

IBM SPSS Forecasting 22

The IBM logo, consisting of the letters 'IBM' in a bold, black, sans-serif font. Each letter is composed of eight horizontal bars of equal height and width, stacked vertically to form the characters. The logo is centered horizontally on the page.

IBM

お願い

本書および本書で紹介する製品をご使用になる前に、37 ページの『特記事項』に記載されている情報をお読みください。

本書は、IBM SPSS Statistics バージョン 22 リリース 0 モディフィケーション 0 および新しい版で明記されない限り、以降のすべてのリリースおよびモディフィケーションに適用されます。

お客様の環境によっては、資料中の円記号がバックスラッシュと表示されたり、バックスラッシュが円記号と表示されたりする場合があります。

原典： IBM SPSS Forecasting 22

発行： 日本アイ・ビー・エム株式会社

担当： トランスレーション・サービス・センター

目次

第 1 章 時系列の概要	1	統計テーブルと予測テーブル	17
時系列データ	1	プロット	18
データの変換	1	出力を最適モデルまたは最不適モデルに限定	19
推定期間と検証期間	2	モデル予測およびモデル指定の保存	19
モデルの構築と予測の生成	2	オプション	20
第 2 章 時系列モデラー	3	TSAPPLY コマンドの追加機能	20
エキスパート・モデラーのオプションの指定	5	第 4 章 季節性の分解	23
モデルの選択とイベントの指定	5	季節性の分解の保存	24
エキスパート・モデラーでの外れ値の処理	5	SEASON コマンドの追加機能	24
カスタム指数平滑法モデル	6	第 5 章 スペクトル・プロット	25
カスタム ARIMA モデル	7	SPECTRA コマンドの追加機能	27
カスタム ARIMA モデルに対するモデル指定	7	第 6 章 適合度測定結果	29
カスタム ARIMA モデルにおける転送関数	8	第 7 章 外れ値のタイプ	31
カスタム ARIMA モデルの外れ値	9	第 8 章 ACF/PACF プロットについて	33
出力	10	特記事項	37
統計テーブルと予測テーブル	10	商標	38
プロット	11	索引	41
出力を最適モデルまたは最不適モデルに限定	11		
モデル予測およびモデル指定の保存	12		
オプション	13		
TSMODEL コマンドの追加機能	14		
第 3 章 時系列モデルの適用	15		
出力	17		

第 1 章 時系列の概要

時系列とは、一定期間にわたって定期的に単一の変数を測定することによって得られる、一連の観測値です。例えば、在庫データの系列では、観測値によって数カ月間の日単位の日在庫レベルが表されることがあります。製品の市場占有率を示す系列が、数年間記録された週単位の市場占有率で構成されることがあります。総売上高の系列が、長年にわたる 1 カ月に 1 つの観測値で構成されることがあります。これらの例のそれぞれに共通するのは、いくつかの変数が特定の期間にわたって、定期的な既知の間隔で観測されていることです。したがって、一般的な時系列のデータ形式は、一定間隔で行われた測定を表す観測値の 1 つのシーケンスまたはリストになります。

表 1. 日単位の在庫の時系列

時間	週	曜日	在庫レベル
t_1	1	月	160
t_2	1	火	135
t_3	1	水	129
t_4	1	木	122
t_5	1	金	108
t_6	2	月	150
		...	
t_{60}	12	金	120

時系列分析を実施する最も重要な理由の 1 つは、この系列の今後の値を予測するためです。過去の値を説明する系列のモデルによって、今後の数個の値が増えるか、減るか、それがどの程度であるかを予測することもできます。このような予測を的確に行う能力が、ビジネス分野または科学分野で重要なのは明らかです。

時系列データ

Forecasting アドオン・モジュールで使用するための時系列データを定義する場合、各系列は個別の変数に対応します。例えば、データ・エディターで時系列を定義するには、「変数ビュー」タブをクリックし、空白行に変数名を入力します。時系列の各観測値はケース (データ・エディターの行) に対応します。

時系列データを含むスプレッドシートを開いた場合は、各系列がスプレッドシートの列に配置される必要があります。時系列が行に配置されたスプレッドシートが既にある場合は、そのスプレッドシートを開き、「データ」メニューの「入れ替え」を使用して行と列を入れ替えることができます。

データの変換

Core System で提供されている多くのデータ変換プロシージャは、時系列分析に役立ちます。

- 「日付の定義」プロシージャ (「データ」メニュー) では、周期の設定や、履歴期間、検証期間および予測期間の区別に使用する日付変数が生成されます。Forecasting は、「日付の定義」プロシージャにより作成された変数を操作するように設計されています。

- 「時系列の作成」プロシージャー（「変換」メニュー）では、新しい時系列変数を既存の時系列変数の関数として作成します。これには、平滑化、平均化、および差分に近接観測値を使用する関数が含まれています。
- 「欠損値の置換」プロシージャー（「変換」メニュー）では、システム欠損値とユーザー欠損値を、いくつかある方法のいずれか 1 つに基づいた推定値に置き換えます。系列の最初または最後に欠損データがあっても特に問題はありません。これらの欠損データにより、使用できる系列の長さが短くなるだけです。系列の途中にあるギャップ（埋め込まれた欠損データ）は、より深刻な問題となる可能性があります。

時系列のデータ変換について詳しくは、「*Core System*・ユーザーズ・ガイド」を参照してください。

推定期間と検証期間

多くの場合、時系列を推定（または履歴）期間と、検証期間に分けると便利です。推定（履歴）期間の観測値に基づいてモデルを作成した上で、検証期間においてそのモデルがどの程度有効に機能するかについての検証を行います。既に把握している点（検証期間内の点）に対してモデルを適用して予測を行うことにより、そのモデルが予測においてどの程度有効であるかを知ることができます。

検証期間のケースは、モデルを構築するプロセスから除外されるため、通常はホールドアウト・ケースと呼ばれます。推定期間は、アクティブなデータ・セットで現在選択されているケースで構成されます。選択された最後のケースより後にある残りのケースはすべて、ホールドアウトとして使用できます。モデルが予測に十分役立つことが確認できたら、推定期間を再定義してホールドアウト・ケースを取り込み、最終モデルを構築することができます。

モデルの構築と予測の生成

Forecasting アドオン・モジュールには、モデルの作成および予測の生成の作業を行うために 2 つのプロシージャーが用意されています。

- 3 ページの『第 2 章 時系列モデラー』のプロシージャーでは、時系列に対するモデルを作成し、予測を生成します。このプロシージャーでは、時系列のそれぞれに対して最適なモデルを自動的に特定するエキスパート・モデラーを使用します。より詳細な制御を必要とする経験豊富なアナリストを対象として、カスタム・モデル構築用のツールも用意されています。
- 15 ページの『第 3 章 時系列モデルの適用』のプロシージャーでは、時系列モデラーで作成した既存の時系列モデルを、アクティブなデータ・セットに適用します。これにより、モデルを再構築することなく、新しいデータまたは変更されたデータに対する系列の予測を得ることができます。モデルが変更されたと判断する根拠がある場合は、時系列モデラーを使用してモデルを再構築できます。

第 2 章 時系列モデラー

時系列モデラーのプロシージャーでは、時系列に対する指数平滑法、1 変量の自己回帰和分移動平均 (ARIMA)、および多変量の ARIMA (または転送関数モデル) モデルを推定し、予測を生成します。このプロシージャーでは、1 つ以上の従属変数系列に対して最適な ARIMA または指数平滑法モデルを自動的に特定および推定するエキスパート・モデラーを使用するため、試行錯誤によって適切なモデルを特定する必要がなくなりました。または、カスタム ARIMA または指数平滑法モデルを指定することもできます。

例: プロダクト・マネージャーとして、100 品目の別個の製品に対する翌月の販売個数および収益の予測を担当しますが、時系列のモデル作成の経験がほとんどないか、まったくありません。100 品目すべての製品の販売個数の履歴データは、1 つの Excel スプレッドシートに保存されています。スプレッドシートを IBM® SPSS® Statistics で開いた後に、エキスパート・モデラーを使用して今後 1 カ月の予測を要求します。エキスパート・モデラーは、各製品の販売個数の最適モデルを探し出し、これらのモデルを使用して予測を生成します。エキスパート・モデラーでは、複数の入力系列を処理できるため、プロシージャーを一度実行するだけで、全製品の予測を取得することができます。予測をアクティブなデータ・セットに保存することを選択した場合は、その結果を簡単に Excel にエクスポートすることができます。

統計: 適合度: 定常 R^2 乗、 R^2 乗 (R^2)、2 乗平均平方根誤差 (RMSE)、平均絶対誤差 (MAE)、平均絶対パーセント誤差 (MAPE)、最大絶対誤差 (MaxAE)、最大絶対パーセント誤差 (MaxAPE)、正規化ベイズ情報量基準 (BIC)。残差: 自己相関関数、偏自己相関関数、Ljung-Box Q 。ARIMA モデルの場合: 従属変数の ARIMA の次数、独立変数の転送関数の次数、外れ値の推定値。さらに、指数平滑法モデルの平滑法パラメーター推定値。

プロット. すべてのモデルの要約プロット: 定常 R^2 乗、 R^2 乗 (R^2)、2 乗平均平方根誤差 (RMSE)、平均絶対誤差 (MAE)、平均絶対パーセント誤差 (MAPE)、最大絶対誤差 (MaxAE)、最大絶対パーセント誤差 (MaxAPE)、正規化ベイズ情報量基準 (BIC) のヒストグラム。残差自己相関および偏自己相関の箱ひげ図。個々のモデルの結果: 予測値、当てはめ値、観測値、信頼限界の上限と下限、残差自己相関および偏自己相関。

時系列モデラー・データの考慮事項

データ: 従属変数および独立変数は数値である必要があります。

仮定: 従属変数および独立変数は時系列として扱われます。つまり、各ケースは時点を表し、連続するケースは一定の時間間隔で区切られています。

- **定常性.** カスタム ARIMA モデルの場合、モデル作成される時系列は定常である必要があります。非定常系列から定常系列への変換で最も有効なのは、差分変換です (「時系列の作成」ダイアログ・ボックスから実行可能です)。
- **予測.** 独立 (予測) 変数を持つモデルを使用して予測を生成するには、予測期間内のすべてのケースについて、これらの変数の値がアクティブなデータ・セットに含まれている必要があります。また、推定期間内の独立変数に欠損値が含まれていないことが必要です。

日付の定義

必須ではありませんが、「日付の定義」ダイアログ・ボックスを使用して、最初のケースに関連付ける日付、および連続するケースの間の時間間隔を指定することをお勧めします。これは、時系列モデラーを使用する前に行い、その結果、各ケースに関連付けられる日付をラベル付けする変数セットが生成されます。ま

た、データの想定周期も設定されます。例えば、連続するケース間の時間間隔が 1 カ月の場合、周期は 12 となります。このような周期は、季節モデルの作成に関心がある場合に必要になります。季節モデルには関心がなく、出力に日付ラベルが必要ない場合は、「日付の定義」ダイアログ・ボックスを省略できます。その場合、各ケースに関連付けられるラベルはケース番号のみです。

時系列モデラーを使用するには

1. メニューから次の項目を選択します。

「分析」 > 「予測」 > 「モデルの作成...」

2. 「変数」タブで、モデル作成される 1 つ以上の従属変数を選択します。

3. 「方法」ドロップダウン・ボックスで、モデル作成の方法を選択します。自動モデル作成の場合は、デフォルトの「エキスパート・モデラー」の方法のままにします。これにより、エキスパート・モデラーが起動され、各従属変数の最適モデルが特定されます。

予測を生成するには:

4. 「オプション」タブをクリックします。

5. 予測期間を指定します。これにより、予測値と観測値を含むグラフが生成されます。

必要に応じて、次のことが可能です。

- 1 つ以上の独立変数を選択する。独立変数は、回帰分析の予測変数とほとんど同様に処理されますが、省略可能です。ARIMA モデルに含めることはできますが、指数平滑法モデルには含めることができません。モデル作成方法として「エキスパート・モデラー」を指定し、独立変数を含めた場合は、ARIMA モデルのみが考慮されます。
- 「基準」をクリックして、モデル作成の詳細を指定する。
- 予測、信頼区間、およびノイズ残差を保存する。
- 推定されたモデルを XML 形式で保存する。保存済みモデルを新しいデータまたは変更されたデータに適用することで、モデルを再作成することなく更新された予測を得ることができます。
- すべての推定モデルから、要約統計量を取得する。
- カスタム ARIMA モデルの独立変数に、転送関数を指定する。
- 外れ値の自動検出を有効にする。
- 特定の時点を、カスタム ARIMA モデルの外れ値としてモデル作成する。

