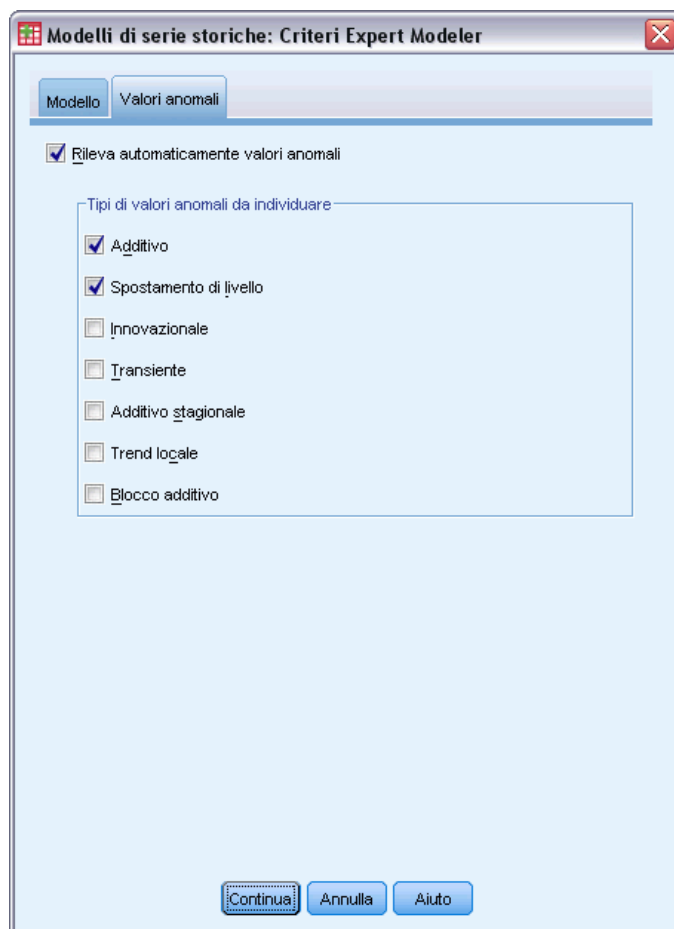
- Dai menu, scegliere:  
Analizza > Previsioni > Crea modelli...

Figura 8-3  
finestra di dialogo Serie storiche



- ▶ Selezionare *Vendite di abbigliamento maschile* come variabile dipendente.
- ▶ Selezionare i campi da *Numero di cataloghi spediti* a *Numero degli operatori dell'assistenza clienti* come variabili indipendenti.
- ▶ Verificare di aver selezionato Expert Modeler nell'elenco a discesa Metodo. Expert Modeler ricerca automaticamente il miglior modello stagionale o non stagionale per la serie di variabili dipendenti.
- ▶ Fare clic su Criteri, quindi scegliere la scheda Valori anomali.

Figura 8-4  
Finestra di dialogo Criteri di Expert Modeler, scheda Valori Anomali

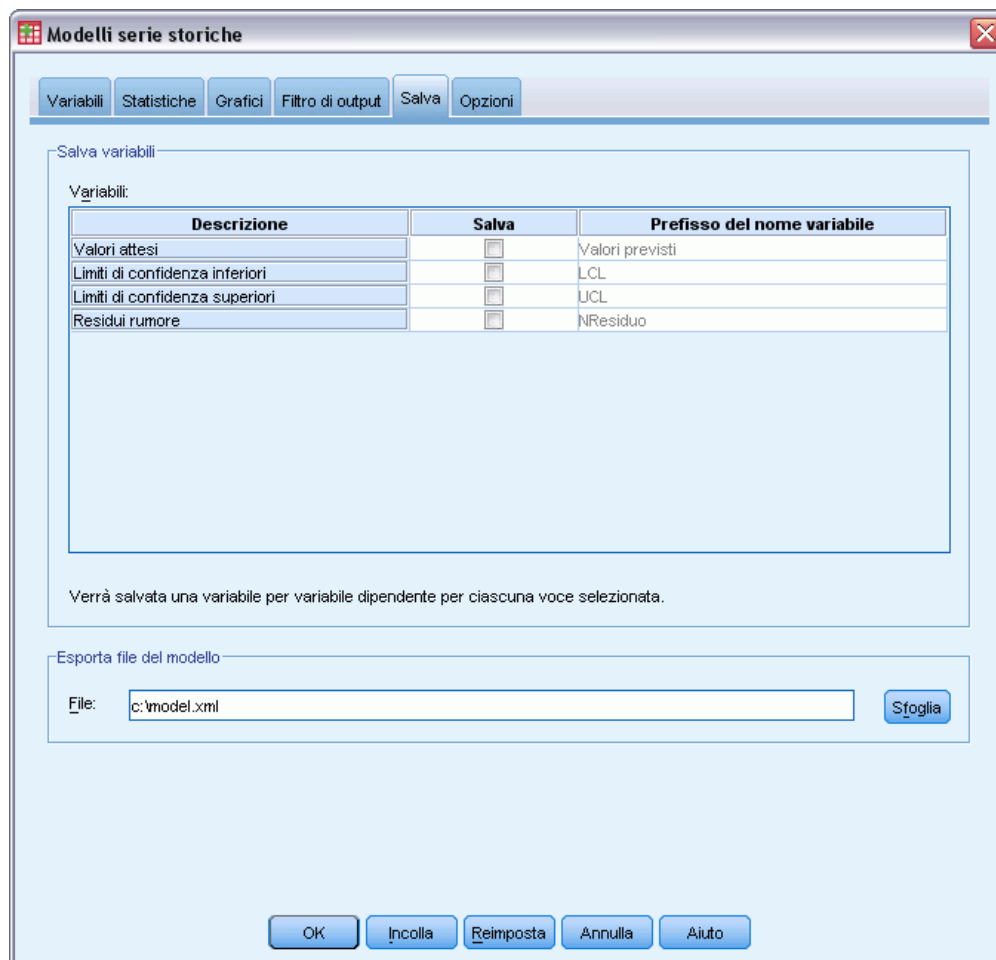


- Selezionare Rileva automaticamente valori anomali e lasciare le impostazioni predefinite per i tipi di valori anomali da rilevare.

L'analisi visiva dei dati ha evidenziato la probabile presenza di valori anomali. Tenendo conto delle impostazioni selezionate, Expert Modeler ricercherà i tipi di valori anomali più comuni e li inserirà nel modello finale. Poiché il rilevamento di valori anomali può incrementare significativamente i tempi di elaborazione di Expert Modeler, questa funzione deve essere usata con molta cautela, soprattutto se si creano contemporaneamente modelli di più serie. Per impostazione predefinita, i valori anomali non vengono generalmente rilevati.

- Fare clic su Continua.
- Fare clic sulla scheda Salva nella finestra di dialogo Time Series Modeler.

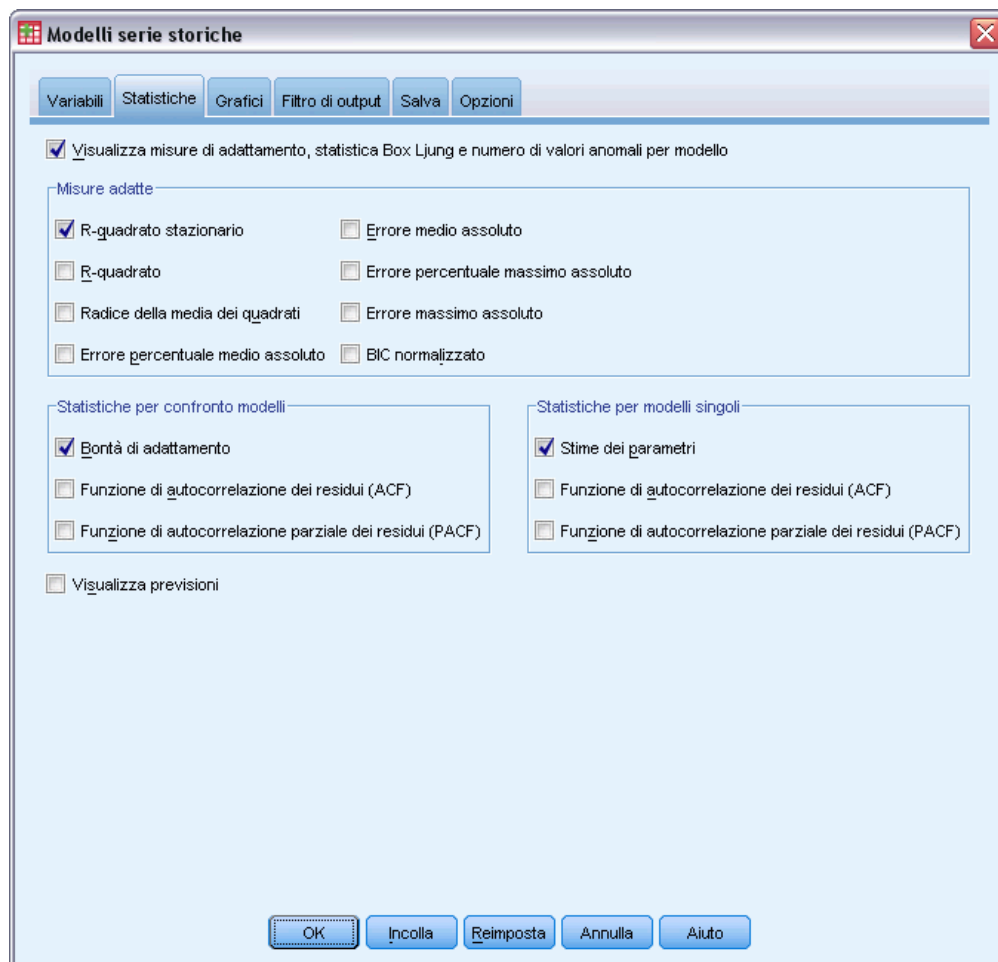
Figura 8-5  
Time Series Modeler, scheda Salva



Può essere utile salvare il modello stimato in un file XML esterno in modo da poter utilizzare predittori diversi — con la procedura Applica modelli delle serie storiche — senza dover ricreare il modello.

- ▶ Fare clic sul pulsante Sfoglia nella scheda Salva.  
Viene aperta la finestra di dialogo standard in cui è possibile salvare il file.
- ▶ Individuare la cartella in cui salvare il file del modello XML, immettere il nome del file, quindi fare clic su Salva.
- ▶ Fare clic sulla scheda Statistiche.

Figura 8-6  
Time Series Modeler, scheda Statistiche

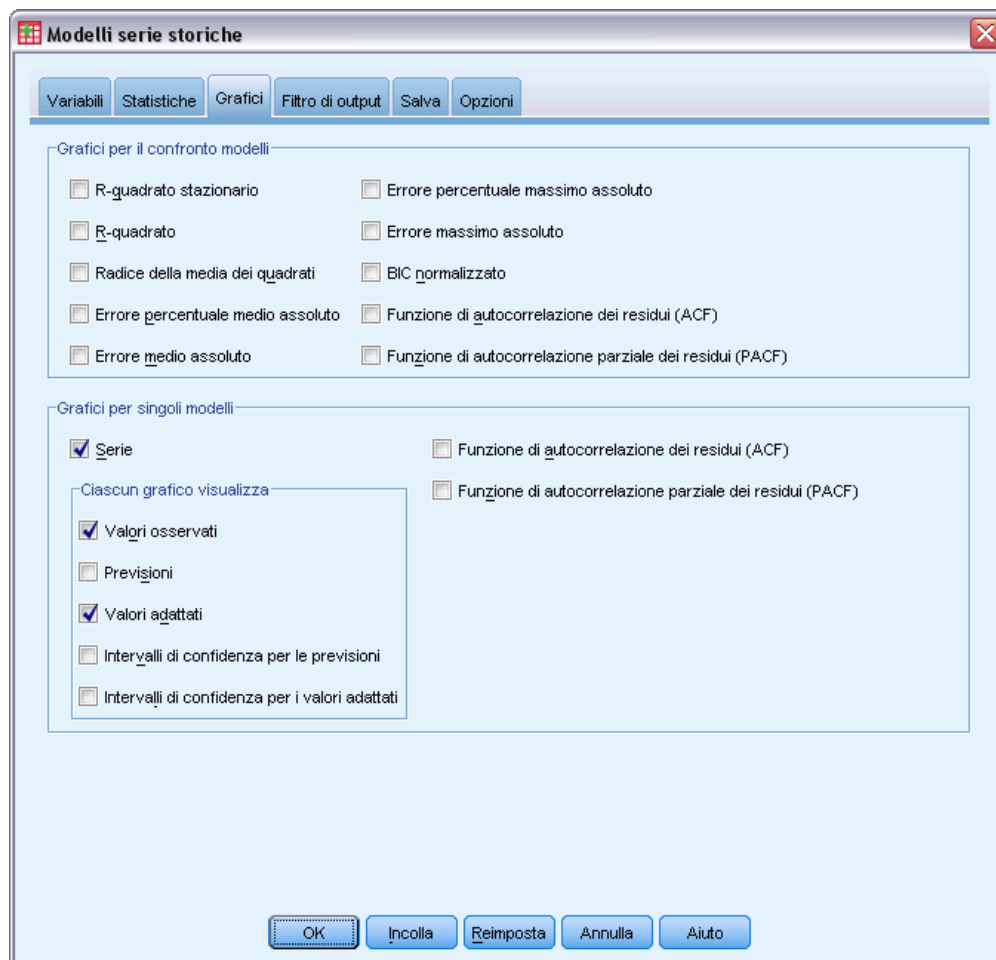


- Selezionare Stime dei parametri.

Questa opzione consente di creare una tabella che visualizza tutti i parametri, compresi i predittori significativi, del modello scelto da Expert Modeler.

- Fare clic sulla scheda Grafici.

Figura 8-7  
Time Series Modeler, scheda Grafici



- Deselezionare Previsioni.

Lo scopo di questo esempio è quello di individuare i predittori significativi e creare un modello, non quello di calcolare previsioni.

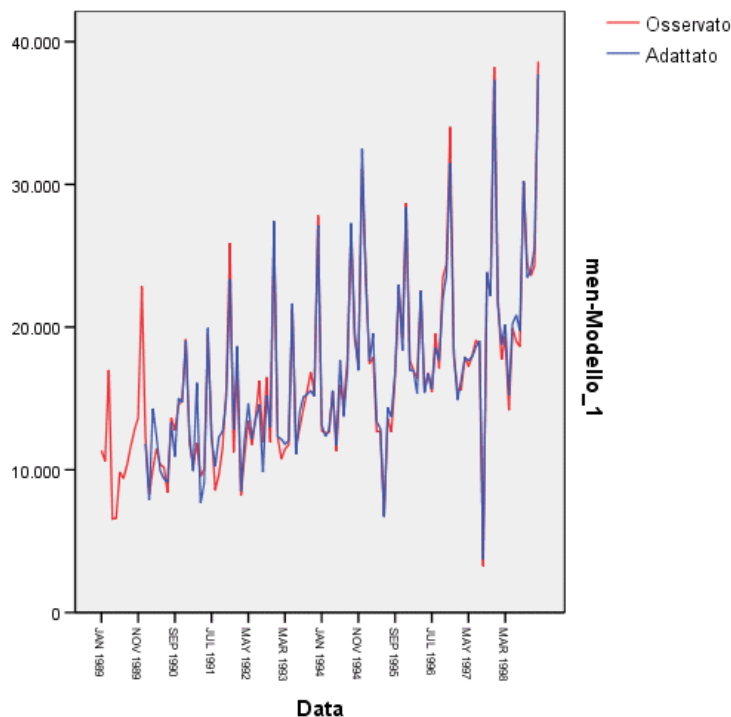
- Selezionare Valori di adattamento.

Questa opzione consente di visualizzare i valori attesi riferiti al periodo usato per stimare il modello. Questo periodo viene chiamato **periodo di stima** e comprende tutti i casi del file dati attivo usato per l'esempio. Questi valori forniscono un'indicazione di come il modello si adatta ai valori osservati e vengono quindi chiamati **valori di adattamento**. Il grafico risultante comprende sia i valori osservati che quelli di adattamento.

- Fare clic su OK nella finestra di dialogo Time Series Modeler.

## Grafico di una serie storica

Figura 8-8  
Valori stimati e osservati



Poiché i valori attesi sono ben allineati ai valori osservati, il modello ha una buona capacità predittiva. Vale la pena notare che il modello è in grado di fornire una buona stima dei picchi stagionali nonché di cogliere il trend di crescita dei dati.

## Tabella descrittiva del modello

Figura 8-9  
Tabella descrittiva del modello

ID modello	men	Modello_1	Tipo di modello
			ARIMA(0,0,0)(0,1,0)

La tabella descrittiva del modello contiene una voce per ciascun modello stimato oltre all'identificatore e al tipo di modello. L'identificatore del modello comprende il nome (o l'etichetta) della variabile dipendente associata e il nome assegnato dal sistema. In questo esempio la variabile dipendente è *Vendite di abbigliamento maschile*, mentre il nome assegnato dal sistema è *Modello\_1*.

Modelli serie storiche supporta sia il livellamento esponenziale che i modelli ARIMA. I tipi di modelli di livellamento esponenziale sono elencati con il loro nome comune come Holt e Additivo di Winters. I tipi di modelli ARIMA vengono elencati utilizzando la notazione standard di ARIMA  $(p,d,q)(P,D,Q)$ , dove  $p$  è l'ordine dell'autoregressione,  $d$  è l'ordine delle differenze (o delle integrazioni) e  $q$  è l'ordine delle medie mobili.  $P,D$  e  $Q$  sono i corrispondenti stagionali.



Expert Modeler ha determinato che il modo migliore per descrivere le vendite di abbigliamento maschile consiste nell'usare un modello ARIMA stagionale con un solo ordine di differenze. La natura stagionale del modello permette di tener conto dei picchi stagionali presenti nel grafico della serie, mentre l'unico ordine di differenze riflette l'andamento in crescita osservato anche nei dati.

