

# 解説 X-12-ARIMA (2002) <sup>1</sup>

国友直人<sup>2</sup>  
(編集責任)

2004年5月

<sup>1</sup>この解説は米国センサス局で開発された X-12-ARIMA Reference Manual (Version 0.2.7) の翻訳に短い解説を付録としたものである。オリジナルの (英文) 原稿ファイルを提供してくれた米国センサス局のデビット・フィンドレー博士 (Dr. David Findley) とブライアン・モンセル博士 (Dr. Brian Monsell) のご好意に特に感謝する。また、未整理の (英文) 原稿の整理・翻訳に協力してくれた東京大学大学院経済学研究科・統計コースの院生 (敬称略で高岡慎・一場知之・松井知也・大和田孝・室井芳史) 諸君にも感謝したい。翻訳作業は COE プロジェクト「市場経済と非市場機構の関連研究拠点」からの援助を受けて実施された。

<sup>2</sup>東京大学大学院経済学研究科教授

## 本解説の概要

この解説は米国センサス局 (U.S.Census Bureau) 発行の X-12-ARIMA Reference Manual (2000年5月16日に発行された Version 0.2.7) の翻訳に短い解説付録をつけたものである。ただし、センサス局では定期的に X-12-ARIMA プログラムを改訂しており、2004年2月の段階でのマニュアルは2002年11月頃に改訂した X-12-ARIMA プログラムの説明である Version 0.2.10 となっているので、気がついた範囲内ではあるがマニュアル Version 0.2.7 の英文原稿に修正を施しておいた。

センサス局時系列研究スタッフにより開発された季節調整法 X-12-ARIMA には統計的時系列解析をはじめ数理統計的手法が多く取り入れられているので、統計学をはじめ統計的時系列分析の研究分野についての専門的予備知識が欠けている場合にはその利用は必ずしも容易でない。季節調整法としての X-12-ARIMA 法の妥当性の評価についてはなお様々な意見があり得るが、官庁統計家など実務関係者にとりまらずはその正確な理解とより適切な利用に資することが必要と判断してこの解説稿を作成した。出来る限り原文の意図の正確な理解を期したつもりであるが、誤解等についてのコメントを歓迎する。

# 目次

<b>第 1 章</b>	<b>はじめに</b>	<b>3</b>
<b>第 2 章</b>	<b>プログラム X-12-ARIMA の実行</b>	<b>6</b>
2.1	入力 (Input) . . . . .	6
2.2	出力 (Output) . . . . .	7
2.3	入力・エラー (Input errors) . . . . .	7
2.4	出力ファイルの設定 . . . . .	8
2.5	複数系列への X-12-ARIMA の適用 . . . . .	8
2.6	ログ・ファイル (Log Files) . . . . .	11
2.7	フラッグ (Flags) . . . . .	11
2.8	プログラム制限 . . . . .	14
<b>第 3 章</b>	<b>RegARIMA モデル分析</b>	<b>18</b>
3.1	モデルの一般型 . . . . .	18
3.2	データ入力と変換 . . . . .	20
3.3	回帰変数の指定 . . . . .	20
3.4	ARIMA モデルの識別と特定化 . . . . .	23
3.5	モデルの推定と統計的推測 . . . . .	24
3.6	外れ値を含む診断 . . . . .	26
3.7	予測 . . . . .	27
<b>第 4 章</b>	<b>RegARIMA モデルの推定に関連する事項</b>	<b>30</b>
4.1	母数の初期値と収束問題への対応 . . . . .	30
4.2	MA 項の反転可能性 . . . . .	31
4.3	AR 項の定常性 . . . . .	31
4.4	AR 項と MA 項の共通因子と過剰識別 . . . . .	32
4.5	モデル選択基準の利用 . . . . .	33
<b>第 5 章</b>	<b>スペック・ファイルと文章構成法</b>	<b>35</b>
5.1	プリントと保存 . . . . .	39
5.2	日付 . . . . .	41
5.3	入力文での一般ルール . . . . .	41

<b>第 6 章 各スペック命令文の解説</b>	<b>45</b>
アリマ (ARIMA) . . . . .	46
自動モデル選択 (automdl) . . . . .	50
集計 (composite) . . . . .	55
チェック (check) . . . . .	62
推定 (ESTIMATE) . . . . .	66
予測 (forecast) . . . . .	73
履歴 (history) . . . . .	77
識別 (identify) . . . . .	86
外れ値 (outlier) . . . . .	90
回帰 (regression) . . . . .	95
系列 (series) . . . . .	110
期間変更の安定性 (slidingspans) . . . . .	120
変換 (transform) . . . . .	129
X11 . . . . .	137
X11 回帰 (X11-REGRESSION) . . . . .	149
<b>参考文献</b>	<b>168</b>
<b>解題</b>	<b>172</b>
A. X-12-ARIMA(2002) とは? . . . . .	172
B. 季節調整法小史 . . . . .	173

## 第1章 はじめに

---

季節調整プログラム X-12-ARIMA はシスキン・ヤング・マスグレーブ (1967) により開発されたセンサス局 11-X11 法の改良版である。X-11 法をより分かりやすく改良するとともに、ユーザーが様々なオプションをもとに季節性や曜日効果を検出し調整する新たな診断方法を提供している。さらにプログラムでは様々な新たな統計的方法を用いて季節調整の問題を解決し、様々な経済データを適度に調整することを可能にした。こうした方法の応用例については Findley=Hood (1999) が説明している。

この新しい方法は主として我々が RegARIMA モデルと呼んでいる統計的時系列分析用モデルをプログラムの中で広範に用いることにより実現されている。この RegARIMA モデルは ARIMA (自己回帰和分移動平均) 過程にしたがう誤差項を持つ回帰モデルである。より正確に云えば、この RegARIMA モデルでは時系列 (あるいはその対数変換値) は回帰変数の線形結合で表され、共分散構造は ARIMA モデルで決められると仮定されている。もし回帰変数を全く用いなければ、期待値ゼロの仮定の下では RegARIMA モデルは ARIMA モデルに帰着される。プログラム X-12-ARIMA の中にはフローやストック系列の取引日効果や休日効果を直接的に推定する回帰変数が組み込まれている。さらに時系列におけるある種の突然の攪乱やレベルの突然の変化がある時には X11 プログラムを使って季節調整を行なう前に回帰変数を用いて処理する。またプログラムで提供されていない問題に対処するためにはモデルのフィットに際してユーザーが自ら回帰変数を定義して利用することも可能である。X-12 プログラムにおける RegARIMA モデル部分はセンサス局の時系列研究グループが開発した RegARIMA プログラムにもとづいている。

時系列の分析において回帰変数の利用が必要となることがあってもなくても、予測に際して RegARIMA モデルを利用することが X-12-ARIMA におけるもっとも重要な改善点であり、この方法により最新時点 (あるいは初期時点) の季節調整値を改善することが可能となっている。時系列の末端部分の扱いにおいて X11 法で用いられているトレンド推定や非対称の季節移動平均過程にまつわる問題の扱いを改善している。こうした重要な技術的改善の試みとしてはカナダ統計局によって開発され利用されている X-11-ARIMA 法が既に知られている。また、こうした方法に関する理論的あるいは実用的利点に関してはこれまでに多くの文献、例えば Geweke (1978), Dagum (1988), Bobbitt=Otto (1990) あるいはそれらが引用している論文もある。

X-12-ARIMA プログラムはパーソナル・コンピュータ (386 あるいはそれ以上の数学演算装置を備えていることが必要) で計算するプログラムとして利用可能であり、DOS (3.0 版以上)、サン・ユニックス・ワークステーション (Sun 4 UNIX)、汎用大型計算機 (VAX/VMS) により実行可能である。他の計算機システムで X-12-ARIMA を実行させるプログラムを作る為にはフォートラン (FORTRAN) ・ソース・コードを利用することもできる。様々な計算機で実行させることが可能

なプログラムの最新テスト版を含むファイル・解説・実例などはインターネット URL

<http://www.census.gov/srd/www/x12a/>

に保管されているので利用者は直接にダウン・ロードすることができる。あるいは利用者は ftp を用いて—pub/ts/x12a—を—ftp.census.gov—からコピーできる。(その際、ログインは “—anonymous—” として自分のメール住所をパスワードとするだけでよい。) また限定された範囲内であるが、通常の郵便や電子郵便による問い合わせに対してセンサス局からのプログラム支援も提供されている。もし特定の入力ファイルを実行させた時に問題が生じたなら入力ファイルと出力ファイルを提供すると問題の所在を識別される可能性が高くなる。

このプログラムの季節調整部分ではシスキン・ヤング・マスグレーブ(1967)及び Dagan(1988)で詳しく説明されている X11 季節調整法を用いている。したがって X-12-ARIMA では X11 法と X11 プログラムにおけるすべての季節調整方法を実行することができ、X11 プログラムと同一の季節移動平均やトレンド移動平均を実行するとともに、X11 法における曜日効果や休日効果の除去方法も実行することもできる。

季節調整部分では幾つかの追加的なオプションを付け加える改善がなされたが、次のような新しいオプションがある。

- (a) 期間変更の安定性 (sliding spans) 法、この分析方法については Findley=Monsell=Shulman=Pugh (1990) が例示している。
- (b) ある系列についての季節調整値の改訂の履歴を計算する能力。
- (c) 新しいヘンダーソン型フィルター法により任意の奇数項ヘンダーソン型フィルターを利用者が選択できるようになったこと。
- (d) 季節フィルターの新しいオプション。
- (e) 季節調整値の不規則要素に対する新しい異常値処理オプション。
- (f) カレンダー上の曜日の要因表。
- (g) 擬加法的季節調整法。

X-12-ARIMA では季節性の伴う経済時系列に対して RegARIMA モデルを用いて統計的分析を行なっている。この目的の為に幾つかの回帰変数を提供しているが、例えばトレンド係数、全平均、固定季節効果、曜日効果、休日効果、一時的効果(加法的外れ値)、水準の変化、一時的変化、傾斜変化などがある。さらに利用者が回帰変数を定義してモデル分析に利用することも容易に可能である。こうしたプログラムは RegARIMA モデル分析を行なうために作られたのでより一般的な統計計算パッケージ用ではないことに注意しておく。特に X-12-ARIMA は時系列の高性能プロットが可能なグラフィック・ソフトとともに用いられるべきであろう。

この X-12-ARIMA プログラムによりモデル分析や季節調整が可能な観測値(データ)は量的変数のみであり、二値変数やカテゴリ変数を扱うには適当ではない。観測値は一定の間隔を持った時系列であることが必要であり、欠則値があることは許されない。またプログラム X-12-ARIMA は一変量時系列モデルのみを扱うことができ、異なる時系列間の関係を推定することはできない。

プログラム X-12-ARIMA では季節 ARIMA モデルに対して標準的な記号  $(p d q) \times (P D Q)_s$  を用いる。ここで  $(p d q)$  はそれぞれ非季節的自己回帰 (AR) 作用素、階差、移動平均 (MA) 作用素を表している。 $(P D Q)_s$  の方は季節自己回帰作用素、季節階差、季節移動平均作用素を表している。

標準的な統計的時系列モデルをデータにフィットしたい利用者にはプログラム X-12-ARIMA により識別・推定・診断という三つのモデル分析を行なうことができる。RegARIMA モデルを特定化して回帰変数を導入することによって誤差項に ARIMA モデル  $(p d q) \times (P D Q)_s$  を持つ回帰モデルを分析に用いることもできる。どのような回帰変数を用いるかはモデル分析を行なう利用者の時系列についての知識にもとづくことになる。回帰モデルの誤差項についての ARIMA モデルの識別はプログラムで提供される標本自己相関や標本偏自己相関などを利用した標準的な方法を用いて行なわれる。RegARIMA モデルが決定されるとプログラムでは一般化最小二乗アルゴリズム (IGLS) を反復する最尤法によりモデルの母数を推定する。さらにモデル診断においては推定した残差によりその適切さを調べることができる。プログラム X-12-ARIMA ではモデルの妥当性を調べる残差についての標準的診断法を提供するとともに、加法的外れ値、水準の変化などについてより複雑な方法を実行することができる。さらに推定されたモデルを用いて点予測値、予測の標準誤差、予測区間なども計算することができる。

こうしたモデル分析の方法に加えて自動モデル選択の機能や利用者が修正 AIC (赤池情報量基準 AIC をデータの長さを考慮して修正した量) を用いて回帰変数 (例えば曜日効果や復活祭の変数など) を分析に用いるか否かを定めることができる。またモデルの尤度統計量 (修正 AIC など) や予測値などの履歴を保存して、競合する統計的モデルを比較することに役だたせることもできる。

これから 5 章をつかってプログラム X-12-ARIMA の能力について詳細に説明する。2 章ではプログラム X-12-ARIMA の実行法と利用者の変更することも可能なプログラム制約について説明する。3 章ではプログラム X-12-ARIMA による RegARIMA モデルと RegARIMA モデルを用いた分析や予測などの技術的要点について議論する。4 章では統計的モデル分析を行なう利用者が知っておくべき時系列モデルの推定や統計的推測についての鍵となる事柄を説明する。5 章では入力ファイル (スペック・ファイルと呼ぶ) を説明するが、特にスペック・ファイルの記述法や関連する問題にも言及しておく。こうして 2 章から 5 章までを使ってプログラム X-12-ARIMA の機能や利用法の概略を説明する。次に 6 章はプログラム X-12-ARIMA の入力スペック・ファイルを作成するとき有用なスペック・コマンドを説明する。スペック・ファイルに書かれる可能性のある各コマンド (命令) を詳細に説明する。利用者は 6 章に書かれている説明により命令文を用いて X-12-ARIMA の実行を制御したり、プログラムで用意されているオプションの選択することができるよう。