モデル作成の方法

使用可能なモデル作成の方法は以下のとおりです。

エキスパート・モデラー。エキスパート・モデラーは、各従属系列の最適モデルを自動的に見つけます。独立 (予測) 変数を指定した場合、エキスパート・モデラーは、従属系列と統計的に有意な関係を持つ独立変数を、ARIMA モデルに含めるために選択します。適切な場合には、差分を使用するか、平方根変換または自然対数変換を使用するか、またはその両方を行ってモデル変数が変換されます。デフォルトでは、エキスパート・モデラーは指数平滑法モデルおよび ARIMA モデルの両方を考慮します。ただし、ARIMA モデルのみを検索するか、指数平滑法モデルのみを検索するように、エキスパート・モデラーに制限を加えることができます。また、外れ値の自動検出も指定できます。

指数平滑法。このオプションは、カスタム指数平滑法モデルを指定する場合に使用します。トレンドと季節性の処理が異なる各種の指数平滑法モデルから選択できます。

ARIMA。このオプションは、カスタム ARIMA モデルを指定する場合に使用します。これを使用する場合は、差分の次数に加えて、自己回帰の次数および移動平均の次数も明示的に指定する必要があります。独立(予測)変数を含めて、その一部または全部に対して転送関数を定義できます。また、外れ値の自動検出を指定するか、外れ値のセットを明示的に指定することもできます。

推定期間と予測期間

推定期間。推定期間により、モデルの特定に使用される一連のケースが定義されます。デフォルトでは、推定期間にはアクティブなデータ・セットのすべてのケースが含まれます。推定期間を設定するには、「ケースの選択」ダイアログ・ボックスで「**時間またはケースの範囲に基づく**」を選択します。使用可能なデータによっては、プロシージャーによって使用される推定期間が従属変数ごとに異なるため、表示される値と異なる場合があります。任意の従属変数に対する本当の推定期間は、指定された推定期間の最初または最後で発生する、変数の連続する欠損値を除去した後に残った期間です。

予測期間。予測期間は、推定期間の後の最初のケースで始まり、デフォルトでは、アクティブなデータ・セット内の最後のケースまでです。予測期間の最後は、「オプション」タブから設定できます。

エキスパート・モデラーのオプションの指定

エキスパート・モデラーでは、候補モデルのセットに制約を加えるオプション、外れ値の処理を指定するオプション、およびイベント変数を含めるオプションが提供されます。

モデルの選択とイベントの指定

「モデル」タブでは、エキスパート・モデラーの考慮対象とするモデルのタイプの指定、およびイベント変数の指定が可能です。

「**モデル・タイプ**」。使用可能なオプションは、以下のとおりです。

- 「**すべてのモデル**」。エキスパート・モデラーは、ARIMA と指数平滑法モデルの両方を考慮します。
- 「**指数平滑法モデルのみ**」。エキスパート・モデラーは、指数平滑法モデルのみを考慮します。
- 「**ARIMA モデルのみ**」。エキスパート・モデラーは ARIMA モデルのみを考慮します。

「**エキスパート・モデラーが季節モデルを考慮する**」。このオプションは、アクティブなデータ・セットに周期が定義されている場合にのみ有効です。このオプションを選択する(チェック・マークを付ける)と、エキスパート・モデラーは季節モデルと非季節モデルの両方を考慮します。このオプションを選択しない場合、エキスパート・モデラーは非季節モデルのみを考慮します。

「**現在の周期**」。アクティブなデータ・セットに現在定義されている周期(ある場合)を示します。現在の周期は、整数値として指定されます。例えば、各ケースが1カ月を表す1年周期の場合は12となります。周期が設定されていない場合は、値「なし」が表示されます。季節モデルには周期が必要です。周期は、「日付の定義」ダイアログ・ボックスから設定できます。

「**イベント**」。イベント変数として扱われる独立変数を選択します。イベント変数に対して値1を持つケースは、従属系列がそのイベントの影響を受けると予想される時間を示します。1以外の値は、影響を受けないことを示します。

エキスパート・モデラーでの外れ値の処理

「外れ値」タブでは、外れ値の自動検出、および検出対象の外れ値のタイプを選択できます。

「自動的に外れ値を検出」。デフォルトでは、外れ値の自動検出は実行されません。外れ値の自動検出を実行するには、このオプションを選択し (チェック・マークを付け)、以下の外れ値のタイプを 1 つ以上選択します。

- 加法
- レベル・シフト
- 革新
- 過渡
- 季節性加法
- 局所トレンド
- 加法パッチ

詳しくは、31 ページの『第 7 章 外れ値のタイプ』のトピックを参照してください。

カスタム指数平滑法モデル

「モデル・タイプ」。指数平滑法モデル¹は、季節または非季節のいずれかに分類されます。季節モデルは、周期がアクティブなデータ・セットに定義されている場合にのみ使用可能です (以下の『現在の周期』を参照)。

- *Simple (単純)*. このモデルは、トレンドも季節性もない系列に適しています。唯一の平滑化パラメーターはレベルです。単純指数平滑法は、0 次の自己回帰、1 次の差分、1 次の移動平均、および定数項なしの ARIMA モデルに非常に類似しています。
- *Holt's linear trend (Holt の線型トレンド)*. このモデルは、線型トレンドがあり、季節性がない系列に適しています。その平滑化パラメーターは水準とトレンドであり、いずれも他方の値を制約しません。Holt モデルは Brown のモデルより一般的ですが、大きな系列の計算に時間がかかる場合があります。Holt の指数平滑法は、0 次の自己回帰、2 次の差分、および 2 次の移動平均の ARIMA モデルに非常に類似しています。
- *Brown's linear trend (Brown の線型トレンド)*. このモデルは、線型トレンドがあり、季節性がない系列に適しています。平滑化パラメーターはレベルとトレンドであり、いずれも等しいと仮定します。したがって、Brown モデルは Holt モデルの特別なケースです。Brown の指数平滑法は、次数 0 の自己回帰、2 次の差分、2 次の移動平均を持ち、2 次移動平均の係数が 1 次の係数の半分の 2 乗に等しい ARIMA モデルに非常に類似しています。
- *Damped trend (減衰トレンド)*. このモデルは、減衰する線型トレンドを持ち、季節性のない系列に適しています。この平滑化パラメーターはレベル、トレンド、および減衰トレンドです。減衰指数平滑法は、1 次の自己回帰、1 次の差分、および 2 次の移動平均の ARIMA モデルに最も類似しています。
- *Simple seasonal (単純季節)*. このモデルは、トレンドがなく、時間が経過しても季節効果が一定である系列に適しています。平滑化パラメーターはレベルと季節です。単純季節指数平滑法は、0 次の自己回帰、1 次の差分、1 次の季節差分、および 1 次、 p 次、 $p + 1$ 次の移動平均の ARIMA モデルに非常に類似しています。ここで、 p は季節区間の期間数です (月次データの場合は $p = 12$)。
- *Winters' additive (Winters 加法)*. このモデルは、線型トレンドを持ち、系列のレベルに依存しない季節効果を持つ系列に適しています。この平滑化パラメーターはレベル、トレンド、および季節です。Winters 加法指数平滑法は、0 次の自己回帰、1 次の差分、1 次の季節差分、および $p + 1$ 次の移動平均の ARIMA モデルに非常に類似しています。ここで、 p は季節区間の期間数です (月次データの場合は $p = 12$)。

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

- *Winters' multiplicative* (*Winters* 乗法). このモデルは、線型トレンドを持ち、系列のレベルに依存する季節効果を持つ系列に適しています。この平滑化パラメーターはレベル、トレンド、および季節です。*Winters* 乗法指数平滑法は、いかなる *ARIMA* モデルにも類似しません。

「現在の周期」。アクティブなデータ・セットに現在定義されている周期 (ある場合) を示します。現在の周期は、整数値として指定されます。例えば、各ケースが 1 カ月を表す 1 年周期の場合は 12 となります。周期が設定されていない場合は、値「なし」が表示されます。季節モデルには周期が必要です。周期は、「日付の定義」ダイアログ・ボックスから設定できます。

「従属変数の変換」。モデル作成の前に、各従属変数に対して実行する変換を指定できます。

- 「なし」。変換は実行されません。
- 「平方根」。平方根変換。
- 「自然対数」。自然対数変換。

カスタム *ARIMA* モデル

時系列モデラーでは、カスタム非季節または季節 *ARIMA* (自己回帰和分移動平均) モデル (Box-Jenkins² モデルとも呼ばれる) を、予測変数の一定のセットの有無に関係なく構築できます。予測変数の一部またはすべての転送関数を定義し、外れ値の自動検出を指定したり、外れ値の明示的なセットを指定したりすることができます。

- 「変数」タブで指定されたすべての従属 (予測) 変数が、明示的にモデルに含まれます。これは、エキスパート・モデラーを使用した場合とは対照的です (この場合、独立変数は従属変数との間に統計的に有意な関係がある場合にのみ含まれます)。

カスタム *ARIMA* モデルに対するモデル指定

「モデル」タブでは、カスタム *ARIMA* モデルの構造を指定できます。

「*ARIMA* の順序」。モデルのさまざまな *ARIMA* 成分の値を、「構造」グリッドの対応するセルに入力します。すべての値が、負でない整数である必要があります。自己回帰および移動平均の成分については、その値が最大次数を表します。それより低い正の次数がすべてモデルに含まれます。例えば、2 を指定すると、モデルには次数 2 および 1 が含まれます。「季節」列のセルは、アクティブなデータ・セットに対して周期が定義されている場合にのみ有効になります (以下の『現在の周期』を参照)。

- 「自己回帰 (*p*)」。モデル内の自己回帰の次数の数値です。自己回帰の次数は、系列のどの過去の値を使用して現在の値を予測するかを指定します。例えば、自己回帰の次数 2 は、過去の 2 期間分の系列値を使用して、現在の値を予測することを示します。
- 「差分 (*d*)」。モデルを推定する前に系列に適用される差分の次数を指定します。差分は、トレンドが存在する場合に必要となり (通常、トレンドの存在する系列は非定常性であり、*ARIMA* モデル作成は定常性を前提とするため)、トレンドの影響を除去する目的に使用されます。差分の次数は、系列のトレンドの次数に対応します (1 次差分は線型トレンドを説明し、2 次差分は 2 次トレンドを説明するなど)。
- 「移動平均 (*q*)」。モデル内の移動平均の次数の数値。移動平均の次数は、過去の値の系列平均からの偏差を、現在の値の予測にどのように使用するかを指定します。例えば、移動平均の次数 1 および 2 は、系列の現在の値を予測するときに、最近の 2 期間のそれぞれから取得した系列の平均値の偏差が考慮されることを示します。

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

季節次数。季節型の自己回帰、移動平均、および差分成分は、対応する非季節の成分と同様の役割を果たします。ただし、季節次数の場合は、現在の系列値が 1 つ以上の季節期間で区切られた過去の系列値の影響を受けます。例えば、月単位のデータ (季節期間 12) の場合、季節次数 1 は、現在の期間より 12 期間前の系列値の影響を受けることを意味します。月単位のデータに関しては、季節次数 1 は、非季節次数 12 を指定するのと同様です。

「現在の周期」。アクティブなデータ・セットに現在定義されている周期 (ある場合) を示します。現在の周期は、整数値として指定されます。例えば、各ケースが 1 カ月を表す 1 年周期の場合は 12 となります。周期が設定されていない場合は、値「なし」が表示されます。季節モデルには周期が必要です。周期は、「日付の定義」ダイアログ・ボックスから設定できます。

「従属変数の変換」。モデル作成の前に、各従属変数に対して実行する変換を指定できます。

- 「なし」。変換は実行されません。
- 「平方根」。平方根変換。
- 「自然対数」。自然対数変換。

「モデルに定数を含める」。系列値の全体平均が 0 だという確信がない限り、通常は定数を含めます。差分を適用する場合は、定数を除外することをお勧めします。

カスタム ARIMA モデルにおける転送関数

「転送関数」タブ (独立変数を指定した場合にのみ表示される) では、「変数」タブで指定した独立変数の一部またはすべての転送関数を定義することができます。転送関数を使用すると、従属系列の今後の値を予測するために独立 (予測) 変数の過去の値の使用方法を指定できます。