## Tabella delle statistiche del modello

Figura 8-10  
Tabella delle statistiche del modello

Modello	Numero di predittori	Statistiche di adattamento del modello	Q di Ljung Box (18)			Numero di valori anomali
		R-quadrato stazionario	Statistiche	DF	Sig.	
men-Modello_1	2	,948	7,589	18	,984	9

La tabella delle statistiche del modello fornisce informazioni riassuntive e statistiche sulla bontà di adattamento per ciascun modello stimato. I risultati di ciascun modello vengono etichettati utilizzando l'identificatore del modello fornito nella tabella descrittiva del modello sesso. Come prima cosa è utile notare che il modello contiene due dei cinque predittori candidati specificati in origine. Ciò indica che Expert Modeler ha probabilmente identificato due variabili indipendenti che potrebbero risultare utili per le previsioni.

Benché Time Series Modeler fornisca diverse statistiche della bontà di adattamento, per l'esempio è stato scelto solo il valore relativo all'*R* quadrato. Questa statistica fornisce una stima della proporzione della variazione totale nella serie, illustrata nel modello, ed è preferibile al valore *R* quadrato standard se la serie contiene un trend o un andamento stagionale, come nel caso dell'esempio. Valori più elevati di *R* quadrato stazionario (fino a un massimo di 1) indicano un migliore adattamento. Un valore di 0,948 indica che il modello illustra perfettamente la variazione osservata nella serie.

La statistica Ljung-Box, nota anche con il nome di statistica di Box-Pierce, indica se il modello è stato correttamente specificato o meno. Un valore di significatività inferiore a 0,05 implica che la struttura della serie osservata non è stata presa in considerazione dal modello. Poiché il valore di 0,984 contenuto nell'esempio non è significativo, il modello selezionato è stato sicuramente correttamente specificato.

Expert Modeler ha rilevato nove punti che possono essere considerati valori anomali. Tuttavia, poiché il modello di ciascuno di questi punti è stato creato in modo corretto, non è necessario rimuoverli dalla serie.

## Tabella dei parametri del modello ARIMA

Figura 8-11  
Tabella dei parametri del modello ARIMA

				Stima	SE	t	Sig.
men-Modello_1	men	Nessuna trasformazione	Differenza stagionale	1			
	mail	Nessuna trasformazione	Numeratore Lag 0	1,549	,071	21,943	,000
			Differenza stagionale	1			
	phone	Nessuna trasformazione	Numeratore Lag 0	315,262	15,298	20,607	,000
			Differenza stagionale	1			

La tabella dei parametri del modello ARIMA visualizza tutti i parametri del modello, con una voce per ciascun modello stimato dall'identificatore del modello. Per questo motivo, elenca tutte le variabili del modello, comprese le variabili dipendenti e indipendenti che Expert Modeler ha individuato come significative. Dalla tabella delle statistiche abbiamo rilevato che sono presenti due predittori significativi. La tabella dei parametri del modello indica che questi predittori sono *Numero di cataloghi spediti* e *Numero di linee attive per la ricezione degli ordini*.

## **Riepilogo**

Nell'esempio si è visto come usare Expert Modeler per creare un modello e identificare i predittori significativi nonché come salvare il modello risultante in un file esterno. A questo punto è possibile usare la procedura *Applica modelli delle serie storiche* per vedere come adattare le serie predittive ai vari scenari e analizzare l'influenza delle singole alternative sulle previsioni di vendita.

## ***Uso dei modelli salvati per l'esecuzione di prove***

Negli esempi precedenti Time Series Modeler è stato usato per creare un modello per i dati e per stabilire quali predittori possono essere usati per le previsioni. Poiché i predittori rappresentano fattori gestibili dall'utente, è possibile effettuare prove con i valori del periodo di previsione per vedere l'effetto risultante sulle previsioni della variabile dipendente. Questa operazione può essere facilmente eseguita con la procedura Applica modelli delle serie storiche, usando il file del modello creato con la procedura Time Series Modeler.

Questo esempio può essere considerato la continuazione naturale dell'esempio precedente, [Uso di Expert Modeler per la determinazione di predittori significativi in il capitolo 8 a pag. 62.](#), ma può essere usato anche in modo indipendente. Scenario: una società che effettua vendite per corrispondenza ha raccolto i dati sulle vendite mensili di abbigliamento maschile dal gennaio 1989 al dicembre 1998, oltre a numerose serie che riteneva potessero essere usate come predittori per le vendite future. Expert Modeler ha stabilito che solo due dei cinque predittori candidati sono significativi: il numero di cataloghi inviati per posta e il numero di linee telefoniche attive per la ricezione degli ordini.

Al momento di pianificare la strategia di vendite per l'anno successivo, si scopre di avere risorse limitate per la stampa dei cataloghi e il mantenimento di linee telefoniche per la ricezione degli ordini. Il budget disponibile per i primi tre mesi del 1999 consente solo di inviare 2000 nuovi cataloghi oppure di aprire altre 5 linee telefoniche rispetto alle proiezioni iniziali. Quale di queste opzioni è più adatta per incrementare le vendite in questo periodo di tre mesi?

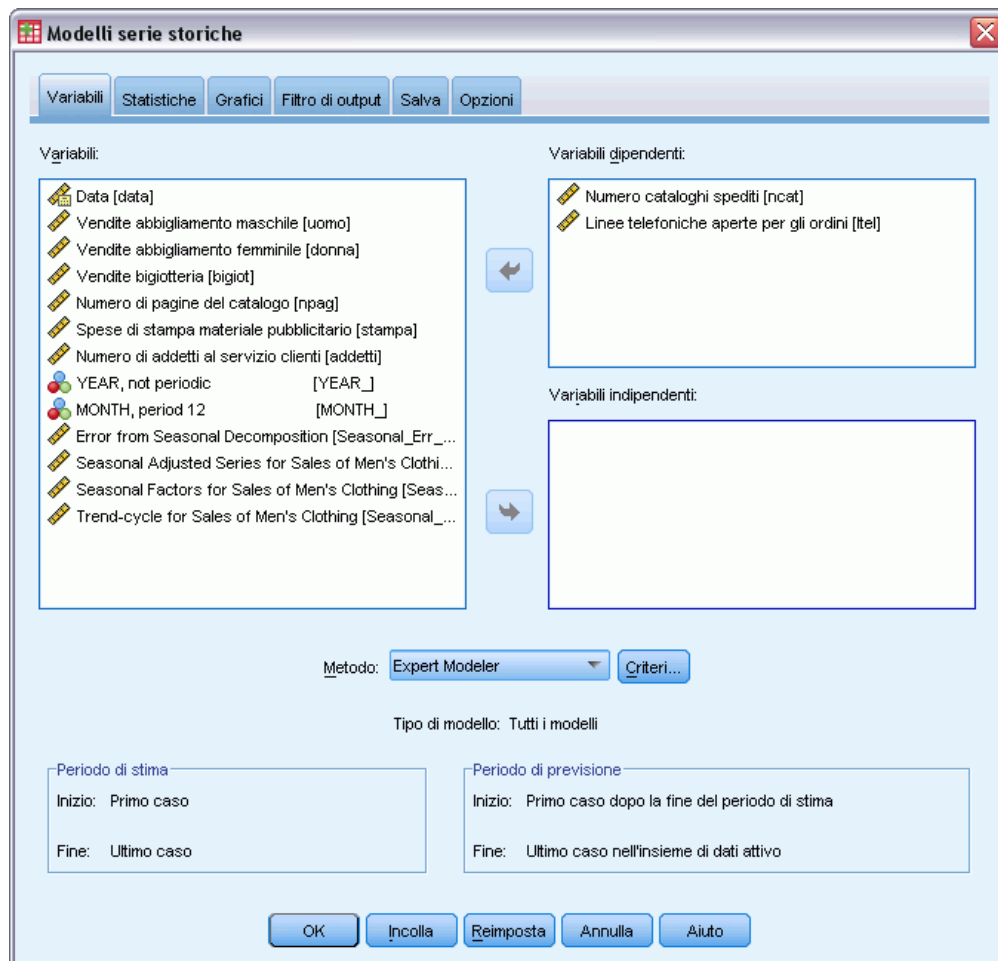
I dati di questo esempio sono contenuti nel file *catalog\_seasfac.sav*, mentre il file *catalog\_model.xml* contiene il modello delle vendite mensili creato con Expert Modeler. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice D in IBM SPSS Forecasting 21.](#) Naturalmente se si utilizza l'esempio precedente e si salva il file del modello, è possibile usare il file salvato anziché *catalog\_model.xml*.

### ***Ampliamento della serie dei predittori***

Quando si creano previsioni per una serie dipendente con predittori, ciascuna serie di predittori deve essere estesa a tutto il periodo di previsione. A meno che non si conoscano esattamente i valori futuri dei predittori, è infatti necessario stimarli. Successivamente è possibile modificare le stime per effettuare più prove con scenari riferiti a predittori diversi. Le proiezioni iniziali possono essere facilmente create con Expert Modeler.

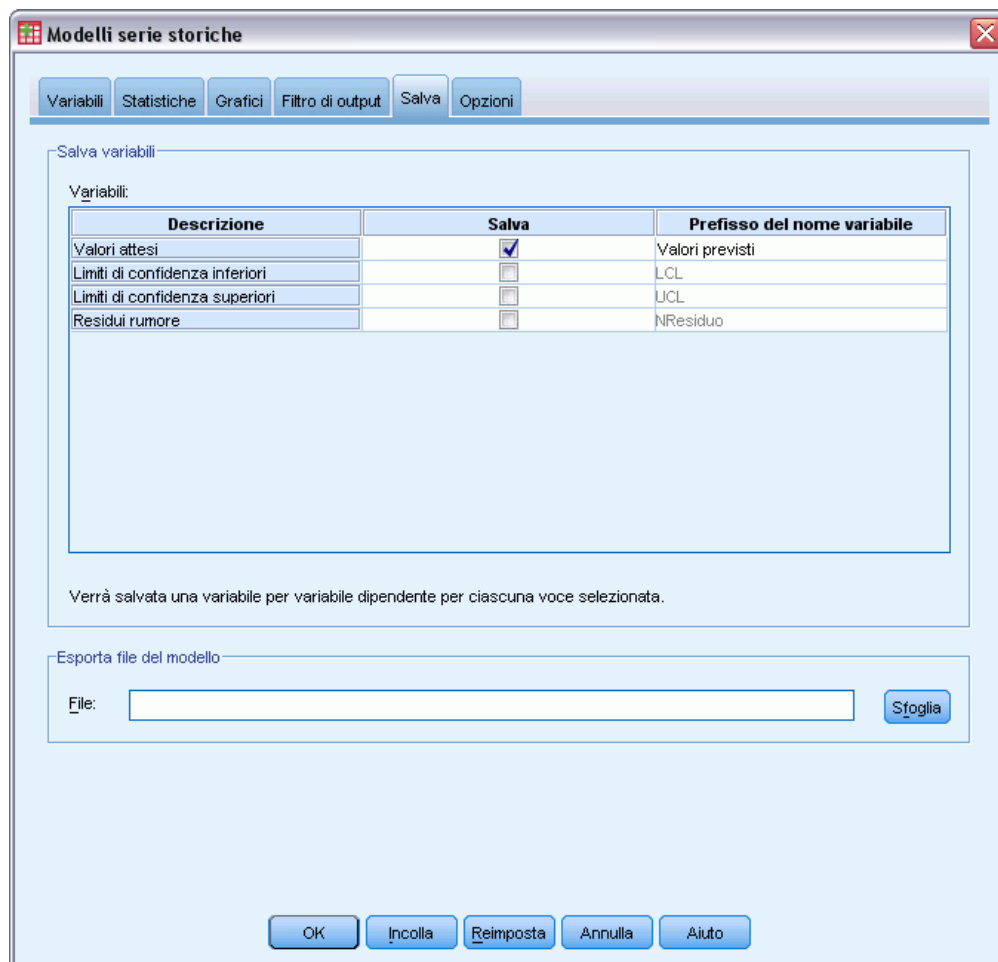
- Dai menu, scegliere:  
Analizza > Previsioni > Crea modelli...

Figura 9-1  
finestra di dialogo Serie storiche



- ▶ Selezionare *Numero di cataloghi inviati* e *Numero di linee attive per la ricezione degli ordini* come variabili dipendenti.
- ▶ Fare clic sulla scheda *Salva*.

Figura 9-2  
Time Series Modeler, scheda Salva



- Nella colonna *Salva* selezionare la voce corrispondente a Valori attesi e lasciare il valore predefinito *Previsto* come prefisso del nome della variabile.
- Fare clic sulla scheda Opzioni.

Figura 9-3  
Time Series Modeler, scheda Opzioni

Modelli serie storiche

Variabili Statistiche Grafici Filtro di output Salva Opzioni

Periodo di previsione

Primo caso tra la fine del periodo di stima e l'ultimo caso dell'insieme di dati attivo

Primo caso tra la fine del periodo di stima e la data specificata

Data:

Anno	Mese
1999	3

Valori mancanti definiti dall'utente

Considera non valido

Considera valido

Larghezza intervallo di confidenza (%): 95

Prefisso per identificatori di modello nell'output: Modello

Numero massimo di ritardi mostrati nell'output ACF e PACF: 24

OK Incolla Reimposta Annulla Aiuto

- Nel gruppo Periodo di previsione selezionare Primo caso tra la fine del periodo di previsione e una data specifica.
- Nella griglia Data immettere 1999 per l'anno e 3 per il mese.  
L'insieme di dati contiene i dati compresi tra il gennaio 1989 e il dicembre 1998, quindi il periodo di previsione relativo alle impostazioni correnti sarà quello compreso tra il gennaio 1999 e il marzo dello stesso anno.
- Fare clic su OK.

Figura 9-4

Nuove variabili contenenti le previsioni per le serie dei predittori

	Predicted_mail_Model_1	Predicted_phone_Model_2
121	11742	51
122	11853	45
123	11965	45

Visualizzazione dati Visualizzazione variabili

L'Editor dei dati mostra le nuove variabili *Modello\_posta\_previsto\_1* e *Modello\_posta\_previsto\_2*, che contengono i valori attesi del modello per il numero di cataloghi inviato e il numero di linee telefoniche attive. Per ampliare la serie dei predittori, è sufficiente utilizzare i valori compresi tra il gennaio e il marzo 1999, ovvero 121 dei 123 casi.

- ▶ Copiare i valori di questi tre casi da *Modello\_posta\_previsto\_1* e aggiungerli alla variabile *posta*.
- ▶ Ripetere questa procedura per *Modello\_posta\_previsto\_2*, copiando gli ultimi tre casi e aggiungendoli alla variabile *telefono*.

Figura 9-5

Serie dei predittori estese al periodo di previsione

	mail	page	phone	print	service	YEAR_	MONTH	DATE_
121	11742	.	51	.	.	1999	1	JAN 1999
122	11853	.	45	.	.	1999	2	FEB 1999
123	11965	.	45	.	.	1999	3	MAR 1999

Visualizzazione dati Visualizzazione variabili

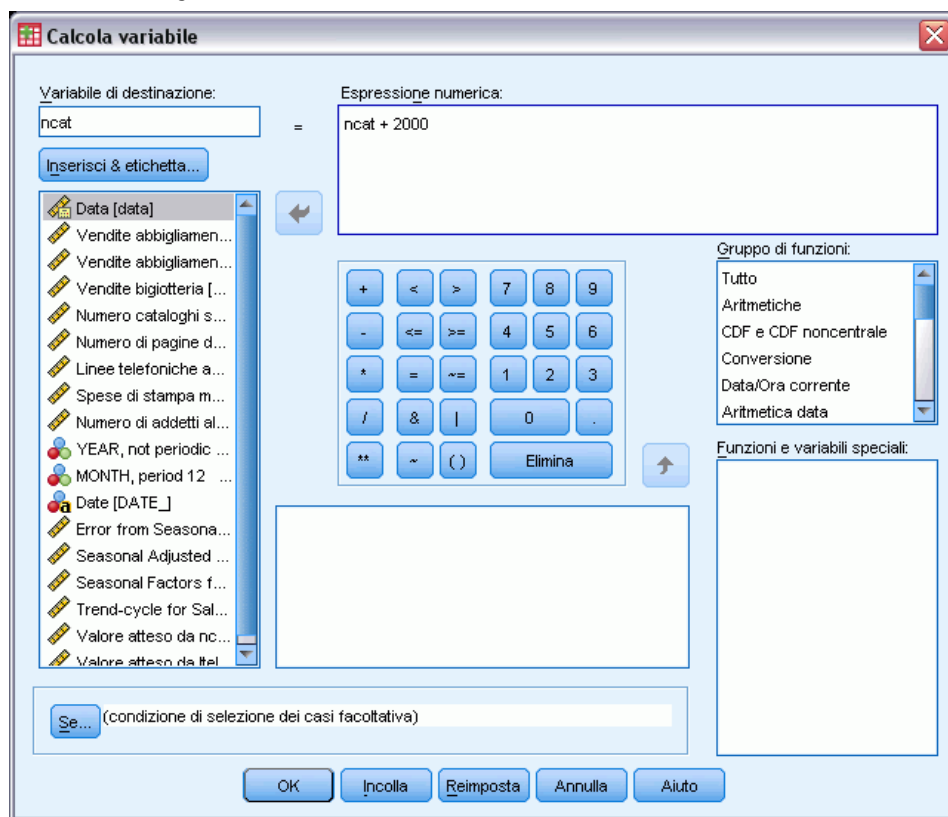
A questo punto i predittori sono stati estesi al periodo di previsione.

## Modifica dei valori dei predittori nel periodo di previsione

Per testare i due scenari relativi all'invio di ulteriori cataloghi o all'aggiunta di nuove linee telefoniche, è necessario modificare le stime dei predittori *posta* o *telefono*. Poiché in questo caso devono essere modificati solo i valori dei predittori di tre casi (mesi), sarebbe sufficiente immettere i nuovi valori nelle celle corrispondenti dell'Editor dei dati. Tuttavia, a scopo di esercitazione, in questo caso useremo la finestra di dialogo Calcola variabile. Questa finestra di dialogo è particolarmente utile quando si tratta di modificare solo qualche valore.

- ▶ Dai menu, scegliere:  
Trasforma > Calcola variabile...