## 第2章 プログラム X-12-ARIMA の実行

---

プログラム X-12-ARIMA をインストールする方法は計算機に依存している。この問題についての情報は本マニュアルに説明している他にインターネット <ftp.census.gov> 上で公開している `/pub/ts/x12a/ReadMe` の中のファイルからも得ることができる。DOS システム上のパソコン上にインストールされていればプログラム X-12-ARIMA を実行するには

```
> path\x12a path\filename
```

とすればよい。この命令文では `path\filename.spc` が入力ファイル (スペック・ファイル) となる。この時プログラム X-12-ARIMA は出力ファイルとして

```
path\filename.out
```

を作り出す。ここで `path` の指定は使用しているディレクトリー (カレント・ディレクトリー) 上にプログラム X-12-ARIMA が存在する必要がある、入力ファイルについても同様である。

ここでファイル名のみを利用者が決めることができることに注意しておく。プログラムは実行時に指定された入力ファイル名を用いてファイルを作り出す。例えばスペック・ファイルが `filename.spc` のときには出力ファイルは `filename.out` という名で書かれ、エラー・メッセージは `filename.err` として書かれることになる。したがって、スペック・ファイル `xuu1.spc` がパソコン DOS 上の現ディレクトリーに置かれていれば

```
> x12a xuu1
```

として改行キー<return> (あるいは<enter>) を押せばプログラムが実行され、ファイル `xuu1.out` と `xuu1.err` がディレクトリー上に作られることになる。

入出力のプログラムは以下で簡単に言及しておくが後でより詳細に説明することとなる。UNIX システム上でプログラムを実行させるにはこれまでのスラッシュを適切に変更するだけでよい。また OS として DOS、UNIX、VAX/VMS の為の簡単なマニュアルも利用可能であり、プログラム X-12-ARIMA を実行するより詳しい説明がなされている。

### 2.1 入力 (Input)

プログラム X-12-ARIMA を実行するには入力ファイル (スペック・ファイルと呼ぶ) を作成する必要がある。このファイルはアスキー形式 (あるいはテキスト形式) でありプログラムを実行する上で必要な時系列、実行したい分析、必要な出力についての情報が要求される。プログラム



はスペック・ファイルが拡張子 `.spc` の形式で書かれていることを想定している。したがって、例えば入力ファイルは `path\filename` となっていればよい。

スペック・ファイル以外でプログラム X-12-ARIMA の入力ファイルとなりうるのは時系列データを含むオプション・ファイルであり、利用者が定義する回帰変数と事前調整要素と自動モデル選択に関するモデル型などがそれにあたる。こうしたファイル・コマンドはスペック・ファイルの中で適当な形式で書かれていなければならない。こうした付加的な入力ファイルが利用可能になったのでスペック・ファイル上で利用可能なデータ値を含めたり自動モデル選択の為のモデルを自動的に設定することができる。

## 2.2 出力 (Output)

標準的には出力ファイルは `path\filename.out` として書かれる。個々のスペック・ファイルの内容に対応して出力ファイルが作られるが、その場合には 5.1 節に言及される出力印刷 (`print`) 変数が使われる。コマンドの保存 (`save`) では例えば残差系列を保存し、グラフ・プログラムでデータをプロットするなど更なる分析の為に他の出力ファイルを作り出す為に使うことができる。ここで保存 (`save`) を利用するときにはプログラム上では 5.1 節に書かれてある標準的な変数名を用いた出力ファイルが作られることに注意しておく。このときもし同一の変数名を用いたファイルが存在している場合には書き換えられることになる。そこでプログラムの利用者が出力を保存する場合には注意を払う必要がある。このことについては 5.1 節の説明をよく読む必要がある。

## 2.3 入力・エラー (Input errors)

プログラムが入力エラーを発見するとその報告の為に適当なエラー・メッセージが書き出される。こうしたエラー・メッセージは `path\filename.err` という名前のファイルの中に保存される。プログラムが実行上でエラーの部分を持定化できればエラーが起きた場所 (^) をスペック・ファイルとともに出力する。そのエラーが全体に影響を及ぼす場合にはメッセージ `ERROR` を個々のエラー・メッセージの前に出力するが、その時にはどの様にプログラムを修正したらよいかなどのメッセージを付けることがある。メッセージ `WARNING` はエラーとはみなされないが注意を喚起する為のものである。

プログラム X-12-ARIMA はまずスペック・ファイル全体を読み、入力エラーがあれば幾つかのエラーとして検出する。このとき利用者は一度プログラムを動かすだけで幾つかのエラーを修正することが可能となる。しかしながら、エラーがある場合にはしばしば最初かあるいは最初の幾つかのエラーの情報だけしか実は意味がなく、はじめの部分のエラーが系列 (`series`) の定義等により他のエラーを引き起こすことも少なくない。プログラムは重大なエラーを検出したときには実行を停止する。メッセージ `WARNING` の場合にはプログラムは停止されないが、利用者は入出力ファイルを調べ望ましい結果が得られているか否か慎重に検討することが望まれる。

## 2.4 出力ファイルの設定

前にも述べたようにスペックファイルを *filename.spc* とするとプログラム X-12-ARIMA では出力ファイルは *filename.out*、エラー・ファイルは *filename.err* 等となる。与えられた原系列を用いて様々な調整を行なったり RegARIMA モデルのオプションを利用したりする目的の為にしばしば入力ファイルで用いた名前ではないファイル名を使いたいことがある。出力ファイルとして他の名前を使うには

```
> path\x12a path\filename path\outname (2.1)
```

とすればよい。この場合にはスペック・ファイルは *filename.spc* であるが出力ファイルは *outname.out*、エラー・ファイルは *outname.err* として保存される。すなわちすべての出力ファイルは通常の入出力ファイル名ではなく利用者が指定するパス名とファイル名に保存されることになる。

## 2.5 複数系列への X-12-ARIMA の適用

実用的にはプログラム X-12-ARIMA の一度の実行で複数の原系列を処理する必要が生じることがある。複数系列に対して一度に X-12-ARIMA を実行する方法としては二つのやり方がある。

- (a) 複数スペック形式 (**multi-spec mode**) : ここでは個々の系列についてそれぞれ入力スペック・ファイルを用意する形式をとる。
- (b) 単一スペック形式 (**single spec mode**) : 各系列に対して一つの入力スペック・ファイルからのオプションを用いて実行する形式をとる。

いずれの形式でプログラムを実行する場合にもまずメタ・ファイル (*metafile*) を作る必要がある。このファイルはアスキー・ファイル形式で X-12-ARIMA が扱うことの出来るファイル名を含んでいる必要がある。X-12-ARIMA が扱う 2 種類のメタ・ファイルとしては入力メタ・ファイルとデータ・メタ・ファイルがある。

メタ・ファイルを実行中にスペック・ファイルの一つでエラーが発生するとプログラムは適当なエラー・メッセージを出力する。その際、エラーが発生したスペック・ファイルの計算は停止するが、他のスペック・ファイルはそのまま継続して実行される。すべての入力ファイルの実行で発生したエラーは X-12-ARIMA のログ・ファイル (2.7 節参照) に書き込まれる。

### 複数スペック・ファイルの実行

複数スペック形式で X-12-ARIMA を実行するにはまず入力メタ・ファイル (*input metafile*) を作っておく必要がある。このファイルはアスキー形式で X-12-ARIMA で読み込むファイル名を含んでいる必要がある。入力メタ・ファイルは 1 行ごとに 2 つまでの内容、すなわちある系列に対する入力スペック・ファイル (必要ならパス名を含む) とその系列に対する出力ファイル名を含む必要がある。(ただし後者を書きこむのはオプションである。) もし出力ファイルが利用者により指定されなければ、出力ファイルのパス名とファイル名は入力ファイル名にしたがって指定される。また入力ファイルは入力メタ・ファイル上に書かれた順にしたがって実行される。

ここで例を用いて以上で述べたことを説明すると3つのスペック・ファイル xuu1.spc, xuu2.spc, xuu3.spc を実行させるには次のようなメタ・ファイルを作ればよい。

```
xuu1
```

```
xuu2
```

```
xuu3
```

ここではすべてのスペック・ファイルは現ディレクトリー上にあると仮定している。もしこれらのスペック・ファイルがDOS上のディレクトリー c:\export\specs に保存されている時にはメタ・ファイルは

```
c:\export\specs\xuu1
```

```
c:\export\specs\xuu2
```

```
c:\export\specs\xuu3
```

となる。入力メタ・ファイルを用いて X-12-ARIMA を実行するには

```
> x12a -m metafile
```

とすればよい。ここでメタ・ファイル (metafile.mta) はメタ・ファイル名であり -m は信号 (フラグ) で X-12-ARIMA の実行にあたってメタ・ファイルの存在を示している。例えば上に定義されたメタ・ファイルが exports.mta として保存されていればスペック・ファイルを実行するには

```
> x12a -m exports
```

と打鍵して改行キーを押せばよい。

ここで上の例のように入力メタ・ファイル名が与えられるときにはファイル名のみ指定されているので、入力メタ・ファイルの拡張子は .mta とする必要がある。また必要であればパス情報も入力メタ・ファイルに付けておくことが望ましい。

X-12-ARIMA で出力ファイルとして使われるファイル名はメタ・ファイルの中のスペック・ファイルから取られるのであってメタ・ファイル名を利用するわけではない。上の例ではメタ・ファイル exports.mta 上の個別の入力スペック・ファイルに対応して出力ファイル名として xuu1.out, xuu2.out, xuu3.out が作られるが、集計した出力ファイルとして exports.out が作られるわけではない。上の例で別の名前での出力ファイルを作らせるには次のように入力メタ・ファイルの各行に希望する出力ファイル名を加えておくだけでよい。

```
c:\export\specs\xuu1 c:\export\output\xuu1
```

```
c:\export\specs\xuu2 c:\export\output\xuu2
```

```
c:\export\specs\xuu3 c:\export\output\xuu3
```

## 単一スペック形式による X-12-ARIMA の実行

X-12-ARIMA を用いて複数の時系列の各系列に対して同一の計算を実行させるにはデータ・メタファイル (*data metafile*) を作る必要がある。データ・メタ・ファイルでは各行は 2 つの内容まで含むことができ、各系列に対するデータ・ファイルの中のファイル名 (必要ならパス名) と出力ファイル名 (必ずしも必要ではない) を書くことになる。もし出力ファイル名が利用者により指定されなければデータ・ファイルのパス名とファイル名が出力ファイルには使われる。

**注意点:** データ・メタ・ファイルでは個別のデータ・ファイルには拡張子を用いない。もしデータ・ファイルが当該ディレクトリーになければ拡張子がパス名とファイル名とともに指定されていなくてはならない。

データ・ファイルはデータ・メタ・ファイルの中に表われる順番にしたがい処理される。個々のデータ・ファイルについてのオプションは実行時において識別される各入力スペック・ファイルにしたがい処理される。このことはデータ・メタ・ファイルのすべてのデータ・ファイルが同一の形式となっていなければならないことを意味している。この場合には X-12-ARIMA でサポートしている幾つかの形式は避けて使わない必要があるが、こうした点について詳しくは 6 章のスペック・コマンドの系列 (*series*) で説明する。

例えば 3 つのデータ・ファイル *xuu1.dat*, *xuu2.dat*, *xuu3.dat* を一度に処理する為にはデータ・メタ・ファイルは次のようになる。

```
xuu1.dat
```

```
xuu2.dat
```

```
xuu3.dat
```

ここではすべてのデータ・ファイルは現ディレクトリーの中にあると想定されていることに注意しておこう。ここで *c:\export\data* の中に保存されているファイルを実行する場合にはメタ・ファイルは次のようになる。

```
c:\export\data\xuu1.dat
```

```
c:\export\data\xuu2.dat
```

```
c:\export\data\xuu3.dat
```

データ・メタ・ファイルを用いて X-12-ARIMA を実行するには命令文

```
> x12a specfile -d metafile
```

を用いる。ここで *metafile.dta* はデータ・メタ・ファイルであり、*-d* は X-12-ARIMA の利用にデータ・メタ・ファイルを用いるという信号 (フラッグ) であり、*specfile.spc* はデータ・メタ・ファイル上の各系列で用いられるスペック・ファイルである。例えば上の例で用いられた 3 つの系列からなるデータ・メタ・ファイルの名が *exports.dta* であれば

```
> x12a default -d exports
```

と打鍵して改行キーを押せば default.spc を入力スペック・ファイルとして X-12-ARIMA が実行される。

この例でデータ・メタ・ファイル名が与えられる時にはファイル名のみが定められその拡張子は定められてはいないので、入力メタ・ファイルには拡張子 .dta が必要となる。さらに必要ならばパス名もデータ・ファイル名に付け加えることになる。

X-12-ARIMA で出力ファイルの作成に使われるファイル名はメタ・ファイル自身ではなく、メタ・ファイルの中のデータ・ファイルからとられる。上の例ではメタ・ファイル exports.dta の中の各データ・ファイルに対応した出力ファイル xuu1.out, xuu2.out, xuu3.out が作られるが、メタ・ファイル名を用いた exports.out ではない。この例で別の出力ファイル名を指定するには次のようにデータ・メタ・ファイルの各行に指定する出力ファイル名を書けばよい。

```
c:\export\data\xuu1.dat  c:\export\output\xuu1
c:\export\data\xuu2.dat  c:\export\output\xuu2
c:\export\data\xuu3.dat  c:\export\output\xuu3
```

## 2.6 ログ・ファイル (Log Files)

プログラム X-12-ARIMA を実行するごとにログ・ファイル (log file) が作り出され、各系列あるいは各スペック・ファイルにたいしてモデル分析の結果や季節調整の診断結果が保存される。X-12-ARIMA が複数あるいは単独のスペック・ファイルにより前節までに説明したように実行するとログ・ファイルはメタ・ファイル (複数スペック形式) あるいはデータ・メタ・ファイル (単独スペック形式) と同一の名前とディレクトリーに拡張子 .log をつけて保存される。例えば

```
> x12a -m exports
```