「転送関数の順序」。転送関数のさまざまな成分の値を、「構造」グリッドの対応するセルに入力します。すべての値が、負でない整数である必要があります。分子および分母の成分については、値が最大次数を表します。それより低い正の次数がすべてモデルに含まれます。また、次数 0 は常に分子成分に含まれます。例えば、分子に 2 を指定した場合、モデルには次数 2、1、および 0 が含まれます。分母に 3 を指定した場合、モデルには次数 3、2、および 1 が含まれます。「季節」列のセルは、アクティブなデータ・セットに周期が定義されている場合にのみ使用可能です (以下の『現在の周期』を参照)。

- 「分子」。転送関数の分子次数です。従属系列の現在の値を予測する場合に、選択された独立 (予測) 系列の過去の値のうち、どれを使用するかを指定します。例えば、分子次数 1 は、過去の 1 期間における独立系列の値 (および独立系列の現在の値) が、各従属系列の現在の値を予測するために使用されることを示します。
- 「分母」。転送関数の分母次数です。従属系列の現在の値を予測する場合に、選択された独立 (予測) 系列の過去の値に対して、系列の平均の偏差をどのように使用するかを指定します。例えば、分母次数 1 は、各従属系列の現在の値を予測する際に、過去の 1 期間における独立系列の平均値の偏差が考慮されることを指定します。
- 「差分」。モデルを推定する前に、選択された独立 (予測) 系列に適用される差分の次数を指定します。トレンドが存在する場合は差分が必要となり、トレンドの影響を除去する目的で使用されます。

季節次数。季節分子、分母、および差分成分は、対応する非季節の成分と同様の役割を果たします。ただし、季節次数の場合は、現在の系列値が 1 つ以上の季節期間で区切られた過去の系列値の影響を受けます。例えば、月単位のデータ (季節期間 12) の場合、季節次数 1 は、現在の期間より 12 期間前の系列値の影響を受けることを意味します。月単位のデータに関しては、季節次数 1 は、非季節次数 12 を指定するのと同様です。

「現在の周期」。アクティブなデータ・セットに現在定義されている周期 (ある場合) を示します。現在の周期は、整数値として指定されます。例えば、各ケースが 1 カ月を表す 1 年周期の場合は 12 となります。周期が設定されていない場合は、値「なし」が表示されます。季節モデルには周期が必要です。周期は、「日付の定義」ダイアログ・ボックスから設定できます。

「遅延」。遅延を設定すると、指定した区間数だけ独立変数の影響が遅延します。例えば、遅延を 5 に設定した場合、 t の時点の独立変数の値は、5 期間が経過するまで $(t + 5)$ 予測には影響を与えません。

「変換」。一連の独立変数に対する転送関数の指定には、それらの変数で実行されるオプションの変換も含まれます。

- 「なし」。変換は実行されません。
- 「平方根」。平方根変換。
- 「自然対数」。自然対数変換。

カスタム ARIMA モデルの外れ値

「外れ値」タブでは、外れ値³の処理として、自動検出、特定のポイントを外れ値として指定、あるいは検出またはモデル化を行わないという選択肢が提供されます。

「外れ値の検出またはモデル化をしない」。デフォルトでは、外れ値の検出とモデル化のどちらも行われません。このオプションを選択すると、外れ値の検出またはモデル化が無効になります。

「自動的に外れ値を検出」。外れ値の自動検出を実行するには、このオプションを選択し、以下の外れ値のタイプを 1 つ以上選択します。

- 加法
- レベル・シフト
- 革新
- 過渡
- 季節性加法
- 局所トレンド
- 加法パッチ

詳しくは、31 ページの『第 7 章 外れ値のタイプ』のトピックを参照してください。

「特定の時点を外れ値としてモデル化」。特定の時点を外れ値として指定するには、このオプションを選択します。外れ値ごとに、「外れ値の定義」グリッドの別の行を使用します。所定の行のすべてのセルに値を入力します。

- 「タイプ」。外れ値のタイプ。サポートされるタイプは、加法 (デフォルト)、レベル・シフト、革新、過渡、季節性加法、および局所トレンドです。

注 1: アクティブなデータ・セットに日付の指定が定義されていない場合、「外れ値の定義」グリッドには「観測」という単一の列が表示されます。外れ値を指定するには、関連するケースの行番号を (データ・エディターの表示どおりに) 入力します。

注 2: 「外れ値の定義」グリッドの「サイクル」列 (ある場合) は、アクティブなデータ・セットの `CYCLE_` 変数の値を参照します。

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

出力

使用可能な出力には、すべてのモデルにわたって計算された結果に加えて、それぞれのモデルの結果が含まれます。個々のモデルの結果は、ユーザー指定の基準に基づいて、最適モデルまたは最不適モデルに制限することができます。

統計テーブルと予測テーブル

「統計」タブでは、モデル作成結果のテーブルを表示するオプションが提供されます。

「モデルごとの適合度、Ljung-Box 統計、および外れ値の数を表示」。選択された適合度、Ljung-Box 値、および外れ値の数を含むテーブルを推定されたモデルごとに表示するには、このオプションを選択します (チェック・マークを付けます)。

「適合度」。推定された各モデルの適合度を示すテーブルに含めるものを、次から 1 つ以上選択できます。

- 定常 R^2 乗
- R^2 乗
- 平方平均誤差平方根
- 平均絶対パーセント誤差
- 平均絶対誤差
- 最大絶対パーセント誤差
- 最大絶対誤差
- 正規化 BIC

詳しくは、29 ページの『第 6 章 適合度測定結果』のトピックを参照してください。

「モデル比較の統計」。この一連のオプションは、推定されたすべてのモデルにわたって計算された統計を含むテーブルの表示を制御します。オプションごとに別のテーブルが生成されます。以下のオプションを 1 つ以上選択できます。

- 「適合度」。定常 R^2 乗、 R^2 乗、2 乗平均平方根誤差、平均絶対パーセント誤差、平均絶対誤差、最大絶対パーセント誤差、最大絶対誤差、および正規化ベイズ情報量基準に関する要約統計量とパーセンタイルのテーブル。
- 「残差自己相関関数 (ACF)」。推定されたすべてのモデルにわたる残差の自己相関に関する要約統計量とパーセンタイルのテーブルです。
- 「残差偏自己相関関数 (PACF)」。推定されたすべてのモデルにわたる残差の偏自己相関に関する要約統計量とパーセンタイルのテーブルです。

「個別モデルの統計」。この一連のオプションは、推定された各モデルの詳細情報を含むテーブルの表示を制御します。オプションごとに別のテーブルが生成されます。以下のオプションを 1 つ以上選択できます。

- 「パラメーター推定値」。推定された各モデルのパラメーター推定値のテーブルを表示します。指数平滑法モデルと ARIMA モデルに対して別個のテーブルが表示されます。外れ値が存在する場合、それに対するパラメーター推定値も別のテーブルに表示されます。
- 「残差自己相関関数 (ACF)」。推定された各モデルについて、ラグごとに残差自己相関のテーブルを表示します。このテーブルには、自己相関に対する信頼区間が含まれます。

- 「残差偏自己相関関数 (PACF)」。推定された各モデルについて、ラグごとの残差偏自己相関のテーブルを表示します。このテーブルには、偏自己相関に対する信頼区間が含まれます。

「予測値の表示」。推定された各モデルのモデル予測値および信頼区間のテーブルを表示します。予測期間は「オプション」タブで設定します。

プロット

「プロット」タブでは、モデル作成結果のプロットを表示するためのオプションが提供されます。

「モデルの比較のプロット」

この一連のオプションは、推定されたすべてのモデルにわたって計算された統計を含むプロットの表示を制御します。オプションごとに別のプロットが生成されます。以下のオプションを 1 つ以上選択できます。

- 定常 R^2 乗
- R^2 乗
- 平方平均誤差平方根
- 平均絶対パーセント誤差
- 平均絶対誤差
- 最大絶対パーセント誤差
- 最大絶対誤差
- 正規化 BIC
- 残差自己相関関数 (ACF)
- 残差偏自己相関関数 (PACF)

詳しくは、29 ページの『第 6 章 適合度測定結果』のトピックを参照してください。

「個別モデルのプロット」

「系列」。推定された各モデルの予測値のプロットを取得するには、このオプションを選択します(チェック・マークを付けます)。プロットに含めるものを、次から 1 つ以上選択できます。

- 「観測値」。従属系列の観測値です。
- 「予測」。予測期間に対するモデルの予測値です。
- 「当てはめ値」。推定期間に対するモデルの予測値です。
- 「予測の信頼区間」。予測期間に対する信頼区間です。
- 「当てはめ値の信頼区間」。推定期間に対する信頼区間です。

「残差自己相関関数 (ACF)」。推定された各モデルの残差自己相関のプロットを表示します。

「残差偏自己相関関数 (PACF)」。推定された各モデルの残差偏自己相関のプロットを表示します。

出力を最適合モデルまたは最不適合モデルに限定

「出力フィルター」タブでは、表形式出力とグラフ出力の両方を、推定されたモデルのサブセットに制限するためのオプションが提供されます。指定した適合基準に従って出力を最適合モデルおよび最不適合モデル、またはそのいずれかに限定することを選択できます。デフォルトでは、推定されたすべてのモデルが出力対象となります。

「**最適モデル**」。出力に最適モデルを含めるには、このオプションを選択します (チェック・マークを付けます)。適合度を選択し、含めるモデルの数を指定します。このオプションを選択しても、最不適合モデルの選択は可能です。その場合、出力には、最不適合モデルと最適モデルが含まれます。

- 「**モデルの固定数**」。 n 個の最適モデルに関する結果が表示されることを指定します。この数が推定されたモデルの数を超える場合、すべてのモデルが表示されます。
- 「**モデルの合計数のパーセント**」。推定されたすべてのモデルの上位 n パーセントの適合度値を持つモデルに関する結果が表示されることを指定します。

「**最不適合モデル**」。出力に最不適合モデルを含めるには、このオプションを選択します (チェック・マークを付けます)。適合度を選択し、含めるモデルの数を指定します。このオプションを選択しても、最適モデルの選択は可能です。その場合、出力には、最適モデルと最不適合モデルが含まれます。

- 「**モデルの固定数**」。 n 個の最不適合モデルに関する結果が表示されることを指定します。この数が推定されたモデルの数を超える場合、すべてのモデルが表示されます。
- 「**モデルの合計数のパーセント**」。推定されたすべてのモデルの下位 n パーセントの適合度値を持つモデルに関する結果が表示されることを指定します。

「**適合度**」。モデルをフィルタリングするために使用する適合度を選択します。デフォルトは定常 R^2 乗です。

モデル予測およびモデル指定の保存

「保存」タブでは、モデル予測を新しい変数としてアクティブなデータ・セット内に保存することと、モデル指定を外部ファイルに XML 形式で保存することができます。

「**変数の保存**」。モデル予測、信頼区間、および残差を新しい変数としてアクティブなデータ・セットに保存できます。従属系列ごとに独自の新しい変数のセットを生成し、新しいそれぞれの変数に推定期間と予測期間の両方の値が含まれます。予測期間が従属変数系列の長さを超える場合、新しいケースが追加されます。それぞれに関連付けられた「保存」チェック・ボックスを選択することにより、新しい変数の保存を選択します。デフォルトでは、新しい変数は保存されません。

- 「**予測値**」。モデルの予測値です。
- 「**信頼限界の下限**」。予測値に対する信頼限界の下限です。
- 「**信頼限界の上限**」。予測値に対する信頼限界の上限です。
- 「**ノイズ残差**」。モデルの残差です。従属変数の変換が実行されたとき (自然対数など)、これらは変換された系列の残差です。
- 「**変数名の接頭辞**」。新しい変数名に使用される接頭辞を指定するか、デフォルトの接頭辞のままにします。変数名は、接頭辞、関連付けられた従属変数の名前、およびモデル識別子で構成されます。この変数名は、変数名の競合を避けるために、必要に応じて拡張されます。接頭辞は、有効な変数名の規則に従っている必要があります。

「**モデル・ファイルのエクスポート**」。推定されたすべてのモデルのモデル指定が、指定したファイルに XML 形式でエクスポートされます。保存されたモデルは、更新された予測を取得するために使用できません。

- 「**XML ファイル**」。モデル指定は、IBM SPSS アプリケーションで使用できる XML ファイルに保存されます。
- 「**PMML ファイル**」。モデル指定は、IBM SPSS アプリケーションなどの PMML 互換アプリケーションで使用できる PMML 互換 XML ファイルに保存されます。

オプション

「オプション」タブでは、予測期間の設定、欠損値の処理の指定、信頼区間の幅の設定、モデル識別子のカスタム接頭辞の指定、自己相関に対して表示されるラグ数の設定を実行できます。