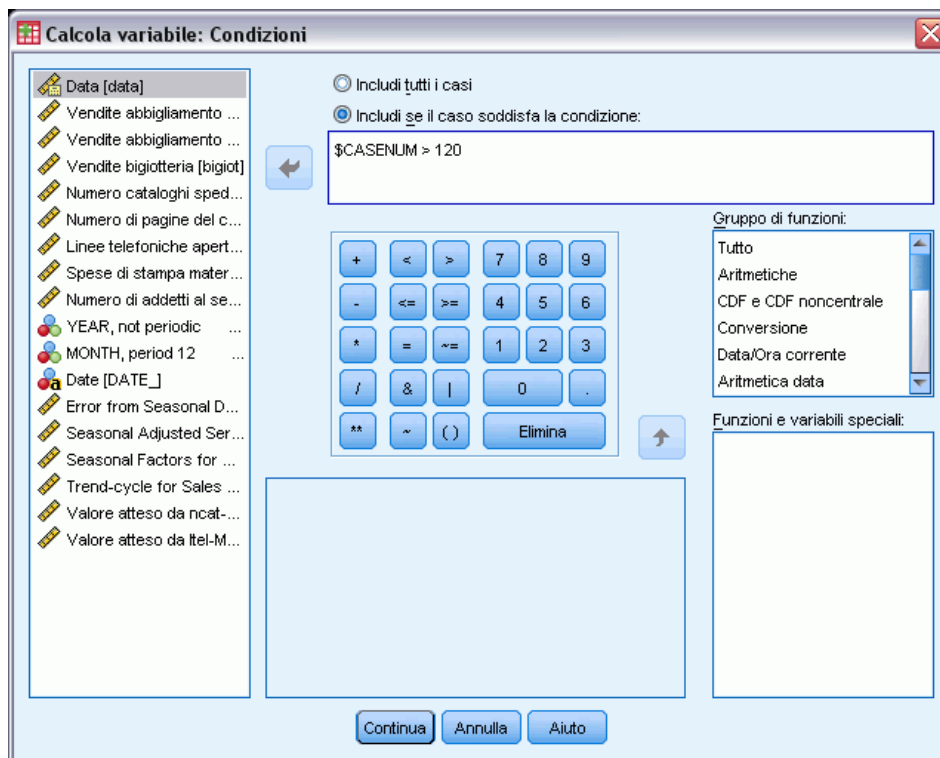
Figura 9-6  
Finestra di dialogo Calcola variabile



- ▶ Immettere posta come variabile di destinazione.
- ▶ Nella casella di testo Espressione numerica, immettere posta + 2000.
- ▶ Fare clic su Se.



Figura 9-7  
Finestra di dialogo Calcola variabile: Condizioni



- ▶ Selezionare Includi se il caso soddisfa la condizione.
- ▶ Nella casella di testo immettere  $\$CASENUM > 120$ .

Questa operazione restringe le modifiche apportate alla variabile *posta* solo ai casi contenuti nel periodo di previsione.

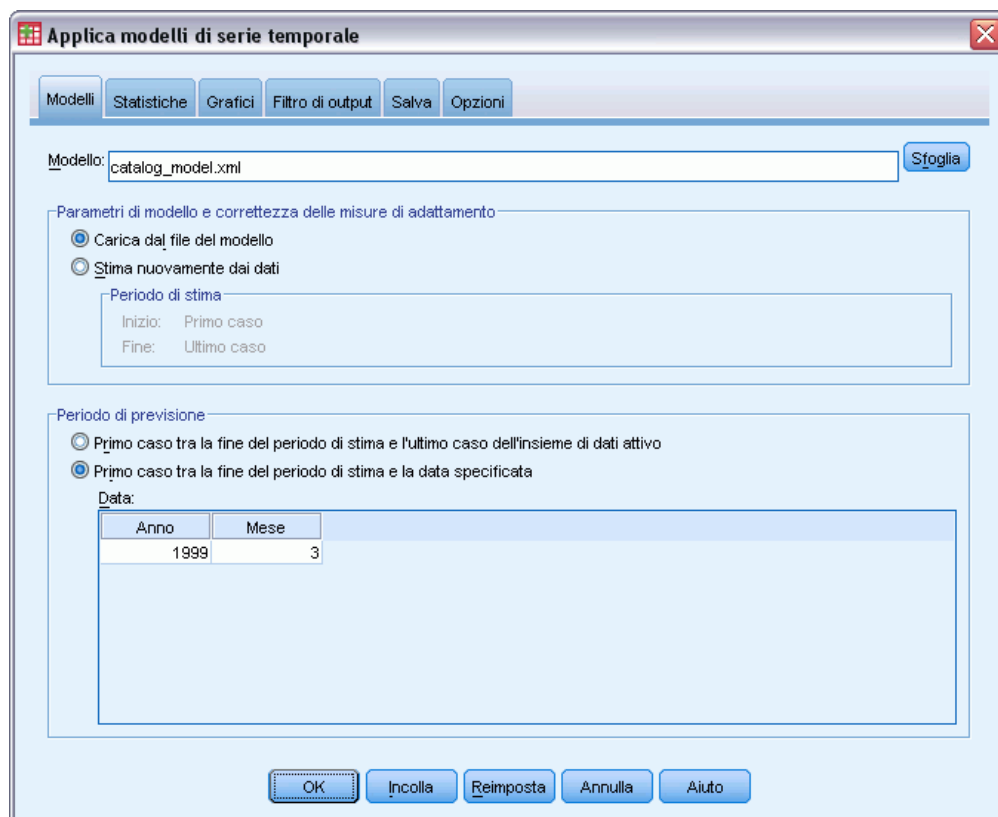
- ▶ Fare clic su Continua.
- ▶ Fare clic su OK nella finestra di dialogo Calcola variabile, quindi fare clic su OK quando viene visualizzato il messaggio che chiede se si desidera modificare la variabile esistente.

Ciò incrementa i valori di *posta*, ovvero del numero di cataloghi inviati, di 2000 per ciascuno dei tre mesi contenuto nel periodo di previsione. A questo punto si dispone di tutti i dati necessari per il primo scenario ed è possibile iniziare ad eseguire l'analisi.

## Esecuzione dell'analisi

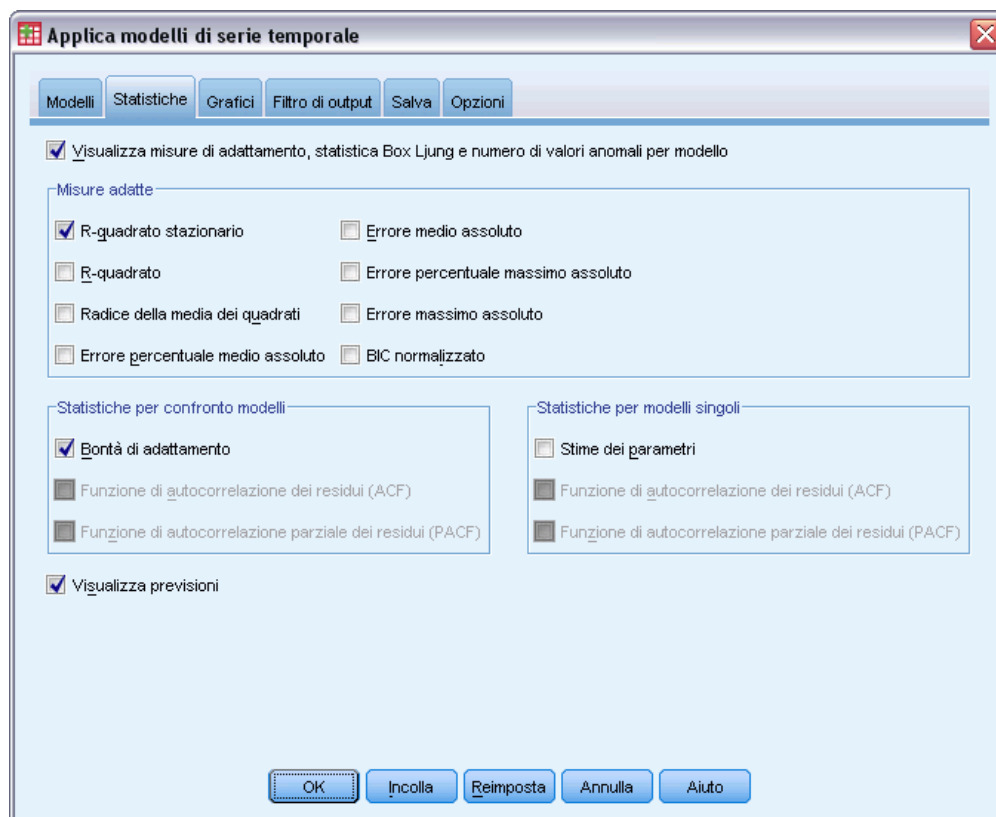
- ▶ Dai menu, scegliere:  
Analizza > Previsioni > Applica modelli...

Figura 9-8  
Finestra di dialogo *Applica modelli delle serie storiche*



- ▶ Fare clic su *Sfogli*, accedere e selezionare *catalog\_model.xml* oppure scegliere il file personale per l'esempio descritto in precedenza. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice D in IBM SPSS Forecasting 21.](#)
- ▶ Nel gruppo *Periodo di previsione* selezionare *Primo caso tra la fine del periodo di previsione e una data specifica*.
- ▶ Nella griglia *Data* immettere 1999 per l'anno e 3 per il mese.
- ▶ Fare clic sulla scheda *Statistiche*.

Figura 9-9  
Applica modelli delle serie storiche, scheda Statistiche



- Selezionare Visualizza previsioni.

Viene creata una tabella dei valori stimati per la variabile dipendente.

- Fare clic su OK nella finestra di dialogo Applica modelli delle serie storiche.

Figura 9-10  
Tabella delle previsioni

Modello		JAN 1999	FEB 1999	MAR 1999
Sales of Men's Clothing-Model_1	Previsione	25279.91	22064.72	21580.96
	UCL	27591.62	24376.42	23892.66
	LCL	22968.21	19753.02	19269.25

La tabella delle previsioni contiene i valori previsti per le serie dipendenti, che tengono conto anche dei valori dei due predittori *posta* e *telefono* del periodo di previsione. La tabella comprende anche il limite di confidenza superiore e il limite di confidenza inferiore delle previsioni.

Viene generata la previsione di vendite per lo scenario relativo all'invio di 2000 nuovi cataloghi al mese. A questo punto occorre preparare i dati per lo scenario relativo all'aggiunta di nuove linee telefoniche. Per effettuare questa operazione, è necessario ripristinare i valori originali della variabile *posta* e aumentare la variabile *telefono* di 5. Per reimpostare la variabile *posta*, è sufficiente copiare i valori di *Modello\_posta\_stimato\_1* nel periodo di previsione e sovrascriverli sui valori correnti di *posta* nel periodo di previsione. Inoltre, è possibile aumentare il numero di

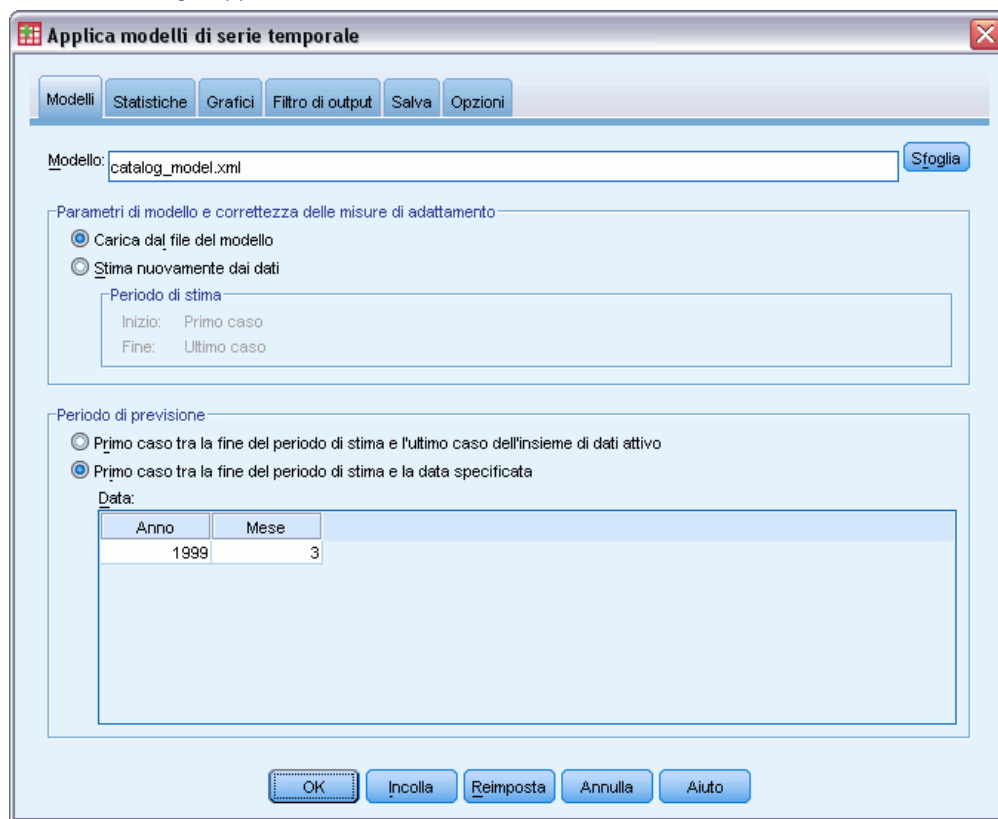
linee telefoniche di 5 per ciascun mese del periodo di previsione, tramite l'Editor dei dati oppure usando la finestra di dialogo Calcola variabile, come nello scenario relativo ai cataloghi.

Per eseguire l'analisi, aprire nuovamente la finestra di dialogo Applica modelli delle serie storiche come descritto di seguito.

- Fare clic sul pulsante Richiama finestra sulla barra degli strumenti.
- Selezionare Applica modelli delle serie storiche.

Figura 9-11

Finestra di dialogo Applica modelli delle serie storiche



- Fare clic su OK nella finestra di dialogo Applica modelli delle serie storiche.

**Figura 9-12**  
 Tabelle delle previsioni per i due scenari

**Previsione con più catalogui**

Modello		JAN 1999	FEB 1999	MAR 1999
Sales of Men's Clothing-Model_1	Previsione	25279.91	22064.72	21580.96
	UCL	27591.62	24376.42	23892.66
	LCL	22968.21	19753.02	19269.25

**Previsione con più linee telefoniche**

Modello		JAN 1999	FEB 1999	MAR 1999
Sales of Men's Clothing-Model_1	Previsione	23757.25	20542.06	20058.29
	UCL	26068.95	22853.76	22370.00
	LCL	21445.55	18230.35	17746.59

La visualizzazione delle tabelle delle previsioni relative ai due scenari mostra che, per ciascuno dei mesi del periodo di previsione, l'incremento del numero di cataloghi inviato consentirà di aumentare le vendite di circa 1500 Dollari rispetto all'aggiunta di nuove linee telefoniche. Quindi, stando all'analisi sembra più opportuno allocare le risorse all'invio di 2000 nuovi cataloghi.

# ***Decomposizione stagionale***

## ***Rimozione della stagionalità dai dati di vendita***

Una società di vendite per corrispondenza è interessata a creare un modello del trend in crescita delle vendite della propria linea di abiti da uomo su un insieme di variabili stimatore (ad esempio il numero di cataloghi spediti e delle linee telefoniche attive per la ricezione degli ordini). Per effettuare questa operazione, è necessario raccogliere i dati sulle vendite mensili di abiti da uomo degli ultimi 10 anni. Tali informazioni vengono raccolte nel file *catalog.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice D in IBM SPSS Forecasting 21.](#)

Per eseguire un'analisi di trend, è necessario rimuovere eventuali variazioni stagionali presenti nei dati. Questa operazione può essere eseguita agevolmente con la procedura di Decomposizione stagionale.

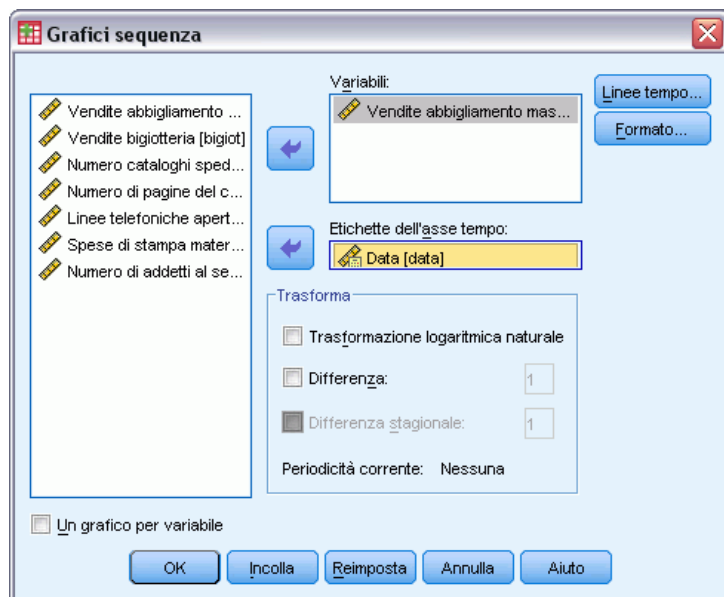
## ***Determinazione e impostazione della periodicità***

La procedura di decomposizione stagionale richiede la presenza di una componente di data periodica nel file di dati attivo—, ad esempio una periodicità 12 all'interno dell'anno (mesi), una periodicità 7 nella settimana (giorni) e così via. È consigliabile tracciare prima le serie storiche, in quanto la visualizzazione di un grafico delle serie storiche consente di farsi un'idea piuttosto precisa della periodicità sottostante.