とすると、メタ・ファイル exports.mta 中のスペック・ファイルが実行され、ファイル exports.log の中に利用者が選んだ診断結果が保存される。ここで一列のみが処理される場合には出力ディレクトリーと出力ファイル名はログ・ファイルを作る拡張子 .log が用いられる。

利用者はログ・ファイルに保存すべき診断結果を指定することができるが、例えばスペック命令である系列 (series), 集計 (composite), 変換 (transform), x11, x11 回帰 (x11 regression), 回帰 (regression), 自動モデル (automdl), 推定値 (estimate), 診断 (check), 期間変更の安定性 (slidingspans), 履歴 (history) などではログ・ファイルや出力ファイルを指定することができる。これら個々のスペック命令とそこでの保存可能なログ・ファイルの詳細は 6 章で説明する。前節でも述べたようにメタ・ファイルのスペック・ファイルの実行時にエラーが発生した場合にはログ・ファイルに入力ファイルとともに書き込まれる。

## 2.7 フラッグ (Flags)

前節の例ではプログラムを実行する為には命令文において信号 (フラッグ) -m と -d が必要であった。命令文の中で入出力オプションを指定するには他のやり方もあり得る。命令文の一般

的な形式は

```
> path\x12a arg1 arg2 ... argN
```

となる。ここで x12a に続く変数は状況により信号 (フラッグ・コマンド) かファイル名となる。表 2.1 には X-12-ARIMA で用いられるフラッグ・コマンドが要約されているが、本節ではこれから少し詳しく各フラッグの意味を説明することとする。このフラッグ命令文の中では順番は問題とはならないが、幾つかのフラッグではその後で適当なファイル名を指定することが必要となる。既に前節ではフラッグ **-m** と **-d** を説明しておいたが、この両方のフラッグを一緒に実行することはできない。

フラッグ **-i** は次の変数が入力スペック・ファイルのパス名とファイル名であることを示している。このフラッグは入力スペック・ファイルが最初の変数である場合には必要ない。したがって二つの命令文 `x12a test` と `x12a -i test` は同等である。二つのフラッグ命令 **-i** と **-m** は同時に実行できない。

フラッグ **-i** と同様にフラッグ **-o** は次の変数が出力のためのパス名とファイル名であることを示す。前に説明した拡張子 (`.out` と `.err`) やコマンドの保存 (save) に関する拡張子などもファイル名に付け加えることができる。このフラッグは入力スペック・ファイルが次に出力ファイル名が 2 行目 (式 2.1 のように) ならば必要ない。したがって次に挙げる命令文はすべて同等となる。

```
x12a test test2
x12a -i test -o test2
x12a -o test2 -i test
```

ここで `x12a -i test test2` とすると、最初の変数がスペック・ファイルでなくフラッグ **-i** なのでエラーとなる。二つのフラッグ **-o** と **-m** は同時に実行することはできない。

フラッグ **-s** は出力の中のある季節調整と RegARIMA モデルの診断結果を主出力ファイルと分けたファイルとして保存する為に必要となる。この結果は出力の時系列表だけではなく出力ファイルの中の表も含んでいるが、必ずしも各時系列表には使われていない形式でも保存できる。フラッグ **-s** が宣言されると X-12-ARIMA は自動的に診断の重要な結果を別のファイルに保存するので診断結果を要約する際に利用することができる。このファイルは季節調整診断ファイル (*seasonal adjustment diagnostics file*) と呼ばれるが、主出力ファイルと同一のパス名とファイル名に拡張子 `.xdg` を付けたものとなる。例えばコマンド

```
> x12a test -s
```

に対する季節調整診断ファイルは `test.xdg` に保存される。またコマンド

```
> x12a test -s -o testout
```

に対する季節調整診断ファイルは `testout.xdg` となる。

季節調整診断ファイルの他にもフラッグ **-s** を用いるとプログラムは RegARIMA モデル分析の結果をモデル診断ファイル (*model diagnostics file*) の中に保存する。このモデル診断ファイルは主出力ファイルと同一のパス名とファイル名を持ち拡張子は `.mdg` となる。

したがって命令文

```
> x12a test -s
```

に対するモデル診断ファイルは test.mdg に保存され、命令文

```
> x12a test -s -o testout
```

に対するモデル診断ファイルは testout.mdg に保存される。

季節調整診断ファイルについてはインターネット上で ftp.census.gov に一つのプログラムが公開されていて、それを使うと季節調整診断ファイルを読んだり要約したりすることができる。そのプログラムはプログラム言語 Icon により書かれている。(Griswold=Griswold (1997) を参照されたい)。

フラッグ -g は次の変数名がディレクトリーのパス名であって別のグラフィックス・プログラムの入力としてその出力を用いることを意味している。この出力には次のようなファイルがある。

- (1) グラフにする診断ファイルでスペック・ファイル (.spc) の中のオプションにより作られるもの、
- (2) ファイル名を含むグラフ・メタ・ファイル、
- (3) 実行される時系列や季節調整の型の情報を含む季節調整・診断・ファイル、
- (4) 時系列に当てはめる RegARIMA モデルの情報を含むモデル診断ファイル。

グラフィック・メタ・ファイルの拡張子は.gmt、季節調整診断ファイルの拡張子は.xdg、またモデル診断ファイルの拡張子は.mdgである。こうしたファイルのファイル名はプログラムの主出力ファイル名が用いられる。例えば利用者が

```
> x12a test -g c:\sagraph
```

と入力すると、グラフィック・メタ・ファイルは c:\sagraph\test.gmt に保存され、季節調整診断ファイルは c:\sagraph\test.xdg に保存され、モデル診断ファイルは c:\sagraph\test.mdg にそれぞれ保存される。命令文

```
> x12a test -g c:\sagraph -o testout
```

に対してはグラフ・メタ・ファイルは c:\sagraph\testout.gmt に保存され、季節調整診断ファイルは c:\sagraph\testout.xdg に、さらにモデル診断ファイルは c:\sagraph\testout.mdg にそれぞれ保存される。この二つの例ではともに季節調整のグラフィックに必要なファイルはディレクトリー c:\sagraph に保存される。ここでの注意点としてはフラッグ -g が入るディレクトリーが出力ファイルとは別に作られている必要があることであるが、同一ディレクトリーの副ディレクトリーであってもよいことである。

グラフィック形式の出力を作るグラフ・ソフトとして SAS/GRAPH(SAS) を使うプログラム X-12-Graph がある。(これについては Hood (1998a)、Hood (1998b) を見よ。) グラフィック形式で

X-12-ARIMA の実行により保存されるファイル一覧とそれらのファイルを記述するグラフィック・メタ・ファイルで用いられるコードについては表 2-2 を見ればよい。

このフラッグ `-g` を用いて作られた季節調整診断ファイルは外部のグラフ・ソフト SAS/GRAPH に必要な季節調整の重要な情報のみを保存するものであり、フラッグ `-s` オプションを使うとファイルの一部のみ保存される。ここで両方のオプションを同時に使うと季節調整診断ファイルすべて（主出力ファイルではなく）は `-g` オプションにより指定されたディレクトリーに保存される。モデル診断ファイルが作られればそのファイルはグラフ・ディレクトリーに保存される。利用者が季節調整診断ファイルを使うとある種の警告がスクリーン上に表われ、またモデル診断ファイルの場合もグラフィック・ディレクトリー警告が表われる。

フラッグ `-n,-w,-p` はどれもプログラムの主出力を制御している。オプション `-n` により利用者は主出力ファイルの表数を制限することができる。プログラム X-12-ARIMA は主出力としてかなりの数の表を作り出すが、X-12-ARIMA では利用者がその出力表を柔軟に決めることができるので、時には幾つかの表に絞った方が便利となる。このことを簡単に実行するには例えばフラッグ `-n` を使えばよく、X-12-ARIMA デフォルトとしての表は主出力ファイルには出なくなり、利用者がスペック・ファイルに書いた表のみが出力されることとなる。

フラッグ `-w` は主出力ファイルの表を横幅のある（132文字）に指定する為に用いられる。このフラッグがなければデフォルトは80文字の形式である。表の正確な文字数は系列の大きさや系列 (series) での精度の指定に依存して決まる。フラッグ `-p` により主出力ファイルにおいて改ページやヘッダーを省略することができる。このオプションを用いなければ改ページが各出力表の先頭に挿入され、実行タイトル、系列名、ページ数も挿入される。

フラッグ `-c` は集計季節調整を入力メタ・ファイル (`-m`) に制限するときのみ使われる。集計季節調整では X-12-ARIMA では通常は構成要素の時系列とともに合成系列（集計量と呼ばれる）の季節調整を行なうが、通常は足し算で集計される。（その詳細は6章のコマンドの集計 (composite) を参照されたい。）この命令を使うときには各系列についてスペック・ファイルが必要となる。フラッグ `-c` を使うときには各系列に対する入力ファイルの中に指定されている季節調整やモデルのオプションは無視され、各系列は集計系列を作る為だけに使われる。このオプションでは集計系列に対する RegARIMA モデルを識別するために利用される。

最後にフラッグ `-v` により X-12-ARIMA を実行するに際しては利用者が入力確認モードを使ってエラーが一つあるいは複数の入力スペック・ファイルにあるか否かを確かめることができる。この方法で利用者は全系列に X-12-ARIMA を実行させることなくエラーのあるプログラム・オプションを調べることができる。このフラッグ `-v` は他のフラッグ `-s,-c,-n,-w,-p` などと同時に用いることは出来ない。

## 2.8 プログラム制限

プログラム X-12-ARIMA には系列の最大数、回帰変数の最大数等々の制限がある。こうした制限の数値は大多数の応用上には十分大きいですが、計算機への負荷があまりに大きくなったり膨大なメモリーを使いすぎたりした結果、プログラムの実行時間が長くなるのを防ぐ為に設定されている。もし必要ならばこうした制限は修正することが可能であるが、そのためにはフォートランで



書かれた原プログラムを変換するなどして新たな制限を設定する必要がある。こうした修正を行なう為のパラメータ設定はファイル model.prm と srslen.prm の中に書かれている。プログラムの制約を修正するときにはオリジナル・ファイルを特にバックアップしておくことを推奨する。表 2-3 には利用者が修正するプログラムの制限に対応したパラメータ変数はファイル model.prm と srslen.prm の中にある。

## 一覧表

表 2-1: X-12-ARIMA プログラム・フラッグ (Flags)

フラッグ	フラッグの説明
-i ファイル名	入力ファイル用の (拡張子なし) ファイル名
-o ファイル名	プログラムの実行により作成されるすべての出力ファイル用の (拡張子なし) ファイル名
-m ファイル名	入力用・メタ・ファイル用の (拡張子なし) ファイル名
-d ファイル名	データ・メタ・ファイル用の (拡張子なし) ファイル名
-g ディレクトリ名	外部のグラフィック・プログラムへ入力するためグラフィック・メタ・ファイルと関連するファイルを保存しておくディレクトリ
-n	(表なし指定) 入力・ファイルに特に指定された表のみを出力
-w	主な出力ファイルで (1 3 2 文字) 拡大表示
-p	主な出力ファイルにおいてページ数を非表示
-s	季節調整の診断 (diagnostics) をファイルに保存
-c	集計される要素の和を計算し、合計値のみにモデル分析や季節調整を実行
-v	入力ファイルに誤りがあるかを調べるが、他は非実行

表 2-2 : X-12-ARIMA グラフィック・メタ・ファイルに関連するコード名

コード (Code)	内容の説明
acf	ACF
acf2	二乗残差の ACF
adjori	事前調整原系列 (Prior-Adjusted Original Series)
aichst	AICC 値の履歴
ao	加法的はずれ値
caf	集計された調整因子 (Combined Adjustment Factors)
cal	不規則変動回帰より得られたカレンダー因子 (Calendar Factors From Irregular Regression)
ccal	不規則変動回帰より得られた集計されたカレンダー因子 (Combined Calendar Factors From Irregular Regression)
cfchst	予測と予測誤差の履歴 (History of Concurrent Forecasts and Forecast Errors)
chol	集計された休日因子 (Combined Holiday Factors)
cmpori	原集計系列 (Original Composite Series)
csahst	季節調整系列のパーセント変化の履歴 (History of the Percent Change in the Seasonally Adjusted Series)
ctd	不規則変動回帰から集計された営業日数因子 (Combined Trading Day Factors From Irregular Regression)
ctrhst	トレンド値のパーセント変化の履歴 (History of the Percent Change in the Trend Values)
fct	原系列と予測値 (Original Series and Forecasts)
fcthst	予測二乗誤差の履歴 (History of the Sum of Squared Forecast Errors)
fincal	集計されたカレンダー因子 (Combined Calendar Factors)
finstst	最終はずれ値の t 統計値 (Final Outlier t-test Statistics)
ftt	変換された原系列と予測値 (Transformed Original Series and Forecasts)
idacf	スペック・識別により生成された ACF 値 (ACFs Generated by Identify Spec)
idpacf	スペック・識別により生成された PACF 値 (PACFs Generated by Identify Spec)
indir	間接的不規則変動 (Indirect Irregular)
indrsi	SI 比への間接的置き換え値 (Indirect Replacement Values for the SI Ratios)
indsa	間接的季節調整値 (Indirect Seasonally Adjusted Series)
indsar	丸め計算を伴う間接的季節調整 (Indirect Seasonally Adjusted with Rounding)
indsat	年集計値を制約した間接的季節調整 (Indirect Seasonally Adj with Forced Annual Totals)
indsf	間接的季節因子 (Indirect Seasonal Factors)
indsi	間接的 S I 比 (Indirect SI Ratios)
indtrn	間接的トレンド (Indirect Trend)
irr	最終的不規則要素 (Final Irregular Component)
irrwt	不規則要素の加重 (Irregular Weights)
isahst	間接的季節調整値の履歴 (History of the Indirect Seasonal Adjustment Values)
ls	水準変化のはずれ値 (Level Shift Outliers)
mdlest	回帰 A R I M A モデルの推定値 (regARIMA Model Estimates)
mvadj	欠損値を置き換えた原系列 (Original Series with Missing Values Replaced)
oadori	はずれ値調整済み原系列 (Outlier-Adjusted Original Series)
odjcmp	はずれ値調整済み集計系列 (Outlier Adjusted Composite Data)
ori	原系列 (Original Series)
otl	はずれ値 (Combined Outliers)
pacf	PACF
prior	事前調整因子 (Prior Adjustment Factors)
ptd	事前的曜日効果因子 (Prior Trading Day Factors)
rgseas	ユーザーが定義する季節回帰因子 (User-Defined Seasonal Regression Factors)
rhof	回帰 A R I M A モデルからの休日因子 (Holiday Factors from regARIMA model)
rsi	置き換え S I 比 (Replacement SI Ratios)
rtf	回帰 A R I M A モデルからの曜日因子 (Trading Day Factors from regARIMA model)
sa	季節調整値 (Seasonally Adjusted Series)
sahst	季節調整値の履歴 (History of the Seasonally Adjusted Series)
sarnd	まるめ操作を伴う季節調整系列 (Seasonally Adjusted Series with Rounding)
satot	年集計値を制約した季節調整値 (Seasonally Adjusted Series with Forced Annual Totals)

