

IBM SPSS Forecasting 23

The IBM logo, consisting of the letters 'IBM' in a bold, black, sans-serif font. Each letter is composed of eight horizontal bars of varying lengths, creating a striped effect. The logo is centered 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、65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 23 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 23

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	モデル化する時系列	32
時系列データ	1	次元の値の選択	33
データの変換	2	観測	33
推定期間と検証期間	2	分析の時間間隔	35
モデルの構築と予測の生成	3	集計と分配 (Aggregation and Distribution)	35
		欠損値	36
第 2 章 時系列モデラー	5	一般的なデータ・オプション	37
エキスパート・モデラーのオプションの指定	7	一般的な作成オプション	37
モデルの選択とイベントの指定	7	表示する系列	37
エキスパート・モデラーでの外れ値の処理	7	出力オプション	38
カスタム指数平滑法モデル	8	推定期間	40
カスタム ARIMA モデル	9	予測	41
カスタム ARIMA モデルに対するモデル指定	9	保存	41
カスタム ARIMA モデルにおける転送関数	10	インタラクティブ出力	41
カスタム ARIMA モデルの外れ値	11		
出力	12	第 7 章 時間的因果モデルの適用	43
統計テーブルと予測テーブル	12	時間的因果モデルの適用	43
プロット	13	時間的因果モデル予測	43
出力を最適モデルまたは最不適合モデルに限定	13	時間的因果モデル予測を使用するには	43
モデル予測およびモデル指定の保存	14	モデルのパラメータおよび予測	44
オプション	14	一般オプション	45
TSMODEL コマンドの追加機能	15	表示する系列	45
		出力オプション	46
		保存	48
第 3 章 時系列モデルの適用	17	時間的因果モデルのシナリオ	49
出力	19	時間的因果モデルのシナリオを実行するには	49
統計テーブルと予測テーブル	19	シナリオ期間の定義	49
プロット	20	シナリオおよびシナリオ グループの追加	50
出力を最適モデルまたは最不適合モデルに限定	21	オプション	54
モデル予測およびモデル指定の保存	21		
オプション	22	第 8 章 適合度測定結果	57
TSAPPLY コマンドの追加機能	22		
		第 9 章 外れ値のタイプ	59
第 4 章 季節性の分解	25		
季節性の分解の保存	26	第 10 章 ACF/PACF プロットについて	61
SEASON コマンドの追加機能	26		
		特記事項	65
第 5 章 スペクトル・プロット	27	商標	67
SPECTRA コマンドの追加機能	29		
		索引	69
第 6 章 時間的因果モデル	31		
時間的因果モデルを取得するには	32		

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

列ベースのデータ

各時系列フィールドに、単一の時系列のデータを含めます。この構造は、時系列モデラーのプロシージャ、季節性の分解のプロシージャ、およびスペクトル プロット のプロシージャで使用される、従来型の時系列データの構造です。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「行と列を入れ替え」を使用して行と列を入れ替えることができます。

多次元データ

多次元データでは、各時系列フィールドに、複数の時系列のデータが含まれます。その場合、特定のフィールド内の別個の時系列は、次元 フィールドと呼ばれるカテゴリ フィールドの値セットによって識別されます。

例えば、異なる地域 (region) およびブランド (brand) の売上 (sales) データを単一の sales フィールドに保管できます。このケースの次元は、region と brand になります。region と brand の各組み合わせにより、sales の特定の時系列が識別されます。例えば、以下のテーブルでは、region が「north」で brand が「brandX」のレコードにより、単一の時系列が定義されます。

表 2. 多次元データ

日付	領域	brand	sales
01/01/2014	north	brandX	82350
01/01/2014	north	brandY	86380
01/01/2014	south	brandX	91375
01/01/2014	south	brandY	70320
01/02/2014	north	brandX	83275
01/02/2014	north	brandY	85260
01/02/2014	south	brandX	94760
01/02/2014	south	brandY	69870

注: IBM® Cognos® TM1® など、OLAP キューブからインポートされたデータは、多次元データとして表されます。

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。これらの変換は、各時系列フィールドに単一の時系列のデータが含まれている列ベースのデータにのみ適用されます。

- 「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を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 (「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 (「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。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 つのプロシージャが用意されています。

- 5 ページの『第 2 章 時系列モデラー』のプロシージャでは、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。
- 17 ページの『第 3 章 時系列モデルの適用』のプロシージャでは、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しようとするエキスパート 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」 > 「時系列」 > 「従来モデルの作成...」

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「日付かケース番号の範囲」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

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

「伝達関数」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。

「伝達関数の順序」。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、59 ページの『第 9 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

「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

「作図」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
- 標準化ベイズ情報量基準
- 残差自己相関関数 (ACF)
- 残差偏自己相関関数 (PACF)

「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す。

- 「**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「日付かケース番号の範囲」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チャーの 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」 > 「時系列」 > 「従来モデルの適用...」

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「**日付かケース番号の範囲**」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

「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

「**作図**」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
- 標準化ベイズ情報量基準
- 残差自己相関関数 (ACF)
- 残差偏自己相関関数 (PACF)

「**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ま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」 > 「時系列」 > 「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」プロシージャーは、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、このプロシージャー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 時間的因果モデル

時間的因果モデリングでは、時系列データ内の重要な因果関係を検出しようとしています。時間的因果モデリングでは、対象系列セットおよびその対象への入力候補セットを指定します。そうすると、プロシージャーにより、各対象の自己回帰時系列モデルが作成され、対象に対して因果関係を持つ入力のみが含まれます。このアプローチは、対象系列の予測変数を明示的に指定する必要がある従来の時系列モデリングとは異なります。時間的因果モデリングでは通常、複数の関連時系列のモデルを作成する必要があるため、結果は、モデル システム と呼びます。

時間的因果モデリングのコンテキストでは、因果 という用語は、グレンジャー因果性を指します。時系列 X が別の時系列 Y の「グレンジャー因果」であると言うのは、 X と Y の両方の過去値の観点における Y の回帰により、 Y の過去値のみの観点における回帰よりも優れた Y のモデルが得られる場合です。

例

ビジネスの意思決定者は、時間的因果モデリングを使用して、ビジネスについて説明する大規模な時間ベースの測定基準のセット内の因果関係を検出できます。分析により、重要業績評価指標に最も大きな影響を与える、いくつかの制御可能な入力が明らかになることがあります。

大規模な IT システムの管理者は、時間的因果モデリングを使用して、相互関係のある大規模運用測定基準セット内の異常を検出できます。その後、因果モデルにより、異常検出に加えて、異常の根本原因として最も可能性が高いものを検出できます。

フィールド要件

1 つ以上の対象がなければなりません。デフォルトでは、事前定義された役割が「なし」のフィールドは使用されません。

データ構造

時間的因果モデリングでは、2 つのタイプのデータ構造がサポートされます。

列ベースのデータ

列ベースのデータの場合、各時系列フィールドには、単一の時系列のデータが含まれます。この構造は、時系列モデラーで使用される、従来型の時系列データの構造です。

多次元データ

多次元データでは、各時系列フィールドに、複数の時系列のデータが含まれます。その場合、特定のフィールド内の別個の時系列は、次元 フィールドと呼ばれるカテゴリ フィールドの値セットによって識別されます。例えば、2 つの異なる販売チャネル (小売と Web) の売上データを単一の *sales* フィールドに保管できます。*channel* という次元フィールド (「retail」 (小売) と「web」 (Web) という値が含まれる) により、2 つの販売チャネルのそれぞれに関連したレコードを識別します。

時間的因果モデルを取得するには

この機能を使用するには Statistics Forecasting オプションが必要です。

メニューから次の項目を選択します。

「分析」 > 「時系列」 > 「時間的因果モデルの作成...」

1. 観測が日付と時刻フィールドによって定義されている場合、そのフィールドを指定します。
2. データが多次元の場合、時系列を識別する次元フィールドを指定します。
 - 次元フィールドの指定順序により、後続のすべてのダイアログおよび出力での次元フィールドの表示順序が決定されます。上矢印ボタンと下矢印ボタンを使用して、次元フィールドを並べ替えます。
 - 次元フィールドの特定の値に分析を制限する次元値のサブセットを指定できます。例えば、地域とブランドの次元がある場合、分析を特定の地域に制限できます。次元のサブセットは、分析で使用されているすべての測定基準フィールドに適用されます。
 - 測定基準フィールド レベルでの役割によって次元値を指定することで、分析をカスタマイズすることもできます。例えば、(値「retail」(小売) と「web」(Web) が含まれた) 販売チャネルの次元と、各チャネルの売上と広告の測定基準がある場合、小売と Web の両方の売上の入力として Web 広告を指定できます。デフォルトでは、このタイプのカスタマイズは有効になっており、各次元フィールドの特定数 (デフォルトでは 250) の異なる値からの選択に制限されています。
3. 「続行」をクリックします。

注: アクティブなデータ セットに日付が指定されている場合、ステップ 1、2 および 3 は適用されません。日付の指定は、「日付の定義」ダイアログまたは DATE コマンドで作成されます。

4. 「フィールド」をクリックして、モデルに含める時系列を指定し、また観測の定義を指定します。少なくとも 1 つのフィールドを、対象として、または入力と対象の両方として、指定する必要があります。
5. 「データ仕様 (Data Specifications)」をクリックして、分析の時間間隔、集計と分配の設定、欠損値の処理などのオプション設定を指定します。
6. 「作成オプション」をクリックして、推定期間を定義し、出力の内容を指定し、対象当たりの最大入力数などの作成設定を指定します。
7. 「モデル オプション」をクリックして、予測の要求、予測の保存、外部ファイルへのモデル システムのエクスポートを行います。
8. 「実行」をクリックしてプロシージャを実行します。

モデル化する時系列

「フィールド」タブで、「時系列」設定を使用して、モデル システムに含める系列を指定します。

列ベースのデータの場合、系列 という用語の意味は、フィールド という用語と同じです。多次元データの場合、時系列が含まれているフィールドは、測定基準 フィールドと呼ばれます。多次元データの時系列は、測定基準フィールドと、各次元フィールドの値によって定義されます。列ベースのデータと多次元データの両方に、以下の考慮事項が適用されます。

- 入力の候補として、または対象と入力の両方として指定された系列は、各対象のモデルに含めるものとして考慮されます。各対象のモデルには常に、対象自体の遅れた値が含まれます。
- 強制入力と指定された系列は常に、各対象のモデルに含められます。
- 対象または対象と入力の両方として、少なくとも 1 つの系列を指定する必要があります。

- 「定義済みの役割を使用」を選択した場合、入力役割を持つフィールドは、入力の候補として設定されます。定義済みの役割は強制入力にマップされません。

多次元データ

多次元データの場合、グリッドで測定基準フィールドおよび関連する役割を指定します。グリッドの各行で、単一の測定基準と役割を指定します。デフォルトでは、モデル システムには、グリッド内の各行に、次元フィールドのすべての組み合わせの系列が含まれます。例えば、*region* (地域) と *brand* (ブランド) の次元がある場合、デフォルトでは、測定基準 *sales* (売上) を対象として指定すると、*region* と *brand* の組み合わせごとに別個の売上対象系列が存在することになります。

グリッド内の各行で、次元の省略符号ボタンをクリックすることで、任意の次元フィールドの値セットをカスタマイズできます。このアクションにより、「次元の値の選択」サブダイアログが開きます。グリッド行の追加、削除、またはコピーを行うこともできます。