Per ottenere un grafico delle vendite di abiti da uomo nel tempo:

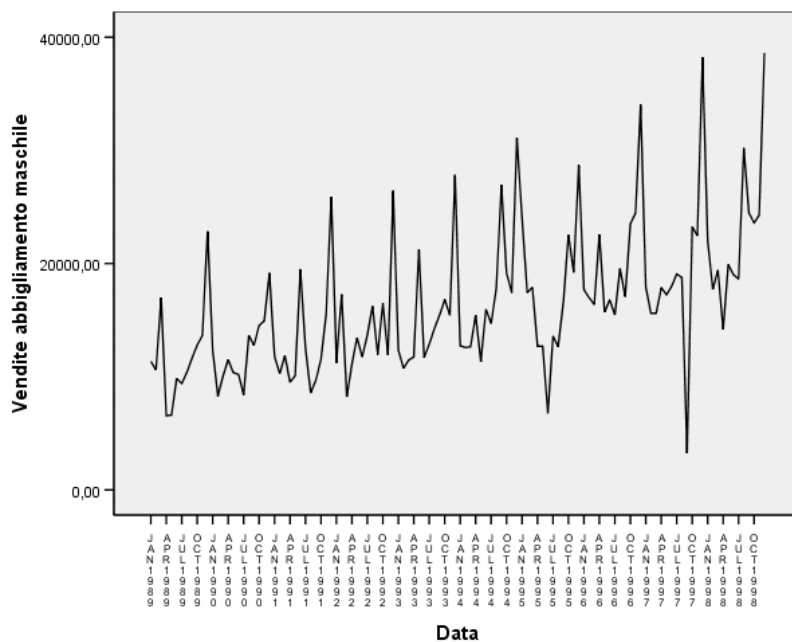
- ▶ Dai menu, scegliere:  
Analizza > Previsioni > Grafici sequenza...

Figura 10-1  
Finestra di dialogo Grafici sequenza



- ▶ Selezionare *Vendite abbigliamento maschile* e spostare i dati nell'elenco Variabili.
- ▶ Selezionare *Data*, quindi passare all'elenco Etichette dell'asse tempo.
- ▶ Fare clic su OK.

Figura 10-2  
*Vendite di abbigliamento maschile (in Dollari USA)*

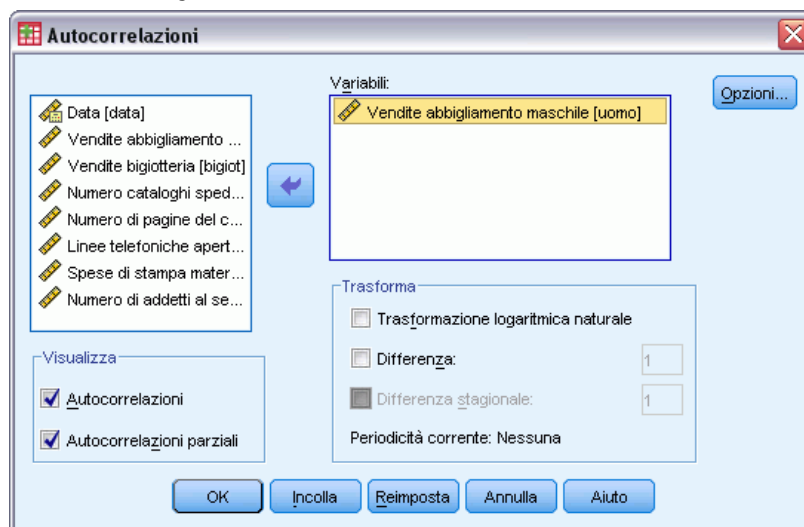


La serie mostra inoltre numerosi picchi, che però non appaiono equamente distanziati. Questo risultato suggerisce che, se la serie ha una componente periodica, possa includere anche fluttuazioni non periodiche: il caso tipico delle serie storiche reali. A parte le fluttuazioni di scala ridotta, i picchi significativi appaiono separati da diversi mesi. Data la natura stagionale delle vendite, con picchi che si verificano normalmente durante le vacanze natalizie, è possibile che le serie storiche abbiano una periodicità annuale. Si noti inoltre che le variazioni stagionali apparentemente crescono insieme al trend verso l'alto della serie, suggerendo che le variazioni stagionali possano essere proporzionali al livello della serie e implicando un modello moltiplicativo anziché additivo.

L'esame delle autocorrelazioni e delle autocorrelazioni parziali di una serie storica offre una conclusione maggiormente quantitativa relativamente alla periodicità sottostante.

- Dai menu, scegliere:  
Analizza > Previsioni > Autocorrelazioni...

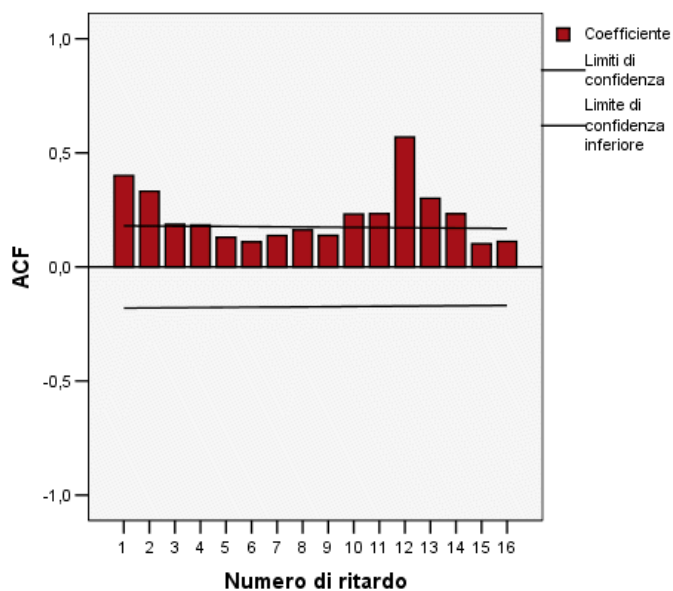
Figura 10-3  
Finestra di dialogo Autocorrelazioni



- Selezionare *Vendite abbigliamento maschile* e spostare i dati nell'elenco Variabili.
- Fare clic su OK.

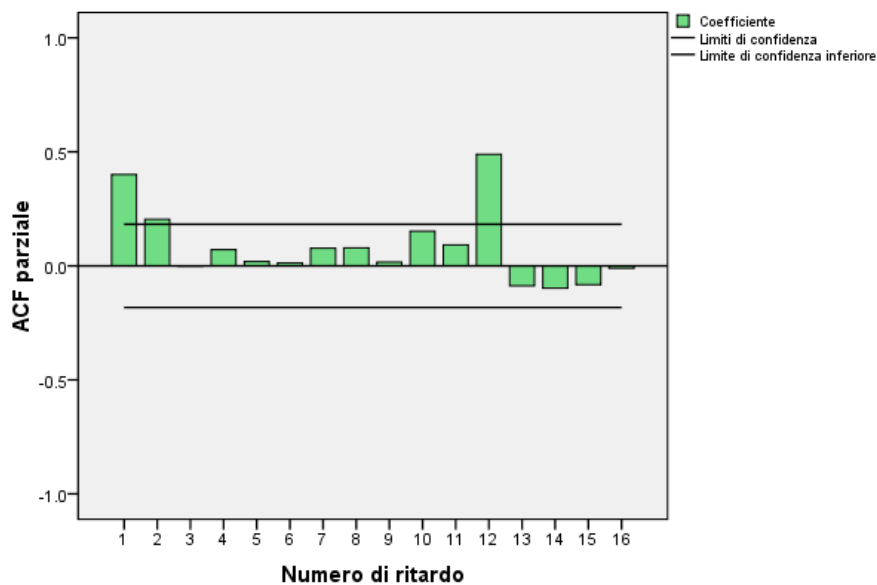


Figura 10-4  
Grafico di autocorrelazione per uomini



La funzione di autocorrelazione mostra un picco significativo in corrispondenza del ritardo 1, con una lunga coda esponenziale, ovvero un modello tipico delle serie storiche. Il picco significativo in corrispondenza del ritardo 12 suggerisce la presenza nei dati di una componente stagionale annuale. L'esame della funzione di autocorrelazione parziale consentirà di trarre una conclusione più definitiva.

Figura 10-5  
Grafico di autocorrelazione parziale per uomini



Il picco significativo in corrispondenza del ritardo 12 nella funzione di autocorrelazione parziale conferma la presenza nei dati di una componente stagionale annuale.

Per impostare una periodicità annuale:

- Dai menu, scegliere:  
Dati > Definisci date...

Figura 10-6  
Finestra di dialogo Definisci date



- Selezionare Anni, mesi nell'elenco I casi sono.
- Immettere 1989 per l'anno e 1 per il mese.

- Fare clic su OK.

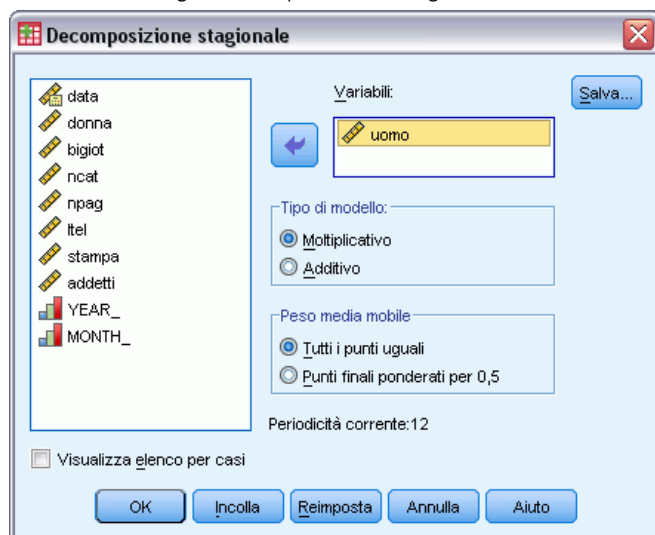
Questo imposta la periodicità su 12 e crea un insieme di variabili di data pensate per l'utilizzo con le procedure Previsioni.

## Esecuzione dell'analisi

Per eseguire la procedura di Decomposizione stagionale:

- Dai menu, scegliere:  
Analizza > Previsioni > Decomposizione stagionale...

Figura 10-7  
Finestra di dialogo Decomposizione stagionale



- Fare clic con il pulsante destro del mouse nell'elenco delle variabili sorgente e dal menu di scelta rapida selezionare Visualizza nomi.
- Selezionare *uomini*, quindi passare all'elenco Variabili.
- Selezionare Moltiplicativo nel gruppo Tipo di modello.
- Fare clic su OK.

## Output

La procedura di Decomposizione stagionale crea quattro nuove variabili per ciascuna delle variabili originali analizzate dalla procedura. Per impostazione predefinita, le nuove variabili vengono aggiunte al file di dati attivo. I nomi delle nuove serie iniziano con i seguenti prefissi:

**SAF.** Fattori di correzione per stagionalità, che rappresentano la variazione stagionale. Per il modello moltiplicativo, il valore 1 rappresenta l'assenza di variazione stagionale; per il modello additivo, è il valore 0 a indicare tale assenza.

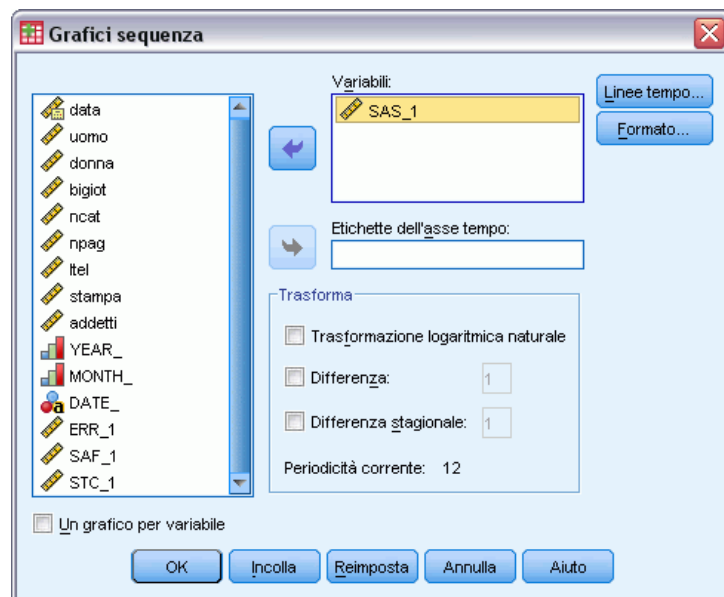
**SAS.** Serie corrette per stagionalità, che rappresenta la serie originale dalla quale sono state rimosse le variazioni stagionali. L'utilizzo di una serie corretta per stagionalità, ad esempio, consente l'isolamento e l'analisi di una componente di trend indipendentemente da eventuali componenti stagionali.

**STC.** Componente di trend-ciclo livellata, ovvero una versione livellata della serie con correzione per stagionalità che mostra le componenti di trend e cicliche.

**ERR.** La componente residua della serie per un'osservazione specifica.

Per il caso attuale, la serie corretta per stagionalità è la più adatta perché rappresenta la serie originale dalla quale sono state rimosse le variazioni stagionali.

Figura 10-8  
Finestra di dialogo Grafici sequenza

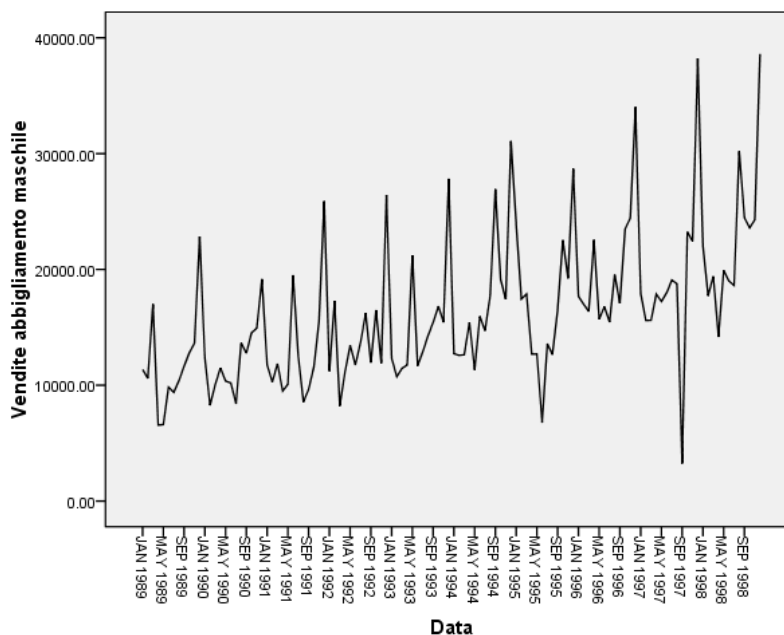


Per tracciare la serie corretta per stagionalità:

- ▶ Viene visualizzata la finestra di dialogo Grafici sequenza.
- ▶ Fare clic su Ripristina per cancellare le selezioni precedenti.
- ▶ Fare clic con il pulsante destro del mouse nell'elenco delle variabili sorgente e dal menu di scelta rapida selezionare Visualizza nomi.

- ▶ Selezionare *SAS\_I*, quindi passare all'elenco Variabili.
- ▶ Fare clic su OK.

Figura 10-9  
Serie corretta per stagionalità



La serie corretta per stagionalità mostra un chiaro trend verso l'alto. Sono evidenti numerosi picchi, che però appaiono a intervalli casuali, senza alcuna indicazione di un modello annuale.

## Riepilogo

La procedura di Decomposizione stagionale ha consentito di rimuovere la componente stagionale di una serie storica periodica, per generare una serie maggiormente adatta all'analisi di trend. L'esame delle autocorrelazioni e delle autocorrelazioni parziali della serie storica ha contribuito a determinare la periodicità sottostante, in questo caso annuale.

## Procedure correlate

La procedura di Decomposizione stagionale è utile per rimuovere una singola componente stagionale da una serie storica periodica.

- Per eseguire un'analisi più approfondita della periodicità di una serie storica, rispetto alla funzione di autocorrelazione parziale, utilizzare la procedura dei Grafici spettrali. Per ulteriori informazioni, vedere [il capitolo 11](#).

# ***Grafici spettrali***

## ***Utilizzo dei grafici spettrali per verificare le aspettative relative alla periodicità***

Le serie storiche che rappresentano dati di vendita al dettaglio hanno generalmente una periodicità annuale sottostante dovuta al normale picco delle vendite durante il periodo natalizio. Generare proiezioni di vendita significa creare un modello della serie storica, ovvero identificare eventuali componenti periodiche. Un grafico della serie storica non sempre evidenzia la periodicità annuale, in quanto le serie storiche contengono fluttuazioni casuali che spesso mascherano la struttura sottostante.

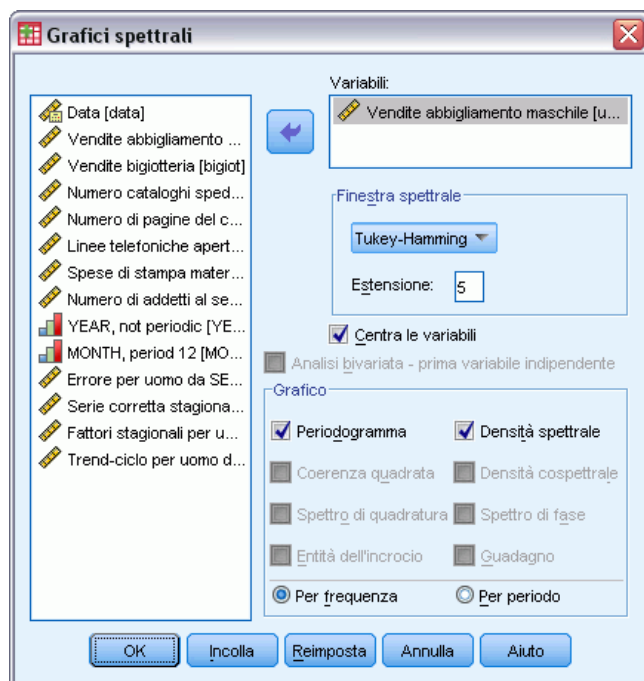
I dati di vendita mensili per una società di vendite per corrispondenza sono memorizzati nel file *catalog.sav*. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice D in IBM SPSS Forecasting 21](#). Prima di procedere con le proiezioni di vendita, è preferibile avere la conferma che i dati di vendita mostrano una periodicità annuale. Un grafico della serie storica mostra vari picchi irregolarmente distanziati, di conseguenza un'eventuale periodicità sottostante non è visibile. Utilizzare la procedura dei Grafici spettrali per identificare eventuali periodicità nei dati di vendita.

### ***Esecuzione dell'analisi***

Per eseguire la procedura dei Grafici spettrali:

- ▶ Dai menu, scegliere:  
Analizza > Previsioni > Analisi spettrale...

Figura 11-1  
Finestra di dialogo Grafici spettrali



- ▶ Selezionare *Vendite abbigliamento maschile* e spostare i dati nell'elenco Variabili.
- ▶ Nel gruppo Grafico fare clic su Densità spettrale.
- ▶ Fare clic su OK.