「**予測期間**」。予測期間は必ず、推定期間 (モデルの特定に使用されるケースのセット) の終了後の最初のケースから始まり、アクティブなデータ・セットまたはユーザー指定の日付の最後のケースまで続きます。デフォルトでは、推定期間の終了はアクティブなデータ・セットの最後のケースですが、「**ケースの選択**」ダイアログ・ボックスで「**時間またはケースの範囲に基づく**」を選択することにより変更できます。

- 「**推定期間の後の最初のケースからアクティブ・データ・セット内の最後のケースまで**」。推定期間の終了がアクティブなデータ・セットの最後のケースより前の場合に、最後のケースまでの予測が必要なときは、このオプションを選択します。このオプションは、通常、ホールドアウト期間の予測を生成して、モデルの予測値と実際の値のサブセットを比較可能にするために使用されます。
- 「**推定期間の後の最初のケースから指定の日付まで**」。予測期間の終了を明示的に指定するには、このオプションを選択します。このオプションは、通常、実際の系列の最後より先の予測を生成する場合に使用されます。「日付」グリッドのすべてのセルに値を入力します。

アクティブなデータ・セットに日付の指定が定義されていない場合、「日付」グリッドには「**観測**」という単一の列が表示されます。予測期間の終了を指定するには、関連するケースの行番号を (データ・エディターの表示どおりに) 入力します。

「日付」グリッドの「**サイクル**」列 (ある場合) は、アクティブなデータ・セットの `CYCLE_` 変数の値を参照します。

「**ユーザー欠損値**」。これらのオプションは、ユーザー欠損値の処理を制御します。

- 「**無効として扱う**」。ユーザー欠損値はシステム欠損値と同様に処理されます。
- 「**有効として扱う**」。ユーザー欠損値は有効なデータとして処理されます。

欠損値のポリシー。モデル作成プロシージャの間、欠損値 (無効として処理されるシステム欠損値およびユーザー欠損値を含む) の処理には次の規則が適用されます。

- 推定期間内に発生する従属変数の欠損値を持つケースは、モデルに含まれます。欠損値の具体的な処理は、推定方法によって異なります。
- 推定期間で独立変数に欠損値が含まれる場合には、警告が発行されます。エキスパート・モデラーでは、独立変数を含むモデルの推定は変数なしで行われます。カスタム ARIMA では、独立変数を含むモデルの推定は行われません。
- 予測期間内で独立変数に欠損値が含まれる場合、このプロシージャは警告を発行し、可能な限り先まで予測します。

「**信頼区間の幅 (%)**」。信頼区間は、モデルの予測と残差自己相関に対して計算されます。100 未満の正の値を指定できます。デフォルトでは、95% の信頼区間が使用されます。

「**出力内のモデル識別子の接頭辞**」。 「**変数**」タブで指定した従属変数によって、それぞれ別個の推定モデルが生成されます。モデルは、カスタマイズ可能な接頭辞と整数の接尾辞で構成される固有の名前で区別されます。接頭辞は、入力するか、デフォルトの `Model` をそのまま使用することができます。

「**ACF および PACF の出力に表示されるラグの最大数**」。自己相関および偏自己相関のテーブルおよびプロットに表示されるラグの最大数を設定できます。

TSMODEL コマンドの追加機能

選択内容をシンタックス・ウィンドウに貼り付け、結果の TSMODEL コマンド・シンタックスを編集すると、時系列のモデル作成をカスタマイズできます。このコマンド・シンタックス言語を使用すると、次のことを実行できます。

- データの季節周期を指定する (AUXILIARY サブコマンドで SEASONLENGTH キーワードを使用)。これによって、アクティブなデータ・セットの現在の周期 (ある場合) が無効になります。
- カスタム ARIMA および転送関数の成分に対して不連続ラグを指定する (ARIMA および TRANSFERFUNCTION サブコマンドを使用)。例えば、次数 1、3、および 6 の自己回帰ラグを使用してカスタム ARIMA モデルを指定したり、次数 2、5、および 8 の分子ラグを使用して転送関数を指定したりすることができます。
- 時系列モデラーのプロシージャラーの 1 回の実行に対して、モデル作成に関する複数セットの指定 (例えば、モデル作成方法、ARIMA の次数、独立変数など) を提供する (MODEL サブコマンドを使用)。

シンタックスの詳細については、「コマンド・シンタックス・リファレンス」を参照してください。

第 3 章 時系列モデルの適用

「時系列モデルの適用」プロシージャーでは、外部ファイルから既存の時系列モデルを読み込み、アクティブなデータ・セットに適用します。このプロシージャーは、新しいデータまたは変更されたデータがある系列に対して、モデルの再構築を行わずに予測を取得する場合に使用できます。モデルは、時系列モデラーのプロシージャーを使用して生成されます。

例: 大手小売店の在庫管理者として、5,000 品目の各製品を担当するとします。各製品の今後 3 カ月の売上を予測するモデルを作成するために、エキスパート・モデラーを使用しました。データウェアハウスは、月単位の最新予測の作成に使用する実際の売上データで毎月更新されます。「時系列モデルの適用」プロシージャーを使用すると、元のモデルを使用し、新しいデータを説明するモデル・パラメーターを単純に再度推定することによって、これを実現することができます。

統計: 適合度: 定常 R^2 乗、 R^2 乗 (R^2)、2 乗平均平方根誤差 (RMSE)、平均絶対誤差 (MAE)、平均絶対パーセント誤差 (MAPE)、最大絶対誤差 (MaxAE)、最大絶対パーセント誤差 (MaxAPE)、正規化ベイズ情報量基準 (BIC)。残差: 自己相関関数、偏自己相関関数、Ljung-Box Q 。

プロット。 すべてのモデルの要約プロット: 定常 R^2 乗、 R^2 乗 (R^2)、2 乗平均平方根誤差 (RMSE)、平均絶対誤差 (MAE)、平均絶対パーセント誤差 (MAPE)、最大絶対誤差 (MaxAE)、最大絶対パーセント誤差 (MaxAPE)、正規化ベイズ情報量基準 (BIC) のヒストグラム。残差自己相関および偏自己相関の箱ひげ図。個々のモデルの結果: 予測値、当てはめ値、観測値、信頼限界の上限と下限、残差自己相関および偏自己相関。

時系列モデルの適用のデータの考慮事項

データ: モデルが適用される変数 (従属変数および独立変数) は数値である必要があります。

仮定: モデルは、モデル内で指定された変数と同じ名前のアクティブなデータ・セット内の変数に適用されます。このような変数はすべて時系列として扱われます。つまり、各ケースは時点を表し、連続するケースは一定の時間間隔で区切られています。

- **予測。** 独立 (予測) 変数を持つモデルを使用して予測を生成するには、予測期間内のすべてのケースについて、これらの変数の値がアクティブなデータ・セットに含まれている必要があります。モデルのパラメーターが再度推定される場合、推定期間内の独立変数に欠損値が含まれていないことが必要です。

日付の定義

「時系列モデルの適用」プロシージャーでは、アクティブなデータ・セットの周期 (ある場合) が、適用されるモデルの周期に一致している必要があります。モデルの構築に使用したデータ・セットと同じもの (新規または変更されたデータを含む場合もある) を使用して単純に予測する場合は、この条件が満たされます。アクティブなデータ・セットに周期が存在しない場合は、「日付の定義」ダイアログ・ボックスに移動して、周期を作成することができます。ただし、周期を指定せずにモデルを作成した場合は、アクティブなデータ・セットも周期なしで作成する必要があります。

モデルを適用するには

1. メニューから次の項目を選択します。

「分析」 > 「予測」 > 「モデルの適用...」

2. モデル・ファイルのファイル指定を入力するか、「参照」をクリックし、モデル・ファイルを選択します (モデル・ファイルは時系列モデラーのプロシージャで作成されます)。

必要に応じて、次のことが可能です。

- アクティブなデータ・セットのデータを使用してモデル・パラメーターを再度推定する。予測は、再度推定されたパラメーターを使用して作成されます。
- 予測、信頼区間、およびノイズ残差を保存する。
- 再度推定されたモデルを XML 形式で保存する。

「モデル・パラメーターと適合度」

「**モデル・ファイルから読み込み**」。予測は、モデル・ファイルから取得したモデル・パラメーターを使用して生成されます (このモデル・パラメーターは再度推定されません)。出力に表示され、モデル (最適または最不適) のフィルタリングに使用される適合度測定結果は、モデル・ファイルから取得され、各モデルの作成時 (または前回更新時) に使用されたデータを反映します。このオプションを使用すると、予測ではアクティブなデータ・セット内の従属変数と独立変数のいずれに対しても、履歴データは考慮されません。履歴データが予測で考慮されるようにするには、「**データから再度推定**」を選択する必要があります。また、予測では、予測期間内の従属系列の値は考慮されませんが、予測期間内の独立変数の値は考慮されます。より最新の従属系列の値があり、それを予測に取り込む場合は、それらの値が含まれるように推定期間を調節して再度推定する必要があります。

「**データから再度推定**」。モデル・パラメーターは、アクティブなデータ・セットのデータを使用して再度推定されます。モデル・パラメーターを再度推定しても、モデルの構造には影響しません。例えば、ARIMA(1,0,1) モデルはそのままですが、自己回帰パラメーターおよび移動平均パラメーターは再度推定されます。再度推定によって新しい外れ値が検出されることはありません。外れ値がある場合は、常にモデル・ファイルから取得されます。

- 「**推定期間**」。推定期間により、モデル・パラメーターを再度推定するのに使用されるケースのセットが定義されます。デフォルトでは、推定期間にはアクティブなデータ・セットのすべてのケースが含まれます。推定期間を設定するには、「**ケースの選択**」ダイアログ・ボックスで「**時間またはケースの範囲に基づく**」を選択します。使用可能なデータによっては、このプロシージャによって使用される推定期間がモデルごとに異なるため、表示される値と異なる場合があります。任意のモデルに対する本当の推定期間は、指定された推定期間の最初または最後で発生する、モデルの従属変数の連続する欠損値を除去した後に残った期間です。

「予測期間」

各モデルの予測期間は必ず、推定期間の終了後の最初のケースから始まり、アクティブなデータ・セットの最後のケースまたはユーザー指定の日付まで続きます。パラメーターが再度推定されない場合 (デフォルト)、各モデルの推定期間は、モデルの作成時 (または前回更新時) に使用されたケースのセットです。

- 「**推定期間の後の最初のケースからアクティブ・データ・セット内の最後のケースまで**」。推定期間の終了がアクティブなデータ・セットの最後のケースより前の場合に、最後のケースまでの予測が必要なときは、このオプションを選択します。
- 「**推定期間の後の最初のケースから指定の日付まで**」。予測期間の終了を明示的に指定するには、このオプションを選択します。「日付」グリッドのすべてのセルに値を入力します。

アクティブなデータ・セットに日付の指定が定義されていない場合、「日付」グリッドには「観測」という単一の列が表示されます。予測期間の終了を指定するには、関連するケースの行番号を (データ・エディターの表示どおりに) 入力します。

「日付」グリッドの「サイクル」列 (ある場合) は、アクティブなデータ・セットの *CYCLE_* 変数の値を参照します。

出力

使用可能な出力には、すべてのモデルにわたる結果に加えてそれぞれのモデルの結果が含まれます。個々のモデルの結果は、ユーザー指定の基準に基づいて、最適モデルまたは最不適合モデルに制限することができます。

統計テーブルと予測テーブル

「統計」タブでは、モデル適合度統計量、モデル・パラメーター、自己相関関数、および予測のテーブルを表示するオプションが提供されます。モデル・パラメーターが再度推定（「モデル」タブの「データから再度推定」）されない限り、表示される適合度の値、Ljung-Box値、およびモデル・パラメーターはモデル・ファイルから取得され、各モデルの作成時（または前回更新時）に使用されたデータを反映します。外れ値の情報は、常にモデル・ファイルから取得されます。

「モデルごとの適合度、Ljung-Box 統計、および外れ値の数を表示」。選択された適合度、Ljung-Box 値、および外れ値の数を含むテーブルをモデルごとに表示するには、このオプションを選択します（チェック・マークを付けます）。

「適合度」。各モデルの適合度を示すテーブルに含めるものを、次から 1 つ以上選択できます。

- 定常 R^2 乗
- R^2 乗
- 平方平均誤差平方根
- 平均絶対パーセント誤差
- 平均絶対誤差
- 最大絶対パーセント誤差
- 最大絶対誤差
- 正規化 BIC

詳しくは、29 ページの『第 6 章 適合度測定結果』のトピックを参照してください。

「モデル比較の統計」。この一連のオプションは、すべてのモデルにわたる統計量を含むテーブルの表示を制御します。オプションごとに別のテーブルが生成されます。以下のオプションを 1 つ以上選択できます。