「系列の数」列には、関連する測定基準について現在指定されている次元値セット数が表示されます。表示されている値は、系列の実際の数 (セットごとに 1 系列) よりも大きい場合があります。この状態が発生するのは、指定した組み合わせの次元値の一部が、関連する測定基準に含まれている系列に対応していない場合です。

次元の値の選択

多次元データの場合、特定の役割を持つ特定の測定基準フィールドに適用される次元値を指定することで、分析をカスタマイズできます。例えば、*sales* (売上) が測定基準フィールドで、*channel* (チャンネル) が値「retail」(小売) と「web」(Web) が含まれた次元である場合、「web」の売上を入力とし、「retail」の売上を対象にするように指定できます。また、分析で使用されているすべての測定基準フィールドに適用される次元サブセットを指定することもできます。例えば、*region* (地域) が地域を示す次元フィールドの場合、分析を特定の地域に制限できます。

すべての値

現在の次元フィールドのすべての値を含めることを指定します。これはデフォルトのオプションです。

含める値または除外する値の選択(L)

このオプションは、現在の次元フィールドの値セットを指定する場合に使用します。「モード」で「含める」を選択すると、「選択済みの値」リストに指定されている値のみが含まれます。「モード」で「除外」を選択すると、「選択済みの値」リストに指定されている値以外のすべての値が含まれます。

選択元の値セットをフィルタに掛けることができます。フィルタ条件を満たしている値が「一致」タブに表示され、フィルタ条件を満たさない値は「選択されていない値」リストの「不一致」タブに表示されます。「すべて」タブには、フィルタ条件に関係なく、選択されていないすべての値がリストされます。

- フィルタの指定時にワイルドカード文字を指定するために、アスタリスク (*) を使用できます。
- 現在のフィルタをクリアするには、「表示された値のフィルタ」ダイアログの検索語として空の値を指定します。

観測

「フィールド」タブの「観測」設定を使用して、観測を定義するフィールドを指定します。

注: アクティブなデータ セットに日付指定がある場合、観測は日付指定によって定義され、時間的因果モデリングのプロシージャで変更できません。日付の指定は、「日付の定義」ダイアログまたは DATE コマンドで作成されます。

日付と時刻で定義された観測

日付、時刻、または日時形式のフィールドまたは日時を表す文字列フィールドによって観測を定義することを指定できます。文字列フィールドでは、YYYY-MM-DD 形式で日付を表すか、HH:MM:SS 形式で時刻を表すか、YYYY-MM-DD HH:MM:SS 形式で日時を表すことができます。文字列表現では、先行ゼロを省略できます。例えば、文字列 2014-9-01 は 2014-09-01 と同等です。

観測を定義するフィールドに加え、観測を記述する適切な時間間隔を選択します。指定した時間間隔に応じて、観測の間隔 (増分) や 1 週間の日数などの他の設定も指定できます。時間間隔には、以下の考慮事項が適用されます。

- 販売注文の処理時刻など、観測値の時間間隔が不規則な場合、値「**不規則**」を使用します。「**不規則**」を選択した場合、「データ仕様 (Data Specifications)」タブの「**時間間隔**」設定から、分析に使用する時間間隔を指定する必要があります。
- 観測値が日時を表し、時間間隔が時間、分、または秒の場合、「**1 日当たりの時間数 (Hours per day)**」、「**1 日当たりの分数 (Minutes per day)**」、または「**1 日当たりの秒数 (Seconds per day)**」を使用します。観測値が、日付を参照せずに時間 (期間) を表していて、時間間隔が時間、分、または秒の場合、「**時間 (非周期) (Hours (non-periodic))**」、「**分 (非周期) (Minutes (non-periodic))**」、または「**秒 (非周期) (Seconds (non-periodic))**」を使用します。
- 選択した時間間隔に基づいて、プロシージャで、欠損している観測値を検出できます。プロシージャでは、すべての観測値が等間隔の時間に存在し、観測値が欠損していないことが想定されるため、欠損している観測値を検出する必要があります。例えば、時間間隔が日で、日付 2014-10-27 の後に 2014-10-29 が続いている場合、2014-10-28 の観測値が欠損しています。欠損している観測値には値が代入されます。欠損値を処理するための設定は、「データ仕様 (Data Specifications)」タブで指定できます。
- 指定した時間間隔により、プロシージャは、集計する必要がある、同じ時間間隔内の複数の観測値を検出したり、月初めなどの間隔の境界で観測値を調整したりして、観測値が等間隔になるようにすることができます。例えば、時間間隔が月の場合、同じ月の複数の日付は集計されます。このタイプの集計は、**グループ化** と呼びます。デフォルトでは、グループ化された観測値は合計されます。「データ仕様 (Data Specifications)」タブの「**集計と分配 (Aggregation and Distribution)**」設定から、観測値の平均など、別のグループ化方法を指定できます。
- 一部の時間間隔では、追加設定により、通常の等間隔内にブレイクを定義できます。例えば、時間間隔が日であるが、平日のみが有効な場合、1 週間に 5 日あり、週が月曜日から始まるというものを指定できます。

期間または周期的な期間で定義された観測

期間または繰り返し期間サイクル (任意の数のサイクル レベル) を表す 1 つ以上の整数フィールドで観測を定義できます。この構造を使用して、標準時間間隔のいずれにも適合しない観測の系列を記述できます。例えば、年を表すサイクル フィールドと、月を表す期間フィールドを使用し、1 サイクルの長さを 10 にすることで、10 カ月しかない年度を記述できます。

周期的な期間を指定するフィールドでは、周期的レベルの階層を定義します。この階層では、最も低いレベルは、「**期間**」フィールドによって定義します。次に高いレベルは、レベルが 1 のサイクル フィールドによって定義し、その次に高いレベルは、レベルが 2 のサイクル フィールドによって定義します。これ以降も同様です。最高レベルを除き、各レベルのフィールド値は、次に高

いレベルを基準として周期的でなければなりません。最高レベルの値を周期的にすることはできません。例えば、10 カ月から成る年度の場合、年内の月は周期ですが、年は周期的ではありません。

- 特定のレベルにおけるサイクルの長さは、次に低いレベルの周期になります。年度の例の場合、次に低いレベルは月を表し、指定年度に 10 カ月あるため、サイクル レベルは 1 つだけであり、サイクルの長さは 10 です。
- 1 から開始しない周期的フィールドの開始値を指定します。欠損値を検出するために、この設定が必要になります。例えば、周期的フィールドが 2 で開始するが、開始値が 1 と指定されている場合、プロシージャは、当該フィールドの各サイクルの最初の期間に欠損値があるものと想定します。

レコード順序で定義された観測

列ベースのデータの場合、レコード順序で観測を定義し、最初のレコードが最初の観測値を表し、2 番目のレコードが 2 番目の観測値を表すといったようにすることを指定できます。その場合、レコードは、等間隔の時間で発生した観測値を表すものと想定されます。

分析の時間間隔

分析で使用される時間間隔は、観測の時間間隔とは異なることがあります。例えば、観測の時間間隔が日の場合に、分析の時間間隔に月を選択できます。この場合、データは、モデルの構築前に、日次から月次のデータに集計されます。長い時間間隔から短い時間間隔にデータを分配することも選択できます。例えば、観測が四半期単位の場合、四半期のデータを月次データに分配できます。

分析が行われる時間間隔に使用可能な選択肢は、観測の定義と観測の時間間隔によって異なります。特に、観測が周期的な期間で定義されている場合、またはアクティブなデータ セットで日付指定が定義されている場合は、集計のみがサポートされます。その場合、分析の時間間隔は、観測の時間間隔以上でなければなりません。

分析用の時間間隔は、「データ仕様 (Data Specifications)」タブの「時間間隔」設定で指定します。データの集計または分配の方法は、「データ仕様 (Data Specifications)」タブの「集計と分配 (Aggregation and Distribution)」設定で指定します。

集計と分配 (Aggregation and Distribution)

集計関数(G)

分析に使用される時間間隔が観測の時間間隔より長い場合、入力データは集計されます。例えば、観測の時間間隔が日で、分析の時間間隔が月の場合、集計が行われます。集計関数 mean、sum、mode、min、または max を使用できます。

分配関数

分析に使用される時間間隔が観測の時間間隔より短い場合、入力データは分配されます。例えば、観測の時間間隔が四半期で、分析の時間間隔が月の場合、分配が行われます。分配関数 mean または sum を使用できます。

グループ化関数(O)

観測が日時で定義されていて、複数の観測が同じ時間間隔で行われる場合、グループ化が適用されます。例えば、観測の時間間隔が月の場合、同じ月内の複数の日付はグループ化され、その発生月に関連付けられます。グループ化関数 mean、sum、mode、min、または max を使用できます。観測が日時で定義されていて、観測の時間間隔が「不規則」と指定されている場合、常にグループ化が行われます。

注: グループ化は集計の 1 つの形式ですが、欠損値の処理前に実行されます。一方、通常集計は、欠損値の処理後に実行されます。観測の時間間隔が不規則と指定されている場合、集計はグループ化関数でのみ実行されます。

日をまたぐ観測を前日に集計(C)

日の境界をまたぐ時間がある観測を前日の値に集計するかどうかを指定します。例えば、20:00 に開始する 8 時間の日がある毎時観測では、この設定により、00:00 から 04:00 までの観測を前日の集計結果に含めるかどうかを指定します。この設定が適用されるのは、観測の時間間隔が 1 日当たりの時間数、1 日当たりの分数、または 1 日当たりの秒数で、分析の時間間隔が日の場合のみです。

指定されたフィールドのカスタム設定

フィールドごとにフィールドに集計関数、分配関数、およびグループ化関数を指定できます。そのような設定により、集計関数、分配関数、およびグループ化関数のデフォルト設定がオーバーライドされます。

欠損値

入力データの欠損値は、代入値で置き換えられます。以下の置換法が使用可能です。

線形補間

線型補間を使用して欠損値を置換します。欠損値の前の最後の有効値および欠損値の後の最初の有効値が補間に使用されます。系列の最初または最後の観測に欠損値がある場合、系列の最初または最後の 2 つの最も近い非欠損値が使用されます。

一連の平均

欠損値を系列全体の平均値に置換します。

隣接ポイントの平均値

欠損値を、有効な周囲の値の平均値に置換します。隣接ポイントのスペンは、その平均値の計算に使用された欠損値の上下の有効値の数です。

隣接ポイントの中央値

欠損値を、有効な周囲の値の中央値に置換します。隣接ポイントのスペンは、その中央値の計算に使用された欠損値の上下の有効値の数です。

線形トレンド

このオプションでは、系列のすべての非欠損観測値を使用して、単純線型回帰モデルを適合させます。その後、そのモデルを使用して、欠損値の代入を行います。

その他の設定:

欠損値の最大パーセンテージ(%)**(X)**

任意の系列で許可される欠損値の最大パーセントを指定します。指定した最大を超える欠損値がある系列は、分析から除外されます。

ユーザー欠損値

このオプションは、ユーザー欠損値を有効なデータとして扱う (それゆえ、系列に含める) かどうかを指定します。デフォルトでは、ユーザー欠損値は除外され、システム欠損値と同様に扱われ、その後、代入されます。