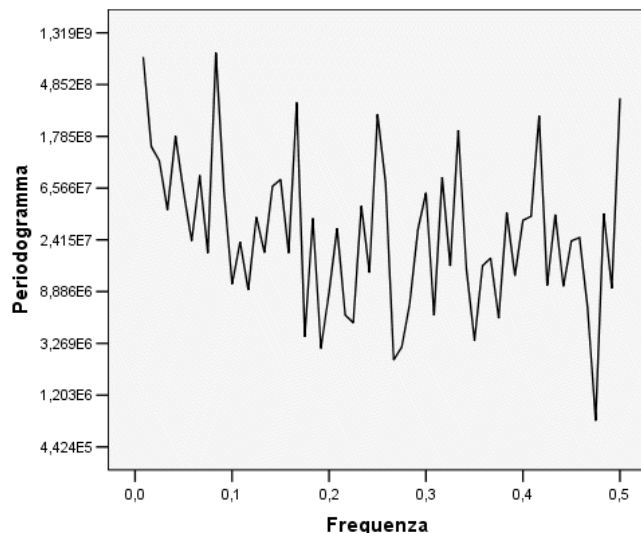
Tali selezioni generano la sintassi seguente:

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

Si noti che, per ottenere la tabella Statistiche univariate nell'output, è necessario trasformare il comando TSET in TSET PRINT=DETAILED.

## Periodogramma e densità spettrale

Figura 11-2  
Periodogramma



Il grafico del periodogramma mostra una sequenza di picchi in evidenza rispetto al disturbo di sfondo, con il picco di frequenza minima in corrispondenza di un valore di frequenza di poco inferiore a 0,1. Poiché si ritiene che i dati possano contenere una componente periodica annuale, si consideri il contributo che una componente annuale potrebbe offrire al periodogramma. Ciascuno dei punti di dati nella serie storica rappresenta un mese, quindi una periodicità annuale corrisponde a un periodo di 12 nell'insieme di dati corrente. Poiché periodo e frequenza sono reciproci tra loro, un periodo di 12 corrisponde a una frequenza di  $1/12$  (o 0,083). Di conseguenza una componente annuale implica un picco nel periodogramma in corrispondenza di 0,083, il che sembra coerente con la presenza del picco poco al di sotto di una frequenza pari a 0,1.

Figura 11-3  
Tabella Statistiche univariate

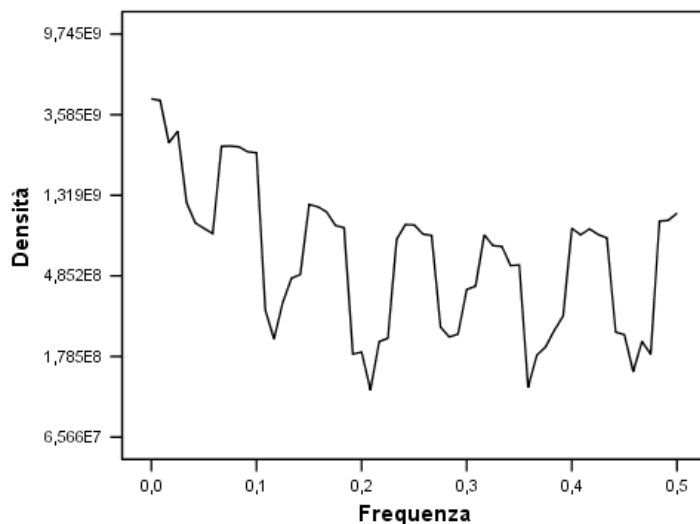
Nome della serie: uomini

	Frequenza	Trasformazione seno	Trasformazione coseno	Periodogramma	Stima di densità spettrale
1	,00000	,000	16242,813	,000	4,361E+09
2	,00833	-3696,643	370,153	828131182	4,278E+09
3	,01667	-1496,277	454,831	146743038	2,526E+09
4	,02500	-1336,400	252,087	110970821	2,921E+09
5	,03333	-662,146	529,734	43143315,6	1,210E+09
6	,04167	-1654,614	501,571	179359225	935924937
7	,05000	-784,814	-636,729	61281367,3	675492029
8	,05833	-335,646	532,062	23744855,0	820375352
9	,06667	-1094,178	-451,489	84064108,9	2,420E+09
10	,07500	264,554	492,876	18774933,5	2,429E+09
11	,08333	-3053,934	2370,483	896742149	2,401E+09
12	,09167	-978,882	-287,035	62435897,6	2,263E+09
13	,10000	-403,128	93,036	10270064,4	2,235E+09



La tabella Statistiche univariate contiene i punti di dati utilizzati per tracciare il periodogramma. Si noti che per frequenze inferiori a 0,1, il valore massimo nella colonna *Periodogramma* ricorre in corrispondenza di una frequenza di 0,08333—esattamente com'è prevedibile in presenza di una componente periodica annuale. Questo dato conferma l'identificazione del picco minimo di frequenza con una componente periodica annuale. Vediamo cosa avviene per quanto riguarda gli altri picchi a frequenze maggiori.

Figura 11-4  
Densità spettrale



È possibile analizzare i picchi restanti in modo ottimale tramite la funzione di densità spettrale, ovvero una versione livellata del periodogramma. Il livellamento fornisce un modo per eliminare il disturbo di sfondo da un periodogramma, consentendo di isolare in modo più chiaro la struttura sottostante.

La densità spettrale è composta da cinque picchi distinti, equamente distanziati. Il picco minimo di frequenza rappresenta semplicemente la versione livellata del picco a 0,08333. Per comprendere la significatività dei quattro picchi di frequenza maggiori, si ricordi che il periodogramma viene calcolato creando un modello della serie storica come somma delle funzioni seno e coseno. Le componenti periodiche che hanno la forma di una funzione di seno o coseno (sinusoidale) nel periodogramma sono indicate da picchi singoli. Le componenti periodiche non sinusoidali sono indicate da serie di picchi equamente distanziati di altezze diverse, con il picco minimo nella serie in corrispondenza della frequenza della componente periodica. Di conseguenza, i quattro picchi maggiori nella densità spettrale indicano semplicemente che la componente periodica annuale non è sinusoidale.

L'intera struttura discernibile del grafico di densità spettrale è ora spiegata ed è possibile concludere che i dati contengono una singola componente periodica con un periodo di 12 mesi.

## **Riepilogo**

La procedura dei Grafici spettrali ha consentito di confermare l'esistenza di una componente periodica annuale in una serie storica e di verificare che non sono presenti altre periodicità significative. La densità spettrale è risultata più utile del periodogramma per evidenziare la struttura sottostante, in quanto livella le fluttuazioni causate dalla componente non periodica dei dati.

## **Procedure correlate**

La procedura dei Grafici spettrali è utile per identificare le componenti periodiche di una serie storica.

- Per rimuovere una componente periodica da una serie storica, ad esempio allo scopo di eseguire un'analisi di trend, utilizzare la procedura Decomposizione stagionale. Vedere [il capitolo 10](#) per ulteriori dettagli.

## ***Misure della bontà di adattamento***

In questa sezione sono riportate le definizioni delle misure della bontà di adattamento utilizzate nei modelli di serie storiche.

- **R-quadrato stazionario.** Una misura che confronta la parte stazionaria del modello con un modello di media semplice. Questa misura è preferibile all'R-quadrato semplice se è presente un trend o un motivo stagionale. L'R-quadrato stazionario è negativo con un intervallo negativo infinito tendente a 1. I valori negativi indicano che il modello esaminato è peggiore del modello di base. I valori positivi indicano che il modello esaminato è migliore del modello di base.
- **R-quadrato.** Stima della proporzione della variazione totale della serie spiegata dal modello. Questa misura è particolarmente utile quando la serie è fissa. L'R-quadrato è negativo con un intervallo negativo infinito tendente a 1. I valori negativi indicano che il modello esaminato è peggiore del modello di base. I valori positivi indicano che il modello esaminato è migliore del modello di base.
- **RMSE.** Acronimo di Root Mean Square Error. La radice 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.
- **MAPE.** Errore percentuale assoluto medio. Indica di quanto la serie dipendente varia rispetto al livello predittivo del modello. L'errore è indipendente dalle unità usate e può quindi essere usato per confrontare le serie con unità diverse.
- **MAE.** Errore assoluto medio. Misura di quanto la serie varia rispetto al livello predittivo del modello. Il valore MAE viene indicato in unità originali della serie.
- **MaxAPE.** Errore percentuale massimo assoluto. Errore stimato più alto, espresso in percentuale. Questa misura è utile per stimare lo scenario peggiore delle 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 informazione bayesiano 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 confrontare più facilmente la statistica con più modelli della stessa serie.

## *Tipi di valori anomali*

In questa sezione sono riportate le definizioni dei tipi di valori anomali utilizzati nei modelli di serie storiche.

- **Additivo.** Un valore anomalo che influisce su una singola osservazione. Ad esempio un valore anomalo additivo può essere rappresentato da un errore di codifica dati.
- **Spostamento di livello.** Valore anomalo che varia tutte le osservazioni in base a una costante, a partire da un punto specifico della serie. Una variazione di livello può essere provocata da una variazione di criteri.
- **Innovazionale.** Valore anomalo che funge come aggiunta al termine rumore in un particolare punto della serie. Nelle serie stazionarie, il valore anomalo innovazionale influisce su svariate osservazioni. Nelle serie non stazionarie, può influire sulle tutte le osservazioni che iniziano da un punto della serie specifico.
- **Transiente.** Un valore anomalo il cui impatto decade esponenzialmente a 0.
- **Additivo stagionale.** Valore anomalo che influisce su un'osservazione specifica e su tutte le osservazioni successive separate da uno o più periodi stagionali. Il valore anomalo influisce nello stesso modo su tutte le osservazioni. Un additivo stagionale può verificarsi se le vendite risultano più alte a gennaio, a partire da un anno specifico.
- **Trend locale.** Un valore anomalo che dà inizio a un trend locale in un punto particolare della serie.
- **Blocco additivo.** Gruppo di due o più valori anomali additivi consecutivi. La selezione di un tipo di valore anomalo permette di rilevare i singoli valori anomali additivi oltre alle relative patch.



## Guida ai grafici ACF/PACF

I grafici qui illustrati si riferiscono a processi ARIMA puri o teorici. Di seguito, vengono fornite alcune linee guida generali utili per identificare i processi.

- Le serie non stazionarie hanno un ACF che rimane significativo per sei o più ritardi anziché scendere rapidamente a 0. Le serie di questo tipo devono essere tenute a parte finché non diventano stazionarie ed è possibile identificare il processo.
- Nei processi autoregressivi, il grafico ACF tende a diminuire in modo esponenziale e le linee di connessione del grafico PACF sono presenti in corrispondenza del primo ritardo o di quelli successivi. Il numero di linee di connessione indica l'ordine dell'autoregressione.
- Nei processi di media mobile, le linee di connessione si trovano in corrispondenza del primo ritardo o di quelli successivi ed il grafico PACF che tende a diminuire in modo esponenziale. Il numero di linee di connessione indica l'ordine della media mobile.
- Nei processi misti (ARIMA) i grafici ACF e PACF tendono generalmente a ridursi in modo esponenziale.

In fase di identificazione non è necessario preoccuparsi del segno dei grafici ACF o PACF né della velocità con cui tali grafici si avvicinano allo zero in modo esponenziale, poiché questi fattori dipendono dal segno e dal valore effettivo dei coefficienti AR e MA. In alcuni casi, un grafico ACF che diminuisce in modo esponenziale presenta un'alternanza tra valori positivi e negativi.

*I grafici ACF e PACF che si riferiscono a dati reali non sono così chiari come i grafici illustrati.* Quindi, è necessario imparare ad identificare gli elementi essenziali dei grafici. Se l'identificazione è errata, è sempre consigliabile controllare i grafici ACF e PACF dei valori residui. Inoltre, è utile ricordare che:

- Nei processi stagionali questi modelli possono essere rilevati in corrispondenza dei ritardi stagionali (multipli del periodo stagionale).
- I valori non significativi possono essere considerati pari a 0. Ciò vuol dire che è possibile ignorare i valori che rientrano nell'intervallo di confidenza dei grafici. Tali valori non possono tuttavia essere ignorati se seguono il modello dei valori statisticamente significativi.
- Un'autocorrelazione occasionale è statisticamente significativa solo per caso. Quindi, è possibile ignorare un'autocorrelazione statisticamente significativa se è isolata, preferibilmente in corrispondenza di un ritardo alto, e si non si verifica in corrispondenza di un ritardo stagionale.

Per una descrizione più dettagliata dei grafici ACF e PACF, consultare qualsiasi testo sull'analisi ARIMA.

Tabella C-1  
 $ARIMA(0,0,1), q>0$

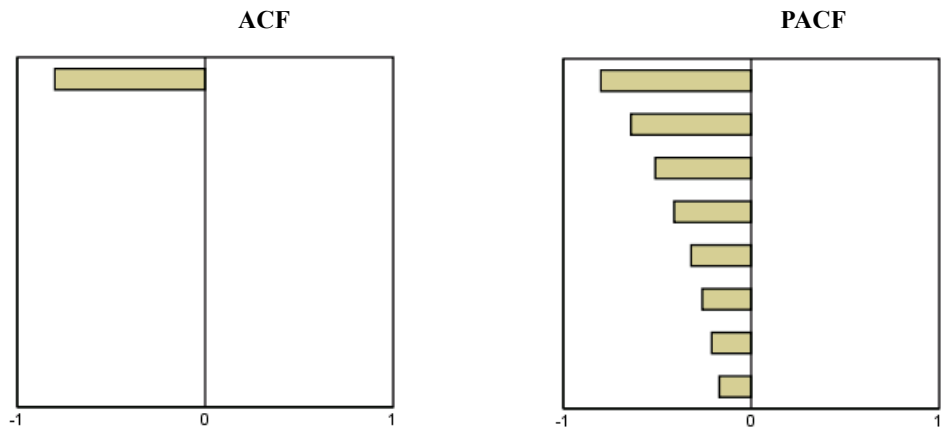
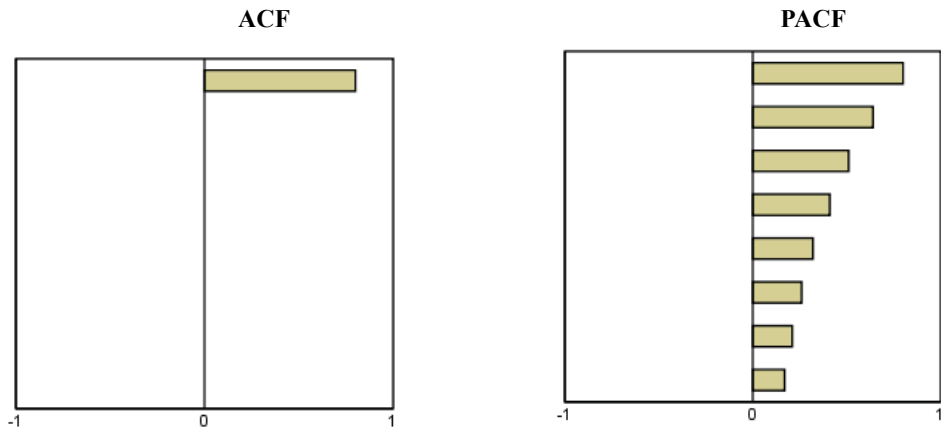


Tabella C-2  
 $ARIMA(0,0,1), q<0$



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

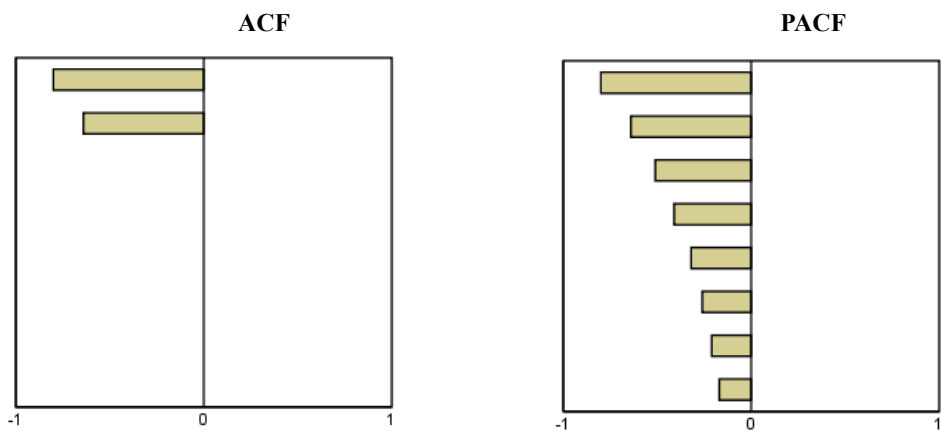


Tabella C-3  
 $ARIMA(1,0,0), f > 0$

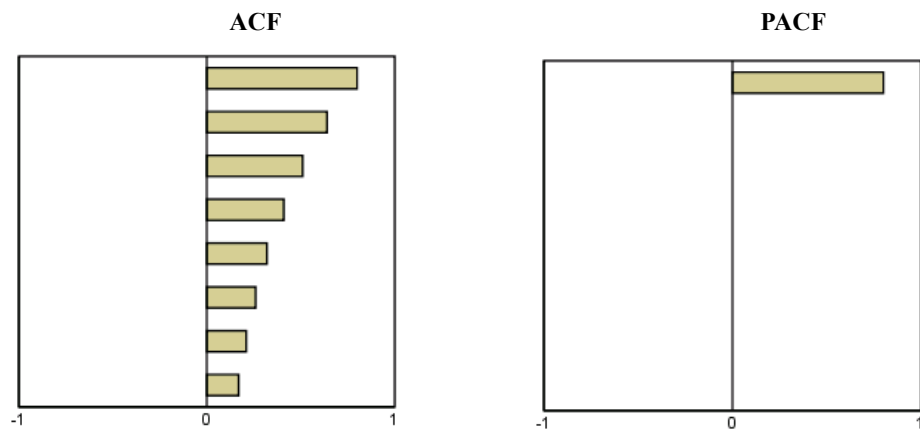
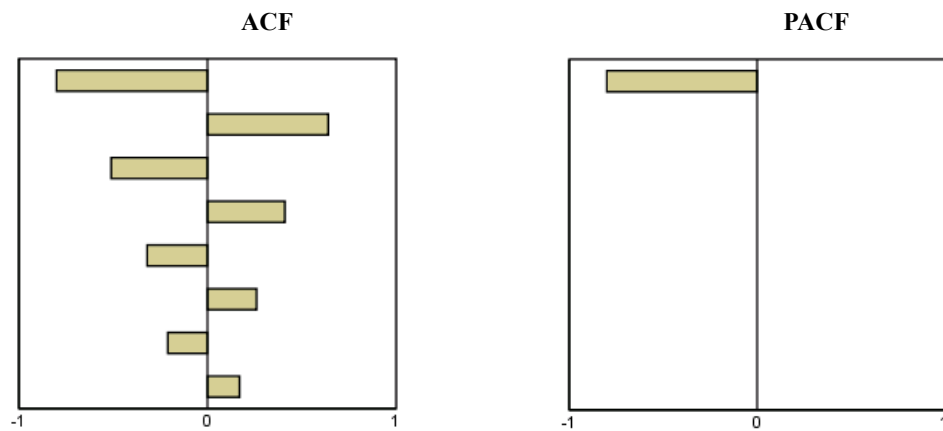
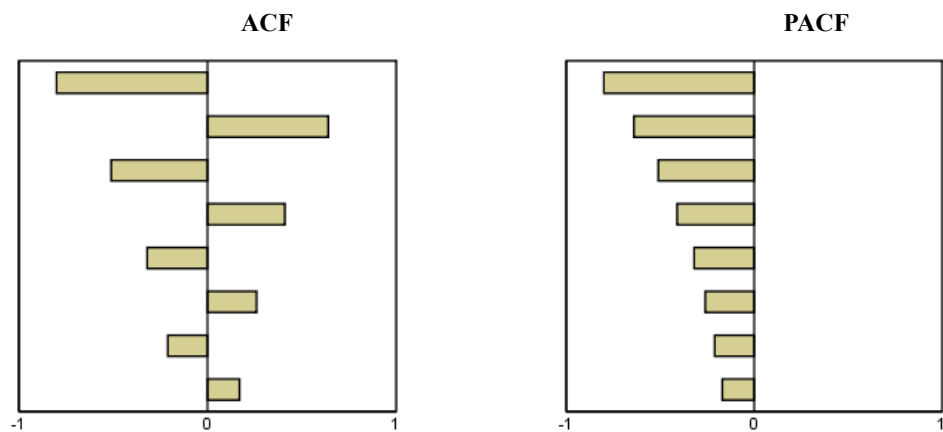