- 「適合度」。定常 R^2 乗、 R^2 乗、2 乗平均平方根誤差、平均絶対パーセント誤差、平均絶対誤差、最大絶対パーセント誤差、最大絶対誤差、および正規化ベイズ情報量基準に関する要約統計量とパーセンタイルのテーブル。
- 「残差自己相関関数 (ACF)」。推定されたすべてのモデルにわたる残差の自己相関に関する要約統計量とパーセンタイルのテーブルです。このテーブルは、モデル・パラメーターが再度推定された場合にのみ使用可能です（「モデル」タブの「データから再度推定」）。
- 「残差偏自己相関関数 (PACF)」。推定されたすべてのモデルにわたる残差の偏自己相関に関する要約統計量とパーセンタイルのテーブルです。このテーブルは、モデル・パラメーターが再度推定された場合にのみ使用可能です（「モデル」タブの「データから再度推定」）。

「個別モデルの統計」。この一連のオプションは、各モデルの詳細情報を含むテーブルの表示を制御します。オプションごとに別のテーブルが生成されます。以下のオプションを 1 つ以上選択できます。

- 「**パラメーター推定値**」。各モデルのパラメーター推定値のテーブルを表示します。指数平滑法モデルと ARIMA モデルに対して別個のテーブルが表示されます。外れ値が存在する場合、それに対するパラメーター推定値も別のテーブルに表示されます。
- 「**残差自己相関関数 (ACF)**」。推定された各モデルについて、ラグごとに残差自己相関のテーブルを表示します。このテーブルには、自己相関に対する信頼区間が含まれます。このテーブルは、モデル・パラメーターが再度推定された場合にのみ使用可能です (「モデル」タブの「**データから再度推定**」)。
- 「**残差偏自己相関関数 (PACF)**」。推定された各モデルについて、ラグごとの残差偏自己相関のテーブルを表示します。このテーブルには、偏自己相関に対する信頼区間が含まれます。このテーブルは、モデル・パラメーターが再度推定された場合にのみ使用可能です (「モデル」タブの「**データから再度推定**」)。

「**予測値の表示**」。各モデルのモデル予測値および信頼区間のテーブルを表示します。

プロット

「プロット」タブでは、モデルの適合度統計量、自己相関関数、および系列値 (予測を含む) のプロットを表示するオプションが提供されます。

「モデルの比較のプロット」

この一連のオプションは、すべてのモデルにわたる統計量を含むプロットの表示を制御します。モデル・パラメーターが再度推定 (「モデル」タブの「**データから再度推定**」) されない限り、表示される値はモデル・ファイルから取得され、各モデルの作成時 (または前回更新時) に使用されたデータを反映します。さらに、自己相関プロットは、モデル・パラメーターが再度推定された場合のみ使用することができます。オプションごとに別のプロットが生成されます。以下のオプションを 1 つ以上選択できます。

- 定常 R^2 乗
- R^2 乗
- 平方平均誤差平方根
- 平均絶対パーセント誤差
- 平均絶対誤差
- 最大絶対パーセント誤差
- 最大絶対誤差
- 正規化 BIC
- 残差自己相関関数 (ACF)
- 残差偏自己相関関数 (PACF)

詳しくは、29 ページの『第 6 章 適合度測定結果』のトピックを参照してください。

「個別モデルのプロット」

「**系列**」。各モデルの予測値のプロットを取得するには、このオプションを選択します (チェック・マークを付けます)。観測値、当てはめ値、当てはめ値の信頼区間、および自己相関は、モデル・パラメーターが再度推定 (「モデル」タブの「**データから再度推定**」) された場合にのみ使用可能です。プロットに含めるものを、次から 1 つ以上選択できます。

- 「**観測値**」。従属系列の観測値です。
- 「**予測**」。予測期間に対するモデルの予測値です。
- 「**当てはめ値**」。推定期間に対するモデルの予測値です。

- 「予測の信頼区間」。予測期間に対する信頼区間です。
- 「当てはめ値の信頼区間」。推定期間に対する信頼区間です。

「残差自己相関関数 (ACF)」。推定された各モデルの残差自己相関のプロットを表示します。

「残差偏自己相関関数 (PACF)」。推定された各モデルの残差偏自己相関のプロットを表示します。

出力を最適モデルまたは最不適合モデルに限定

「出力フィルター」タブでは、表形式出力とグラフ出力の両方を、モデルのサブセットに制限するためのオプションが提供されます。指定した適合基準に従って出力を最適モデルおよび最不適合モデル、またはそのいずれかに限定することを選択できます。デフォルトでは、すべてのモデルが出力対象となります。モデル・パラメーターが再度推定（「モデル」タブの「データから再度推定」）されない限り、モデルのフィルタリングに使用される適合度の値はモデル・ファイルから取得され、各モデルの作成時（または前回更新時）に使用されたデータを反映します。

「最適モデル」。出力に最適モデルを含めるには、このオプションを選択します（チェック・マークを付けます）。適合度を選択し、含めるモデルの数を指定します。このオプションを選択しても、最不適合モデルの選択は可能です。その場合、出力には、最不適合モデルと最適モデルが含まれます。

- 「モデルの固定数」。 n 個の最適モデルに関する結果が表示されることを指定します。この数がモデルの合計数を超える場合、すべてのモデルが表示されます。
- 「モデルの合計数のパーセント」。すべてのモデルの上位 n パーセントの適合度値を持つモデルに関する結果が表示されることを指定します。

「最不適合モデル」。出力に最不適合モデルを含めるには、このオプションを選択します（チェック・マークを付けます）。適合度を選択し、含めるモデルの数を指定します。このオプションを選択しても、最適モデルの選択は可能です。その場合、出力には、最適モデルと最不適合モデルが含まれます。

- 「モデルの固定数」。 n 個の最不適合モデルに関する結果が表示されることを指定します。この数がモデルの合計数を超える場合、すべてのモデルが表示されます。
- 「モデルの合計数のパーセント」。すべてのモデルの下位 n パーセントの適合度値を持つモデルに関する結果が表示されることを指定します。

「適合度」。モデルをフィルタリングするために使用する適合度を選択します。デフォルトは定常 R^2 乗です。

モデル予測およびモデル指定の保存

「保存」タブでは、モデル予測を新しい変数としてアクティブなデータ・セット内に保存することと、モデル指定を外部ファイルに XML 形式で保存することができます。

「変数の保存」。モデル予測、信頼区間、および残差を新しい変数としてアクティブなデータ・セットに保存できます。モデルごとに、独自の新しい変数のセットが生成されます。予測期間がモデルに関連付けられている従属変数系列の長さを超える場合、新しいケースが追加されます。モデル・パラメーターが再度推定（「モデル」タブの「データから再度推定」）されない限り、予測値および信頼限界は予測期間に対してのみ作成されます。それぞれに関連付けられた「保存」チェック・ボックスを選択することにより、新しい変数の保存を選択します。デフォルトでは、新しい変数は保存されません。

- 「予測値」。モデルの予測値です。
- 「信頼限界の下限」。予測値に対する信頼限界の下限です。
- 「信頼限界の上限」。予測値に対する信頼限界の上限です。

- 「ノイズ残差」。モデルの残差です。従属変数の変換が実行されたとき (自然対数など)、これらは変換された系列の残差です。この選択項目は、モデル・パラメーターが再度推定された場合にのみ使用可能です (「モデル」タブの「データから再度推定」)。
- 「変数名の接頭辞」。新しい変数名に使用される接頭辞を指定するか、デフォルトの接頭辞のままにします。変数名は、接頭辞、関連付けられた従属変数の名前、およびモデル識別子で構成されます。この変数名は、変数名の競合を避けるために、必要に応じて拡張されます。接頭辞は、有効な変数名の規則に従っている必要があります。

「モデル・ファイルのエクスポート」。モデル指定 (再度推定されたパラメーターおよび適合度統計量を含む) が、指定したファイルに XML 形式でエクスポートされます。このオプションは、モデル・パラメーターが再度推定された場合にのみ使用可能です (「モデル」タブの「データから再度推定」)。

- 「XML ファイル」。モデル指定は、IBM SPSS アプリケーションで使用できる XML ファイルに保存されます。
- 「PMML ファイル」。モデル指定は、IBM SPSS アプリケーションなどの PMML 互換アプリケーションで使用できる PMML 互換 XML ファイルに保存されます。

オプション

「オプション」タブでは、欠損値の処理を指定したり、信頼区間の幅を設定したり、自己相関に対して表示されるラグの数を設定したりできます。

「ユーザー欠損値」。これらのオプションは、ユーザー欠損値の処理を制御します。

- 「無効として扱う」。ユーザー欠損値はシステム欠損値と同様に処理されます。
- 「有効として扱う」。ユーザー欠損値は有効なデータとして処理されます。

欠損値のポリシー。欠損値 (無効として処理されるシステム欠損値およびユーザー欠損値を含む) の処理に次の規則が適用されます。

- 推定期間内に発生する従属変数の欠損値を持つケースは、モデルに含まれます。欠損値の具体的な処理は、推定方法によって異なります。
- ARIMA モデルでは、推定期間内で予測値に欠損値が含まれる場合、警告が発行されます。予測値を伴うモデルは、再度推定されません。
- 予測期間内で独立変数に欠損値が含まれる場合、このプロシージャは警告を発行し、可能な限り先まで予測します。

「信頼区間の幅 (%)」。信頼区間は、モデルの予測と残差自己相関に対して計算されます。100 未満の正の値を指定できます。デフォルトでは、95% の信頼区間が使用されます。

「ACF および PACF の出力に表示されるラグの最大数」。自己相関および偏自己相関のテーブルおよびプロットに表示されるラグの最大数を設定できます。このオプションは、モデル・パラメーターが再度推定された場合にのみ使用可能です (「モデル」タブの「データから再度推定」)。

TSAPPLY コマンドの追加機能

選択内容をシンタックス・ウィンドウに貼り付け、結果の TSAPPLY コマンド・シンタックスを編集すると、追加の機能が使用可能になります。このコマンド・シンタックス言語を使用すると、次のことを実行できます。

- 1 つのモデル・ファイル内のモデルのサブセットのみをアクティブなデータ・セットに適用することを指定する (MODEL サブコマンドで DROP および KEEP キーワードを使用)。

- 2 つ以上のモデル・ファイルからのモデルをデータに適用する (MODEL サブコマンドを使用)。例えば、あるモデル・ファイルに販売個数を表す系列のモデルが含まれていて、別のモデル・ファイルには収益を表す系列のモデルが含まれる場合があります。

シンタックスの詳細については、「コマンド・シンタックス・リファレンス」を参照してください。

第 4 章 季節性の分解

「季節性の分解」プロシージャーでは、系列が季節成分、結合されたトレンドとサイクル成分、および「誤差」成分に分解されます。このプロシージャーは、国勢調査局が開発した Method I の実装で、移動平均比率法としても知られます。

例: ある科学者は、特定の気象観測所で月に 1 回測定されるオゾン・レベルの分析に関心があります。目標は、データになんらかのトレンドが存在するかどうかを判断することです。実際になんらかのトレンドを見つけるには、まず科学者は季節的影響による測定値の変動を説明する必要があります。「季節性の分解」プロシージャーを使用すると、系統的な季節変動を除去できます。その後、季節調整系列でトレンド分析を実行します。

統計: 一連の季節因子です。

季節性の分解のデータの考慮事項

データ: 変数は数値である必要があります。

仮定: 変数は埋め込まれた欠損データを含めないようにする必要があります。少なくとも 1 つの周期的な日付成分を定義する必要があります。

季節因子の推定

1. メニューから次の項目を選択します。

「分析」 > 「予測」 > 「季節性の分解...」

2. 選択可能なリストから 1 つ以上の変数を選択し、「変数」リストに移動します。このリストに含まれるのは数値変数のみであることに注意してください。

「モデル・タイプ」。「季節性の分解」プロシージャーでは、季節因子のモデル作成用に、乗法と加法という 2 つの異なる方法が用意されています。

- 「乗法」。季節成分は、元の系列を得るために季節調整系列に乗じる係数です。実質的に、系列の全体的レベルに比例する季節成分です。季節変動のない観測では季節成分が 1 になります。
- 「加法」。この季節調整は、観測値を得るために季節調整系列に加えられます。この調整は、季節成分に覆い隠されている可能性のある重要な他の特性を調べるために、系列から季節効果を取り除きます。実質的に、系列の全体的レベルに依存しない季節成分です。季節変動のない観測では季節成分が 0 になります。