一般的なデータ・オプション

次元フィールドあたりの異なる数値の最大数(M)

この設定は、多次元データに適用され、1つの次元フィールドに許可される異なる値の最大数を指定します。デフォルトでは、この制限は10000に設定されていますが、任意の大きい数に増やすことができます。

一般的な作成オプション

信頼区間の幅 (%) (C)

この設定は、予測とモデルのパラメータの両方の信頼区間を制御します。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。

対象ごとの入力フィールドの最大数(M)

この設定は、各対象のモデルで許可される最大入力数を指定します。1から20までの範囲の整数を指定できます。各対象のモデルには常に、それ自体の遅れた値が含まれるため、この値を1に設定すると、唯一の入力が対象自体になることが指定されます。

モデル許容度(O)

この設定は、各対象の最適な入力セットを判別するために使用される反復プロセスを制御します。0より大きな値を指定できます。デフォルトは0.001です。

外れ値のしきい値 (%) (T)

観測値には、モデルから計算された、それが外れ値である確率がこのしきい値を超えた場合、外れ値のフラグが立てられます。50から100の範囲の値を指定することができます。

各入力のラグ数 (Number of Lags for Each Input)

この設定は、各対象のモデル内の各入力のラグ項目数を指定します。デフォルトでは、ラグ項目数は分析に使用される時間間隔で自動的に決まります。例えば、時間間隔が月の場合(1カ月の増分を使用)、ラグ数は12です。必要に応じて、明示的にラグ数を指定できます。指定値は1から20までの範囲の整数でなければなりません。

表示する系列

これらのオプションでは、出力を表示する系列(対象または入力)を指定します。指定した系列の出力の内容は、「出力オプション」設定によって決定されます。

最適モデルに関連付けられた対象の表示

デフォルトでは、R2乗の値によって決定される、上位10個の最適モデルに関連した対象の出力が表示されます。最適モデルの別の固定数を指定したり、最適モデルのパーセントを指定したりすることができます。また、以下の適合度を選択できます。

R2乗 線型モデルの適合度。決定係数とも呼びます。ターゲット変数の変動のうち、モデルによって説明される割合です。値の範囲は0から1までです。値が小さい場合は、モデルが十分にデータに適合していないことを示します。

平均平方根パーセント誤差

モデル予測値が系列の観測値とどれほど異なるかを示す指標。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。

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。

BIC ベイズ情報量基準。-2 対数尤度に基づいてモデルを選択し、比較するための指標。値が小さいほどモデルが良好であることを示します。BIC もパラメータが過多のモデル (例えば、大量の入力がある複雑なモデル) にペナルティを科しますが、AIC よりも厳密にそれを行います。

AIC 赤池情報量基準。-2 対数尤度に基づいてモデルを選択し、比較するための指標。値が小さいほどモデルが良好であることを示します。AIC は、パラメータが過多のモデル (例えば、大量の入力がある複雑なモデル) にペナルティを科します。

個別の系列の指定

出力する個別の系列を指定できます。

- 列ベースのデータの場合、必要な系列が含まれているフィールドを指定します。指定したフィールドの順序によって、出力に表示される順序が決定されます。
- 多次元データの場合、系列が含まれている測定基準フィールドのグリッドに項目を追加することで、特定の系列を指定します。その後、系列を定義する次元フィールドの値を指定します。
 - グリッドに直接各次元フィールドの値を入力するか、使用可能な次元値のリストから選択することができます。使用可能な次元値のリストから選択するには、必要な次元のセルの省略符号ボタンをクリックします。このアクションにより、「次元の値の選択」サブダイアログが開きます。
 - 「次元の値の選択」サブダイアログで、双眼鏡アイコンをクリックして検索語を指定することで、次元値のリストを検索できます。スペースは、検索語の一部として扱われます。検索語内のアスタリスク (*) は、ワイルドカード文字を示しません。
 - グリッド内の系列の順序によって、出力に表示される順序が決定されます。

列ベースのデータと多次元データの両方で、出力は 30 系列までに制限されています。この制限には、指定した個別の系列 (入力または対象) および最適モデルに関連した対象が含まれます。個別に指定された系列は、最適モデルに関連した対象よりも優先されます。

出力オプション

これらのオプションは、出力の内容を指定します。「**対象の出力**」グループのオプションは、「**表示する系列**」設定の最適モデルに関連した対象の出力を生成します。「**系列の出力**」グループのオプションは、「**表示する系列**」設定で指定された個別系列の出力を生成します。

全体のモデル システム(O)

モデル システム内の系列間の因果関係のグラフィカル表現を表示します。表示対象のモデル適合統計量と外れ値の両方のテーブルが、出力項目の一部として含まれます。「**系列の出力**」グループでこのオプションを選択した場合、「**表示する系列**」設定で指定されている個別系列ごとに別個の出力項目が作成されます。

系列間の因果関係には、関連する有意水準が含まれており、有意水準が小さいほど、より有意な関連を示しています。特定の値より大きい有意水準の関係を非表示にするように選択できます。

モデルの適合度の統計量と外れ値

表示するように選択した対象系列のモデル適合統計量と外れ値のテーブル。これらのテーブルには、全体のモデル システムの視覚化のテーブルと同じ情報が含まれます。これらのテーブルでは、テーブルのピボットおよび編集を行うための標準機能がすべてサポートされます。

モデル効果とモデルのパラメータ

表示するように選択した対象系列のモデル効果検定とモデルパラメータのテーブル。モデル効果検定には、モデルに含まれている各入力 of F 統計量と関連する有意確率値が含まれます。

影響図(D)

対象系列とそれが影響を与える他の系列またはそれに影響する他の系列との間の因果関係のグラフィカル表現を表示します。対象系列に影響する系列は、**原因** と呼びます。「**効果**」を選択すると、効果を表示するために初期化された影響図が生成されます。「**原因**」を選択すると、原因を表示するために初期化された影響図が生成されます。「**原因と効果の両方**」を選択すると、2 つの別個の影響図 (原因のために初期化された影響図と効果のために初期化された影響図) が生成されます。影響図を表示する出力項目の原因と効果をインタラクティブに切り替えることができます。

表示する原因または効果のレベル数を指定できます。最初のレベルは、対象系列のみです。レベルが増すごとに、対象系列のさらに間接的な原因または効果が表示されます。例えば、効果の表示における 3 番目のレベルは、直接入力として 2 番目のレベルの系列が含まれている系列で構成されます。この場合、対象系列が 2 番目のレベルの系列に対する直接入力であるため、3 番目のレベルの系列は、間接的に対象系列の影響を受けます。

系列プロット

表示するように選択した対象系列の観測値と予測値の作図。予測が要求されると、作図には、予測値と予測の信頼区間も表示されます。

残差プロット

表示するように選択された対象系列のモデル残差の作図。

上位入力

時系列に沿った各表示対象、および対象の上位 3 位までの入力の作図。上位の入力とは、有意確率値が小さい入力です。入力と対象に別のスケールを表示するために、y 軸は、各系列の Z 得点を表します。

予測テーブル(F)

表示するように選択された対象系列の予測値と予測の信頼区間のテーブル。

外れ値の根本原因分析(O)

対象系列の各外れ値の原因である可能性が高い系列を判別します。外れ値の根本原因分析は、「**表示する系列**」設定の個別系列のリストに含まれている対象系列ごとに行われます。

出力

対話式の外れ値のテーブルと図表

各対象系列の外れ値とその外れ値の根本原因のテーブルと図表。テーブルには、外れ値ごとに 1 つの行が含まれています。図表は、影響図です。テーブル内の行を選択すると、影響図内で、対象系列から関連する外れ値の原因である可能性が高い系列までのパスが強調表示されます。

外れ値のピボット テーブル

各対象系列の外れ値とその外れ値の根本原因のテーブル。このテーブルには、インタラクティブ表示のテーブルと同じ情報が含まれます。このテーブルでは、テーブルのピボットおよび編集を行うための標準機能がすべてサポートされます。

因果関係レベル

根本原因の検索で含めるレベル数を指定できます。ここで使用されるレベルの概念は、影響図で説明したものと同じです。

すべてのモデル間でのモデルの適合度

すべてのモデルのモデルの適合度および選択した適合度の統計量のヒストグラム。以下の適合度の統計量を使用することができます。

R2 乗 線型モデルの適合度。決定係数とも呼びます。ターゲット変数の変動のうち、モデルによって説明される割合です。値の範囲は 0 から 1 までです。値が小さい場合は、モデルが十分にデータに適合していないことを示します。

平均平方根パーセント誤差

モデル予測値が系列の観測値とどれほど異なるかを示す指標。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。

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。

BIC ベイズ情報量基準。-2 対数尤度に基づいてモデルを選択し、比較するための指標。値が小さいほどモデルが良好であることを示します。BIC もパラメータが過多のモデル (例えば、大量の入力がある複雑なモデル) にペナルティを科しますが、AIC よりも厳密にそれを行います。

AIC 赤池情報量基準。-2 対数尤度に基づいてモデルを選択し、比較するための指標。値が小さいほどモデルが良好であることを示します。AIC は、パラメータが過多のモデル (例えば、大量の入力がある複雑なモデル) にペナルティを科します。

長期間の外れ値(V)

推定期間内の時間間隔ごとの、すべての対象における、外れ値の数の棒グラフ。

系列の変換(T)

モデル システム内の系列に適用された変換のテーブル。可能な変換は、欠損値の代入、集計、および分配です。

推定期間

デフォルトでは、推定期間は、最も早い観測の時刻から始まり、すべての系列における最も遅い観測の時刻に終わります。

開始時刻と終了時刻で指定(B)

推定期間の開始と終了の両方を指定するか、開始または終了だけを指定することができます。推定期間の開始または終了を省略した場合、デフォルト値が使用されます。

- 観測が日付と時刻フィールドで定義されている場合、日付と時刻フィールドで使用されているのと同じ形式で、開始および終了の値を入力します。
- 周期的な期間で定義された観測の場合、周期的な期間フィールドのそれぞれの値を指定します。各フィールドは個別の列で表示されます。
- アクティブなデータ セットに有効な日付指定がある場合、日付指定の各構成要素 (月など) の値を指定する必要があります。各構成要素は個別の列で表示されます。
- 観測がレコード順序で定義されている場合、推定期間の開始および終了は、関連ケースの行番号 (データ エディタに表示) で定義されます。

最新の時間間隔または最も早い時間間隔で指定 (By latest or earliest time intervals)

データにおける最も早い時間間隔から開始するか、最新の時間間隔で終了する特定の数の時間間隔 (およびオプションとしてオフセット) として推定期間を定義します。このコンテキストでは、時間間隔とは、分析の時間間隔を指します。例えば、観測が月次であるが、分析の時間間隔が四半期で

あるものとします。「最新」を指定し、「時間間隔数 (Number of time intervals)」に値 24 を指定すると、最新の 24 四半期を指定したことになります。

オプションとして、特定の数の時間間隔を除外できます。例えば、最新の 24 個の時間間隔を指定し、除外する数として 1 を指定すると、推定期間は、最後の間隔の前の 24 個の間隔で構成されます。

予測

「レコードの将来への拡張」のオプションでは、推定期間の終了後に予測する時間間隔数を設定します。この場合の時間間隔は分析の時間間隔であり、「データ仕様 (Data Specifications)」タブで指定されます。予測が要求されると、対象でもないすべての入力系列に対して、自己回帰モデルが自動的に作成されます。その後、これらのモデルを使用して、予測期間における当該入力系列の値が生成されます。