表 2-2: X-12-ARIMA グラフィック・メタ・ファイルに伴うコード (つづき)

コード (Code)	内容の説明
sf	季節因子 (Seasonal Factors)
sfhst	季節因子推定値の履歴 (History of the Seasonal Factor Values)
sfr	ユーザー定義の回帰分析による季節因子 (Seasonal Factors with User-Defined Regression)
si	S I 比 (SI Ratios)
siox	はずれ値と極値ラベル付き S I 比 (SI Ratios, with Labels for Outliers and Extreme Values)
spccmp	集計系列のスペクトラム (Spectrum of the Composite Series)
spciir	間接修正済み不規則変動のスペクトラム (Spectrum of the Indirect Modified Irregular)
spcirr	修正済み不規則変動のスペクトラム (Spectrum of the Modified Irregular)
spcisa	間接的季節修正済み系列のスペクトラム (Spectrum of the Indirect Seasonally Adjusted Series)
spcori	原系列のスペクトラム (Spectrum of the Original Series)
spcrsd	回帰 ARIMA モデルの残差スペクトラム (Spectrum of the regARIMA Model Residuals)
spcsa	季節調整済み系列のスペクトラム (Spectrum of the Seasonally Adjusted Series)
tc	一時的変化はずれ値 (Temporary Change Outliers)
trn	最終的トレンド・循環要素 (Final Trend-Cycle Component)
trnhst	トレンド推定値の履歴 (History of the Trend Values)
usrdef	ユーザーが定義した回帰因子 (User-Defined Regression Factors)
xeastr	イースター因子 (Easter Factors)
xhol	不規則変動回帰における休日因子 (Holiday Factors From Irregular Regression)
xtd	不規則変動回帰における曜日効果因子 (Trading Day Factors From Irregular Regression)
xtrm	最終的極値修正因子 (Final Extreme Value Adjustment Factors)

表 2-3: X-12-ARIMA プログラム上の制限

パラメーター 変数	制限値 (limit)	内容の説明
pobs	600	入力系列の最大長、The number, pobs + pfcst (以下を参照), はユーザーが指定する回帰変数や事前調整要素の為の入力系列の最大長、加えて pfcst 値は回帰変数値や予測期間の調整要素にあわせることもできる
pyrs	60	予測と逆予測で拡張された系列の最大年数
pfcst	60	予測の最大値 (maximum number of forecasts)
pb	80	回帰モデルの変数の最大値 (あらかじめ決まっている変数やユーザーが指定する変数、および自動はずれ値探索や A I C テストで生成される変数を含む)
pureg	52	ユーザーが定義できる回帰変数の最大値 (maximum number of user-defined regression variables)
porder	36	AR 要素と MA 要素の最大次数 (maximum lag corresponding to any AR or MA parameter)
pdflg	3	ARIMA 要素 (非季節・季節) における階差の最大値 (maximum number of differences in any ARIMA factor (nonseasonal or seasonal))

## 第3章 RegARIMA モデル分析

3.1 節ではプログラム X-12-ARIMA で扱う一般的な統計的モデルを説明する。3.2 節から 3.7 節では RegARIMA モデル分析と予測の様々な段階、すなわちデータ入力、変換、回帰、説明変数選択、ARIMA モデルの識別と選択、モデルの推定と統計的推測、外れ値を含む診断、予測、などにおける X-12-ARIMA の能力について説明する。こうした課題の実行を制御する入力（スペック・ファイルの書き方）についても以下の節で説明する。スペック・ファイルのより細かなことについては 6 章で説明する。

RegARIMA モデルを作る際には高性能の時系列プロットを用いることを強く勧めたい。プロットすることにより季節性のパターンをはじめ外れ値の可能性や確率的非定常性といった貴重な情報が得られる。さらに系列の変換や様々な階差変換などの影響を調べることも有力な情報を提供する。X-12-ARIMA 自体はそうしたプロット能力を持っていないので他のソフトウェアを用いる必要がある。

### 3.1 モデルの一般型

季節性をもつ時系列についてもボックス・ジェンキンス (1976) により議論されている ARIMA モデルはしばしば利用されている。時系列  $z_t$  に対する乗法型季節 ARIMA モデルの一般型は

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D z_t = \theta(B)\Theta(B^s)a_t \quad (1)$$

と表現することができる。ここで  $B$  は遅れ作用素 (backshift operator  $Bz_t = z_{t-1}$ )、 $s$  は季節周期、 $\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$  は非季節自己回帰 (AR) 作用素、 $\Phi(B^s) = (1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps})$  は季節自己回帰作用素、 $\theta(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$  は非季節移動平均 (MA) 作用素、 $\Theta(B^s) = (1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs})$  は季節移動平均作用素、 $a_t$  項は独立・同一分布にしたがい、期待値ゼロで分散一定  $\sigma^2$  のホワイト・ノイズである。さらに  $(1-B)^d(1-B^s)^D$  は次数  $d$  の非季節階差と次数  $D$  の季節階差を表している。もし  $d = D = 0$  (階差なし) であれば、通常は (1) 式の  $z_t$  をその平均からの差、すなわち  $z_t - \mu$  (ここで  $\mu = E[z_t]$  とする) で置き換える。

この ARIMA モデルを拡張して時間に依存する期待値を回帰モデルを組み込むことで有益な統計的モデルを得ることができる。より具体的には時系列  $y_t$  を

$$y_t = \sum_i \beta_i x_{it} + z_t \quad (2)$$

と表現する。ここで  $y_t$  は (被説明) 時系列であり、 $x_{it}$  は  $y_t$  とともに観察可能な回帰変数、 $\beta_i$  は回帰母数、 $z_t = y_t - \sum \beta_i x_{it}$  であり、回帰の誤差項は (1) 式で与えられる ARIMA モデルにしたが

うと仮定される。時系列  $z_t$  について ARIMA モデルを利用するので標準的な回帰分析を時系列に用いることによる基本的問題が生じる。標準的回帰分析では回帰誤差項（すなわち (2) 式の  $z_t$ ）は時間について無相関であることが仮定される。ところが時系列データでは (2) 式の誤差項は通常は自己相関がある上にしばしば階差操作を行なう必要もある。そうした状況で確率変数系列  $z_t$  が互いに無相関であることを仮定すると、通常は適切でない結果をもたらしがちである。

プログラム X-12-ARIMA において用いられる RegARIMA モデルは (1) 式と (2) 式で定められる。この二つの式を組み合わせると統計的モデルは

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D\left(y_t - \sum_i \beta_i x_{it}\right) = \theta(B)\Theta(B^s)a_t. \quad (3)$$

と表現されることになる。この RegARIMA モデルは (1) 式で与えられる純粋な ARIMA モデルに回帰関数 ( $\sum \beta_i x_{it}$ ) を加えるか、あるいは (2) 式で与えられる回帰モデルの誤差項  $z_t$  に (1) 式で与えられる ARIMA モデルを加えることにより得ることができる。いずれにしても RegARIMA モデルではまず時系列  $y_t$  から回帰効果を引き去り、期待値ゼロの系列  $z_t$  を得た後に、誤差系列  $z_t$  を階差操作することにより定常時系列  $w_t$  を求めることになる。この確率変数系列  $w_t$  は定常 ARMA モデル  $\phi(B)\Phi(B^s)w_t = \theta(B)\Theta(B^s)a_t$  にしたがうと仮定される。したがって (3) 式で与えられる RegARIMA モデルを別の形で表現すると

$$(1-B)^d(1-B^s)^D y_t = \sum_i \beta_i (1-B)^d(1-B^s)^D x_{it} + w_t. \quad (4)$$

となる。ここで  $w_t$  は定常 ARMA モデルにしたがっている確率変数系列である。(4) 式の形では RegARIMA モデルにおける回帰変数  $x_{it}$  や原系列  $y_t$  に対して階差作用素  $(1-B)^d(1-B^s)^D$  を用いる ARIMA モデルを強調した形式となっている。

ここで (3) 式で表現された RegARIMA モデルでは回帰変数  $x_{it}$  は同時点の被説明変数系列のみに影響すると仮定していることに注意しておく。すなわちモデル (3) では  $\beta x_{i,t-1}$  など遅れをもつ効果は明示的には考慮されていないのである。むしろ遅れのある回帰効果を X-12-ARIMA に含めることは可能であるが、それには利用者がラグ（遅れを持つ）変数を定義する必要がある。

プログラム X-12-ARIMA では RegARIMA モデルの ARIMA 部分を定式化する際に様々な柔軟性を備えているが、例えば次の事などが挙げられる。(i) 二つ以上の乗法的 ARIMA 因子、(ii) AR 多項式と MA 多項式でのラグの一部分の省略、(iii) モデルを推定した時に各 AR 項と MA 項の母数を利用者が特定の値に設定できること、(iv) トレンドの定数項を導入し、階差系列  $((1-B)^d(1-B^s)^D y_t)$  に対してゼロでない平均を設定する。こうした RegARIMA モデルの定式化についての詳細は 6 章で議論し、かつ例も示される。

時系列 ARIMA モデル分析自体に関してはボックス・ジェンキンス (1976) の古典的研究に詳しく説明されているが、他の統計的時系列分析の教科書、例えば Abraham=Ledolter (1983), Vandaele (1983) などでも説明されている。また Bell (1999) は特に RegARIMA モデルについて議論しているが、ちょうどプログラム X-12-ARIMA の議論に適している。

## 3.2 データ入力と変換

分析が必要な原系列の観測値は系列 (series) を使ってプログラムに読み込まれる。データは系列 (series) に含まれるか、あるいは別のファイルとして置くかいずれかの形をとる。分析に使うデータの長さを制御するには期間 (span) やモデル期間 (modelspace) を使うが、原系列の最初や最後の幾つかのデータを分析に含めないことも可能である。系列 (series) を使って初期時点、(必要なら) 季節周期、時系列のタイトルを指定することもできる。

コマンドの変換 **transform** によりデータの非線型変換及び事前要素による修正を行なうことができる。利用できる非線型変換として対数変換や平方根変換などボックス・コックス (Box-Cox) 変換やロジット変換 (0 と 1 の間の時系列に役立つ) などが含まれている。事前調整を定義すると月の長さ (*length of month*) (あるいは四半期系列の場合には四半期の長さ (*length of quarter*)) により観測値を月次系列を分割したり、月次 (四半期) 平均でスケール変換することも可能である。さらに、2月についてはうるう年要素により調整することもできる。また一連の利用者が定義することにより原系列に割ったり引いたりする事前調整を行なうこともできる。

## 3.3 回帰変数の指定

RgARIMA モデルを定式化するには回帰変数 ((2) 式の  $x_{it}$ ) 及び回帰の誤差項  $z_t$  に対する (1) 式で与えられる ARIMA モデルを定める必要がある。最初の問題は回帰 (regression) を使って行ない、後者はアリマ (arima) (3.4 節で議論する) を用いて行なう。どのような回帰変数を選択するかは利用者がモデル分析しようとする時系列についての事前知識を使って定めることになる。経済季節時系列のモデル分析を行なう際にしばしば利用されている回帰変数は X-12-ARIMA の中に組み込まれているので簡単に利用できる。その方法についてはこれから説明するが、回帰変数リストは本章の表 3.1 にまとめられている。これらの変数の定義と利用法については 6 章の回帰 (regression) の項目で説明される。さらに利用者が入力するデータ変数 (利用者の定義する回帰変数) を使うこともできる。統計モデルの推定 (3.5 節) に際しては X-12-ARIMA は各説明変数の統計的有意性を評価する  $t$ -統計量や曜日効果など) 特別の効果全体に対する回帰係数の有意性を  $\chi^2$ -統計量を使って調べている。