Tabella C-4  
 $ARIMA(1,0,0), f < 0$



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



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

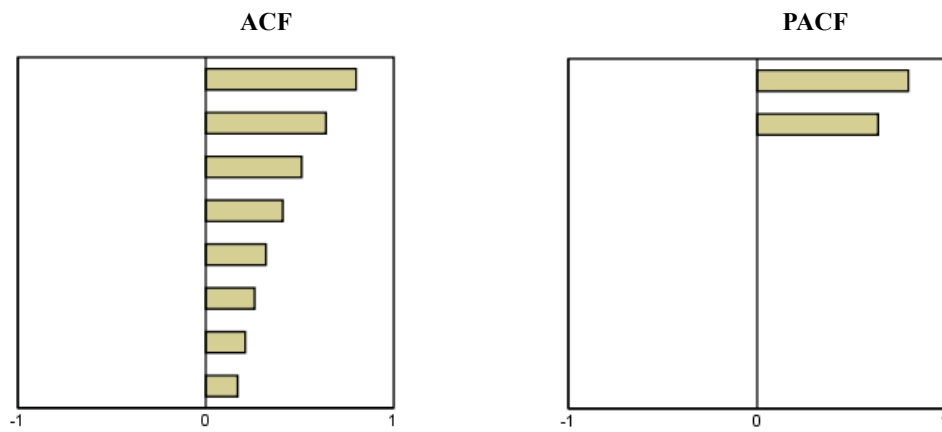
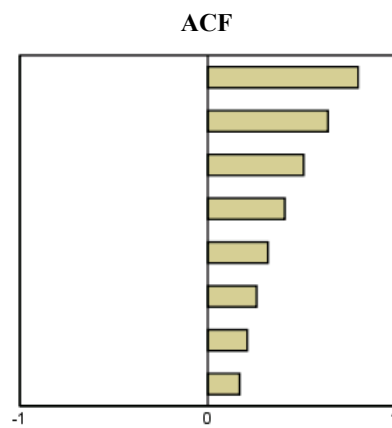


Tabella C-5  
ARIMA(0,1,0) (serie integrata)





## File di esempio

Il file di esempio installato con il prodotto si trova nella sottodirectory *Samples* della directory di installazione. La sottodirectory *Samples* contiene cartelle separate per ciascuna delle seguenti lingue: Inglese, Francese, Tedesco, Italiano, Giapponese, Coreano, Polacco, Russo, Cinese semplificato, Spagnolo e Cinese tradizionale.

Non tutti i file di esempio sono disponibili in tutte le lingue. Se un file di esempio non è disponibile in una lingua, la cartella di tale lingua contiene una versione inglese del file.

### Descrizioni

Questa sezione contiene brevi descrizioni dei file di esempio utilizzati negli esempi riportati in tutta la documentazione.

- **accidents.sav.** File di dati ipotetici che prende in esame una compagnia di assicurazioni impegnata nello studio dei fattori di rischio correlati all'età e al sesso per gli incidenti automobilistici che si verificano in una determinata regione. Ciascun caso corrisponde a una classificazione incrociata della categoria relativa età e del sesso.
- **adl.sav.** File di dati ipotetici che prende in esame l'impegno richiesto per determinare i vantaggi di un tipo di terapia proposto per i pazienti con problemi di cuore. I medici hanno assegnato in modo casuale i pazienti con problemi di cuore di sesso femminile a uno di due gruppi. Al primo gruppo è stata assegnata la terapia fisica standard; al secondo gruppo, un'ulteriore terapia di supporto psicologico. Dopo tre mesi di trattamenti, a ciascuna capacità dei pazienti che consente di riprendere le normali attività giornaliere è stato assegnato un punteggio come variabile ordinale.
- **advert.sav.** File di dati ipotetici che prende in esame l'impegno di un rivenditore al dettaglio che desidera esaminare la relazione tra il denaro speso per la pubblicità e le vendite risultanti. Finora sono stati raccolti i dati delle vendite precedenti e i relativi costi pubblicitari.
- **aflatoxin.sav.** File di dati ipotetici che prende in esame il test di raccolti di mais con presenza di Aflatossina, un veleno la cui concentrazione varia notevolmente nei raccolti. Una macchina per la lavorazione dei cereali ha ricevuto 16 campioni da ciascuno degli otto raccolti di mais e ha misurato i livelli di Aflatossina in parti per miliardo (PPB).
- **anorectic.sav.** Per trovare una sintomatologia standardizzata del comportamento anoressico/bulimico, i ricercatori (Van der Ham, Meulman, Van Strien, e Van Engeland, 1997) hanno condotto uno studio basato su 55 adolescenti affetti da disturbi alimentari conosciuti. Ogni paziente è stato visitato quattro volte in quattro anni, per un totale di 220 visite. Durante ogni visita, ai pazienti sono stati assegnati punteggi per ciascuno dei 16 sintomi. I punteggi relativi ai sintomi sono assenti per il paziente 71 alla visita 2, il paziente 76 alla visita 2 e il paziente 47 alla visita 3, con 217 osservazioni valide.
- **bankloan.sav.** File di dati ipotetici che prende in esame l'impegno di una banca nel tentativo di ridurre il tasso di inadempienza nel rimborso di un prestito. Il file contiene informazioni finanziarie e demografiche su 850 vecchi e potenziali clienti. I primi 700 casi riguardano i

clienti a cui sono stati concessi dei prestiti precedentemente. Gli ultimi 150 casi riguardano i potenziali clienti che la banca deve classificare come rischi di credito positivi o negativi.

- **bankloan\_binning.sav.** File di dati ipotetici che contiene informazioni finanziarie e demografiche su 5000 vecchi clienti.
- **behavior.sav.** In un classico esempio (Prezzo e Bouffard, 1974), è stato chiesto a 52 studenti di classificare una combinazione di 15 situazioni e 15 comportamenti utilizzando una scala da 0=“molto appropriato” a 9=“molto inadeguato”. I valori medi riferiti ai partecipanti sono stati considerati dissimilarità.
- **behavior\_ini.sav.** Questo file di dati contiene la configurazione iniziale di una soluzione a due dimensioni per *behavior.sav*.
- **brakes.sav.** File di dati ipotetici che prende in esame il controllo di qualità di un'industria che produce freni a disco per automobili con elevate prestazioni. Il file di dati contiene le misurazioni del diametro di 16 dischi da ciascuna delle otto macchine di produzione. L'obiettivo finale è ottenere un diametro dei dischi pari a 322 millimetri.
- **breakfast.sav.** In uno studio classico (Green e Rao, 1972), è stato chiesto a 21 studenti MBA della Wharton School e ai loro consorti di classificare 15 cibi da colazione in ordine di preferenza, dove il valore 1 corrispondeva all'alimento preferito in assoluto e il valore 15 a quello meno preferito. Le loro preferenze sono state registrate per sei diversi scenari, che comprendevano tutti gli scenari compresi tra “Preferenza generale” e “Solo snack con bibita”.
- **breakfast-overall.sav.** Questo file contiene le preferenze degli alimenti della colazione solo per il primo scenario, “Preferenza generale”.
- **broadband\_1.sav.** File di dati ipotetici che contiene il numero di sottoscrittori, per area, di un provider di servizi a banda larga nazionale. Il file di dati contiene il numero dei sottoscrittori mensili di 85 aree in un periodo di quattro anni.
- **broadband\_2.sav.** Questo file è identico al file *broadband\_1.sav*, ma contiene i dati per ulteriori tre mesi.
- **car\_insurance\_claims.sav.** Un insieme di dati presentato e analizzato altrove (McCullagh e Nelder, 1989) riguarda le richieste di risarcimento auto. La quantità media di richieste di risarcimento può essere adattata come avente una distribuzione gamma, utilizzando una funzione di collegamento inverso per correlare la media della variabile dipendente a una combinazione lineare di età del contraente della polizza e tipo e anni del veicolo. Il numero delle richieste di risarcimento specificato può essere utilizzato come peso scalato.
- **car\_sales.sav.** Questo file di dati ipotetici contiene le stime sulle vendite, i prezzi di listino e le specifiche fisiche di numerose marche e modelli di veicoli. I prezzi di listino e le specifiche fisiche sono state ottenute dal sito *edmunds.com* e dai siti dei produttori.
- **car\_sales\_uprepared.sav.** Questa è una versione modificata di *car\_sales.sav* che non comprende versioni trasformate dei campi.
- **carpet.sav.** Come esempio tipico (Green e Wind, 1973), un'azienda interessata alla commercializzazione di un nuovo battitappeto desidera esaminare l'influenza di cinque fattori sulle preferenze del consumatore, ovvero design della confezione, marca, prezzo, la presenza di un *marchio di qualità* e una garanzia “Soddisfatti o rimborsati”. Esistono tre livelli di fattore per il design della confezione, che differiscono per la posizione della spazzola dell'applicatore; tre marchi (*K2R*, *Glory* e *Bissell*); tre livelli di prezzo e due livelli (no o sì) per ciascuno degli ultimi due fattori. Dieci consumatori sono classificati in 22 profili definiti

da questi fattori. La variabile *Preferenza* include il rango delle classificazioni medie per ogni profilo. Classificazioni basse corrispondono a una preferenza elevata. La variabile riflette una misura globale della preferenza per ogni profilo.

- **carpet\_prefs.sav.** Questo file di dati si basa sullo stesso esempio del file *carpet.sav*, ma contiene le classificazioni effettive raccolte da ciascuno dei 10 clienti. Ai clienti è stato chiesto di classificare 22 profili di prodotti in ordine di preferenza. Le variabili da *PREF1* a *PREF22* contengono gli ID dei profili associati, come definito nel file *carpet\_plan.sav*.
- **catalog.sav.** File di dati ipotetico che contiene le cifre sulle vendite mensili di tre prodotti venduti da una società di vendita per corrispondenza. Il file include anche i dati di cinque possibili variabili predittore.
- **catalog\_seasfac.sav.** Questo file di dati è uguale al file *catalog.sav* con l'eccezione che contiene un insieme di fattori stagionali calcolati dalla procedura Decomposizionale stagionale insieme a variabili di dati.
- **cellular.sav.** File di dati ipotetici che prende in esame l'impegno di un'azienda di telefonia cellulare nel tentativo di ridurre il churn, ovvero l'abbandono dei clienti. Agli account vengono applicati i punteggi relativi alla propensione al churn, con valori compresi tra 0 e 100. Gli account con punteggio pari a 50 o superiore è probabile che stiano cercando nuovi provider.
- **ceramics.sav.** File di dati ipotetici che prende in esame l'impegno di un produttore che desidera stabilire se una nuova lega premium ha una maggiore resistenza al calore rispetto alla lega standard. Ciascun caso rappresenta il test separato di una delle leghe. È indicata la temperatura massima alla quale può essere sottoposto il cuscinetto.
- **cereal.sav.** File di dati ipotetici che prende in esame le preferenze relative agli alimenti della colazione di un campione di 880 persone. Il file riporta anche l'età, il sesso e lo stato civile del campione e se le persone conducono uno stile di vita attivo (in base a un'attività sportiva con frequenza di due volte alla settimana). Ogni caso rappresenta un rispondente separato.
- **clothing\_defects.sav.** File di dati ipotetici che prende in esame il processo di controllo di qualità di un'industria di abbigliamento. Per ciascun lotto prodotto nella fabbrica, gli ispettori prelevano un campione di abiti per contare il numero dei capi che non sono accettabili per la vendita.
- **coffee.sav.** Questo file di dati contiene informazioni sulle immagini percepite di sei marche di caffè freddo (Kennedy, Riquier, e Sharp, 1996). Per ciascuno dei 23 attributi dell'immagine del caffè freddo, sono state selezionate tutte le marche descritte da tale attributo. Le sei marche sono indicate dalle sigle AA, BB, CC, DD, EE e FF per tutelare la confidenzialità dei dati.
- **contacts.sav.** File di dati ipotetici che prende in esame l'elenco dei contatti di un gruppo di rappresentanti di vendita di computer aziendali. Ciascun contatto è classificato in base al reparto della società in cui lavora e dalle relative categorie aziendali. Il file riporta anche l'importo dell'ultima vendita effettuata, il tempo trascorso dall'ultima vendita e le dimensioni della società del contatto.
- **creditpromo.sav.** File di dati ipotetici che prende in esame l'impegno di un grande magazzino nel tentativo di valutare l'efficacia di una recente promozione con carta di credito. A tale scopo, sono stati selezionati 500 titolari di carta in modo casuale. Alla metà di questi è stato inviato un annuncio promozionale che comunica la riduzione del tasso d'interesse nel caso di acquisti effettuati entro i tre mesi successivi. All'altra metà è stato inviato un annuncio stagionale standard.

- **customer\_dbase.sav.** File di dati ipotetico che prende in esame l'impegno di una società nel tentativo di utilizzare le informazioni contenute nel proprio database dei dati per creare offerte speciali per i clienti che più probabilmente risponderanno all'offerta. È stato selezionato in modo casuale un sottoinsieme della base dei clienti a cui è stata inviata l'offerta speciale e sono state registrate le risposte ricevute.
- **customer\_information.sav.** File di dati ipotetici contenente le informazioni postali del cliente, ad esempio il nome e l'indirizzo.
- **customer\_subset.sav.** Un sottoinsieme di 80 casi da *customer\_dbase.sav*.
- **debate.sav.** File di dati ipotetici che prende in esame le risposte appaite a un'indagine da parte dei partecipanti a un dibattito politico prima e dopo il dibattito. Ogni caso rappresenta un rispondente separato.
- **debate\_aggregate.sav.** File di dati ipotetici che aggrega le risposte contenute nel file *debate.sav*. Ciascun caso corrisponde a una classificazione incrociata della preferenza prima e dopo il dibattito.
- **demo.sav.** File di dati ipotetici che prende in esame un database di clienti che hanno fatto acquisti al fine di inviare offerte mensili tramite il metodo del direct mailing. Viene registrata la risposta dei clienti, sia che abbiano aderito all'offerta o meno, insieme a diverse informazioni demografiche.
- **demo\_cs\_1.sav.** File di dati ipotetici che prende in esame il primo passo che una società intraprende per compilare un database con informazioni ricavate dai sondaggi. Ogni caso rappresenta una diversa città. Sono registrate anche le informazioni sulla regione, provincia, distretto e città.
- **demo\_cs\_2.sav.** File di dati ipotetici che prende in esame il secondo passo che una società intraprende per compilare un database con informazioni ricavate dai sondaggi. Ogni caso rappresenta una diversa unità di abitazione, ricavata dalle città selezionate nel primo passo. Sono registrate anche le informazioni sulla regione, provincia, distretto, città, suddivisione e unità. Il file include inoltre informazioni sul campionamento ottenute dai primi due stadi del disegno.
- **demo\_cs.sav.** File di dati ipotetici che contiene informazioni sulle indagini raccolte utilizzando un disegno di campionamento complesso. Ogni caso rappresenta una diversa unità di abitazione. Sono registrate diverse informazioni demografiche e sul campionamento.
- **dmdata.sav.** File di dati ipotetici che contiene informazioni demografiche e di acquisto di una società di direct marketing. *dmdata2.sav* contiene informazioni su un sottoinsieme di contatti che hanno ricevuto un mailing di prova e *dmdata3.sav* contiene informazioni sui contatti rimanenti che non hanno ricevuto il mailing di prova.
- **dietstudy.sav.** File di dati ipotetici che contiene il risultato di uno studio ipotetico sulla dieta chiamato "Stillman diet" (Rickman, Mitchell, Dingman, e Dalen, 1974). Ogni caso rappresenta un diverso soggetto e ne riporta il peso prima e dopo la dieta in libbre e i livelli dei trigliceridi in mg/100 ml.
- **dvdplayer.sav.** File di dati ipotetici che prende in esame lo sviluppo di un nuovo lettore DVD. Utilizzando un prototipo, il personale addetto al marketing ha raccolto dati sui gruppi di interesse. Ogni caso rappresenta un diverso utente che è stato sottoposto all'indagine e include informazioni demografiche personali dell'utente e sulle risposte che ha fornito riguardo al prototipo.