保存

宛先オプション

データの変換 (集計や欠損値の代入など) と新しい変数 (「対象の保存」設定で指定されたもの) の両方を IBM SPSS Statistics データ ファイルまたは現行セッション内の新規データ セットに保存できます。保存データ内の日付と時刻の値は、各時間間隔の最初 (月初など) に合わせられ、モデル システムの分析の時間間隔を表します。観測が日付指定またはレコードの順序で定義されていて、データが集計されていない場合のみ、任意の新しい変数をアクティブなデータ セットに保存できます。

対象の保存

モデル予測、信頼区間、および残差を新しい変数として保存できます。保存するように指定した対象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。予測値、信頼区間、およびノイズ残差の場合、新しい変数の接頭辞として使用するルート名を指定できます。完全な変数名は、ルート名と、対象系列が含まれているフィールドの名前を連結したものです。ルート名は、有効な変数名の規則に従っている必要があります。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。

予測値を含むケースを示す

レコードに予測データが含まれているかどうかを示す変数を作成します。変数名を指定することもできます。デフォルトは *ForecastIndicator* です。

保存する対象

新しい変数をモデル システム内のすべての対象系列に対して作成するのか、「表示する系列」設定で指定された対象系列に対してのみ作成するのかを指定します。

モデル システムのエクスポート(X)

モデル システムを圧縮ファイル アーカイブ (.zip ファイル) に保存します。モデル システム ファイルは、時間的因果モデル予測のプロシージャで、更新された予測を取得するため、または使用可能な出力のいずれかを生成するために使用できます。また、時間的因果モデル シナリオのプロシージャで、シナリオ分析を実行するためにも使用できます。

インタラクティブ出力

時間的因果モデリングの出力には、多数のインタラクティブ出力オブジェクトが含まれます。インタラクティブ機能を使用するには、出力ビューアーでオブジェクトをアクティブ化 (ダブルクリック) します。

全体のモデル システム(O)

モデル システム内の系列間の因果関係を表示します。同じ特定のターゲットをその入力に接続する線には、すべて同じ色が使用されます。線の太さは因果関係の有意性を示しており、線が太いほど有意性のより高い接続を表します。ターゲットでもない入力は黒い四角形で示されます。

- 上位のモデル、特定の系列、すべての系列、または入力がないモデルの関係を表示できます。上位のモデルとは、「表示する系列」設定で最適合モデルに指定された基準を満たすモデルのことです。
- 図表内の系列名を選択して右クリックし、コンテキスト メニューから「影響図の作成」を選択することで、1 つ以上の系列の影響図を生成できます。
- 特定の値より大きい有意水準の因果関係を非表示にするように選択できます。有意水準が小さければ小さいほど、因果関係が有意であることを示します。
- 図表内の系列名を選択して右クリックし、次にコンテキスト メニューから「系列の関係を強調表示」を選択することで、特定の系列の関係を表示できます。

影響図(D)

対象系列とそれが影響を与える他の系列またはそれに影響する他の系列との間の因果関係のグラフィカル表現を表示します。対象系列に影響する系列は、原因 と呼びます。

- 対象とする系列の名前を指定することで、対象系列を変更できます。影響図内の任意のノードをダブルクリックすると、そのノードに関連した系列に対象系列が変更されます。
- 原因と効果の間で表示を切り替えることができ、また表示する原因または効果のレベル数を変更できます。
- 任意のノードをシングル クリックすると、そのノードに関連した系列の詳細系列ダイアグラムが開きます。

外れ値の根本原因分析(O)

対象系列の各外れ値の原因である可能性が高い系列を判別します。

- 「外れ値」テーブルで外れ値の行を選択することで、任意の外れ値の根本原因を表示できます。時系列グラフ内の外れ値のアイコンをクリックすることでも、根本原因を表示できます。
- 任意のノードをシングル クリックすると、そのノードに関連した系列の詳細系列ダイアグラムが開きます。

全体のモデル品質

すべてのモデルのモデルの適合度および特定の適合度の統計量のヒストグラム。棒グラフ内の棒をクリックすると、ドット プロットがフィルタに掛けられ、選択した棒に関連したモデルのみが表示されます。ドット プロット内の特定の対象系列のモデルを見つけるには、系列名を指定します。

外れ値分布

推定期間内の時間間隔ごとの、すべての対象における、外れ値の数の棒グラフ。棒グラフ内の棒をクリックすると、ドット プロットがフィルタに掛けられ、選択した棒に関連した外れ値のみが表示されます。

注:

- 時間的因果モデリングからのインタラクティブ出力を含んだ出力ドキュメントを保存する場合に、インタラクティブ機能を保持するのであれば、「出力を名前を付けて保存」ダイアログで「必要なモデル情報を出力ドキュメントとともに保管」が選択されていることを確認してください。
- 一部のインタラクティブ機能では、アクティブなデータセットは時間的因果モデル システムの構築に使用されたデータでなければなりません。

第 7 章 時間的因果モデルの適用

時間的因果モデルの適用

時間的因果モデルのプロシージャで作成されたモデルを適用するために使用可能な 2 つのプロシージャがあります。両方のプロシージャともに、モデル システム ファイルを必要とします。このファイルは、時間的因果モデルのプロシージャの一部として保存できます。

時間的因果モデル予測

このプロシージャを使用して、モデルを再構築することなく、より新しいデータが使用可能な系列の予測を取得できます。また、時間的因果モデリングのプロシージャで使用可能な任意の出力を生成できます。

時間的因果モデルのシナリオ

このプロシージャは、モデル システムにおいて特定の時系列の特定の値が、因果関係がある時系列の予測値にどのように影響するかを調査する場合に使用します。

時間的因果モデル予測

時間的因果モデル予測のプロシージャでは、時間的因果モデリングのプロシージャで作成したモデル システム ファイルをロードし、モデルをアクティブなデータ セットに適用します。このプロシージャを使用して、モデルを再構築することなく、より新しいデータが使用可能な系列の予測を取得できます。また、時間的因果モデリングのプロシージャで使用可能な任意の出力を生成できます。

仮定

- 列ベースまたは多次元のいずれかのアクティブなデータ セット内のデータの構造は、モデル システムの構築時に使用されたのと同じ構造でなければなりません。多次元データの場合、次元フィールドは、モデル システムの構築に使用されたのと同じものでなければなりません。また、モデル システムの構築に使用された次元値が、アクティブなデータ セット内に存在していなければなりません。
- モデルは、モデル システム内で指定されたフィールドと同じ名前のアクティブなデータ セット内のフィールドに適用されます。
- モデル システムの構築時に観測を定義したフィールドが、アクティブなデータ セット内に存在していなければなりません。観測間の時間間隔は、モデルの構築時と同じものであると想定されます。観測が日付の指定によって定義された場合、同じ日付の指定がアクティブなデータ セット内に存在していなければなりません。日付の指定は、「日付の定義」ダイアログまたは DATE コマンドで作成されます。
- 分析の時間間隔、および集計、分配、欠損値のすべての設定は、モデルの構築時と同じものでなければなりません。

時間的因果モデル予測を使用するには

この機能を使用するには Statistics Forecasting オプションが必要です。

メニューから次の項目を選択します。

「分析」 > 「時系列」 > 「時間的因果モデルの適用...」

1. モデル システム ファイルのファイル指定を入力するか、「参照」をクリックしてモデル システム ファイルを選択します。モデル システム ファイルは、時間的因果モデリングのプロシージャで作成されます。
2. モデルの再度推定、予測の作成、出力の生成のためのオプションをクリックします。
3. 「続行」をクリックします。
4. 既存のモデル パラメータを使用するのか、アクティブなデータ セットのデータからモデル パラメータを再度推定するのかを指定します。
5. いつまで予測するのかを指定するか、予測しないことを指定します。
6. 「オプション」をクリックして、出力の内容を指定します。
7. 「保存」をクリックして予測を保存し、モデル パラメータを再度推定した場合は、更新したモデル システムを外部ファイルにエクスポートします。
8. 「実行」をクリックしてプロシージャを実行します。

モデルのパラメータおよび予測

モデル ファイルを読み込み(M)

モデル システム ファイルのモデルのパラメータおよびアクティブなデータ セットのデータを使用して、モデルのパラメータを再度推定することなく、予測は生成されます。出力で表示され、最適モデルの選択に使用される適合度指標は、モデル システム ファイルから取得されます。その後、適合度指標は、各モデルの作成 (または最終更新) 時に使用されたデータを反映します。このオプションは、モデル システムの構築に使用されたデータから予測および出力を生成する場合に適しています。

データから再度推定(D)

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。モデルのパラメータの再度推定は、各対象についてモデルに含める入力に影響しません。このオプションは、元の推定期間範囲外の新しいデータがあり、更新されたデータを使用して予測または他の出力を生成する場合に適しています。

すべての観測(A)

推定期間がすべての系列における最も早い観測の時刻から始まり、最も遅い観測の時刻に終わることを指定します。

開始時刻と終了時刻で指定(B)

推定期間の開始と終了の両方を指定するか、開始または終了だけを指定することができます。推定期間の開始または終了を省略した場合、デフォルト値が使用されます。

- 観測が日付と時刻フィールドで定義されている場合、日付と時刻フィールドで使用されているのと同じ形式で、開始および終了の値を入力します。
- 周期的な期間で定義された観測の場合、周期的な期間フィールドのそれぞれの値を指定します。各フィールドは個別の列で表示されます。
- アクティブなデータ セットに有効な日付指定がある場合、日付指定の各構成要素 (月など) の値を指定する必要があります。各構成要素は個別の列で表示されます。
- 観測がレコード順序で定義されている場合、推定期間の開始および終了は、関連ケースの行番号 (データ エディタに表示) で定義されます。

最新の時間間隔または最も早い時間間隔で指定 (By latest or earliest time intervals)

データにおける最も早い時間間隔から開始するか、最新の時間間隔で終了する特定の数の時間間隔 (およびオプションとしてオフセット) として推定期間を定義します。このコンテ

キストでは、時間間隔とは、分析の時間間隔を指します。例えば、観測が月次であるが、分析の時間間隔が四半期であるものとします。「最新」を指定し、「時間間隔数 (Number of time intervals)」に値 24 を指定すると、最新の 24 四半期を指定したことになります。

オプションとして、特定の数の時間間隔を除外できます。例えば、最新の 24 個の時間間隔を指定し、除外する数として 1 を指定すると、推定期間は、最後の間隔の前の 24 個の間隔で構成されます。

レコードの将来への拡張(X)

推定期間の終わりを超えて予測する時間間隔数を設定します。この場合の時間間隔は分析の時間間隔です。予測が要求されると、対象でもないすべての入力系列に対して、自己回帰モデルが自動的に作成されます。その後、これらのモデルを使用して、当該入力の対象の予測を取得するために、予測期間における当該入力系列の値が生成されます。

一般オプション

信頼区間の幅 (%) (C)

この設定は、予測とモデルのパラメータの両方の信頼区間を制御します。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。

外れ値のしきい値 (%) (T)

観測値には、モデルから計算された、それが外れ値である確率がこのしきい値を超えた場合、外れ値のフラグが立てられます。50 から 100 の範囲の値を指定することができます。