もっとも基本的な回帰変数としては定数項がある。もし ARIMA モデルが階差操作を含まなければ定数項は通常の切片をあらわす回帰変数であり、他の変数を含まなければ定常時系列の期待値 (平均) を表現する。もし ARIMA モデルが階差操作を含んでいれば (4) 式で与えられる ARIMA モデルにもとづく階差に対して定数 1 からなる変数をプログラム X-12-ARIMA は作り出す。この変数に対する母数はトレンド定数と呼ばれるが、これは階差の次数と同じ多項式トレンドを作るからである。例えば非季節階差が正 ( $d > 0$ ) で季節階差がゼロ ( $D = 0$ ) とすると、もとの多項式トレンドは変数  $t^d$  に比例する。ここで注意しておく必要があるのはより低次の多項式  $t^j$  ( $0 \leq j < d$ ) は階差  $(1 - B)^d$  をとるとゼロになり係数を推定できないので回帰変数に含まれないことである。むしろ、トレンド定数が回帰に含まれるか否かにかかわらずモデル (3)(あるいは (4)) では階差操作をとおして低次の多項式トレンド項を含んでいる。ここで季節階差操作 ( $D > 0$ ) を含む時には原系列でのトレンド定数項はより複雑になり、次数  $d + D$  の多項式を含むことを意味する。トレンド定数を含まなければ次数  $d + D - 1$  の多項式トレンドを含んでいることを意味している。

月次系列においては固定的季節効果 (*Fixed seasonal effects*) は各月に対応する 12 個のダミー変数によって表すことができる。ただし、こうしたダミー変数の和は 1 になるので全平均効果と区別できない。このことから 2 つの問題が生じる。階差なしのモデルでは定数項との多重共線性、12 個の変数を階差を取ると和がゼロになるので階差モデルでは積率行列の非正則性が生じる。この場合のよく知られた対処法としては、12 個のダミー変数の代わりに 11 個のダミー変数を使うことでモデルの母数を変換すればよい。あるいは別の方法としては 11 個の変数を使って固定的月次季節性をフーリエ (三角関数) 表現の中に組み込むことも考えられる。このような母数の変換に用いられる変数は表 3-1 に与えられている。プログラム X-12-ARIMA ではこうしたオプションも利用できるし、特定の周波数のみに三角関数を利用することも可能である。四半期系列や他の季節周期系列に対しては X-12-ARIMA により必要な変数をつく出すこともできる。ただし、こうした変数はすべて階差をとるとゼロになる場合には季節階差を含むモデルで利用することはできない。

曜日効果 (*Trading-day effects*) は時系列の変動が相異なる年の同月の曜日の構成に影響される時に生じる。曜日効果としては 7 個の変数でおのおの各曜日数 (月曜の数) - (火曜の数), ..., (土曜の数) - (日曜の数) をあらわす変数、月の長さ (1om) の変数やその季節性を除いた変数、あるいはうろう年変数 (1pyear) などである。X-12-ARIMA では 6 つの曜日変数は `tdnolpyear` と呼ばれている。7 個目の回帰変数を追加して用いるよりもより簡単で賢く乗法的にうろう年効果を扱う方法としては原系列に対して変換  $\bar{m}_{Feb} Y_t / m_t$  (ここで  $Y_t$  は変換前の原系列、 $m_t$  は  $t$  月の長さ (28 か 29)、 $\bar{m}_{Feb} = 28.25$  は 2 月の平均長さ) として 2 月の値  $Y_t$  のみをスケール変換してしまうことであろう。もし RegARIMA モデルが季節効果を含んでいる場合には 2 月の周期を除いてすでに月の長さの効果は含まれているので、曜日効果はうろう年効果のみを扱えばよいことになろう。この場合には変数 `tdnolpyear` をモデルに組み込むだけでよい。X-12-ARIMA ではどちらのやり方も可能であり、さらにオプション (`td`) があり、月の長さの効果を自動的に扱ってくれるが、この方法については回帰 (regression) の説明を参照されたい。

ここまでの説明では時系列は (観測できない) ある日次系列を月別に集計した作られていると仮定していた。このような系列は月次のフロー (*flow*) 系列と呼ばれている。これに対して系列がある日次系列の月末での値となっている時にはには ストック (*stock*) 系列と呼ばれるが、この場合には他の回帰変数を用いる方が適当であろう。月末ストック系列の曜日効果は月末値の曜日を 7 つのダミー変数を使ってモデル分析することが考えられよう。ここでダミー変数の和は常に 1 なので積率行列が非正則性となるのを防ぐ為に 6 個のダミー変数だけ用いればよいことに注意しておく。(表 3-1 を参照のこと。) X-12-ARIMA では月末でない他に日 (例えば月初め日) による計算される月次ストック・データについても適当な回帰変数を利用できる。

四半期フロー系列に対しても X-12-ARIMA では月次系列と同一の曜日効果のオプションを用意している。ただし、四半期データにおける曜日効果は比較的まれにしか検出されない。四半期におけるカレンダーの構成は月次系列に比較して比率の意味では時間の経過とともにそれほど変化しない。月次と四半期以外の季節周期を持つフロー系列及び月次以外のストック系列に対しては曜日効果の説明変数は利用できない。

さらに X-12-ARIMA では月次系列と四半期系列における曜日効果変動を簡単に説明する一つの

回帰変数 (平日・週末対比変数)

$$I_t = (\text{平日日数}) - \frac{5}{2}(\text{土曜} \cdot \text{日曜数})$$

を使うことができる。

このモデルでは月曜から金曜までの全ての曜日は同一の効果を持ち、土曜と日曜は同一の効果を持つと仮定している。X-12-ARIMA ではこのモデルは次の二つのやり方で推定できる。まず利用者がオプション `td` と同様に月の長さの効果をモデル分析するためにプログラムを書くのであればオプション `td1coef` を規定すればよい。また月の長さ効果を導入するオプション `tdnolpyear` とともにオプション `td1nolpyear` を使うこともできる。

月次フロー系列における 休日効果 (*Holiday effects*) は系列で計測する経済活動が休日の日付とともに変動することから生じるが、例えば (i) 経済活動が休日の日付の周辺で増加したり減少したりすること、あるいは (ii) 各年に変化する休日の日付により 2 ヶ月以上にわたり異なる効果がありうること、などが挙げられる。クリスマスのように固定された日付の休日は固定季節効果と識別可能ではない。復活祭効果 (*Easter effects*) は米国の経済時系列ではもっともよく観察されている休日効果である。というのは復活祭の日付は 3 月 22 日から 4 月 25 日まで変動するからである。レーバー日 (*Labor Day*) や感謝祭日 (*Thanksgiving*) の効果がある可能性もあるが、これらの休日効果はそれほど一般的には観察されない。X-12-ARIMA における復活祭日やレーバー日の基本モデルではある特定の整数を  $w$  として経済活動の水準がその日の  $w$  日前から変化し、休日当日までその水準にとどまると仮定している。感謝祭日のモデルではその日の前後固定された日数だけ経済活動の水準が変化し、12 月 24 日までその水準にしたがうと仮定されている。こうした休日効果を扱う回帰変数では所与の月  $t$  とその効果が持続する月数により構成される。(実際、表 3-1 で示されているようにこうした回帰変数は長期の全月平均からの乖離幅として基準化されている。) 原理的には同じ復活祭効果の変数は四半期フロー系列にたいしても適用可能であるが、レーバー日と感謝祭日効果の方は四半期系列には利用可能でない。X-12-ARIMA ではストック変数については休日効果に対処する変数は用意していない。

プログラム X-12-ARIMA には系列の水準の (一時的ないし恒久的) 急激な変化を扱う 4 種類の回帰変数が用意されている。これらの変数は加法的外れ値 (*additive outliers*) (AO), 水準変化 (*level shifts*) (LS), 一時的変化 (*temporary changes*) (TC), 傾斜的水準変化 (*ramps*) とそれぞれ呼ばれている。変数 AO では時系列の特定の観察値のみの変化を表現し、変数 LS ではある時点から後の全ての時系列に一定値の増加・減少することを表現し、変数 TC では系列の水準が特定時点で変化した後に指数的に急速にもとの水準に戻ることを表現し、変数 *ramp* では一定の期間に線形的に増加あるいは減少することを表現している。こうした効果をモデル分析する為に利用できる回帰変数は表 3-1 にまとめられている。(回帰変数 LS は -1 と 0 をとるように定義し、はじめにゼロをとり後に 1 をとる変数としていないのは、全体の回帰関数の平均水準の予測値を時系列の最新時点の水準と矛盾しないようにしたことによるものである。同様にして変数 *ramp* も定義している。)

コマンドの回帰 (*regression*) を使うと事前知識として時系列に既知の時刻に変化があると考えられる場合に変数 AO, 変数 LS, 変数 TC, 変数 *ramps* などを用いて対処することができる。しかしながら、しばしば多くの季節変動では水準の変化が起こったか否かを識別することが困難であ



ろう。そこで潜在的な外れ値の場所や性質を決めることが外れ値 (outlier) による外れ値の検出方法の目的となる。この問題については 3.6 節か 6 章の外れ値 (outlier) を参照されたい。この方法では変数 AO, 変数 TC, 変数 LS (変数 ramp は除かれる) の効果を検出し、検出した効果をそのまま回帰変数として自動的に付け加えることができる。

ここで説明した変数 AO, LS, TC, ramp は Box=Tiao (1975) がインターベンション (interventions) と呼んでいた時系列分析の簡単な形である。X-12-ARIMA は Box=Tiao (1975) が動学的インターベンション効果として議論している分析をすべて扱えるわけではないが、ここで言及した変数 AO, 変数 LS, 変数 TC, 変数 ramp を適当に選んでつないでいけばかなり複雑な動学的インターベンション分析を行なうことができ、しかも各ステップで生じる追加的な母数の数は 1 か 2 の増加で押さえられる。同様にして利用者が設定する変数を含む (確率過程としての時系列) RegARIMA モデルとボックス・ジェンキンス (Box=Jenkins (1976)) の 10 章と 11 章で議論した動学的トランスファー関数モデルと解釈することもできよう。したがって RegARIMA モデルはしばしばより一般的な動学的トランスファー関数モデルとして用いることができる。他方、トランスファー・モデルの方は一般的には確率的説明変数が未知なので予測については特別な扱いが必要という問題がある。(この問題については Box=Jenkins (1976) の 11.5 節を参照されたい。)

### 3.4 ARIMA モデルの識別と特定化

RegARIMA モデルにおける ARIMA 部分は (1) 式における次数とその構造、すなわち 3 つ組  $(p\ d\ q)$ ,  $(P\ D\ Q)$ ,  $s$  により決定される。もし回帰変数がモデルに含まなければ通常の ARIMA モデルについての次数決定 (すなわち ARIMA モデルの識別と呼ばれている) を通常によく知られた方法で行なえばよい。すなわち、時系列  $y_t$  とその階差についての標本自己相関関数 (ACF) や標本偏自己相関関数 (PACF) に基づく方法である。RegARIMA モデルでは回帰効果が存在することから ACF や PACF を見かけ上ゆがめるので、それを修正する必要がある。典型的には階差次数は時系列  $y_t$  とその階差系列を検討してモデルを識別することができる。次に階差データの回帰変数の階差系列への回帰残差を得ることができる。こうして得られた残差の ACF と PACF により RegARIMA モデルにおける誤差項の自己回帰項 (AR) と移動平均項 (MA) の次数の識別を行なえば良い。こうした RegARIMA モデルの識別は Bell=Hillmer(1983) や Bell(1999) に議論され例示されている方法である。

RegARIMA を用いて分析方法を行なう鍵となるのは命令 (コマンド) の識別 (identify) である。これを月次時系列の例で説明してみよう。階差次数を決めることはまず時系列  $y_t$ ,  $(1-B)y_t$ ,  $(1-B^{12})y_t$ ,  $(1-B)(1-B^{12})y_t$  の自己相関関数 ACF を利用する。これは識別 (identify) を使えば一度に実行することができる。階差次数が決まれば識別 (identify) と回帰 (regression) を使って (i) 時系列  $y_t$  の階差系列を回帰説明変数の階差系列に回帰, (ii) 回帰残差の ACF と PACF を使って自己回帰成分と移動平均成分を識別する、などを行なう。例えば 1 次の非季節階差と 1 次の季節階差を指定 ( $d = 1$  かつ  $D = 1$ ) して、識別 (identify) と回帰 (regression) を使って ARIMA モデル

$$(1-B)(1-B^{12})y_t = \sum_i \beta_i (1-B)(1-B^{12})x_{it} + w_t \quad (5)$$

を最小二乗法 (OLS) で当てはめれば (5) 式の回帰残差の ACF と PACF が計算される。

ここでもし候補となる階差（非季節階差と季節階差ともに）の最大値をあらかじめ指定して良ければ必ずしもプログラム X-12-ARIMA を二度も実行させる必要はない。例えばこの最大階差の次数を  $d = 1$  と  $D = 1$  と想定してみよう。このとき識別 (identify) と回帰 (regression) を用いると (i) (5) 式に OLS により回帰を行ない母数  $\tilde{\beta}_i$  の推定値を求め, (ii) 推定された (原系列からの) 回帰残差系列  $\tilde{z}_t = y_t - \sum_i \tilde{\beta}_i x_{it}$  を計算し, さらに (iii)  $\tilde{z}_t, (1-B)\tilde{z}_t, (1-B^2)\tilde{z}_t, (1-B)(1-B^2)\tilde{z}_t$  の ACF と PACF を計算してくれる。