「移動平均の重み付け」。「移動平均の重み付け」オプションを使用すると、移動平均を計算する際の系列の処理方法を指定できます。これらのオプションは、系列の周期が偶数の場合にだけ使用できます。周期が奇数の場合は、すべての点が同じ重みになります。

- *All points equal* (全ての点が等しい (季節性の分解)). 移動平均は、周期に等しいスパンを使用し、均等に重み付けしたすべての点を使用して計算されます。周期が奇数の場合は、常にこの方法を使用できません。
- *Endpoints weighted by .5* (終点の重み付けが 0.5). 周期が偶数の系列の移動平均を、周期に 1 を加えた値に等しいスパンと 0.5 で重み付けしたスパンの終点で計算します。

必要に応じて、次のことが可能です。

- 「保存」をクリックして、新しい変数の保存方法を指定する。

季節性の分解の保存

「変数の作成」。新しい変数を処理する方法を選択できます。

- *Add to file* (ファイルに追加). 季節性の分解によって作成された新しい系列を、アクティブ・データ・セットに通常の変数として保存します。変数名は、3 文字の接頭辞、下線、および数字で構成されます。
- *Replace existing* (変数を置換 (季節性の分解)). 季節性の分解によって作成された新しい系列を、アクティブ・データ・セットに一時変数として保存します。同時に、予測プロシージャによって作成された一時変数をすべて破棄します。変数名は、3 文字の接頭辞、ポンド記号 (#)、および数字で構成されます。
- *Do not create* (作成しない). 新しい系列をアクティブ・データ・セットに追加しません。

新しい変数名

「季節性の分解」プロシージャにより、4 つの新しい変数 (系列) が作成され、指定した系列ごとに次の 3 文字の接頭辞が付けられます。

SAF. 季節調整因子 です。この値は、各期間が系列のレベルに与える影響を示します。

SAS. 季節調整系列 です。これは、系列の季節変動を除去した後に取得された値です。

STC. 平滑化されたトレンド・サイクル成分 です。この値は、系列に存在するトレンドおよび周期的性質を示します。

ERR. 残差または「誤差」値 です。季節成分、トレンド成分、およびサイクル成分が系列から除去された後の残った値です。

SEASON コマンドの追加機能

このコマンド・シンタックス言語では、次の作業を行うこともできます。

- 「日付の定義」プロシージャで提供される選択肢のいずれかを選択するのではなく、SEASON コマンドに周期を指定する。

シンタックスの詳細については、「コマンド・シンタックス・リファレンス」を参照してください。

第 5 章 スペクトル・プロット

「スペクトル・プロット」プロシージャは、時系列での周期的な性質を判別するために使用します。ある時点から次の時点までの変動を分析するのではなく、系列の変動を全体として分析して、さまざまな周波数の周期成分を調べます。平滑化系列では低い周波数に強い周期成分が存在し、ランダムな変動 (ホワイト・ノイズ) では成分の強度がすべての周波数に拡散されます。

欠損データが含まれる系列は、このプロシージャでは分析できません。

例: 住宅の新築率は、経済状態の重要な指標です。一般に、住宅着工件数のデータは強い季節成分を提示します。ただし、現在の数値を評価した場合、アナリストが気付く必要のあるような長いサイクルがそのデータに見られるでしょうか。

統計: 周波数または周期成分ごとのサイン変換およびコサイン変換、ピリオドグラム値、およびスペクトル密度の推定値。2 変量分析を選択した場合: 周波数または周期成分ごとのクロスピリオドグラムの実数部と虚数部、共スペクトル密度、直角位相スペクトル、ゲイン、平方コヒーレンス、および位相スペクトル。

プロット。 1 変量および 2 変量分析の場合: ピリオドグラムおよびスペクトル密度。2 変量分析の場合: 平方コヒーレンス、直角位相スペクトル、クロス振幅、共スペクトル密度、位相スペクトル、およびゲイン。

スペクトル・プロットのデータの考慮事項

データ: 変数は数値である必要があります。

仮定: 変数は埋め込まれた欠損データを含めないようにする必要があります。分析する時系列は定常的である必要があります、ゼロ以外の平均値を系列から差し引く必要があります。

- 定常。ARIMA モデルを適合させる時系列が満たす必要がある条件。純粹 MA 系列は定常的ですが、AR 系列と ARMA 系列は定常的でない場合があります。定常系列は、時間が経過しても平均値と分散が一定です。

スペクトル分析の取得

1. メニューから次の項目を選択します。

「分析」 > 「時系列」 > 「スペクトル分析...」

2. 選択可能なリストから 1 つ以上の変数を選択し、「変数」リストに移動します。このリストに含まれるのは数値変数のみであることに注意してください。
3. 「スペクトル・ウィンドウ」オプションのいずれかを選択して、スペクトル密度の推定値を取得するためのピリオドグラムの平滑化方法を選択します。使用可能な平滑化オプションは、「Tukey-Hamming」、「Tukey」、「Parzen」、「Bartlett」、「Daniell (単位)」、および「なし」です。
 - *Tukey-Hamming*。 $k = 0, \dots, p$ のとき、重みを $W_k = .54D_p(2 \pi f_k) + .23D_p(2 \pi f_k + \pi/p) + .23D_p(2 \pi f_k - \pi/p)$ とします。ここで、 p はスパンを 2 で割った値の整数部、 D_p は次数 p の Dirichlet 核です。
 - *Tukey*。 $k = 0, \dots, p$ のとき、重みを $W_k = 0.5D_p(2 \pi f_k) + 0.25D_p(2 \pi f_k + \pi/p) + 0.25D_p(2 \pi f_k - \pi/p)$ とします。ここで、 p はスパンを 2 で割った値の整数部、 D_p は次数 p の Dirichlet 核です。

- *Parzen*. $k = 0, \dots, p$ のとき、重みは $W_k = 1/p(2 + \cos(2\pi fk)) (F[p/2] (2\pi fk))^{**2}$ です。ここで、 p はスパンの $1/2$ の整数部分、 $F[p/2]$ は次数 $p/2$ のフェイェール核です。
- *Bartlett*. $k = 0, \dots, p$ のとき、ウィンドウの上半分の重みが $W_k = F_p (2\pi fk)$ と計算されるスペクトル・ウィンドウの形状。ここで、 p はスパンの $1/2$ の整数部分、 F_p は次数 p のフェイェール核です。下半分は上半分と対称です。
- *Daniell* (単位)。重みがすべて 1 に等しいスペクトル・ウィンドウの形状。
- なし。平滑化を行いません。このオプションを選択した場合、スペクトル密度の推定値はピリオドグラムと同じになります。

「スパン」。平滑化を実行する連続した値の範囲。一般に奇数の整数を使用します。スパンが大きいほどスペクトル密度プロットが平滑になります。

Center variables (センター変数)。スペクトルを計算する前に平均値をゼロにするため、および系列の平均値に関連する可能性がある大きな項を除去するために、系列を調整します。

「2 変量分析 (最初の変数と残りの変数)」。2 つ以上の変数を選択した場合、このオプションを選択して、2 変量スペクトル分析を要求できます。

- 「変数」リストの最初の変数は、独立変数として処理され、その他のすべての変数は従属変数として処理されます。
- 最初の系列より後の各系列は、指定された他の系列から独立して、最初の系列を使用して分析されます。各系列の 1 変量分析も実行されます。

「プロット」。1 変量および 2 変量分析の両方に対して、ピリオドグラムおよびスペクトル密度を使用できます。他のすべての選択肢は、2 変量分析でのみ使用できます。

- *Periodogram* (ペリオドグラム)。周波数または周期を基準にして作成した、スペクトルの大きさのプロット。対数スケールを使用して平滑化せずにプロットします。変動の周波数が低い場合は、系列が平滑であるという特徴があります。すべての周波数にわたって変動が一様に分布している場合は「ホワイト・ノイズ」を示します。
- *Squared coherency* (平方コヒーレンス)。2 つの系列のゲインの積。
- *Quadrature spectrum* (直角位相)。クロスピリオドグラムの虚数部。2 つの時系列の異相周波数成分の相関を測定します。この成分は、 $\pi/2$ ラジアンだけ位相がずれています。
- *Cross amplitude* (クロス振幅)。平方共スペクトル密度と平方直角位相スペクトルの和の平方根。
- 「スペクトル密度」。平滑化して不規則な変動を除去したピリオドグラム。
- *Cospectral density* (共スペクトル密度)。クロスピリオドグラムの実部。2 つの時系列の内部位相周波数成分の相関を測定します。
- *Phase spectrum* (位相スペクトル)。ある系列の各周波数成分が他の成分よりどの程度進んでいるか遅れているかの指標。
- *Gain* (ゲイン)。いずれかの系列のスペクトル密度でクロス振幅を割った商。2 つの系列の各々がそれぞれのゲイン値を持ちます。

By frequency (周波数別 (スペクトル))。すべてのプロットは、周波数 0 (定数項または平均値項) から周波数 0.5 (2 つの観測のサイクルに対応する項) までの範囲で周波数別に作成されます。

By period (周期別 (スペクトル))。すべてのプロットは、 2 (2 つの観測のサイクルに対応する項) から観測数に等しい期間 (定数項または平均値項) までの範囲で周期別に作成されます。周期は対数スケールで表示されます。

SPECTRA コマンドの追加機能

このコマンド・シンタックス言語では、次の作業を行うこともできます。

- 計算されたスペクトル分析変数を、後で使用できるようにアクティブなデータ・セットに保存する。
- スペクトル・ウィンドウのカスタムの重みを指定する。
- 度数と期間の両方を基準にプロットを生成する。
- プロットに示される各値のすべてのリストを表示する。

シンタックスの詳細については、「コマンド・シンタックス・リファレンス」を参照してください。

第 6 章 適合度測定結果

この項では、時系列モデル作成で使用される適合度測定結果の定義を示します。

- *Stationary R-squared* (定常 R^2 乗). モデルの定常部分を単純平均モデルと比較する指標。トレンド・パターンまたは季節パターンがある場合は、通常の R^2 乗よりもこの指標を推奨します。定常 R^2 乗は負になる場合があり、範囲は負の無限大から 1 までです。負の値は、検討中のモデルがベースライン・モデルより悪いことを意味します。正の値は、検討中のモデルがベースライン・モデルより良いことを意味します。
- *R-squared* (R^2 乗). 系列の全変動のうち、モデルで説明される割合の推定値。この指標は、系列が定常的な場合に非常に有用です。 R^2 乗は負になる場合があり、範囲は負の無限大から 1 までです。負の値は、検討中のモデルがベースライン・モデルより悪いことを意味します。正の値は、検討中のモデルがベースライン・モデルより良いことを意味します。
- *RMSE*. 平方平均誤差平方根。平均平方誤差の平方根。モデルによって予測されるレベルから従属系列がどの程度外れているかを、従属系列と同じ単位を使用して表した指標。
- *MAPE*. 平均絶対パーセント誤差。モデルによって予測されるレベルから従属系列がどの程度外れているかの指標。使用する単位に依存しないので、異なる単位の系列との比較に使用することができます。
- *MAE*. 平均絶対誤差。モデルによって予測されるレベルから系列がどの程度外れているかを測定します。MAE は、元の系列単位で報告されます。
- *MaxAPE*. 最大絶対パーセント誤差。最大予測誤差であり、パーセント単位で表します。この指標は、予測に対する最悪のシナリオを想定する場合に有用です。
- *MaxAE*. 最大絶対誤差。最大予測誤差であり、従属系列と同じ単位で表します。MaxAPE と同様に、予測に対する最悪のシナリオを想定する場合に有用です。最大絶対値誤差と最大絶対パーセント誤差は、異なる系列ポイントで生じる場合があります。例えば、大きな系列値の絶対誤差が小さな系列値の絶対誤差よりわずかに大きい場合が挙げられます。その場合、最大絶対誤差は大きい側の系列値で発生し、最大絶対パーセント誤差は小さい側の系列値で発生します。
- *Normalized BIC* (正規化 *BIC*). 正規化ベイズ情報量基準。モデルの複雑さを説明しようとする、モデルの全体的適合度の一般的な指標。平方平均誤差に基づいたスコアであり、モデルおよび系列の長さのパラメーターの数に応じたペナルティーを含みます。ペナルティーにより、パラメーターが多いモデルの利点が減殺されますが、同じ系列の各種のモデルにわたる統計量の比較が容易になります。