- **german\_credit.sav.** Questo file di dati contiene informazioni ricavate dall'insieme di dati "German Credit" del Repository of Machine Learning Databases (Blake e Merz, 1998) presso la University of California, Irvine.
- **grocery\_1month.sav.** Questo file di dati ipotetici corrisponde al file di dati *grocery\_coupons.sav* con gli acquisti settimanali organizzati in modo che ogni caso corrisponda a un cliente separato. Alcune delle variabili che cambiano settimanalmente non vengono riportate nei risultati; l'importo speso registrato corrisponde ora alla somma degli importi spesi durante le quattro settimane dello studio.
- **grocery\_coupons.sav.** File di dati ipotetici che contiene i dati sui sondaggi raccolti da una catena di drogherie interessata alle abitudini di acquisto dei suoi clienti. Ciascun cliente viene seguito per quattro settimane e ciascun caso corrisponde a una settimana per cliente con informazioni sul luogo degli acquisti e i tipi di acquisti, incluso l'importo speso nelle drogherie durante la settimana.
- **guttman.sav.** Bell (Bell, 1961) ha presentato una tabella per illustrare i possibili gruppi sociali. Guttman (Guttman, 1968) ha utilizzato una parte di tale tabella, in cui cinque variabili che descrivono elementi come l'interazione sociale, i sentimenti di appartenenza a un gruppo, la vicinanza fisica dei membri e il grado di formalità della relazione, sono state incrociate con cinque gruppi sociali teorici, compresi folla (ad esempio, le persone presenti a una partita di calcio), uditorio (ad esempio, di uno spettacolo teatrale o di una lezione universitaria), pubblico (ad esempio televisivo), calca (come una folla, ma con un'interazione molto maggiore), gruppi primari (intimi), gruppi secondari (volontari) e la comunità moderna (unione non stretta derivante da una vicinanza fisica elevata e dall'esigenza di servizi specializzati).
- **health\_funding.sav.** File di dati ipotetici che contiene i dati sui fondi di assistenza sanitaria (importo per 100 persone), sui tassi di malattie (tasso per 10.000 persone) e sulle visite ai fornitori di assistenza sanitaria (tasso per 10.000 persone). Ogni caso rappresenta una diversa città.
- **hivassay.sav.** File di dati ipotetici che prende in esame l'impegno di un'industria farmaceutica nel tentativo di sviluppare un'analisi che riesca a rilevare in tempi brevi l'infezione da virus HIV. I risultati dell'analisi sono otto sfumature di colore rosso sempre più intenso; le sfumature più intense indicano la maggiore probabilità di infezione. Un esperimento di laboratorio è stato condotto su 2000 campioni di sangue. La metà di questi è risultata infetta al virus HIV, l'altra metà non è risultata infetta.
- **hourlywagedata.sav.** File di dati ipotetici che prende in esame la paga oraria degli infermieri occupati presso uffici e ospedali e in base ai diversi livelli di esperienza.
- **insurance\_claims.sav.** File di dati ipotetici che prende in esame una compagnia di assicurazioni impegnata nella creazione di un modello per contrassegnare le richieste di risarcimento sospette e potenzialmente fraudolente. Ogni caso rappresenta una richiesta di risarcimento separata.
- **insure.sav.** File di dati ipotetici che prende in esame una compagnia di assicurazioni impegnata nello studio dei fattori di rischio, che indicano l'eventualità che un cliente presenti una domanda di indennizzo in un contratto assicurativo sulla vita della durata di dieci anni. Ogni caso nel file di dati rappresenta una coppia di contratti. In un contratto sono contenute informazioni su una richiesta di risarcimento, l'altro sull'età e sul sesso.

- **judges.sav.** File di dati ipotetici che prende in esame il punteggio assegnato, da giurie qualificate (più un appassionato) a 300 prestazioni sportive. Ciascuna riga rappresenta una diversa prestazione; i giudici hanno esaminato le stesse prestazioni.
- **kinship\_dat.sav.** Rosenberg e Kim (Rosenberg e Kim, 1975) si prefiggono di analizzare 15 termini indicanti parentela (zia, fratello, cugino, padre, nipote femmina, di nonni, nonno, nonna, nipote maschio di nonni, madre, nipote maschio di zii), nipote femmina di zii, sorella, figlio, zio). Hanno richiesto a quattro gruppi di studenti universitari (due composti da femmine e due da maschi) di ordinare questi termini in base alla similitudine. A due gruppi (uno femminile e uno maschile) è stato richiesto di effettuare l'ordinamento due volte, con il secondo ordinamento basato su un criterio diverso rispetto al primo. Di conseguenza, sono state ottenute sei "sorgenti" in totale. Ogni sorgente corrisponde a una matrice di prossimità  $15 \times 15$ , le cui celle sono uguali al numero delle persone in una sorgente meno il numero di volte in cui gli oggetti sono stati ripartiti insieme nella sorgente.
- **kinship\_ini.sav.** Questo file di dati contiene la configurazione iniziale di una soluzione a tre dimensioni per *kinship\_dat.sav*.
- **kinship\_var.sav.** Questo file di dati contiene variabili indipendenti relative a *sexo*, *generazione* e *grado* di separazione che possono essere utilizzate per interpretare le dimensioni di una soluzione per *kinship\_dat.sav*. In modo specifico, tali variabili possono essere utilizzate per limitare lo spazio della soluzione a una combinazione lineare di tali variabili.
- **marketvalues.sav.** File di dati che prende in esame le vendite di abitazioni in un nuovo centro abitato in Algonquin, Ill., durante gli anni 1999–2000. Tali vendite sono una questione di dominio pubblico.
- **nhis2000\_subset.sav.** Il National Health Interview Survey (NHIS) è un sondaggio di grandi dimensioni condotto sulla popolazione civile americana. Le interviste vengono realizzate di persona e si basano su un campione rappresentativo di famiglie a livello nazionale. Per ogni membro di una famiglia vengono raccolte osservazioni e informazioni di carattere demografico relative allo stato di salute. Questo file di dati contiene un sottoinsieme delle informazioni ottenute dall'indagine del 2000. National Center for Health Statistics. National Health Interview Survey, 2000. File di dati e documentazione di dominio pubblico. [ftp://ftp.cdc.gov/pub/Health\\_Statistics/NCHS/Datasets/NHIS/2000/](ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/). Accesso 2003.
- **ozone.sav** I dati includono 330 osservazioni basate su sei variabili meteorologiche per quantificare la concentrazione dell'ozono dalle variabili rimanenti. I precedenti ricercatori, (Breiman e Friedman, 1985) e (Hastie e Tibshirani, 1990), hanno rilevato non linearità tra queste variabili, che impediscono un approccio di regressione standard.
- **pain\_medication.sav.** File di dati ipotetici che contiene i risultati di un test clinico per stabilire la cura antinfiammatoria per il trattamento del dolore generato dall'artrite cronica. Di particolare interesse, il test ha evidenziato il tempo che impiega il farmaco ad avere effetto e il confronto con altri farmaci esistenti.
- **patient\_los.sav.** File di dati ipotetici che contiene informazioni sul trattamento dei pazienti ricoverati per sospetto di infarto del miocardio. Ogni caso corrisponde a un diverso paziente e contiene diverse variabili correlate alla degenza nell'ospedale.
- **patlos\_sample.sav.** File di dati ipotetici che contiene informazioni sul trattamento di un campione di pazienti curato con trombolitici durante la degenza per infarto del miocardio. Ogni caso corrisponde a un diverso paziente e contiene diverse variabili correlate alla degenza nell'ospedale.

- **poll\_cs.sav.** File di dati ipotetici che prende in esame i sondaggi per stabilire il livello di sostegno pubblico nei confronti di un disegno di legge prima che diventi una legge vera e propria. I casi corrispondono ai votanti registrati. Ciascun caso riporta informazioni sulla contea, sul comune e sul quartiere in cui vive il votante.
- **poll\_cs\_sample.sav.** File di dati ipotetici che contiene un campione dei votanti elencati nel file *poll\_cs.sav*. Il campione è stato selezionato in base al disegno specificato nel file di piano *poll.csplan* e questo file di dati contiene le probabilità di inclusione e i pesi del campione. Tuttavia, notare che poiché fa uso del metodo PPS (probability-proportional-to-size, probabilità proporzionale alla dimensione), esiste anche un file contenente le probabilità di selezione congiunte (*poll\_jointprob.sav*). Le ulteriori variabili corrispondenti ai dati demografici dei votanti e alla loro opinione sul disegno di legge, sono state raccolte e aggiunte al file di dati dopo aver acquisito il campione.
- **property\_assess.sav.** File di dati ipotetici che prende in esame l'impegno di un perito di una contea nel tentativo di mantenere gli accertamenti sui valori delle proprietà aggiornati in base alle risorse limitate. I casi rappresentano le proprietà vendute nella contea nello scorso anno. Ogni caso nel file di dati contiene informazioni sul comune in cui si trova la proprietà, il perito che per ultimo ha visitato la proprietà, il tempo trascorso dall'accertamento, la valutazione fatta in tale momento e il valore di vendita della proprietà.
- **property\_assess\_cs.sav.** File di dati ipotetici che prende in esame l'impegno di un perito di uno stato nel tentativo di mantenere aggiornati gli accertamenti sui valori delle proprietà in base alle risorse limitate. I casi corrispondono alle proprietà nello stato. Ogni caso nel file di dati include informazioni sulla contea, il comune e il quartiere in cui risiede la proprietà, la data dell'ultimo accertamento e la valutazione fatta in tale data.
- **property\_assess\_cs\_sample.sav.** File di dati ipotetici che contiene un campione delle proprietà elencate nel file *property\_assess\_cs.sav*. Il campione è stato selezionato in base al disegno specificato nel file di piano *property\_assess.csplan* e questo file di dati contiene le probabilità di inclusione e i pesi del campione. L'ulteriore variabile *Valore corrente* è stata raccolta e aggiunta al file di dati dopo aver acquisito il campione.
- **recidivism.sav.** File di dati ipotetici che prende in esame l'impegno delle Forze dell'Ordine nel tentativo di valutare il tasso di recidività nella propria area di giurisdizione. Ogni caso corrisponde a un precedente trasgressore e include le informazioni demografiche, alcuni dettagli sul primo crimine, il tempo trascorso fino al secondo arresto e se tale arresto è avvenuto entro due anni dal primo.
- **recidivism\_cs\_sample.sav.** File di dati ipotetici che prende in esame l'impegno delle Forze dell'Ordine nel tentativo di valutare il tasso di recidività nella propria area di giurisdizione. Ogni caso corrisponde a un trasgressore precedente, rilasciato dopo il primo arresto durante il mese di giugno del 2003 e registra le relative informazioni demografiche, alcuni dettagli sul primo crimine commesso e i dati del secondo arresto, se si è verificato prima della fine di giugno del 2006. I trasgressori sono stati selezionati dai dipartimenti sottoposti a campione in base al piano di campionamento specificato nel file *recidivism\_cs.csplan*. Poiché viene utilizzato un metodo PPS (Probability-Proportional-to-Size, probabilità proporzionale alla dimensione), esiste anche un file contenente le probabilità di selezione congiunte (*recidivism\_cs\_jointprob.sav*).
- **rfm\_transactions.sav.** File di dati ipotetici contenente i dati delle transazioni di acquisto, inclusa la data di acquisto, gli articoli acquistati e il valore monetario di ciascuna transazione.

- **salesperformance.sav.** File di dati ipotetici che prende in esame la valutazione di due nuovi corsi di formazione alle vendite. Sessanta dipendenti, divisi in tre gruppi, ricevono tutti la formazione standard. In più, al gruppo 2 viene assegnato un corso di formazione tecnica e al gruppo 3 un'esercitazione pratica. Alla fine del corso di formazione, ciascun dipendente viene sottoposto a un esame e il punteggio conseguito viene registrato. Ciascun caso nel file di dati rappresenta un diverso partecipante. Il file di dati include il gruppo a cui è assegnato il partecipante e il punteggio conseguito all'esame finale.
- **satisf.sav.** File di dati ipotetico che prende in esame un'indagine sulla soddisfazione dei clienti condotta da una società di vendita al dettaglio presso 4 negozi. Sono stati intervistati 582 clienti e ciascun caso rappresenta le risposte ottenute da un singolo cliente.
- **screws.sav.** Questo file di dati contiene informazioni sulle caratteristiche di viti, bulloni, dadi e puntine (Hartigan, 1975).
- **shampoo\_ph.sav.** File di dati ipotetici che prende in esame il processo di controllo di qualità di un'industria di prodotti per capelli. A intervalli di tempo regolari, vengono misurati sei diversi lotti prodotti e ne viene registrato il relativo pH. I valori accettati sono compresi tra 4,5 e 5,5.
- **ships.sav.** Ad esempio, un insieme di dati presentato e analizzato altrove (McCullagh et al., 1989) riguarda i danni subiti dalle navi da carico a causa delle onde. I conteggi degli incidenti possono essere presentati con un tasso di Poisson in base al tipo di nave, al periodo di costruzione e al periodo di servizio. I mesi di servizio aggregati di ciascuna cella della tabella generata dalla classificazione incrociata dei fattori fornisce i valori di esposizione al rischio.
- **site.sav.** File di dati ipotetici che prende in esame l'impegno di una società nella scelta di nuovi siti in cui espandere la propria presenza. La società ha incaricato due consulenti separati che, oltre a valutare i siti e presentare un report completo, devono classificarli come potenzialmente "molto adatti", "adatti" o "poco adatti".
- **smokers.sav.** Questo file di dati è un estratto del 1998 National Household Survey of Drug Abuse e rappresenta un campione probabile di famiglie americane. (<http://dx.doi.org/10.3886/ICPSR02934>) Il primo passo nell'analisi di questo file di dati consiste quindi nel pesare i dati per rispecchiare le tendenze della popolazione.
- **stocks.sav** Questo file di dati ipotetici contiene i prezzi e i volumi delle scorte riferiti a un anno.
- **stroke\_clean.sav.** File di dati ipotetici che riporta lo stato di un database medico dopo averne eseguito la pulizia utilizzando le procedure del modulo Data Preparation.
- **stroke\_invalid.sav.** File di dati ipotetici che riporta lo stato iniziale di un database medico e contiene numerosi errori di immissione dati.
- **stroke\_survival.** Questo file di dati ipotetici riguarda i tempi di sopravvivenza per i pazienti che, dopo avere completato un programma riabilitativo in seguito a un ictus postischemico, affrontano alcune sfide. Dopo l'attacco, viene annotata l'occorrenza dell'infarto miocardico, dell'ictus ischemico o emorragico e viene registrata l'ora dell'evento. Questo campione viene troncato a sinistra perché include solo i pazienti che sono sopravvissuti fino alla fine del programma riabilitativo post-ictus.
- **stroke\_valid.sav.** File di dati ipotetici che riporta lo stato di un database medico dopo il controllo dei valori eseguito con la procedura Convalida i dati. Il database contiene comunque casi potenzialmente anomali.



- **survey\_sample.sav.** File di dati che contiene i dati dell'indagine, compresi i dati demografici e varie misure dell'atteggiamento. Si basa su un sottoinsieme di variabili tratte dal 1998 NORC General Social Survey, benché i valori di alcuni dati siano stati modificati e siano state aggiunte variabili fittizie a scopo dimostrativo.
- **telco.sav.** File di dati ipotetici che prende in esame l'impegno di un'azienda di telecomunicazioni nel tentativo di ridurre il churn, ovvero l'abbandono dei propri clienti. Ciascun caso rappresenta un cliente separato e riporta diverse informazioni demografiche e sull'uso del servizio.
- **telco\_extra.sav.** Questo file di dati è simile al file *telco.sav*, ma le variabili "tenure" e spesa del cliente trasformata tramite logaritmo sono state sostituite dalle variabili di spesa del cliente trasformata tramite logaritmo standardizzate.
- **telco\_missing.sav.** Questo file di dati è un sottoinsieme del file di dati *telco.sav*, ma alcuni dei valori di dati demografici sono stati sostituiti con valori mancanti.
- **testmarket.sav.** File di dati ipotetici che prende in esame i piani di una catena di fast food per aggiungere un nuovo prodotto al proprio menu. Sono previste tre campagne promozionali del nuovo prodotto. Il prodotto viene introdotto in diversi mercati selezionati in modo casuale. Per ogni sede viene utilizzata una promozione differente registrando le vendite settimanali della nuova voce per le prime quattro settimane. Ogni caso rappresenta un luogo e una settimana diversi.
- **testmarket\_1month.sav.** Questo file di dati ipotetici corrisponde al file *testmarket.sav* con le vendite settimanali organizzate in modo che ogni caso corrisponda a un luogo separato. Alcune delle variabili che cambiano settimanalmente non vengono riportate nei risultati; le vendite registrate corrispondono ora alla somma delle vendite conseguite durante le quattro settimane dello studio.
- **tree\_car.sav.** File di dati ipotetici che contiene dati demografici e sul prezzo di acquisto dei veicoli.
- **tree\_credit.sav.** File di dati ipotetici che contiene dati demografici e sulla cronologia dei mutui di una banca.
- **tree\_missing\_data.sav.** File di dati ipotetici che contiene dati demografici e sulla cronologia dei mutui di una banca con un numero elevato di valori mancanti.
- **tree\_score\_car.sav.** File di dati ipotetici che contiene dati demografici e sul prezzo di acquisto dei veicoli.
- **tree\_textdata.sav.** File di dati semplice con due variabili destinato principalmente per mostrare lo stato predefinito delle variabili prima dell'assegnazione dei livelli di misurazione e delle etichette dei valori.
- **tv-survey.sav.** File di dati ipotetici che prende in esame un sondaggio condotto da una emittente televisiva che deve stabilire se estendere la durata di un programma di successo. A un campione di 906 intervistati è stato chiesto se preferisce guardare il programma con diverse condizioni. Ciascuna riga rappresenta un diverso intervistato e ciascuna colonna una diversa condizione.
- **ulcer\_recurrence.sav.** Questo file contiene informazioni parziali su uno studio svolto per mettere a confronto l'efficacia di due terapie preventive per la recidiva delle ulcere. Fornisce un ottimo esempio di dati acquisiti a intervalli ed è stato presentato e analizzato in altri luoghi (Collett, 2003).