表示する系列

これらのオプションでは、出力を表示する系列 (対象または入力) を指定します。指定した系列の出力の内容は、「出力オプション」設定によって決定されます。

最適モデルに関連付けられた対象の表示

デフォルトでは、R2 乗の値によって決定される、上位 10 個の最適モデルに関連した対象の出力が表示されます。最適モデルの別の固定数を指定したり、最適モデルのパーセントを指定したりすることができます。また、以下の適合度を選択できます。

R2 乗 線型モデルの適合度。決定係数とも呼びます。ターゲット変数の変動のうち、モデルによって説明される割合です。値の範囲は 0 から 1 までです。値が小さい場合は、モデルが十分にデータに適合していないことを示します。

平均平方根パーセント誤差

モデル予測値が系列の観測値とどれほど異なるかを示す指標。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。

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。

BIC ベイズ情報量基準。-2 対数尤度に基づいてモデルを選択し、比較するための指標。値が小さいほどモデルが良好であることを示します。BIC もパラメータが過多のモデル (例えば、大量の入力がある複雑なモデル) にペナルティを科しますが、AIC よりも厳密にそれを行います。

AIC 赤池情報量基準。-2 対数尤度に基づいてモデルを選択し、比較するための指標。値が小さいほどモデルが良好であることを示します。AIC は、パラメータが過多のモデル (例えば、大量の入力がある複雑なモデル) にペナルティを科します。

個別の系列の指定

出力する個別の系列を指定できます。

- 列ベースのデータの場合、必要な系列が含まれているフィールドを指定します。指定したフィールドの順序によって、出力に表示される順序が決定されます。
- 多次元データの場合、系列が含まれている測定基準フィールドのグリッドに項目を追加することで、特定の系列を指定します。その後、系列を定義する次元フィールドの値を指定します。
 - グリッドに直接各次元フィールドの値を入力するか、使用可能な次元値のリストから選択することができます。使用可能な次元値のリストから選択するには、必要な次元のセルの省略符号ボタンをクリックします。このアクションにより、「次元の値の選択」サブダイアログが開きます。
 - 「次元の値の選択」サブダイアログで、双眼鏡アイコンをクリックして検索語を指定することで、次元値のリストを検索できます。スペースは、検索語の一部として扱われます。検索語内のアスタリスク (*) は、ワイルドカード文字を示しません。
 - グリッド内の系列の順序によって、出力に表示される順序が決定されます。

列ベースのデータと多次元データの両方で、出力は 30 系列までに制限されています。この制限には、指定した個別の系列 (入力または対象) および最適モデルに関連した対象が含まれます。個別に指定された系列は、最適モデルに関連した対象よりも優先されます。

出力オプション

これらのオプションは、出力の内容を指定します。「**対象の出力**」グループのオプションは、「**表示する系列**」設定の最適モデルに関連した対象の出力を生成します。「**系列の出力**」グループのオプションは、「**表示する系列**」設定で指定された個別系列の出力を生成します。

全体のモデル システム(O)

モデル システム内の系列間の因果関係のグラフィカル表現を表示します。表示対象のモデル適合統計量と外れ値の両方のテーブルが、出力項目の一部として含まれます。「**系列の出力**」グループでこのオプションを選択した場合、「**表示する系列**」設定で指定されている個別系列ごとに別個の出力項目が作成されます。

系列間の因果関係には、関連する有意水準が含まれており、有意水準が小さいほど、より有意な関連を示しています。特定の値より大きい有意水準の関係を非表示にするように選択できます。

モデルの適合度の統計量と外れ値

表示するように選択した対象系列のモデル適合統計量と外れ値のテーブル。これらのテーブルには、全体のモデル システムの視覚化のテーブルと同じ情報が含まれます。これらのテーブルでは、テーブルのピボットおよび編集を行うための標準機能がすべてサポートされます。

モデル効果とモデルのパラメータ

表示するように選択した対象系列のモデル効果検定とモデル パラメータのテーブル。モデル効果検定には、モデルに含まれている各入力の F 統計量と関連する有意確率値が含まれます。

影響図(D)

対象系列とそれが影響を与える他の系列またはそれに影響する他の系列との間の因果関係のグラフィカル表現を表示します。対象系列に影響する系列は、**原因** と呼びます。「**効果**」を選択すると、効果を表示するために初期化された影響図が生成されます。「**原因**」を選択すると、原因を表示するために初期化された影響図が生成されます。「**原因と効果の両方**」を選択すると、2 つの別個の影響図 (原因のために初期化された影響図と効果のために初期化された影響図) が生成されます。影響図を表示する出力項目の原因と効果をインタラクティブに切り替えることができます。

表示する原因または効果のレベル数を指定できます。最初のレベルは、対象系列のみです。レベルが増すごとに、対象系列のさらに間接的な原因または効果が表示されます。例えば、効果の表示における 3 番目のレベルは、直接入力として 2 番目のレベルの系列が含まれている系列で構成されます。この場合、対象系列が 2 番目のレベルの系列に対する直接入力であるため、3 番目のレベルの系列は、間接的に対象系列の影響を受けます。

系列プロット

表示するように選択した対象系列の観測値と予測値の作図。予測が要求されると、作図には、予測値と予測の信頼区間も表示されます。

残差プロット

表示するように選択された対象系列のモデル残差の作図。

上位入力

時系列に沿った各表示対象、および対象の上位 3 位までの入力の作図。上位の入力とは、有意確率値が小さい入力です。入力と対象に別のスケールを表示するために、y 軸は、各系列の Z 得点を表します。

予測テーブル(F)

表示するように選択された対象系列の予測値と予測の信頼区間のテーブル。

外れ値の根本原因分析(O)

対象系列の各外れ値の原因である可能性が高い系列を判別します。外れ値の根本原因分析は、「表示する系列」設定の個別系列のリストに含まれている対象系列ごとに行われます。

出力

対話式の外れ値のテーブルと図表

各対象系列の外れ値とその外れ値の根本原因のテーブルと図表。テーブルには、外れ値ごとに 1 つの行が含まれています。図表は、影響図です。テーブル内の行を選択すると、影響図内で、対象系列から関連する外れ値の原因である可能性が高い系列までのパスが強調表示されます。

外れ値のピボット テーブル

各対象系列の外れ値とその外れ値の根本原因のテーブル。このテーブルには、インタラクティブ表示のテーブルと同じ情報が含まれます。このテーブルでは、テーブルのピボットおよび編集を行うための標準機能がすべてサポートされます。

因果関係レベル

根本原因の検索で含めるレベル数を指定できます。ここで使用されるレベルの概念は、影響図で説明したものと同じです。

すべてのモデル間でのモデルの適合度

すべてのモデルのモデルの適合度および選択した適合度の統計量のヒストグラム。以下の適合度の統計量を使用することができます。

R2 乗 線型モデルの適合度。決定係数とも呼びます。ターゲット変数の変動のうち、モデルによって説明される割合です。値の範囲は 0 から 1 までです。値が小さい場合は、モデルが十分にデータに適合していないことを示します。

平均平方根パーセント誤差

モデル予測値が系列の観測値とどれほど異なるかを示す指標。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。

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。

BIC ベイズ情報量基準。-2 対数尤度に基づいてモデルを選択し、比較するための指標。値が小さいほどモデルが良好であることを示します。BIC もパラメータが過多のモデル (例えば、大量の入力がある複雑なモデル) にペナルティを科しますが、AIC よりも厳密にそれを行います。

AIC 赤池情報量基準。-2 対数尤度に基づいてモデルを選択し、比較するための指標。値が小さいほどモデルが良好であることを示します。AIC は、パラメータが過多のモデル (例えば、大量の入力がある複雑なモデル) にペナルティを科します。

長期間の外れ値(V)

推定期間内の時間間隔ごとの、すべての対象における、外れ値の数の棒グラフ。

系列の変換(T)

モデル システム内の系列に適用された変換のテーブル。可能な変換は、欠損値の代入、集計、および分配です。

保存

対象の保存

モデル予測、信頼区間、および残差を新しい変数として保存できます。保存するように指定した対象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。予測値、信頼区間、およびノイズ残差の場合、新しい変数の接頭辞として使用するルート名を指定できます。完全な変数名は、ルート名と、対象系列が含まれているフィールドの名前を連結したものです。ルート名は、有効な変数名の規則に従っている必要があります。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。

予測値を含むケースを示す

レコードに予測データが含まれているかどうかを示す変数を作成します。変数名を指定することもできます。デフォルトは *ForecastIndicator* です。

保存する対象

新しい変数をモデル システム内のすべての対象系列に対して作成するのか、「表示する系列」設定で指定された対象系列に対してのみ作成するのかを指定します。

宛先オプション

データの変換 (集計や欠損値の代入など) と新しい変数 (「**対象の保存**」設定で指定されたもの) の両方を IBM SPSS Statistics データ ファイルまたは現行セッション内の新規データ セットに保存できます。保存データ内の日付と時刻の値は、各時間間隔の最初 (月初など) に合わせられ、モデル システムの分析の時間間隔を表します。観測が日付指定またはレコードの順序で定義されていて、データが集計されていない場合のみ、任意の新しい変数をアクティブなデータ セットに保存できます。

モデル システムのエクスポート(X)

モデル システムを圧縮ファイル アーカイブ (.zip ファイル) に保存します。モデル システム ファイルは、このプロシージャで再利用できます。また、時間的因果モデル シナリオのプロシージャで、シナリオ分析を実行するためにも使用できます。このオプションが使用可能なのは、モデルのパラメータが再度推定された場合のみです。

時間的因果モデルのシナリオ

時間的因果モデルのシナリオのプロシージャでは、アクティブなデータ セットのデータを使用して、時間的因果モデル システムのユーザー定義のシナリオを実行します。シナリオ は、ルート系列 と呼ばれる時系列と、特定の時間範囲におけるその系列のユーザー定義値セットによって定義されます。その後、指定された値を使用して、ルート系列の影響を受ける時系列の予測が生成されます。このプロシージャでは、時間的因果モデリングのプロシージャで作成したモデル システム ファイルが必要です。アクティブなデータ セットが、モデル システム ファイルの作成に使用されたのと同じデータであるものと想定されます。

例

時間的因果モデリングのプロシージャを使用して、あるビジネスの意思決定者が、多数の重要なパフォーマンス指標に影響する重要な測定基準を検出しました。測定基準は制御可能であるため、意思決定者は、次の四半期における測定基準の各種値セットの効果を調査したいと考えます。モデル システム ファイルを時間的因果モデルのシナリオのプロシージャにロードし、重要な測定基準の値セットを指定することで、この調査を簡単に行えます。

時間的因果モデルのシナリオを実行するには

この機能を使用するには Statistics Forecasting オプションが必要です。

メニューから次の項目を選択します。

「分析」 > 「時系列」 > 「時間的因果モデルの適用...」

1. モデル システム ファイルのファイル指定を入力するか、「参照」をクリックしてモデル システム ファイルを選択します。モデル システム ファイルは、時間的因果モデリングのプロシージャで作成されます。
2. シナリオを実行するオプションをクリックします。
3. 「続行」をクリックします。
4. (「時間的因果モデルのシナリオ」ダイアログの)「シナリオ」タブで、「シナリオ期間の定義」をクリックし、シナリオ期間を指定します。
5. 列ベースのデータの場合、「シナリオの追加」をクリックして、各シナリオを定義します。多次元データの場合、「シナリオの追加」をクリックして個別のシナリオをそれぞれ定義し、「シナリオ グループの追加」をクリックして各シナリオ グループを定義します。
6. 「オプション」をクリックして、出力の内容を指定し、シナリオの影響を受ける系列の幅を指定します。
7. 「実行」をクリックしてプロシージャを実行します。