第 7 章 外れ値のタイプ

この項では、時系列モデル作成で使用される外れ値のタイプの定義を示します。

- *Additive* (加法). 単一の観測値に影響する外れ値。例えば、データ・コーディング・エラーは加法的外れ値として識別される場合があります。
- *Level shift* (レベル シフト). 特定の系列点から開始し、すべての観測値を定数分だけシフトさせる外れ値。レベル・シフトは方針の変更に起因することがあります。
- *Innovational* (攪乱). 特定の系列ポイントでノイズ項への追加の役割をする外れ値。定常系列では、攪乱的外れ値がいくつかの観測値に影響します。非定常系列では、特定の系列ポイントで始まるすべての観測値に影響する場合があります。
- *Transient* (一時的). 影響が指数関数的に 0 まで減衰する外れ値。
- *Seasonal additive* (季節加法). 特定の観測値と、季節期間が 1 つ以上離れた以降のすべての観測値に影響する外れ値。これらの観察にはすべて等しい影響が及びます。特定の年以降、毎年 1 月に売上が高くなる場合は、加法的季節外れ値が発生している可能性があります。
- *Local trend* (局所トレンド). 特定の系列ポイントで局所トレンドを開始する外れ値。
- *Additive patch* (加法パッチ). 2 つ以上の連続する加法的な外れ値のグループ。この外れ値の種類を選択すると、外れ値のグループに加えて、個々の加法外れ値の検出につながります。

第 8 章 ACF/PACF プロットについて

ここでは、純粋な ARIMA プロセスまたは理論上の ARIMA プロセスのプロットを示します。プロセスを特定するための一般的なガイドラインがいくつかあります。

- 非定常系列では、すぐに 0 に減少するのではなく、6 つ以上のラグで有意なままである ACF があります。プロセスを特定するには、このような系列が定常になるまで差分を取る必要があります。
- 自己回帰プロセスは、指数関数的に減少する ACF と、PACF の先頭から 1 つ以上のラグに垂線を持ちます。垂線の数は、自己回帰の次数を示しています。
- 移動平均プロセスは、ACF の先頭から 1 つ以上のラグに垂線を持ち、指数関数的に減少する PACF を持ちます。垂線の数は、移動平均の次数を示しています。
- 混合 (ARMA) プロセスは、通常、ACF と PACF の両方で指数関数的な減少を示します。

特定する段階では、ACF または PACF の符号について、または指数関数的に減少する ACF または PACF が 0 に達する速度について気にする必要はありません。これらは、AR および MA 係数の符号および実際の値によって異なります。場合によっては、指数関数的に減少する ACF の値が正と負で交互に入れ替わることがあります。

実際のデータからの ACF および PACF プロットは、ここで示すプロットのように整っていることは決してありません。任意のプロットに不可欠なデータを抽出する必要がある場合があります。特定が誤っている場合に備えて、常に残差の ACF および PACF を確認してください。次の点に注意してください。

- 季節性のプロセスは、季節ラグでこのようなパターンを示します (季節期間の倍数)。
- 有意でない値を 0 として処理できます。つまり、プロットの信頼区間内にある値は無視できます。ただし、値が統計上有意な値のパターンを維持している場合は、その値を無視する必要はありません。
- 偶発的な自己相関は、偶然以外では統計的に有意になりません。統計的に有意な自己相関は、それが孤立している場合 (望ましいのは高いラグで)、および季節ラグで発生していない場合は無視できます。

ACF および PACF プロットの詳細は、ARIMA 分析に関するテキストを参照してください。

表 2. ARIMA(0,0,1), $q>0$

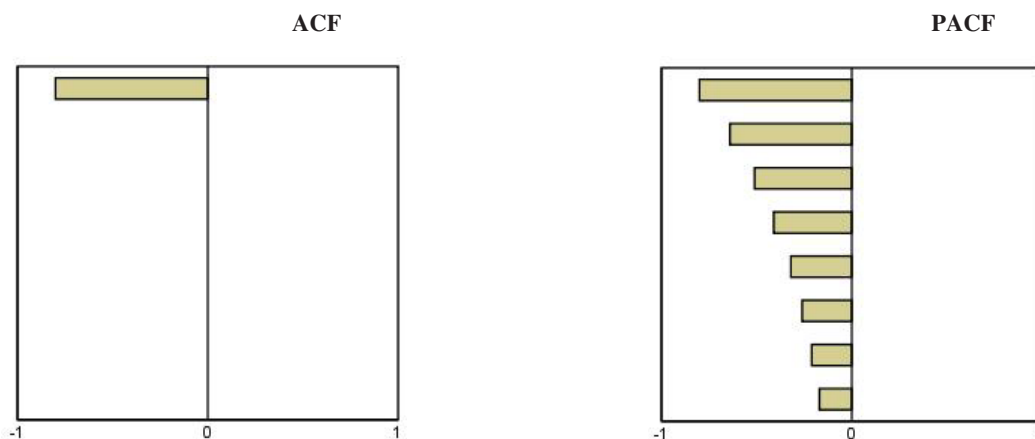
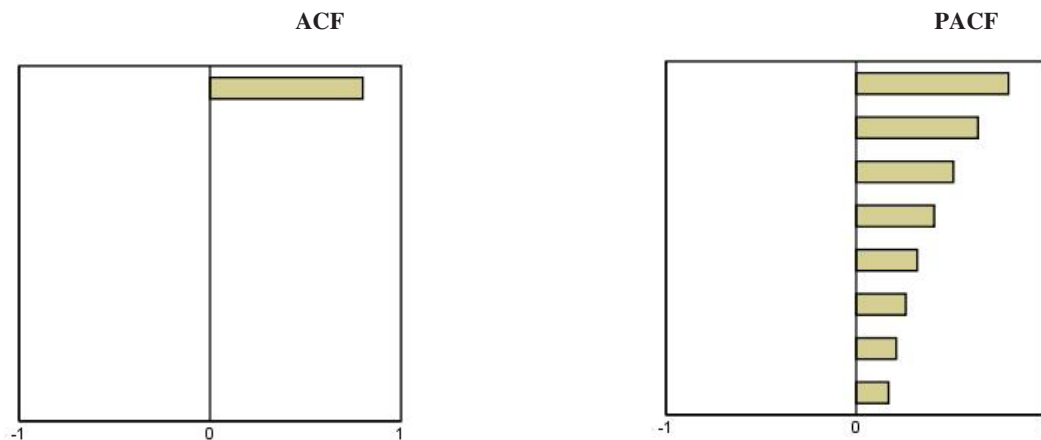


表 3. $ARIMA(0,0,1), q < 0$



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

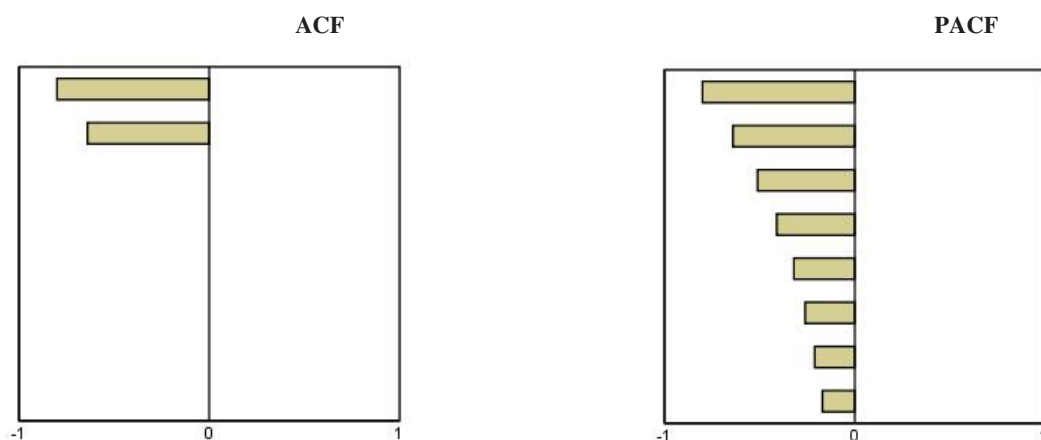


表 4. $ARIMA(1,0,0), f > 0$

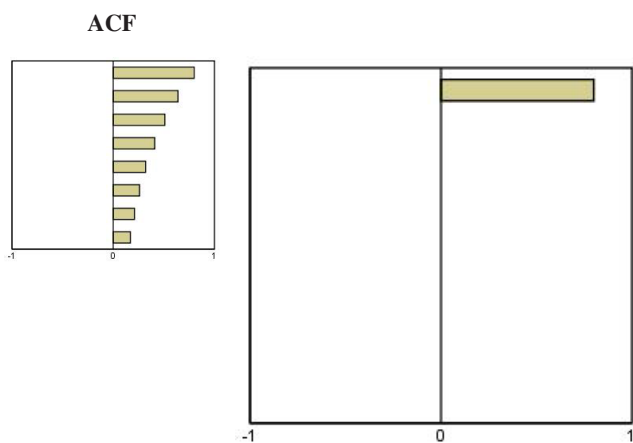
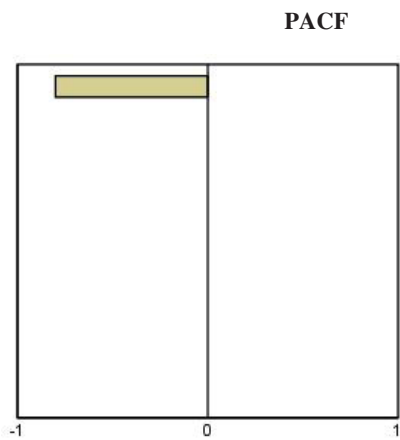
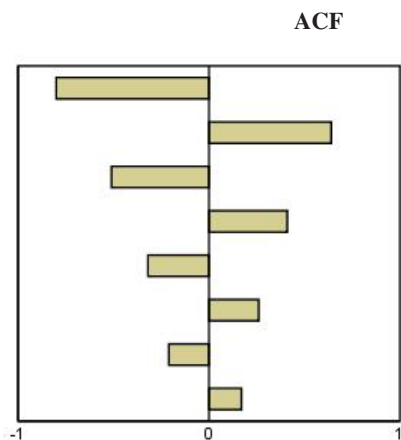
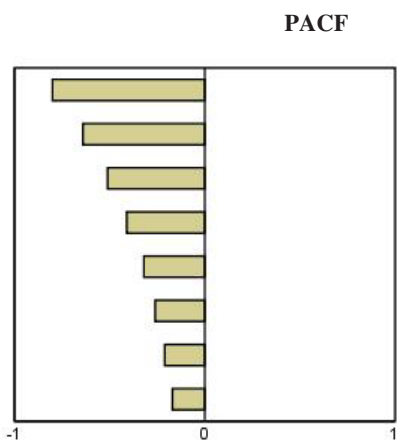
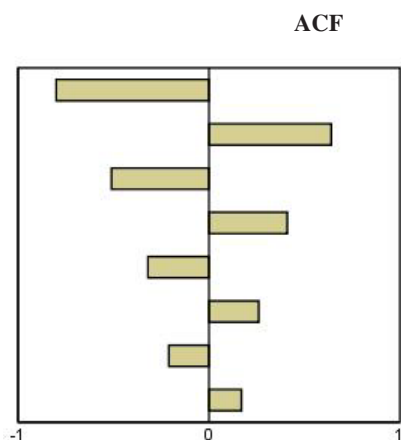


表 5. ARIMA(1,0,0), $f < 0$



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



ARIMA(2,0,0), $\rho_1 \rho_2 > 0$

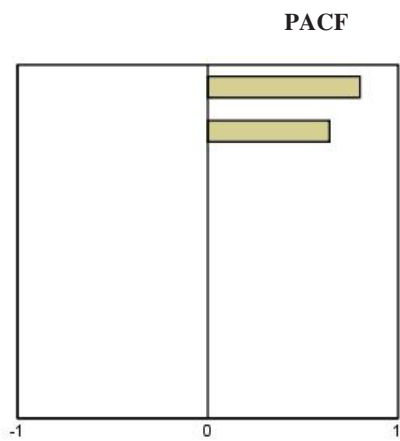
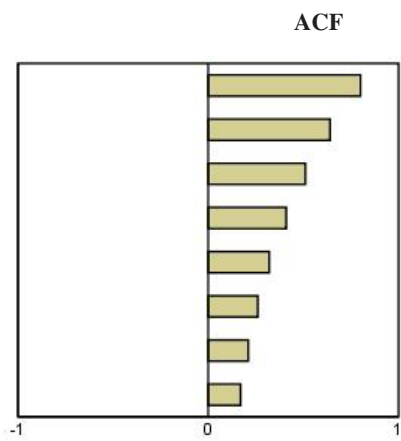
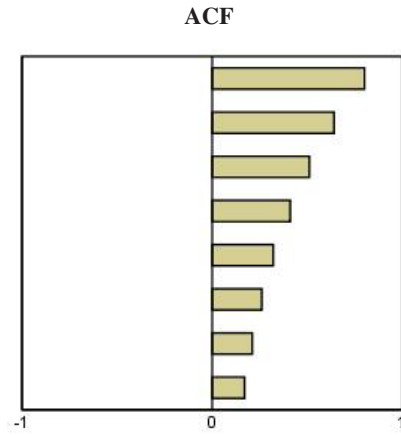