これまでの説明について例外が一つあるので注意しておく。回帰 (regression) で定数項を設定すると、(5) 式にもとづく最小二乗回帰では定数項は含まれるが、データから回帰効果は取り除かれない。したがって、 $\tilde{\beta}_1 x_{1t}$  がトレンド項の場合には  $\tilde{z}_t = y_t - \sum_{i \geq 2} \tilde{\beta}_i x_{it}$  となる。このように扱う理由を (5) 式により説明すると、3.3 節で述べたように (5) 式のトレンド定数項は二次多項式を意味するが、階差操作  $(1-B)(1-B^2)$  により定数項と一次トレンド項を  $t^0 \equiv 1$  と  $t$  について含む形にしている。定数項と一次トレンド項の係数を推定することができないので系列  $y_t$  から完全な形で多項式効果を取り除くことは出来ないからである。そこで X-12-ARIMA では階差のない回帰残差  $\tilde{z}_t$  を推定する際には二次多項式項  $t^2$  を取り除くのではなく推定されるトレンド定数を無視することで解決を図っている。(4) 式で表されるより一般 ARIMA モデルにおいても同様に扱われている。識別 (identify) におけるトレンド定数項を入れる効果は (4) 式の  $i \geq 2$  に対する係数  $\tilde{\beta}_i$  の回帰による推定値に影響が生じるこののみである。

### 3.5 モデル推定と統計的推測

回帰 (regression) とアリマ (arima) により RegARIMA モデルを定式化することができる。次に推定 (estimate) により厳密な最尤法、あるいは条件付最尤法 (Box=Jenkins (1976) の 209 項-212 項の説明を見ればよいがしばしば条件付最小二乗法とも呼ばれている) の一種により時系列モデルの母数を推定できる。利用者は完全な厳密尤度、自己回帰 (AR) 部分には条件付尤度で移動平均 (MA) 部分は厳密尤度、あるいは両方を条件付推定、といったいずれの最大化を指定することができる。自己回帰 (AR) 母数については完全な厳密最尤推定値と条件付尤度推定値の差は一般的に小さい上にどちらの方法が良いかについての説明も可能なことがある。移動平均 (MA) 母数については厳密尤度と条件付尤度の違いはより本質的であり、厳密尤度の方が推奨される。移動平均母数の条件付最尤推定はオプションとして与えられているが、主として他のソフトウェアとの比較や厳密尤度の最大化に際して収束問題が発生した時に初期値を設定する為に利用することができる。(この収束問題については 4.1 節を参照されたい。) 特別の指定が無ければ AR 母数と MA 母数の推定では厳密最尤法を用いる。

いずれの推定方法を採用するにせよ純粋な ARIMA モデルの尤度関数は二乗和関数に帰着されるので非線形最小二乗 (More=Garbow=Hillstrom(1980) で説明されている MINPACK) プログラムにより最小化される。RegARIMA モデルの全体についての尤度関数の最大化は Burman=Otto=Bell(1987) が用いた繰り返し一般化最小二乗 (IGLS) アルゴリズムを用いている。このアルゴリズムは次の二つのステップから構成されている。すなわち

(i) 所与の AR と MA 母数にたいして回帰係数を一般化最小二乗 (GLS) 法で推定するが、その際には ARIMA モデルで決まる共分散構造を利用する。

(ii) 回帰モデルの母数  $\beta_i$  の値を所与として最尤法により ARIMA モデルを回帰残差  $z_t = y_t - \sum \beta_i x_{it}$  に適用する。

このように IGLS ではこの二つのステップを収束が達成されるまで繰り返す。(ここでコマンド推定 (estimate) のオプションを使うと希望するなら反復推定の途中結果を表示することもできる。) 厳密 (あるいは条件付) 尤度関数の評価法については Box=Jenkins(1976, 7 章), Ljung=Box(1979), Hillmer=Tiao(1979), Wilson(1983)などを参考にしたが、アルゴリズムの詳細は別の技術メモに説明しておいた。また 4 章では全ての利用者が母数推定について注意すべき事項を述べている。

RegARIMA モデルの統計的推測については ARIMA モデルの推定及び RegARIMA モデルの推定についての漸近理論 (Box=Jenkins(1976) の 7 章, Brockwell=Davis(1987) の 8 章, Pierce(1971), Bell(1999)) を利用することができよう。漸近理論によれば適当な仮定の下では母数の推定量は近似的に正規分布にしたがい、その平均は真の母数であり、その共分散行列は推定可能である。ここで適当な仮定とは例えば正しい統計的モデルの下で自己回帰係数は定常性条件を満たし、移動平均係数は反転条件を満たし、時系列データが漸近理論が適用できるほど十分長く存在していることなどである。こうした理論的結果にもとづき X-12-ARIMA では ARMA モデルの母数推定値や回帰の母数推定値などの標準偏差や場合により ARMA 母数と回帰母数全体の推定値の相関 (共分散) 行列を計算している。(漸近的には回帰推定値と ARMA 推定値は無相関となる。) こうした結果により 3. 3 節で述べたように母数推定に関する標準的な正規分布にもとづく議論に利用することもできるし  $t$ -統計量や  $\chi^2$ -統計量を使って個別の回帰母数の有意性や回帰効果を検定するために幾つかの母数をまとめた有意性の検定に利用することもできる。これに加えて X-12-ARIMA では対数尤度関数の最大値も出力するので様々な統計モデルを推定した時に各種の尤度比検定などを行なうことが可能である。

X-12-ARIMA では誤差分散の推定量  $\sigma^2$  として  $\hat{\sigma}^2 = SS/(n-d-s \cdot D)$  を用いている。ここで SS とは残差平方和であり、 $n-d-s \cdot D$  は階差操作の後の実効的な観測数である。(尤度関数が AR 母数について条件付きで計算されている場合には  $n-d-s \cdot D$  は  $n-p-d-s \cdot P-s \cdot D$  で置き換える必要がある。) ここでは推定された母数について自由度を調整していないことに注意しておく。このためにもしプログラム X-12-ARIMA により純粋な回帰モデル (つまり ARIMA(0 0 0) モデル) を推定すると推定値  $\hat{\sigma}^2$  は通常回帰分析における分散の不偏推定値とは異なる。したがって推定の標準偏差、 $t$ -統計量、 $\chi^2$ -統計量なども標準的な回帰分析プログラムとは若干のずれが生じることになっている。

統計的推測についてはもう一つの別の方法として尤度にもとづくモデル選択基準を用いることが X-12-ARIMA で可能であり、AIC (赤池情報量基準), AICC(F-修正 AIC), Hannan-Quinn の基準, シュワルツの BIC などが計算される。これらの統計量を用いると、それぞれより小さな値をとるモデルを選択する傾向になることが知られている。こうした基準よりもむしろ伝統的な基準である  $t$ -統計量、 $\chi^2$ -統計量、尤度比検定統計量などと比較すると非入れ子型 (nonnested) モデル (すなわち一つのモデルの母数を取り除くなどすることにより他のモデルが導かれないモデル選択の状況で、例えば AR(1) モデルと MA(1) モデルの比較が例である) における選択が可能となりうることを挙げることができよう。ただし、モデル選択基準を用いるときにはかなり注意を払わねばならない。4.5 節では RegARIMA モデル分析で生じる状況で標準的な尤度比検定が適当でないばかりか、こうしたモデル選択基準も適当でない場合について言及する。

### 3.6 外れ値を含む診断

RegARIMA モデルの診断はモデル推定から得られる残差の分析から行なわれるが、いずれも真の残差 (3) 式の  $a_t$  がホワイト・ノイズと見なせるか、すなわち  $N(0, \sigma^2)$  にしたがう独立・同一分布の確率変数か否かを調べるのがその目的となる。(ここで分布の正規性は標本からの推定の漸近論や統計的推測には必ずしも必要でないが、予測における予測区間の正当化には極めて重要となることに注意しておく。) チェック (check) を使えば当てはめた統計モデルから残差にもとづく様々な診断のための統計量が作られる。またこのコマンドでは自己相関関数を調べる為に残差の ACF と PACF やその標準偏差推定値とともに求め、リュンゲン・ボックス (Ljung=Box(1978)) の Q-統計量も計算する。さらに X-12-ARIMA では残差の記述統計量や標準化残差のヒストグラムも計算するが、求めた残差は他のソフトウェアによるグラフィックスを含むさらなる分析の為にファイルに書くこともできる。

時系列モデルにおける診断の一つの重要な要素として外れ値の検出がある。外れ値 (outlier) を用いることで自動的に加法的な外れ値を変数 AO で、一時的な変化としての外れ値を変数 I' で C、さらに水準の変化を変数 LS により検出することができる。こうした外れ値のタイプ (AO, TC, LS) やそれに関連する回帰変数については 3.3 節で既に説明した。X-12-ARIMA における外れ値の検出法は基本的には Chang=Tiao(1983) にもとづいている (Chang=Tiao=Chen(1988) を見よ) がさらに Bell(1983, 1999) や Otto=Bell(1990) により拡張された方法にもよっている。その一般的な方法は GLS 回帰の反復法と同様であるが、利用する回帰変数としては変数 AO, LS, TC などを用いて外れ値検出に関するすべての期間について行なう、すなわち  $n$  の長さに時系列に対して変数 AO, LS, TC の検出を実行する。実際には  $3n$  個の変数を使うのではなく少なめの変数を用いる事になるが、その理由については 6 章の外れ値 (outlier) の詳細 (DETAILS) に議論されている。

簡単に要約すると、この方法では各時点において計算した  $t$  統計量が有意となるとモデルに対応する変数 AO, LS, TC を回帰変数に組み込む操作を行なう。あまりに過度の計算負荷を避けるために AR 項と MA 項の母数は各時点で外れ値に対する  $t$  統計量を計算する際には固定しておく。X-12-ARIMA ではこうした一般的な外れ値検出について二つのやり方が可能である。外れ値推定 1 (addone 法と呼ぶ) では各外れ値をモデルに加えるごとに全体のモデルを再推定する方法であり、外れ値推定 2 (addall 法と呼ぶ) では検出された幾つかの外れ値を加えるごとにモデルを再推定する手続きをとる。こうした二つの方法については 6 章の外れ値 (outlier) で説明されているがより詳しい形で Findley=Monsell=Bell=Otto=Chen(1997) の付録 B に説明されている。

外れ値の検出を行なっている際には残差の標準偏差についての頑健な推定値  $1.48 \times$  残差メディアン絶対偏差 (Hampel et al. (1986)105 項による) により推定を行なっている。外れ値は有意な外れ値を全ての (あるいは一部の) 観測時点について探すので通常の正規分布に基づく臨界点 (すなわちほぼ 2.0) では外れ値の有意性検出には低すぎる。プログラムでは 3.8 を標準的臨界点としているが利用者が変えることも可能である。

モデルが複数の水準の変化を含む場合は外れ値の検出や回帰 (regression) を使うときに起こり得るが、X-12-ARIMA ではオプションとして 2 個のシフト、3 個のシフト等々が一時的な水準の変化 (temporary level shift) として隠れていたという帰無仮説を検定する  $t$  統計量を作り出す。もし一つの変化が別の変化と相殺していれば 2 個の連続する水準の変化が一時的な水準変化を覆い隠すことがあり、このときには二つに回帰母数の和はゼロとなる。同様のもし 3 つの回帰母数の和が

ゼロならば三個の水準変化が相殺することもあり得る。(こうした連続する水準変化の個数には上限がある。) こうした  $t$  統計量は水準変化の母数推定値の和をその標準偏差推定値で割って計算される。水準の変化について有意でない  $t$  統計量 (例えば絶対値が 2 よりも小さい) 水準の連続的变化が一時的水準変化を隠している仮説は棄却される。こうした検定により利用者はモデルにおける水準の変化のインパクトを評価することが可能となる。もろんこうした統計量の一つか複数が有意となれば利用者は一時的レベル変化の変数の代わりに適当な水準の変化をあらゆる回帰変数により再度定式化すればよい。(こうしたことは利用者が変数を定義することでも、あるいは水準の変化に対応する変数の母係数の和がゼロになるように設定しても実行可能である。) 水準の変化 (これは系列の水準が恒久的に変化することに対応する) で記述するか、あるいは一時的な水準の変化で定式化するかはデータ期間の末端部分において系列の水準変化を伴う予測と関連する重要な問題である。

### 3.7 予測

X-12-ARIMA により RegARIMA の母数が推定されれば予測 (forecast) により点予測値、予測誤差、予測区間を求めることができる。点予測値は真のモデルを用いていると仮定し、現在・過去の時系列  $y_t$  の値を所与として将来の線形予測の中で平均二乗誤差 MMSE を最小化することで求められる。すなわちここでは RegARIMA モデルが正しく、正しい回帰変数を用い、追加的な外れ値や水準の変化が予測期間に起こらず、ARIMA モデルが正しく、母数推定値が真の値に一致していることを意味している。こうした仮定は標準的であるが当然ながら実際の応用では現実的ではないかもしれない。実際に期待できることとしては RegARIMA モデルが真の状況にかなり近い近似となっていて、真のモデルがわからなくともその結果が近似的に適切となっていることであろう。ここでは 2 つのやり方で予測誤差を計算することができる。まずすべての母数の値を既知と仮定して予測値を計算することが考えられる。さらに、回帰母数を推定することから追加的な予測誤差が発生するが AR 項と MA 項の母数は既知とする方法も考えられよう。かなり長い時系列ではボックス・ジェンキンス (1976, 267 項-269 項) によれば AR 項と MA 項の推定誤差は小さいので予測の標準偏差を計算するに際してはその誤差を無視できるとの正当化がありえよう。

もし時系列が変数変換されているときには予測結果はまず変換したスケールで求め、次に原系列の予測に変換しなおすことが考えられる。例えば原系列  $Y_t$  に対数変換  $y_t = \log(Y_t)$  を施してモデル (3) を当てはめたとすると、まず系列  $y_t$  に対して予測を行ない、求めた点予測値と予測区間を原系列のスケールに変換しなおすことが考えられる。ただし、標準的な仮定の下で予測と述べたのは変換後の変数  $y_t$  についての MMSE の点予測値であって原系列ではないことに注意しておく。同様の予測方法を X-12-ARIMA が用意している変数変換についても行なうことができる。もし事前調整がなされていれば点予測値と区間予測値はもとのスケールに変換されることになる。