シナリオ期間の定義

シナリオ期間は、シナリオの実行に使用される値を指定した期間です。推定期間の終わりの前または後に開始できます。オプションとして、シナリオ期間の終わりを超えて予測するように指定できます。デフォルトでは、シナリオ期間の終わりまでの予測が生成されます。すべてのシナリオでは、同じシナリオ期間といつまで予測するのかの仕様を使用します。

注: 予測はシナリオ期間が開始した後の最初の期間に開始されます。例えば、シナリオ期間が 2014-11-01 に開始され、時間間隔が月の場合、最初の予測の対象は 2014-12-01 です。

開始時刻、終了時刻、および指定による予測の時刻で指定(C)

- 観測が日付と時刻フィールドで定義されている場合、日付と時刻フィールドで使用されているのと同じ形式で、開始、終了、指定による予測の値を入力します。日付と時刻フィールドの値は、関連する時間間隔の先頭に合わせられます。例えば、分析の時間間隔が月の場合、値 10/10/2014 は、月初めである 10/01/2014 に調整されます。
- 周期的な期間で定義された観測の場合、周期的な期間フィールドのそれぞれの値を指定します。各フィールドは個別の列で表示されます。
- アクティブなデータ セットに有効な日付指定がある場合、日付指定の各構成要素 (月など) の値を指定する必要があります。各構成要素は個別の列で表示されます。
- 観測がレコード順序で定義されている場合、開始、終了、および指定による予測は、関連ケースの行番号 (データ エディタに表示) で定義されます。

推定期間の終了からの相対的時間間隔で指定(F)

推定期間の終わりに相対的な時間間隔数で開始および終了を定義します。ここで、時間間隔とは、分析の時間間隔です。推定期間の終わりは、時間間隔 0 と定義されます。推定期間の終わりより前の時間間隔は負の値で、推定期間の終わりより後の間隔は正の値になります。また、シナリオ期間の終わりを超えて予測する間隔数も指定できます。デフォルトは 0 です。

例えば、分析の時間間隔が月で、開始間隔に 1、終了間隔に 3、およびそれを超えていつまで予測するかに 1 を指定したものとします。この場合、シナリオ期間は、推定期間の終わりに続く 3 カ月間になります。予測は、シナリオ期間の 2 カ月目と 3 カ月目、およびシナリオ期間の終わりを超えてさらに 1 カ月分、生成されます。

シナリオおよびシナリオ グループの追加

「シナリオ」タブでは、実行するシナリオを指定します。シナリオを定義するには、まず「シナリオ期間の定義」をクリックしてシナリオ期間を定義する必要があります。シナリオおよびシナリオ グループ (多次元データの場合にのみ適用) を作成するには、関連する「シナリオの追加」ボタンまたは「シナリオ グループの追加」ボタンをクリックします。関連するグリッドで特定のシナリオまたはシナリオ グループを選択することで、そのシナリオまたはシナリオ グループの編集、コピー、または削除を行うことができます。

列ベースのデータ

グリッド内の「ルート フィールド」列では、値をシナリオ値に置き換える時系列フィールドを指定します。「シナリオ値」列には、指定したシナリオ値が古いもの順に表示されます。シナリオ値が式で定義されている場合、この列には式が表示されます。

多次元データ

個別のシナリオ

「個別シナリオ」グリッド内の各行では、指定したシナリオ値で値が置き換えられる時系列を指定します。系列は、「ルート メトリック」列で指定されているフィールドと各次元フィールドの指定値の組み合わせによって定義されます。「シナリオ値」列の内容は、列ベースのデータのものと同じです。

シナリオ グループ

シナリオ グループ は、単一のルート測定基準フィールドと複数の次元値セットに基づいたシナリオ セットを定義します。指定した測定基準フィールドについて、各次元値セット (次元フィールドごとに 1 つの値) により、時系列が定義されます。その後、このような時系列ごとに、個別のシナリオが生成されます。時系列の値は、シナリオ値で置き換えられます。シナリオ グループのシナリオ値は式によって指定され、その式はグループ内の各時系列に適用されます。

「**系列の数**」列には、シナリオ グループに関連した次元値セット数が表示されます。表示されている値は、シナリオ グループに関連した時系列の実際の数 (セットごとに 1 系列) よりも大きい場合があります。この状態が発生するのは、指定した組み合わせの次元値の一部が、グループのルート測定基準に含まれている系列に対応していない場合です。

シナリオ グループの例として、測定基準フィールド *advertising* (広告) および 2 つの次元フィールド *region* (地域) と *brand* (ブランド) を考えます。ルート測定基準 *advertising* に基づき、*region* と *brand* のすべての組み合わせが含まれたシナリオ グループを定義できます。この場合、*advertising* フィールドに関連付けられている各時系列について *advertising* を 20 % 増やした場合の効果を調査するために、式として *advertising*1.2* を指定できます。*region* の値が 4 つあり、*brand* の値が 2 つある場合、時系列は 8 つになるため、このグループでは 8 つのシナリオが定義されます。

シナリオ定義

シナリオを定義する設定は、データが列ベースなのか多次元なのかによって異なります。

ルート系列

シナリオのルート系列を指定します。各シナリオは、単一のルート系列に基づきます。列ベースのデータの場合、ルート系列を定義するフィールドを選択します。多次元データの場合、ルート系列が含まれている測定基準フィールドのグリッドに項目を追加することで、ルート系列を指定します。その後、ルート系列を定義する次元フィールドの値を指定します。以下は、次元値の指定に適用されます。

- グリッドに直接各次元フィールドの値を入力するか、使用可能な次元値のリストから選択することができます。使用可能な次元値のリストから選択するには、必要な次元のセルの省略符号ボタンをクリックします。このアクションにより、「次元の値の選択」サブダイアログが開きます。
- 「次元の値の選択」サブダイアログで、双眼鏡アイコンをクリックして検索語を指定することで、次元値のリストを検索できます。スペースは、検索語の一部として扱われます。検索語内のアスタリスク (*) は、ワイルドカード文字を示しません。

影響を受ける対象の指定(C)

このオプションは、ルート系列の影響を受ける特定の対象が分かっている、その対象に対する効果のみを調査する場合に使用します。デフォルトでは、ルート系列の影響を受ける対象は、自動的に判別されます。「オプション」タブの設定を使用して、シナリオの影響を受ける系列の幅を指定できます。

列ベースのデータの場合、必要な対象を選択します。多次元データの場合、対象系列が含まれている対象測定基準フィールドのグリッドに項目を追加することで、対象系列を指定します。デフォルトでは、指定した測定基準フィールドに含まれているすべての系列が含まれます。1 つ以上の次元フィールドの含める値をカスタマイズすることで、含める系列セットをカスタマイズできます。含まれる次元値をカスタマイズするには、必要な次元の省略符号ボタンをクリックします。このアクションにより、「次元の値の選択」ダイアログが開きます。

「**系列の数**」列 (多次元データの場合) には、関連する対象測定基準について現在指定されている次元値セット数が表示されます。表示されている値は、影響を受ける対象系列の実際の数 (セットごとに 1 系列) よりも大きい場合があります。この状態が発生するのは、指定した組み合わせの次元値の一部が、関連する対象測定基準に含まれている系列に対応していない場合です。

シナリオ ID

各シナリオに固有の ID が必要です。この ID は、シナリオに関連した出力に表示されます。ID の値には、固有性以外の制限はありません。

ルート系列のシナリオ値の指定

このオプションは、シナリオ期間のルート系列に明示的な値を指定する場合に使用します。グリッドにリストされている時間間隔ごとに数値を指定する必要があります。シナリオ期間内の各間隔のルート系列の値 (実際の値または予測値) を取得するには、「読み取り」、「予測」、または「読み取り/予測」をクリックします。

ルート系列のシナリオ値の式を指定

シナリオ期間のルート系列の値を計算するための式を定義できます。式を直接入力するか、電卓ボタンをクリックし、シナリオ値の式ビルダーで式を作成できます。

- 式にはモデル システムの任意の対象または入力を含めることができます。
- シナリオ期間が既存データを超えて延長された場合、式は、式内のフィールドの予測値に適用されます。
- 多次元データの場合、式内の各フィールドでは、当該フィールドと、ルート測定基準で指定された次元値によって定義される時系列を指定します。これらの時系列を使用して式を評価します。

例として、ルート フィールドが *advertising* (広告) であり、式が $advertising * 1.2$ であるものとします。シナリオで使用されている *advertising* の値は、既存の値と比較して 20 % 増加します。

注: シナリオを作成するには、「シナリオ」タブで「シナリオの追加」をクリックします。

シナリオ値の式ビルダー: シナリオ値の式ビルダーは、単一のシナリオまたはシナリオ グループのシナリオ値の計算に使用される式を作成する場合に使用します。式を作成するには、「式」フィールドに成分を貼り付けるか、「式」フィールドに直接入力します。

- 式にはモデル システムの任意の対象または入力を含めることができます。
- 関数を貼り付けるには、「関数グループ」リストからグループを選択し、「関数」リストで関数をダブルクリックします (または、関数を選択し、「関数グループ」リストの隣にある矢印をクリックします)。疑問符で示されたパラメーターを入力します。「すべて」というラベルの付いた関数グループには、使用可能なすべての関数がリストされます。現在選択されている関数の短い説明が、ダイアログ・ボックスの予約領域に表示されます。
- 文字定数は、引用符で囲む必要があります。
- 値に小数が含まれている場合、小数点にはピリオド (.) を使用する必要があります。

シナリオ値の式ビルダーにアクセスするには、「シナリオ定義」ダイアログまたは「シナリオ グループ定義」ダイアログで電卓ボタンをクリックします。

次元の値の選択: 多次元データの場合、シナリオまたはシナリオ グループの影響を受ける対象を定義する次元値をカスタマイズできます。シナリオ グループのルート系列セットを定義する次元値をカスタマイズすることもできます。

すべての値

現在の次元フィールドのすべての値を含めることを指定します。これはデフォルトのオプションです。

値の選択

このオプションは、現在の次元フィールドの値セットを指定する場合に使用します。選択元の値セットをフィルタに掛けることができます。フィルタ条件を満たしている値が「一致」タブに表示され、フィルタ条件を満たさない値は「選択されていない値」リストの「不一致」タブに表示されます。「すべて」タブには、フィルタ条件に関係なく、選択されていないすべての値がリストされます。

- フィルタの指定時にワイルドカード文字を指定するために、アスタリスク (*) を使用できます。

- 現在のフィルタをクリアするには、「表示された値のフィルタ」ダイアログの検索語として空の値を指定します。

影響を受ける対象の次元値をカスタマイズするには、以下のようにします。

1. 「シナリオ定義」または「シナリオ グループ定義」ダイアログで、次元値をカスタマイズする対象測定基準を選択します。
2. カスタマイズする次元の列にある省略符号ボタンをクリックします。