表 6. $ARIMA(0,1,0)$ (和分系列)



特記事項

本書は米国 IBM が提供する製品およびサービスについて作成したものです。

本書に記載の製品、サービス、または機能が日本においては提供されていない場合があります。日本で利用可能な製品、サービス、および機能については、日本 IBM の営業担当員にお尋ねください。本書で IBM 製品、プログラム、またはサービスに言及していても、その IBM 製品、プログラム、またはサービスのみが使用可能であることを意味するものではありません。これらに代えて、IBM の知的所有権を侵害することのない、機能的に同等の製品、プログラム、またはサービスを使用することができます。ただし、IBM 以外の製品とプログラムの操作またはサービスの評価および検証は、お客様の責任で行っていただきます。

IBM は、本書に記載されている内容に関して特許権 (特許出願中のものを含む) を保有している場合があります。本書の提供は、お客様にこれらの特許権について実施権を許諾することを意味するものではありません。実施権についてのお問い合わせは、書面にて下記宛先にお送りください。

〒103-8510

東京都中央区日本橋箱崎町19番21号

日本アイ・ビー・エム株式会社

法務・知的財産

知的財産権ライセンス渉外

以下の保証は、国または地域の法律に沿わない場合は、適用されません。IBM およびその直接または間接の子会社は、本書を特定物として現存するままの状態を提供し、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任を負わないものとします。国または地域によっては、法律の強行規定により、保証責任の制限が禁じられる場合、強行規定の制限を受けるものとします。

この情報には、技術的に不適切な記述や誤植を含む場合があります。本書は定期的に見直され、必要な変更は本書の次版に組み込まれます。IBM は予告なしに、随時、この文書に記載されている製品またはプログラムに対して、改良または変更を行うことがあります。

本書において IBM 以外の Web サイトに言及している場合がありますが、便宜のため記載しただけであり、決してそれらの Web サイトを推奨するものではありません。それらの Web サイトにある資料は、この IBM 製品の資料の一部ではありません。それらの Web サイトは、お客様の責任でご使用ください。

IBM は、お客様が提供するいかなる情報も、お客様に対してなんら義務も負うことのない、自ら適切と信ずる方法で、使用もしくは配布することができるものとします。

本プログラムのライセンス保持者で、(i) 独自に作成したプログラムとその他のプログラム (本プログラムを含む) との間での情報交換、および (ii) 交換された情報の相互利用を可能にすることを目的として、本プログラムに関する情報を必要とする方は、下記に連絡してください。

IBM Software Group

ATTN: Licensing

200 W. Madison St.

Chicago, IL; 60606

U.S.A.

本プログラムに関する上記の情報は、適切な使用条件の下で使用することができますが、有償の場合もあります。

本書で説明されているライセンス・プログラムまたはその他のライセンス資料は、IBM 所定のプログラム契約の契約条項、IBM プログラムのご使用条件、またはそれと同等の条項に基づいて、IBM より提供されます。

この文書に含まれるいかなるパフォーマンス・データも、管理環境下で決定されたものです。そのため、他の操作環境で得られた結果は、異なる可能性があります。一部の測定が、開発レベルのシステムで行われた可能性があります。その測定値が、一般に利用可能なシステムのものと同じである保証はありません。さらに、一部の測定値が、推定値である可能性があります。実際の結果は、異なる可能性があります。お客様は、お客様の特定の環境に適したデータを確かめる必要があります。

IBM 以外の製品に関する情報は、その製品の供給者、出版物、もしくはその他の公に利用可能なソースから入手したものです。IBM は、それらの製品のテストは行っておりません。したがって、他社製品に関する実行性、互換性、またはその他の要求については確認できません。IBM 以外の製品の性能に関する質問は、それらの製品の供給者をお願いします。

IBM の将来の方向または意向に関する記述については、予告なしに変更または撤回される場合があります、単に目標を示しているものです。

本書には、日常の業務処理で用いられるデータや報告書の例が含まれています。より具体性を与えるために、それらの例には、個人、企業、ブランド、あるいは製品などの名前が含まれている場合があります。これらの名称はすべて架空のものであり、名称や住所が類似する企業が実在しているとしても、それは偶然にすぎません。

それぞれの複製物、サンプル・プログラムのいかなる部分、またはすべての派生的創作物にも、次のように、著作権表示を入れていただく必要があります。

本書には、日常の業務処理で用いられるデータや報告書の例が含まれています。より具体性を与えるために、それらの例には、個人、企業、ブランド、あるいは製品などの名前が含まれている場合があります。これらの名称はすべて架空のものであり、名称や住所が類似する企業が実在しているとしても、それは偶然にすぎません。

それぞれの複製物、サンプル・プログラムのいかなる部分、またはすべての派生的創作物にも、次のように、著作権表示を入れていただく必要があります。

© (お客様の会社名) (西暦年). このコードの一部は、IBM Corp. のサンプル・プログラムから取られています。

© Copyright IBM Corp. _年を入れる_. All rights reserved.

商標

IBM、IBM ロゴおよび ibm.com は、世界の多くの国で登録された International Business Machines Corporation の商標です。他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。現時点での IBM の商標リストについては、<http://www.ibm.com/legal/copytrade.shtml> をご覧ください。

Adobe、Adobe ロゴ、PostScript、PostScript ロゴは、Adobe Systems Incorporated の米国およびその他の国における登録商標または商標です。

インテル、Intel、Intel ロゴ、Intel Inside、Intel Inside ロゴ、Centrino、Intel Centrino ロゴ、Celeron、Xeon、Intel SpeedStep、Itanium、および Pentium は、Intel Corporation または子会社の米国およびその他の国における商標または登録商標です。

Linux は、Linus Torvalds の米国およびその他の国における登録商標です。

Microsoft、Windows、Windows NT および Windows ロゴは、Microsoft Corporation の米国およびその他の国における商標です。

UNIX は The Open Group の米国およびその他の国における登録商標です。

Java およびすべての Java 関連の商標およびロゴは Oracle やその関連会社の米国およびその他の国における商標または登録商標です。

索引

日本語, 数字, 英字, 特殊文字の順に配列されています。なお, 濁音と半濁音は清音と同等に扱われています。

[ア行]

当てはめ値

- 時系列モデラー 11
- 時系列モデルの適用 18

イベント 5

- 時系列モデラー 5

エキスパート・モデラー 3

- 外れ値 5
- モデル・スペースの制限 5

[カ行]

外れ値

- エキスパート・モデラー 5
- 定義 31
- ARIMA モデル 9

過渡的な外れ値 31

- 時系列モデラー 5, 9

加法の外れ値 31

- 時系列モデラー 5, 9

加法バッチの外れ値 31

- 時系列モデラー 5, 9

技術革新的外れ値 31

- 時系列モデラー 5, 9

季節性加法の外れ値 31

- 時系列モデラー 5, 9

季節性の分解 23, 24

- 移動平均の計算 23

仮定 23

新変数の保存 24

変数の作成 24

モデル 23

局所トレンドの外れ値 31

- 時系列モデラー 5, 9

欠損値

- 時系列モデラー 13
- 時系列モデルの適用 20

検証期間 2

減衰指数平滑法モデル 6

[サ行]

最大絶対誤差 29

- 時系列モデラー 10, 11

最大絶対誤差 (続き)

- 時系列モデルの適用 17, 18

最大絶対パーセント誤差 29

- 時系列モデラー 10, 11

時系列モデルの適用 17, 18

残差

- 時系列モデラー 10, 11

時系列モデルの適用 17, 18

時系列モデラー 3

新しい変数名 12

当てはめ値 11

イベント 5

エキスパート・モデラー 3

外れ値 5, 9

系列変換 6, 7, 8

欠損値 13

最適モデルと最不適モデル 11

残差自己相関関数 10, 11

残差偏自己相関関数 10, 11

指数平滑法 3, 6

周期 5, 6, 7, 8

信頼区間 11, 13

推定期間 3

すべてのモデルにわたる統計量 10, 11

適合度統計量 10, 11

転送関数 8

モデル指定を XML 形式で保存 12

モデル名 13

モデル・パラメーター 10

予測 10, 11

予測期間 3, 13

予測の保存 12

ARIMA 3, 7

Box-Ljung 統計量 10

時系列モデルの適用 15

新しい変数名 19

当てはめ値 18

欠損値 20

最適モデルと最不適モデル 19

再度推定されたモデルを XML 形式で保存 19

残差自己相関関数 17, 18

残差偏自己相関関数 17, 18

信頼区間 18, 20

推定期間 15

すべてのモデルにわたる統計量 17, 18

適合度統計量 17, 18

モデル・パラメーター 17

モデル・パラメーターの再度推定 15

時系列モデルの適用 (続き)

- 予測 17, 18

予測期間 15

予測の保存 19

Box-Ljung 統計量 17

自己相関関数

- 時系列モデラー 10, 11

時系列モデルの適用 17, 18

純粋な ARIMA プロセスのプロット

33

指数平滑法モデル 3, 6

自然対数変換

- 時系列モデラー 6, 7, 8

周期

- 時系列モデラー 5, 6, 7, 8

信頼区間

- 時系列モデラー 11, 13

時系列モデルの適用 18, 20

推定期間 2

- 時系列モデラー 3

時系列モデルの適用 15

スペクトル・プロット 25, 27

仮定 25

スペクトル・ウィンドウ 25

中心化変換 25

2 変量スペクトル分析 25

正規化 BIC (バイズ情報量基準) 29

- 時系列モデラー 10, 11

時系列モデルの適用 17, 18

[タ行]

対数変換

- 時系列モデラー 6, 7, 8

単純季節指数平滑法モデル 6

単純指数平滑法モデル 6

調和分析 25

定常 R2 29

- 時系列モデラー 10, 11

時系列モデルの適用 17, 18

適合度

- 時系列モデラー 10, 11

時系列モデルの適用 17, 18

定義 29

転送関数 8

季節次数 8

差分次数 8

遅延 8

分子次数 8

分母次数 8

[ハ行]

- 平均絶対誤差 29
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18
- 平均絶対パーセント誤差 29
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18
- 平方根変換
 - 時系列モデラー 6, 7, 8
- 偏自己相関関数
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18
 - 純粋な ARIMA プロセスのプロット 33
- 変数名
 - 時系列モデラー 12
 - 時系列モデルの適用 19
- ホールドアウト・ケース 2
- 保存
 - 新しい変数名 12, 19
 - モデルの予測 12, 19
 - XML 形式でのモデル指定 12
 - XML 形式の再度推定されたモデル 19

[マ行]

- モデル
 - エキスパート・モデラー 3
 - 指数平滑法 3, 6
 - ARIMA 3
- モデル名
 - 時系列モデラー 13
- モデル・パラメーター
 - 時系列モデラー 10
 - 時系列モデルの適用 17
- モデル・パラメーターの再度推定
 - 時系列モデルの適用 15

[ヤ行]

- 予測
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18
- 予測期間
 - 時系列モデラー 3, 13
 - 時系列モデルの適用 15

[ラ行]

- 履歴期間 2
- 履歴データ
 - 時系列モデラー 11
 - 時系列モデルの適用 18

- レベル・シフトの外れ値 31
- 時系列モデラー 5, 9

[数字]

- 2 乗平均平方根誤差 29
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18

A

- ACF
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18
 - 純粋な ARIMA プロセスのプロット 33
- ARIMA モデル 3
 - 外れ値 9
 - 転送関数 8

B

- Box-Ljung 統計量
 - 時系列モデラー 10
 - 時系列モデルの適用 17
- Brown の指数平滑法モデル 6

H

- Holt の指数平滑法モデル 6

M

- MAE 29
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18
- MAPE 29
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18
- MaxAE 29
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18
- MaxAPE 29
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18

P

- PACF
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18
 - 純粋な ARIMA プロセスのプロット 33

R

- R2 29
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18
- RMSE 29
 - 時系列モデラー 10, 11
 - 時系列モデルの適用 17, 18

W

- Winters の指数平滑法モデル
 - 加法 6
 - 乗法 6

X

- XML
 - 再度推定されたモデルを XML 形式で保存 19
 - 時系列モデルを XML 形式で保存 12



Printed in Japan