さて利用者が定義した回帰の説明変数が含まれていると X-12-ARIMA では予測区間におけるその値を与えることが必要となる。既成の回帰変数については必要な将来の値はプログラム上で作り出す。利用者が自ら事前調整要素を定義する場合には予測期間についてもその値を定めることが必要となる。

表 3-1: X-12-ARIMA に組み込まれている回帰変数

変数名	定義と説明
トレンド const	$(1-B)^{-d}(1-B^s)^{-D}I(t \geq 1)$ ただし $I(t \geq 1) = \begin{cases} 1 & \text{for } t \geq 1 \\ 0 & \text{for } t < 1 \end{cases}$
<sup>2</sup> 固定的季節性 seasonal	$M_{1,t} = \begin{cases} 1 & \text{1月} \\ -1 & \text{12月} \\ 0 & \text{その他} \end{cases}, \dots, M_{11,t} = \begin{cases} 1 & \text{11月} \\ -1 & \text{12月} \\ 0 & \text{その他} \end{cases}$
<sup>2</sup> 固定的季節性 sincos[ ]	$\sin(\omega_j t), \cos(\omega_j t)$ , ただし $\omega_j = 2\pi j/12, 1 \leq j \leq 6$ ( $\sin(\omega_6 t) \equiv 0$ )
曜日効果変数 (月次、または四半期) tdnolpyear, <sup>3</sup> td 係数タイプ 曜日効果変数 (月次、あるいは四半期) td1nolpyear, <sup>4</sup> td1coef	$T_{1,t} = (\text{月曜の数}) - (\text{日曜の数}), \dots$ $T_{6,t} = (\text{土曜の数}) - (\text{日曜の数})$  $(\text{平日の数}) - \frac{5}{2}(\text{土曜と日曜の数})$
月の長さ (月次) lom	$m_t - \bar{m}$ , ただし $m_t = \text{月の長さ } t \text{ (日数)}$ および $\bar{m} = 30.4375$ (月の平均長)
四半期の長さ (四半期) loq	$q_t - \bar{q}$ , ただし $q_t = \text{四半期の長さ } t \text{ (日数)}$ および $\bar{q} = 91.3125$ (四半期の平均長)
閏年 (月次および四半期) lpyear	$LY_t = \begin{cases} 0.75 & \text{閏年の2月 (第1四半期)} \\ -0.25 & \text{他の2月 (第一四半期)} \\ 0 & \text{その他} \end{cases}$
曜日効果 (ストック) (月次ストック) tdstock[w]	$D_{1,t} = \begin{cases} 1 & \bar{w}^{\text{日}} \text{ 月中日 } t \text{ 月曜日} \\ -1 & \bar{w}^{\text{日}} \text{ 月中日 } t \text{ 日曜日} \\ 0 & \text{その他} \end{cases}, \dots$  $D_{6,t} = \begin{cases} 1 & \bar{w}^{\text{日}} \text{ 月中日 } t \text{ 土曜日} \\ -1 & \bar{w}^{\text{日}} \text{ 月中日 } t \text{ 日曜日} \\ 0 & \text{その他} \end{cases}$ , ただし $\bar{w}$ は最小の $w$ 月の長さは $t$ . 月末ストック系列では $w$ を 31 とする, i.e., すなわち $\text{tdstock}[31]$ .
カナダ・イースター変数 (月次あるいは四半期) sceaster[ w ]	イースターが4月前なら $w$ , 変数 $n_E$ をイースター前の3月の $w$ 日数とする, $E(w, t) = \begin{cases} n_E/w & \text{3月} \\ -n_E/w & \text{4月} \\ 0 & \text{その他} \end{cases}$ . イースターが4月 $w$ であると, $E(w, t) = 0$ とする. (注意: この変数は3月と4月 (あるいは第1四半期と第2四半期) を除いて0である.)

<sup>1</sup> 制約がある場合には括弧に記した。それぞれの項目にはスベック回帰、すなわち `regression { variables=const }` において、回帰変数として用いられる変数名が与えられている。<sup>2</sup> 変数は月次系列に対して示されている。他の季節周期にも対応する変数は利用可能である。<sup>3</sup> これら6つの変数に加えてオプション `td` では `lpyear` 回帰変数 (変換なしの系列に対し)、あるいはこの変数は2月の数値  $Y_t$  を  $\bar{m}_{Feb} Y_t / m_t$  に変換する。ただし、(変換される) 原系列  $Y_t$  に対して  $\bar{m}_{Feb} = 28.25$  (2月の平均長) としている。四半期の `td1coef` も同様である。

表 3-1: X-12-ARIMA に組み込まれている回帰変数 (つづき)

1 回帰項	変数の定義と説明
<sup>4</sup> イースター変数 (月次、あるいは四半期) easter[w]	$E(w, t) = \frac{1}{w} \times [t \text{ 月 (四半期) においてイースター以前の } w \text{ 日数}]$ . (注 意: この変数は 2 月・3 月・4 月 (第一四半期あるいは第二四半期) 以外は 0 となる。2 月には $w > 22$ の時に限りゼロではない。)
<sup>5</sup> Labor-Day 変数 (月次) labor[w]	$L(w, t) = \frac{1}{w} \times [t \text{ 月に Labor Day 前の } w \text{ 日数}]$ . (注意: 8 月と 9 月を除 き 0 となる。)
<sup>5</sup> 感謝祭変数 (月次) thank[w]	$ThC(w, t) = t \text{ 月に 1 2 月 2 4 日までに感謝祭前の } w \text{ 日数 (負値は感謝祭}$ $\text{後の日数を意味する。)}.$ (注意: この変数は 1 1 月と 1 2 月を除き 0 であ る。)
$t_0$ 時の加法的はずれ値 ao 日付 <sub>0</sub>	$AO_t^{(t_0)} = \begin{cases} 1 & t = t_0 \text{ の時} \\ 0 & t \neq t_0 \text{ の時} \end{cases}$ (日付 $t_0$ は時刻 $t_0$ に対応する日時)
$t_0$ 時のレベル・シフト ls 日付 <sub>0</sub>	$LS_t^{(t_0)} = \begin{cases} -1 & t < t_0 \text{ の時} \\ 0 & t \geq t_0 \text{ の時} \end{cases}$
$t_0$ 時の一時的変化 tc 日付 <sub>0</sub>	$TC_t^{(t_0)} = \begin{cases} 0 & t < t_0 \text{ の時} \\ \alpha^{t-t_0} & t \geq t_0 \text{ の時} \end{cases}$ , ただし $\alpha$ は前の水準へ戻る割合を表す。 ( $0 < \alpha < 1$ ).
傾斜変化 (Ramp) ( $t_0$ から $t_1$ まで) rp 日付 <sub>0</sub> -日付 <sub>1</sub>	$RP_t^{(t_0, t_1)} = \begin{cases} -1 & t \leq t_0 \text{ の時} \\ (t - t_0)/(t_1 - t_0) - 1 & t_0 < t < t_1 \text{ の時} \\ 0 & t \geq t_1 \text{ の時} \end{cases}$

<sup>4</sup> 制約がある場合には括弧に記す。それぞれの項目にはスベック回帰、すなわち回帰 (regression) { variables=const } において回帰変数として用いられる変数名が与えられる。

<sup>5</sup> 月次のイースター効果で実際に用いられる変数は  $E(w, t) - \bar{E}(w, t)$  である ( $\bar{E}(w, t)$  は  $E(w, t)$  の長期的月次平均であり、グレゴリア歴 1583 年-1982 年の 400 年で計算されている。この数値はより長いイースター日のサイクルから計算される平均の近似になっている) が、詳しくは Montes (1998) を参照のこと。2 月・3 月・4 月のみに 0 でない。同じような季節性を除いた変数は Labor Day、Thanksgiving 効果、四半期 Easter 効果でも使われている。

## 第4章 RegARIMAモデルの推定に関連する事項

---

プログラム X-12-ARIMA で使っている IGLS アルゴリズムと非線形最小二乗法は RegARIMA モデルの最尤推定値を見つける為には信頼できる方法であるが、時には推定上に問題が発生することがある。統計的モデルを推定する上で生じる可能性のある幾つかの問題について以下では発生した時の対処の方法とともに議論する。さらに X-12-ARIMA におけるモデル選択基準の利用に関して重要な注意点も説明する。

### 4.1 母数の初期値と収束問題への対応

利用者は繰り返しによる最尤推定の最適化において AR 母数と MA 母数の初期値を設定することができる。しかしながらこの設定はめったに必要としないし、一般的には設定するようにも推奨もできない。何も指定しなければ X-12-ARIMA では AR 母数と MA 母数の全てについて 0.1 が設定される。(回帰モデル部分については GLS 回帰で推定するので初期値は必要でない。) X-12-ARIMA で自動設定されている値は多くの場合には適当であるように判断される。より適切な初期値を設定したからと言って(例えば条件付最尤推定により推定した値を使うなどして)初期値をわざわざ推定することを正当化するような収束スピードの向上などは期待できそうもない。ただし、例外的としては例えば時系列に幾つかの観測値を追加するなど初期値としてかなり正確な値がわかっているときには設定することでメリットが生じよう。しかしながら初期値を設定することを可能にした真の理由は推定が困難な場合に収束問題に対処するためにあることに言及しておこう。

X-12-ARIMA における繰り返し推定により収束しない場合には幾つかの対処法がある。プログラムが繰り返しの最大数に達して収束を止めた場合には(この場合には警告のメッセージが出て最後の繰り返しで得られた値が出力される)最後の繰り返しで得られた値を初期値と設定してプログラムを再び実行すると収束する可能性がある。より簡単ではあるが計算時間はかかる方法としては繰り返し制約数を増加させプログラムを実行することである。もしプログラムが収束以前あるいは最大の繰り返し数以前にトラブルを発生させた場合にはまず条件付尤度によりモデルを推定し、そこで得られた推定値を初期値として利用して厳密な最尤推定を行なうことができる。他方、我々の経験からは収束問題はしばしば推定するモデルがかなり複雑な場合(例えば高次のモデル等)や適切でない条件のもとで起こりがちである。そのような場合には推定結果を検討し、より単純なモデルを選択することが適切な対処と言えよう。4.2 節から 4.4 節にかけては生じうる推定問題とその場合の適切なモデルの選択についての幾つかの状況について説明する。



## 4.2 MA 項の反転可能性

移動平均 MA 項の多項式  $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$  は方程式  $\theta(B) = 0$  のすべての根  $G_1, \dots, G_q$  が単位円の外にあるとき ( $|G_j| > 1$ ) に反転可能と呼ばれる。例えば Brockwell=Davis (1991, 123 項-125 項) が説明しているように、ARIMA モデルにおける MA 作用素の一つ以上が反転可能でないと同一の自己相関構造を持ち同一の条件付尤度を持つ反転不能な一つ以上の MA 作用素が存在し得る。したがってデータだけでは反転可能なモデルと反転不能なモデルを区別することは出来ないので、反転可能なモデルを選択した方がよい。この問題は予測において特に重要であり、反転不能なモデルを用いると大域的に正しくない予測を行なうことになってしまう。一つの例外としては、MA 多項式の根が単位円上に存在する場合 ( $|G_j| = 1$  となる時) であるが、これは反転可能領域の境界であり MA モデルの厳密な最尤推定を行なうなど適切に対処すれば予測における問題は生じない。

X-12-ARIMA による推定では尤度関数を非線形最適化を反復するとき MA 母数について反転可能性の制約条件を課している。したがって正確に云うと X-12-ARIMA により推定されたモデルは反転可能である。もし最尤推定値 (MLE) が境界上にあれば (すなわち MA 作用素の根が単位円上に厳密にある場合) X-12-ARIMA では非線形最適化により内側から反転可能領域に近づき収束基準、あるいは繰り返し最大数により境界領域に近づく方法をとっている。これにより X-12-ARIMA では結果として反転境界上のモデルを推定できている。このような状況では母数空間の制約の境界上での尤度の最大化は困難な最適化となるので繰り返し収束には時間がかかる。より重要な点として反転可能境界上の推定値の収束にかかわる状況はもとのモデルが適切でないことが多く、利用者へ結果 (及び繰り返し推定における細かな出力) を調べることを警告しているともとれる。4.4 節で不適切なモデル化のもっとも重要な原因となる AR 因子と MA 因子が共通因子を持つ状況や過剰階差問題とその適切な対処法について説明する。

推定するモデルが季節階差や季節 MA 項を含む時に反転不能な結果がもたらされることが少なくない。(すなわち、母数  $\theta$  の MLE が 1 のときに  $1 - \theta B^s$  は反転可能ではない。) また季節 ARIMA モデルは季節的経済時系列ではよく用いられるので利用者は反転可能性問題について注意し、問題が生じたときには 4.4 節で説明している対処法を理解しておく必要がある。

## 4.3 AR 項の定常性

自己回帰 AR 多項式  $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  において  $\phi(B) = 0$  の根がすべて単位円外にあれば定常時系列であるが、そうでなければ非定常となる。ここでより正確に述べると (1) 式において  $\phi(B)\Phi(B^s)w_t = \theta(B)\Theta(B^s)a_t$  にしたがう確率過程  $w_t = (1 - B)^d(1 - B^s)^D z_t$  は AR 多項式のゼロ根すべてが単位円外にあるとき定常過程となる。AR 項を含む厳密な尤度関数はすべての AR 作用素が定常性の条件を満たすものと仮定して導かれる。すなわち AR 項を含む厳密な尤度関数の評価、推定、予測など他の分析はすべてこの AR 項についての条件が満たされていることを仮定して行なわれる。したがって AR 項を含む尤度関数の厳密な評価ではプログラム X-12-ARIMA では推定上で母数の制約条件を課している。次節で述べるように AR 項と MA 項に共通因子が存在する状況でなければ X-12-ARIMA においては尤度が  $-\infty$  に近くなる定常性の境界に近づく可能性は低い。