シナリオ グループのルート系列の次元値をカスタマイズするには、以下のようにします。

1. 「シナリオ グループ定義」ダイアログで、カスタマイズする次元の (ルート系列グリッド内にある) 省略符号ボタンをクリックします。

シナリオ グループ定義

ルート系列

シナリオ グループのルート系列セットを指定します。セット内の時系列ごとに個別のシナリオが生成されます。ルート系列を指定するには、必要な系列が含まれている測定基準フィールドのグリッドに項目を追加します。その後、セットを定義する次元フィールドの値を指定します。デフォルトでは、指定したルート測定基準フィールドに含まれているすべての系列が含まれます。1 つ以上の次元フィールドの含める値をカスタマイズすることで、含める系列セットをカスタマイズできます。含まれる次元値をカスタマイズするには、次元の省略符号ボタンをクリックします。このアクションにより、「次元の値の選択」ダイアログが開きます。

「系列の数」列には、関連するルート測定基準について現在含まれている次元値セット数が表示されます。表示されている値は、シナリオ グループのルート系列の実際の数 (セットごとに 1 系列) よりも大きい場合があります。この状態が発生するのは、指定した組み合わせの次元値の一部が、ルート測定基準に含まれている系列に対応していない場合です。

影響を受ける対象系列の指定

このオプションは、ルート系列セットの影響を受ける特定の対象が分かっている、その対象に対する効果のみを調査する場合に使用します。デフォルトでは、各ルート系列の影響を受ける対象は、自動的に判別されます。「オプション」タブの設定を使用して、各個別シナリオの影響を受ける系列の幅を指定できます。

対象系列を指定するには、その系列が含まれている測定基準フィールドのグリッドに項目を追加します。デフォルトでは、指定した測定基準フィールドに含まれているすべての系列が含まれます。1 つ以上の次元フィールドの含める値をカスタマイズすることで、含める系列セットをカスタマイズできます。含まれる次元値をカスタマイズするには、必要な次元の省略符号ボタンをクリックします。このアクションにより、「次元の値の選択」ダイアログが開きます。

「系列の数」列には、関連する対象測定基準について現在指定されている次元値セット数が表示されます。表示されている値は、影響を受ける対象系列の実際の数 (セットごとに 1 系列) よりも大きい場合があります。この状態が発生するのは、指定した組み合わせの次元値の一部が、関連する対象測定基準に含まれている系列に対応していない場合です。

シナリオ ID 接頭辞

シナリオ グループの接頭辞は一意である必要があります。接頭辞を使用して、シナリオ グループ内の各個別シナリオに関連付けられた出力に表示する識別子を構成します。個別シナリオの識別子は、接頭辞の後に下線を付け、その後にルート系列を識別する各次元フィールドの値を続けたものです。複数の次元値は、下線で区切られます。接頭辞の値には、固有性以外の制限はありません。

ルート系列のシナリオ値の式

シナリオ グループのシナリオ値は式によって指定され、その式を使用してグループ内の各ルート系列の値を計算します。式を直接入力するか、電卓ボタンをクリックし、シナリオ値の式ビルダーで式を作成できます。

- 式にはモデル システムの任意の対象または入力を含めることができます。
- シナリオ期間が既存データを超えて延長された場合、式は、式内のフィールドの予測値に適用されます。
- グループ内の各ルート系列について、式内のフィールドでは、当該フィールドおよびルート系列を定義する次元値によって定義された時系列を指定します。これらの時系列を使用して式を評価します。例えば、ルート系列が `region='north'` および `brand='X'` によって定義される場合、式で使用される時系列は、同じ次元値によって定義されます。

例として、ルート測定基準フィールドが *advertising* (広告) であり、2 つの次元フィールド *region* (地域) と *brand* (ブランド) があるものとします。また、シナリオ グループに次元フィールド値のすべての組み合わせが含まれているものとします。この場合、*advertising* フィールドに関連付けられている各時系列について *advertising* を 20 % 増やした場合の効果を調査するために、式として `advertising*1.2` を指定できます。

注: シナリオ グループは多次元データにのみ適用され、「シナリオ」タブの「シナリオ グループの追加」をクリックすることで作成します。

オプション

影響を受ける対象の最大レベル(L)

影響を受ける対象のレベルの最大数を指定します。レベルを辿るごとに (最大で 5 まで)、より間接的にルート系列の影響を受ける対象が含まれます。具体的には、1 番目のレベルには、直接入力としてルート系列を持つ対象が含まれます。2 番目のレベルの対象には、直接入力として 1 番目のレベルの対象が含まれます。3 番目以降も同様です。この設定の値を大きくすると、計算が複雑化するため、パフォーマンスに影響が出る可能性があります。

自動検出される対象の最大数

ルート系列ごとに自動的に検出される、影響を受ける対象の最大数を指定します。この設定の値を大きくすると、計算が複雑化するため、パフォーマンスに影響が出る可能性があります。

影響図(D)

各シナリオのルート系列とそれが影響を与える対象系列間の因果関係のグラフィカル表現を表示します。影響を受ける対象のシナリオ値と予測値の両方のテーブルが、出力項目の一部として含まれます。グラフには、影響を受ける対象の予測値の作図が含まれます。影響図内の任意のノードをシングル クリックすると、そのノードに関連した系列の詳細系列ダイアグラムが開きます。シナリオごとに別個の影響図が生成されます。

系列図(E)

シナリオごとに影響を受ける各対象の予測値の系列図を生成します。

予測とシナリオのテーブル(F)

予測値と各シナリオのシナリオ値のテーブル。これらのテーブルには、影響図内のテーブルと同じ情報が含まれます。これらのテーブルでは、テーブルのピボットおよび編集を行うための標準機能がすべてサポートされます。

プロットとテーブルに信頼区間を含める(N)

シナリオ予測の信頼区間を図表とテーブルの両方の出力に含めるかどうかを指定します。

信頼区間の幅 (%) (C)

この設定は、シナリオ予測の信頼区間を制御します。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。

第 8 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 9 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 10 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、その値は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 3. ARIMA(0,0,1), $q>0$

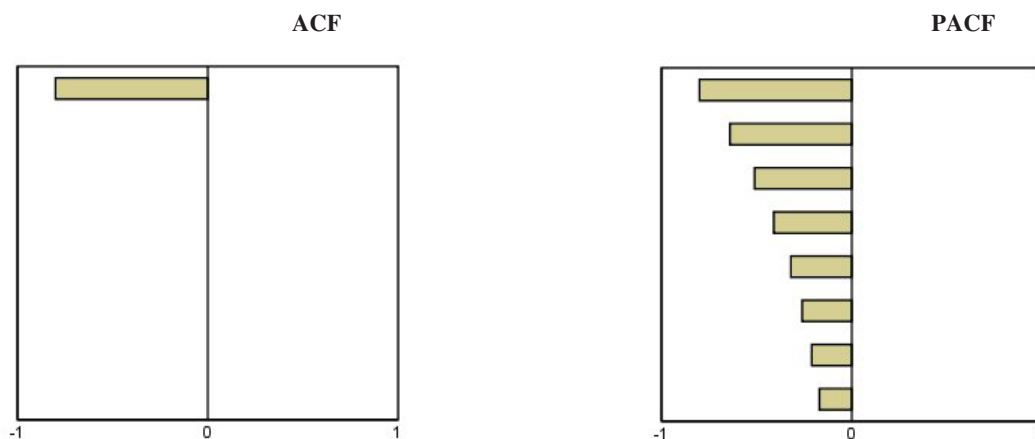
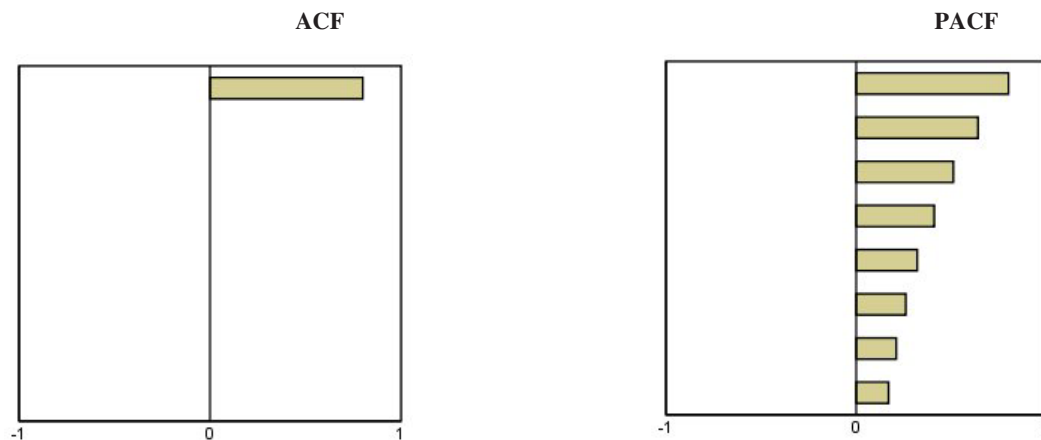