- **ulcer\_recurrence\_recoded.sav.** In questo file sono contenute le informazioni del file *ulcer\_recurrence.sav* riorganizzate per consentire di presentare la probabilità degli eventi per ciascun intervallo dello studio, anziché solo alla fine. È stato presentato e analizzato in altri luoghi (Collett et al., 2003).
- **verd1985.sav.** Questo file di dati prende in esame un'indagine (Verdegaal, 1985). Sono state registrate le risposte di quindici soggetti a otto variabili. Le variabili di interesse sono suddivise in tre insiemi. L'insieme 1 include *età* e *statociv*, l'insieme 2 include *andom* e *giornale* e l'insieme 3 include *musica* e *vicinato*. *Andom* viene scalata come nominale multipla ed *età* come ordinale; tutte le altre variabili vengono scalate come nominali singole.
- **virus.sav.** File di dati ipotetici che prende in esame l'impegno di un ISP (Internet Service Provider) nel tentativo di determinare gli effetti che un virus può generare nelle sue reti. Si è tenuta traccia della percentuale (approssimativa) di traffico e-mail infettato da virus sulla rete in un lasso di tempo, dal momento dell'individuazione fino alla soppressione della minaccia.
- **wheeze\_steubenville.sav.** Questo file è un sottoinsieme di uno studio longitudinale degli effetti che l'inquinamento provoca sulla salute dei bambini (Ware, Dockery, Spiro III, Speizer, e Ferris Jr., 1984). I dati contengono misure binarie ripetute del livello di asma dei bambini della città di Steubenville, Ohio, di 7, 8, 9 e 10 anni. I dati indicano anche se la mamma dei bambini era fumatrice durante il primo anno dello studio.
- **workprog.sav.** File di dati ipotetici che prende in esame un programma di lavoro governativo il cui obiettivo è fornire attività più adatte alle persone diversamente abili. È stato seguito un campione di potenziali partecipanti al programma, alcuni dei quali sono stati selezionati in modo casuale e altri no. Ogni caso rappresenta un diverso partecipante al programma.
- **worldsales.sav** Questo file di dati ipotetici contiene i ricavi suddivisi per continenti e prodotti.

## Note

Queste informazioni sono state preparate per prodotti e servizi offerti in tutto il mondo.

IBM potrebbe non offrire i prodotti, i servizi o le funzionalità di cui si tratta nel presente documento in altri paesi. Contattare il rappresentante IBM locale per informazioni sui prodotti e i servizi attualmente disponibili nella propria zona. Qualsiasi riferimento a un prodotto, programma o servizio IBM non intende dichiarare o implicare che sia possibile utilizzare esclusivamente tale prodotto, programma o servizio IBM. Potrà invece essere utilizzato qualsiasi prodotto, programma o servizio con funzionalità equivalente e che non violi i diritti di proprietà intellettuale di IBM. Tuttavia, è responsabilità dell'utente valutare e verificare il funzionamento di qualsiasi prodotto, programma o servizio non IBM.

IBM può essere titolare di brevetti o domande di brevetto relativi alla materia oggetto del presente documento. La consegna del presente documento non conferisce alcuna licenza rispetto a questi brevetti. Rivolgere per iscritto i quesiti sulle licenze a:

*IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785, U.S.A.*

Per richieste di informazioni sulle licenze riguardanti il set di caratteri a byte doppio (DBCS), contattare l'Intellectual Property Department di IBM del proprio paese, oppure inviare le richieste in forma scritta all'indirizzo:

*Intellectual Property Licensing, Legal and Intellectual Property Law, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 242-8502 Giappone.*

**Il seguente paragrafo non si applica per il Regno Unito o altri paesi in cui le presenti disposizioni non sono conformi alle leggi locali:** INTERNATIONAL BUSINESS MACHINES FORNISCE QUESTA PUBBLICAZIONE “COSÌ COM'È” SENZA GARANZIA DI ALCUN TIPO, SIA ESSA ESPRESSA O IMPLICITA, INCLUSE, MA NON LIMITATE A, LE GARANZIE IMPLICITE DI NON VIOLAZIONE, COMMERCIALIZZABILITÀ O IDONEITÀ A UNO SCOPO SPECIFICO. Alcuni stati non consentono limitazioni di garanzie espresse o implicite in determinate transazioni, pertanto quanto sopra potrebbe non essere applicabile.

Le presenti informazioni possono includere imprecisioni tecniche o errori tipografici. Le modifiche periodiche apportate alle informazioni contenute in questa pubblicazione verranno inserite nelle nuove edizioni della pubblicazione. IBM può apportare miglioramenti e/o modifiche al/ai prodotto/i e/o al/ai programma/i descritti nella presente pubblicazione in qualsiasi momento senza preavviso.

Qualsiasi riferimento nelle presenti informazioni a siti Web non IBM viene fornito esclusivamente per facilitare la consultazione e non rappresenta in alcun modo un'approvazione o sostegno da parte nostra di tali siti Web. I materiali contenuti in tali siti Web non fanno parte dei materiali di questo prodotto IBM e il loro utilizzo è esclusivamente a rischio dell'utente.

IBM può utilizzare o distribuire eventuali informazioni fornite dall'utente nei modi che ritiene appropriati senza incorrere in alcun obbligo nei confronti dell'utente.

I licenziatari del programma che desiderassero informazioni su di esso allo scopo di abilitare: (i) lo scambio di informazioni tra programmi creati indipendentemente e altri programmi (questo compreso) e (ii) l'utilizzo in comune delle informazioni scambiate, dovranno rivolgersi a:

*IBM Software Group, All'attenzione di: Licensing, 233 S. Wacker Dr., Chicago, IL 60606, USA.*

Tali informazioni saranno fornite in conformità ai termini e alle condizioni in vigore e, in alcuni casi, dietro pagamento.

Il programma concesso in licenza descritto nel presente documento e tutto il materiale correlato disponibile sono forniti da IBM in base ai termini del contratto di licenza cliente IBM, del contratto di licenza internazionale IBM o del contratto equivalente esistente tra le parti.

le informazioni relative a prodotti non IBM sono state ottenute dai fornitori di tali prodotti, da loro annunci pubblicati e da altre fonti disponibili al pubblico. IBM non ha verificato tali prodotti e non può confermare l'accuratezza delle prestazioni, la compatibilità o qualsiasi altra dichiarazione relativa a prodotti non IBM. Eventuali domande in merito alle funzionalità dei prodotti non IBM vanno indirizzate ai fornitori di tali prodotti.

Le presenti informazioni includono esempi di dati e report utilizzati in operazioni aziendali quotidiane. Per fornire una descrizione il più possibile esaustiva, gli esempi includono nomi di persone, società, marchi e prodotti. Tutti questi nomi sono fittizi e ogni somiglianza a nomi e indirizzi utilizzati da aziende reali è puramente casuale.

Per chi visualizza queste informazioni a video: le fotografie e le illustrazioni a colori potrebbero non essere disponibili.

### ***Marchi***

IBM, il logo IBM, ibm.com e SPSS sono marchi di IBM Corporation, registrati in numerose giurisdizioni nel mondo. Un elenco aggiornato dei marchi IBM è disponibile sul Web all'indirizzo <http://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.

Java e tutti i marchi e i logo basati su Java sono marchi di Sun Microsystems, Inc. negli Stati Uniti e/o negli 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 registrato di The Open Group negli Stati Uniti e in altri paesi.

Questo prodotto utilizza WinWrap Basic, Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com>.

Altri nomi di prodotti e servizi possono essere marchi commerciali di IBM o di altre aziende.

---

Le schermate dei prodotti Adobe sono state ristampate su autorizzazione di Adobe Systems Incorporated.

Le schermate dei prodotti Microsoft sono state ristampate su autorizzazione di Microsoft Corporation.



---

# Bibliografia

- Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.
- Blake, C. L., e C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mlearn/MLRepository.html>.
- Breiman, L., e J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, .
- Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.
- Gardner, E. S. 1985. Exponential smoothing: The state of the art. *Journal of Forecasting*, 4, .
- Green, P. E., e V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.
- Green, P. E., e Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.
- Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, .
- Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.
- Hastie, T., e R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.
- Kennedy, R., C. Riquier, e B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement and Analysis for Marketing*, 5, .
- McCullagh, P., e J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.
- Pena, D., G. C. Tiao, e R. S. Tsay, ed. 2001. *A course in time series analysis*. New York: John Wiley and Sons.
- Prezzo, R. H., e D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, .
- Rickman, R., N. Mitchell, J. Dingman, e J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228, .
- Rosenberg, S., e M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, .
- Scatola, G. E. P., G. M. Jenkins, e G. C. Reinsel. 1994. *Time series analysis: Forecasting and control*, 3rd ed. Englewood Cliffs, N.J.: Prentice Hall.
- Van der Ham, T., J. J. Meulman, D. C. Van Strien, e H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, .
- Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (in Dutch)*. Leiden: Department of Data Theory, University of Leiden.
- Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, e B. G. Ferris Jr.. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases*, 129, .

- definizioni, 98
- ACF
  - grafici per i processi ARIMA puri, 99
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- analisi armonica, 41
- Applica modelli di serie storiche, 27, 57, 73
  - funzione di autocorrelazione dei residui, 30, 32
  - funzione di autocorrelazione parziale dei residui, 30, 32
  - intervalli di confidenza, 32, 36
  - modelli migliori e peggiori, 34
  - nuova stima dei parametri del modello, 29, 58
  - nuovi nomi delle variabili, 35, 61
  - parametri del modello, 30
  - periodo di previsione, 29, 58, 80
  - periodo di stima, 29
  - previsioni, 30, 32, 81
  - salvataggio dei modelli stimati di nuovo in formato XML, 35
  - salvataggio delle previsioni, 35, 59
  - statistica di Box-Ljung, 30
  - statistiche su tutti i modelli, 30, 32, 60
  - statistiche sulla bontà dell'adattamento., 30, 32, 60
  - tabella Adattamento del modello, 60
  - tabella delle previsioni, 81
  - valori di adattamento, 32
  - valori mancanti, 36
- autoregressione
  - modelli ARIMA, 12
- BIC normalizzato (Criterio di informazione bayesiano), 97
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- bontà di adattamento
  - definizioni, 97
  - in Applica modelli delle serie storiche, 30, 32, 60
  - in Time Series Modeler, 18, 20, 52
- casi di controllo, 2
- dati storici
  - in Applica modelli delle serie storiche, 32
  - in Time Series Modeler, 20
- Decomposizione stagionale, 38–40
  - assunzioni, 38
  - calcolo delle medie mobili, 38
  - componente di data periodica, 84
  - creazione di variabili, 39
  - modelli, 38
  - nuove variabili, 90
  - procedure correlate, 91
  - salvataggio di nuove variabili, 39
- errore massimo assoluto, 97
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- errore medio assoluto, 97
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- errore percentuale massimo assoluto, 97
  - in Applica modelli delle serie storiche, 30, 32, 61
  - in Time Series Modeler, 18, 20, 53
- errore percentuale medio assoluto, 97
  - in Applica modelli delle serie storiche, 30, 32, 61
  - in Time Series Modeler, 18, 20, 53
- events, 8
  - in Time Series Modeler, 7
- Expert Modeler, 6, 46
  - limitazione dello spazio del modello, 7, 49
  - valori anomali, 9, 66
- file di esempio
  - posizione, 103
- funzione di autocorrelazione
  - grafici per i processi ARIMA puri, 99
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- funzione di autocorrelazione parziale
  - grafici per i processi ARIMA puri, 99
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- funzioni di trasferimento, 14
  - ordini della differenza, 14
  - ordini denominatore, 14
  - ordini numeratore, 14
  - ordini stagionali, 14
  - ritardo, 14
- Grafici spettrali, 41, 43
  - analisi spettrale bivariata, 43
  - assunzioni, 41
  - centratura trasformazione, 43
  - densità spettrale, 94
  - finestre spettrali, 41
  - periodogramma, 94
  - procedure correlate, 96
- integrazione
  - modelli ARIMA, 12
- intervalli di confidenza
  - in Applica modelli delle serie storiche, 32, 36
  - in Time Series Modeler, 20, 24
- MAE, 97
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- MAPE, 97
  - in Applica modelli delle serie storiche, 30, 32, 61

- in Time Series Modeler, 18, 20, 53
- marchi, 114
- MaxAE, 97
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- MaxAPE, 97
  - in Applica modelli delle serie storiche, 30, 32, 61
  - in Time Series Modeler, 18, 20, 53
- media mobile
  - modelli ARIMA, 12
- modelli
  - ARIMA, 6, 12
  - Expert Modeler, 6
  - livellamento esponenziale, 6, 10
- modelli ARIMA, 6
  - costante, 12
  - funzioni di trasferimento, 14
  - ordini autoregressivi, 12
  - ordini di differenziazione, 12
  - ordini di media mobile, 12
  - ordini stagionali, 12
  - valori anomali, 16
- modelli di livellamento esponenziale, 6, 10
- Modelli serie storiche, 3
  - ARIMA, 6, 12
  - events, 7
  - Expert Modeler, 6, 46, 62
  - funzione di autocorrelazione dei residui, 18, 20
  - funzione di autocorrelazione parziale dei residui, 18, 20
  - funzioni di trasferimento, 14
  - intervalli di confidenza, 20, 24
  - livellamento esponenziale, 6, 10
  - modelli migliori e peggiori, 22
  - nomi dei modelli, 24
  - nuovi nomi delle variabili, 23, 55
  - parametri del modello, 18, 68
  - periodicità, 7, 10, 12, 14
  - periodo di previsione, 6, 24, 49–50
  - periodo di stima, 6, 49
  - previsioni, 18, 20, 52
  - salvataggio delle previsioni, 23, 51
  - salvataggio delle specifiche dei modelli in XML, 23, 51, 67
  - statistica di Box-Ljung, 18
  - statistiche su tutti i modelli, 18, 20, 52, 54
  - statistiche sulla bontà dell'adattamento., 18, 20, 52, 71
  - tabella dei parametri del modello ARIMA, 71
  - tabella delle previsioni, 56
  - tabella delle statiche del modello, 71
  - tabella descrittiva del modello, 70
  - trasformazione della serie, 10, 12, 14
  - valori anomali, 9, 16, 66
  - valori di adattamento, 20, 69
  - valori mancanti, 24
- modello di livellamento esponenziale di Brown, 10
- modello di livellamento esponenziale di Holt, 10
- modello di livellamento esponenziale di Winters
  - additivo, 10
  - moltiplicativo, 10
- modello di livellamento esponenziale semplice, 10
- modello di livellamento esponenziale smorzato, 10
- modello di livellamento esponenziale stagionale semplice, 10
- nomi dei modelli
  - in Time Series Modeler, 24
- nomi di variabili
  - in Applica modelli delle serie storiche, 35
  - in Time Series Modeler, 23
- note legali, 113
- nuova stima dei parametri del modello
  - in Applica modelli delle serie storiche, 29, 58
- ordini stagionali
  - modelli ARIMA, 12
- PACF
  - grafici per i processi ARIMA puri, 99
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- parametri del modello
  - in Applica modelli delle serie storiche, 30
  - in Time Series Modeler, 18, 68
- periodicità
  - in Time Series Modeler, 7, 10, 12, 14
- periodo di convalida, 2
- periodo di previsione
  - in Applica modelli delle serie storiche, 29, 58, 80
  - in Time Series Modeler, 6, 24, 49–50
- periodo di stima, 2
  - in Applica modelli delle serie storiche, 29
  - in Time Series Modeler, 6, 49
- periodo storico, 2
- previsioni
  - in Applica modelli delle serie storiche, 30, 32, 81
  - in Time Series Modeler, 18, 20, 52
- $R^2$ , 97
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- radice della media dei quadrati, 97
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- residui
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- RMSE, 97
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20
- $R$ quadrato stazionario, 97
  - in Applica modelli delle serie storiche, 30, 32
  - in Time Series Modeler, 18, 20, 71



- 
- Salva
    - modelli stimati di nuovo in formato XML, 35
    - nuovi nomi delle variabili, 23, 35
    - specifiche del modello in formato XML, 23
    - valori stimati del modello, 23, 35
  - statistica di Box-Ljung
    - in Applica modelli delle serie storiche, 30
    - in Time Series Modeler, 18, 71
  
  - tabella Adattamento del modello
    - in Applica modelli delle serie storiche, 60
  - tabella dei parametri del modello ARIMA
    - in Time Series Modeler, 71
  - tabella delle previsioni
    - in Applica modelli delle serie storiche, 81
    - in Time Series Modeler, 56
  - tabella delle statiche del modello
    - in Time Series Modeler, 71
  - tabella descrittiva del modello
    - in Time Series Modeler, 70
  - trasformazione a radice quadrata
    - in Time Series Modeler, 10, 12, 14
  - trasformazione della differenza
    - modelli ARIMA, 12
  - trasformazione della differenza stagionale
    - modelli ARIMA, 12
  - trasformazione logaritmica
    - in Time Series Modeler, 10, 12, 14
  - trasformazione logaritmica naturale
    - in Time Series Modeler, 10, 12, 14
  
  - valore anomalo additivo stagionale, 98
    - in Time Series Modeler, 9, 16
  - valore anomalo blocco additivo, 98
    - in Time Series Modeler, 9, 16
  - valore anomalo dello spostamento di livello, 98
    - in Time Series Modeler, 9, 16
  - valore anomalo innovazionale, 98
    - in Time Series Modeler, 9, 16
  - valore anomalo transiente, 98
    - in Time Series Modeler, 9, 16
  - valore anomalo trend locale, 98
    - in Time Series Modeler, 9, 16
  - valori anomali
    - Expert Modeler, 9, 66
    - modelli ARIMA, 16
  - valori anomali additivi, 98
    - in Time Series Modeler, 9, 16
  - valori di adattamento
    - in Applica modelli delle serie storiche, 32
    - in Time Series Modeler, 20, 69
  - valori mancanti
    - in Applica modelli delle serie storiche, 36
    - in Time Series Modeler, 24
  
  - XML
    - salvataggio dei modelli stimati di nuovo in formato XML, 35
    - salvataggio delle serie storiche in XML, 23, 51, 67