表 4. $ARIMA(0,0,1), q < 0$



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

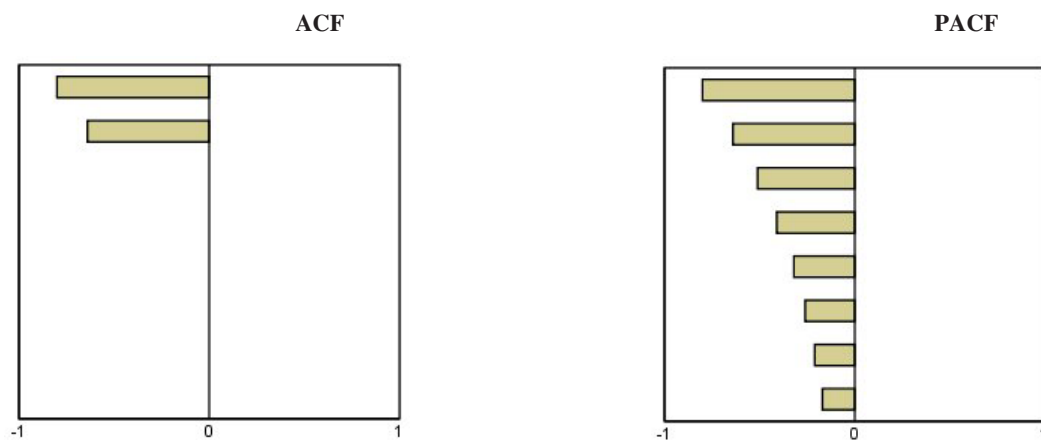


表 5. $ARIMA(1,0,0), f > 0$

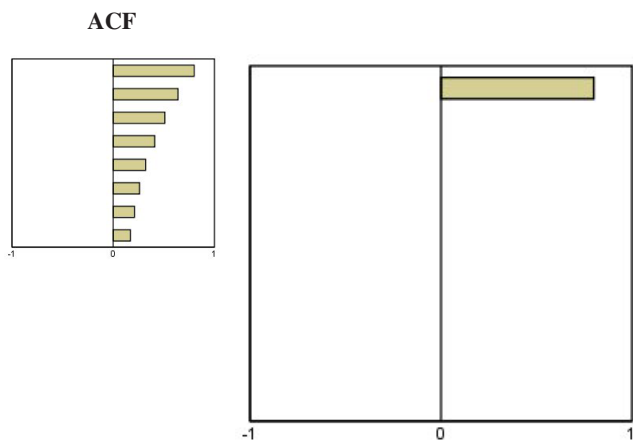
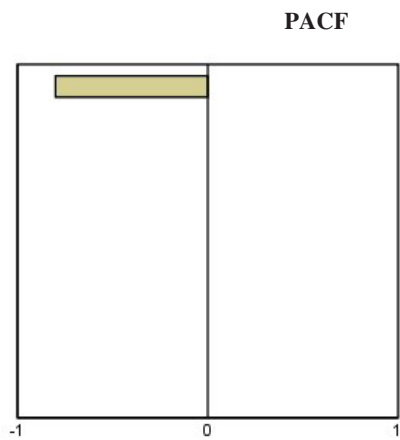
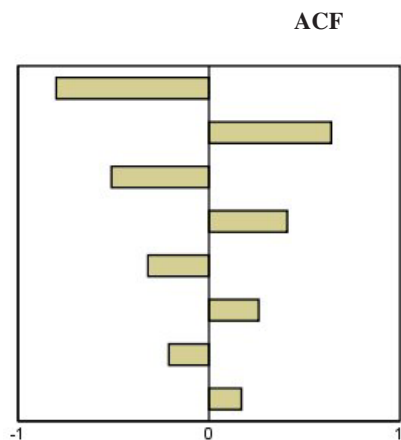
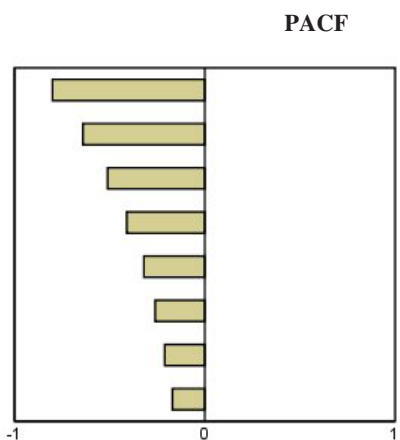
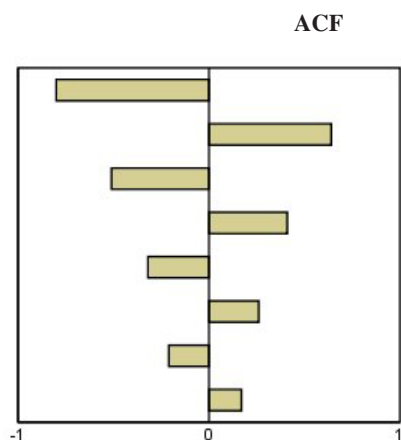


表 6. ARIMA(1,0,0), $f < 0$



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



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

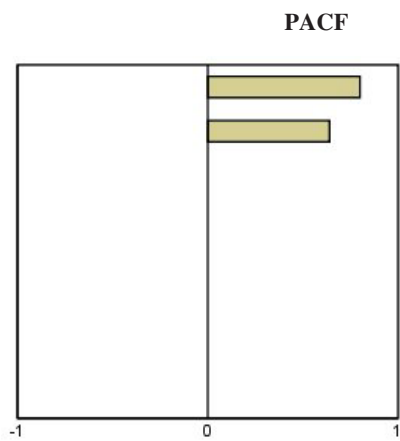
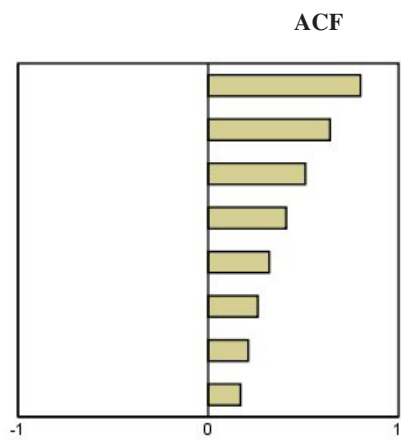
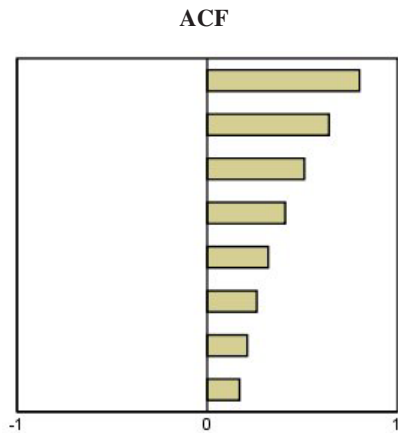


表 7. $ARIMA(0,1,0)$ (和分系列)



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。

著作権使用許諾:

本書には、様々なオペレーティング・プラットフォームでのプログラミング手法を例示するサンプル・アプリケーション・プログラムがソース言語で掲載されています。お客様は、サンプル・プログラムが書かれているオペレーティング・プラットフォームのアプリケーション・プログラミング・インターフェースに準拠したアプリケーション・プログラムの開発、使用、販売、配布を目的として、いかなる形式においても、IBM に対価を支払うことなくこれを複製し、改変し、配布することができます。このサンプル・プログラムは、あらゆる条件下における完全なテストを経ていません。従って IBM は、これらのサンプル・プログラムについて信頼性、利便性もしくは機能性があることをほのめかしたり、保証することはできません。これらのサンプル・プログラムは特定物として現存するままの状態を提供されるものであり、いかなる保証も提供されません。IBM は、お客様の当該サンプル・プログラムの使用から生ずるいかなる損害に対しても一切の責任を負いません。

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 13
- 時系列モデルの適用 20

イベント 7

- 時系列モデラー 7

エキスパート・モデラー 5

- 外れ値 7
- モデル・スペースの制限 7

[カ行]

外れ値

- エキスパート・モデラー 7
- 定義 59

ARIMA モデル 11

過渡的な外れ値 59

- 時系列モデラー 7, 11

加法の外れ値 59

- 時系列モデラー 7, 11

加法バッチの外れ値 59

- 時系列モデラー 7, 11

技術革新的外れ値 59

- 時系列モデラー 7, 11

季節性加法の外れ値 59

- 時系列モデラー 7, 11

季節性の分解 25, 26

- 移動平均の計算 25

仮定 25

新変数の保存 26

変数の作成 26

モデル 25

局所トレンドの外れ値 59

- 時系列モデラー 7, 11

欠損値

- 時系列モデラー 14
- 時系列モデルの適用 22

検証期間 2

減衰指数平滑法モデル 8

[サ行]

最大絶対誤差 57

- 時系列モデラー 12, 13

最大絶対誤差 (続き)

- 時系列モデルの適用 19, 20

最大絶対パーセント誤差 57

- 時系列モデラー 12, 13

時系列モデルの適用 19, 20

残差

- 時系列モデラー 12, 13

時系列モデルの適用 19, 20

時間的因果モデル 31, 32, 33, 35, 36, 37,

38, 40, 41

時間的因果モデルのシナリオ 49, 50, 51,

52, 53, 54

時間的因果モデル予測 43, 44, 45, 46, 48

時系列分析

- 時間的因果モデル 31

時系列モデラー 5

新しい変数名 14

当てはめ値 13

イベント 7

エキスパート・モデラー 5

外れ値 7, 11

系列変換 8, 9, 10

欠損値 14

最適モデルと最不適モデル 13

残差自己相関関数 12, 13

残差偏自己相関関数 12, 13

指数平滑法 5, 8

周期 7, 8, 9, 10

信頼区間 13, 14

推定期間 5

すべてのモデルにわたる統計量 12, 13

適合度統計量 12, 13

転送関数 10

モデル指定を XML 形式で保存 14

モデル名 14

モデル・パラメーター 12

予測 12, 13

予測期間 5, 14

予測の保存 14

ARIMA 5, 9

Box-Ljung 統計量 12

時系列モデルの適用 17

新しい変数名 21

当てはめ値 20

欠損値 22

最適モデルと最不適モデル 21

再度推定されたモデルを XML 形式で保存 21

残差自己相関関数 19, 20

残差偏自己相関関数 19, 20

時系列モデルの適用 (続き)

信頼区間 20, 22

推定期間 17

すべてのモデルにわたる統計量 19, 20

適合度統計量 19, 20

モデル・パラメーター 19

モデル・パラメーターの再度推定 17

予測 19, 20

予測期間 17

予測の保存 21

Box-Ljung 統計量 19

自己相関関数

時系列モデラー 12, 13

時系列モデルの適用 19, 20

純粋な ARIMA プロセスのプロット 61

指数平滑法モデル 5, 8

自然対数変換

時系列モデラー 8, 9, 10

周期

時系列モデラー 7, 8, 9, 10

信頼区間

時系列モデラー 13, 14

時系列モデルの適用 20, 22

推定期間 2

時系列モデラー 5

時系列モデルの適用 17

スペクトル・プロット 27, 29

仮定 27

スペクトル・ウィンドウ 27

中心化変換 27

2 変量スペクトル分析 27

正規化 BIC (バイズ情報量基準) 57

時系列モデラー 12, 13

時系列モデルの適用 19, 20

[タ行]

対数変換

時系列モデラー 8, 9, 10

単純季節指数平滑法モデル 8

単純指数平滑法モデル 8

調和分析 27

定常 R2 57

時系列モデラー 12, 13

時系列モデルの適用 19, 20

適合度

時系列モデラー 12, 13

時系列モデルの適用 19, 20

定義 57

転送関数 10
 季節次数 10
 差分次数 10
 遅延 10
 分子次数 10
 分母次数 10

[ハ行]

平均絶対誤差 57
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20
平均絶対パーセント誤差 57
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20
平方根変換
 時系列モデラー 8, 9, 10
偏自己相関関数
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20
 純粋な ARIMA プロセスのプロット
 61
変数名
 時系列モデラー 14
 時系列モデルの適用 21
ホールドアウト・ケース 2
保存
 新しい変数名 14, 21
 モデルの予測 14, 21
 XML 形式でのモデル指定 14
 XML 形式の再度推定されたモデル
 21

[マ行]

モデル
 エキスパート・モデラー 5
 指数平滑法 5, 8
 ARIMA 5
モデル名
 時系列モデラー 14
モデル・パラメーター
 時系列モデラー 12
 時系列モデルの適用 19
モデル・パラメーターの再度推定
 時系列モデルの適用 17

[ヤ行]

予測
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20
予測期間
 時系列モデラー 5, 14
 時系列モデルの適用 17

[ラ行]

履歴期間 2
履歴データ
 時系列モデラー 13
 時系列モデルの適用 20
レベル・シフトの外れ値 59
 時系列モデラー 7, 11

[数字]

2 乗平均平方根誤差 57
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20

A

ACF
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20
 純粋な ARIMA プロセスのプロット
 61
ARIMA モデル 5
 外れ値 11
 転送関数 10

B

Box-Ljung 統計量
 時系列モデラー 12
 時系列モデルの適用 19
Brown の指数平滑法モデル 8

H

Holt の指数平滑法モデル 8

M

MAE 57
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20
MAPE 57
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20
MaxAE 57
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20
MaxAPE 57
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20

P

PACF
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20
 純粋な ARIMA プロセスのプロット
 61

R

R2 57
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20
RMSE 57
 時系列モデラー 12, 13
 時系列モデルの適用 19, 20

W

Winters の指数平滑法モデル
 加法 8
 乗法 8

X

XML
 再度推定されたモデルを XML 形式で
 保存 21
 時系列モデルを XML 形式で保存 14



Printed in Japan