尤度を AR 母数について条件付で定義すれば X-12-ARIMA でも定常性を仮定したり制約にする必要は無い。この場合にはモデルの推定や予測などは母数の定常性領域に制約されずに実行できる。しかしながら 3.5 節で説明した統計的推測の議論は成立しない。定常条件が無ければ AR 母数の推測については例えば Fuller(1976) の 8.5 節にあるような様々な込入った議論を用いる必要が生じる。

#### 4.4 AR 項と MA 項の共通因子と過剰識別

自己回帰移動平均 (ARMA) モデルを推定する際には AR 項と MA 項が相殺する可能性が生じる。(1) 式 (あるいは (3) 式) において  $p > 0$  かつ  $q > 0$ 、あるいは  $P > 0$  かつ  $Q > 0$  のいずれかが満たされる時に ARMA モデルと呼ばれている。理論的には  $p > 0$  かつ  $Q > 0$ 、あるいは  $P > 0$  かつ  $q > 0$  の場合も ARMA モデルとなるがそのような場合にはこの AR 項と MA 項の相殺問題はほぼ生じない。この問題のもっとも簡単な例としては ARMA(1,1) モデル  $(1 - \phi B)z_t = (1 - \theta B)a_t$  において  $\phi = \theta$  となる状況である。両辺の  $(1 - \phi B)$  要素をとれば  $z_t = a_t$  という簡単なモデルである。したがってこの場合には尤度関数は母数空間上の  $\phi = \theta$  周辺でほぼ一定となる。こうした状況では ARMA(1,1) モデルの MLE がほぼ制約  $\hat{\phi} = \hat{\theta}$  を満たしているため非線形推定の収束が困難となる。また同様の問題は AR 多項式と MA 多項式が共通のゼロ根を持つときに生じることになる。この問題についてのより詳細な議論についてはボックス・ジェンキンス (1976, 248 項-250 項) を参照されたい。

X-12-ARIMA プログラムによるこの混合 ARMA モデルの推定で収束に問題が発生するときにはこの AR 項と MA 項の相殺効果が原因となることが少なくない。どちらにしても AR 多項式と MA 多項式のゼロ根を調べることにより簡単にこの問題を検討できる (コマンド推定 (estimate) 中の print=roots を使えばよい)。仮に共通ゼロ根が見つければモデルを修正して共通因子を相殺すればよいので AR 多項式と MA 多項式の次数を落として再推定すれば良い。この共通因子は厳密な意味では必ずしもなく、AR 項のゼロ根と MA 項のゼロ根がほぼ共通であっても問題が生じるので対処する必要がある。

ところで推定された MA 多項式は階差作用素と相殺する可能性がある。特に非季節階差をとって推定した MA モデルが  $(1 - B)$  因子を含んでいたり季節階差をとって推定した季節 MA モデルが  $(1 - B^s)$  因子を含む場合などに発生することがある。例えばモデル  $(1 - B)(1 - B^s)z_t = (1 - \theta B)(1 - \theta B^s)a_t$  では  $\theta$  か  $\theta$  が 1 に近く推定されていれば問題が生じる。この種の共通因子は過剰階差問題と呼ばれているが、これは時系列を定常化するために必要以上に階差操作を行なったことから生じていると考えられる。階差操作と MA 項に過剰階差が生じてモデルを単純化したときには利用者は階差作用素により前には消されていた非確率的な時間の関数をモデルの回帰項を加えておく必要が生じることに注意しておこう。このことは非季節階差が  $\hat{\theta} = 1$  付近で MA 要素  $(1 - \theta B)$  と相殺する時にはより簡略化したモデルにはトレンド定数項 (あるいは元のモデルが 1 次の階差であれば全平均) を含めるべきということを意味する。もし季節階差が  $\hat{\theta} = 1$  付近で季節 MA 項  $(1 - \theta B^s)$  とが相殺される時にはより簡略化したモデルにはトレンド定数と固定的季節要素を含める必要があることになる。過剰階差問題については Abraham=Box(1978) や Bell(1987) が議論している。

過剰階差を含むモデルでは推定結果が収束するときには、相殺している階差操作と MA 要素を取り除き、適当な回帰変数を付け加えてモデルを推定するのは常によいとは限らないが、それはこの共通因子の問題が必ず推定問題や他の問題を（MA 母数について厳密な尤度関数を利用すると仮定しても）もたらすとは限らないからである。特に予測結果は過剰階差を含むモデルとそれを修正したモデルでは同一の結果をもたらす、回帰や ARMA 母数の推定値のその標準偏差はほぼ同一となる。（むしろ次節で述べるように対数尤度の値や対応するモデル選択基準は異なる。）このことはここでの状況が AR 項と MA 項の共通因子の問題とはかなり異なることを意味している。AR 項と MA 項の共通因子問題では推定の収束問題を生じさせるために発生すれば必ずモデルから取り除き、モデルを再推定することが必要である。

#### 4.5 モデル選択基準の利用

X-12-ARIMA プログラムは次に挙げるモデル選択基準を提供している。すなわち、AIC (赤池情報量基準 Akaike(1973), Findley(1985),(1999) を参照), AICC (Hurvich=Tsai(1989) を参照)、Hannan=Quinn(1979), BIC (Schwarz(1978)) などを提供している。ホワイト・ノイズの分散を含め推定する母数の数を  $n_p$  としよう。さらに階差や季節階差を適用した後のデータ数を  $N$ 、推定により得られた厳密な尤度関数の最大値を  $L_N$  とするとこれらのモデル選択基準は次の式で与えられる：

$$\begin{aligned} AIC_N &= -2L_N + 2n_p \\ AICC_N &= -2L_N + 2n_p \left\{ \frac{N}{N - \frac{n_p+1}{N}} \right\} \\ Hannan - Quinn_N &= -2L_N + 2n_p \log \log N \\ BIC_N &= -2L_N + n_p \log N. \end{aligned}$$

各基準の中では与えられた時系列に対してその値がもっとも小さくなるモデルが選択されることになる。こうしたモデル選択基準を用いる場合には注意すべき幾つかの事項がある。第一にこうしたモデル選択が可能なのは同一の次数の階差モデルに対して正当化が可能なことである。したがって、モデル選択基準を階差操作の次数選択に使うことはできないことになる。第二にはモデル選択を可能とするには AR 母数と MA 母数について厳密な最尤法を使わねばならないことであるが、X-12-ARIMA ではこの推定方法が自動選択される。このことから X-12-ARIMA では条件付尤度関数を使った場合にはモデル選択基準の値は出力されない。

さらにモデル選択基準は異なる外れ値 (変数 AO, LS, TS, ramp) の回帰変数を含む RegARIMA モデルには利用すべきでないことである。特にどの外れ値をモデルに含めるか否かを定める際には用いるべきでない。この点についての理論的問題としてモデル選択基準を導出するにあたってはモデルの未知母数は将来データが得られるにつれてより正確に推定できることが仮定されているが、外れ値の母数に関してはこのことは成り立たないことから生じる。

なお同一の外れ値変数を含むモデルでモデルを比較するときには経験的か直感的には外れ値変数に

よる回帰推定では外れ値は最尤推定値に対しては同一の方向で影響を与える（つまりその影響は対数尤度の差を考える限り相殺する）ことが云えるようである。例えば AO 変数を時刻  $t$  に使うと時刻  $t$  の観測値が欠落しているかのように扱ってもほぼ同様の効果があり（Gomez=Maravall=Pena(1999)を参照）、本質的には比較する複数のモデルにおいて当該観測値を欠測値と扱ってかのような状況になることが知られている。これに対して、変数 AO を回帰変数とすると実際には異なるデータ・セットからモデルを推定するような状況となり、対数尤度関数やその対数尤度から計算される他の統計量を比較することについての正当化の根拠はほとんどない。（この点は二つのモデルを比較する際に季節階差と非季節階差の次数を同一に取る必要があることと本質的に同じである。）尤度が階差データから計算される時には階差の次数が異なるとデータは本質的に異なると見なければならぬ。）

実際に異なる外れ値変数を含むモデルを比較してみると通常は多くのこうした変数を含むモデルが選択される傾向にあるが、予測はより悪くなる可能性がある。さらに選択された外れ値のかなりが 1 年後、あるいは二つのデータが加わると外れ値として識別されないことも少なくない。すなわちモデル識別がかなり不安定となる可能性がある。

変数変換や事前調整はモデル選択基準に影響を与える。例えば変換 (transform) の中に変数として含まれない関数 (function) による変数があったり  $\lambda \neq 1$  であるべき (power) 変換があると変換コマンドは原系列に関する尤度関数に影響を及ぼし、その結果モデル選択基準も影響される。このことから X-12-ARIMA では原時系列に対する対数尤度を計算する際には変数変換調整を計算し、その変数変換調整済み対数尤度によりモデル選択基準が計算される。このように与えられた時系列に対して異なる変数変換、あるいは事前調整などの要素が処理されている。したがって X-12-ARIMA にかける入力前に時系列が変数変換されていると（つまり例えば原系列の対数変換値が X-12-ARIMA 読み込まれたりすると）変数変換を X-12-ARIMA の中で行なったときは異なるモデル選択基準の値となることに注意しておく。同様のことが事前調整要素についても当てはまることにも注意しておく。したがって、モデル選択基準を用いる場合には同一の入力系列に対する異なる統計モデルの比較に限定すべきである。このことから系列 (series) の中で期間 (span) を用いて系列の異なる部分に対してモデル選択基準を用いることができないことも明らかであろう。

## 第5章 スペック・ファイルと文章構成法

---

X-12-ARIMA の主たる入力はスペック・ファイルと呼ばれている特別な入力ファイルで作られる。このファイルは一連の命令文 (コマンド)・スペック (specs) によりデータ、利用したい季節調整オプション、出力、利用する時系列等々の情報を X-12-ARIMA に与える。ここで幾つかのスペック命令文としては次のようなものがある。

**系列 (series)** : このスペック命令は集計値の調整を除き常に必要である。このスペックにより時系列データ、開始時点、季節周期、分析期間、系列のタイトルが指定される。

**集計 (composite)** : このスペック命令により直接的あるいは間接的な集計値の調整が指定され、スペック命令である系列 (series) の代わりに用いられる。

**変換 (transform)** : このスペック・コマンド (命令) ではデータ系列の変換や事前調整を指定する。

**x11** : このスペック命令は季節調整オプションを指定し、季節調整の形式、季節成分とトレンド成分へのフィルター、復活祭調整オプション、季節調整の診断オプション等を指定する。

**x11 回帰 (x11 regression)** : このスペック命令は非標準的回帰オプション、例えば回帰変数の指定や不規則要素に対する頑健な回帰を行なうために指定する。

**自動選択 (automdl)** : このスペック命令は自動的モデル選択の方法を指定する。

**アリマ (arima)** : RegARIMA モデルの ARIMA モデル部分を指定する。

**回帰 (regression)** : このスペック命令は RegARIMA モデルの回帰部分で利用する説明変数を指定し、識別 (identify) により取り除く回帰効果を指定する。

**推定 (estimate)** : このスペック命令では回帰 (regression) とアリマ (arima) により特定化されたモデルの推定と尤度の評価、それから推定におけるオプションを指定する。

**チェック (check)** : このスペック命令は推定されたモデルを診断する際に有用な統計量を指定する。

**予測 (forecast)** : このスペック命令は推定したモデルに基づく予測を行なうことを指定する。

**外れ値 (outlier)** : このスペック命令は加法的な外れ値や水準の変化の自動検出を指定する。さらに、一時的な水準変化の検定が利用可能である。

**識別 (identify)** : このスペック命令により決められたデータの階差次数を用いて ARIMA モデル識別の為に回帰効果を含んだ (回帰 (regression) により別に定める) 自己回帰関数と偏自己回帰関数を計算する。

**期間変更の安定性 (slidingspans)** : このスペック命令により季節調整値の安定性について期間変更の安定性分析を行なう。

**履歴 (history)** : このスペック命令により季節調整の改定値の履歴や関連する RegARIMA モデルの統計量を求めることができる。

各スペック命令はその名前によりスペック・ファイルとして定義され、それに続いて変数 (arguments) と指定された値 (values) を含んだ {} との形をとる。変数と指定された値は変数 (argument) = 値 (value) あるいは複数の数値が必要な時には変数 (argument) = (値<sub>1</sub>, 値<sub>2</sub>, ...) という形になる。ここで様々な値 (value) の形があり、タイトル、変数名、鍵言葉、数値、それに日付等がある。こうした例は 6 章の個々のスペック命令文についての説明で定義されたり例示されることになる。幾つかのスペック命令においてしばしば表われるので 5.1 ではスペック命令である印刷 (print) と保存 (save) における変数 (arguments) について、5.2 節では日付 (date) の変数について説明する。

以下で説明するが系列 (series) あるいは集計 (composite) を除いては決められた変数は必要としない。ほとんどの変数はあらかじめ指定された値を持っていてその値は各スペック命令の解説に書かれている。特に指定が無ければ利用する変数やパラメータ値すべてについてあらかじめ指定した値が使われる。

一般的にはある一つのスペック・ファイルにおいて多くのスペック命令は必要としない。実際、ほとんどすべての X-12-ARIMA の計算 (複合的計算は除かれるが) は一つだけスペック・ファイルとして必要な命令の系列 (series) があれば実行される。このスペック命令ではデータ (data) あるいはファイル (file) 変数が必要である。ただし、例外としてはデータ・メタ・ファイルを使う場合があるが、この場合についてより詳しくは 2.5.2 節を参照されたい。したがって X-12-ARIMA にとり最低限のスペック・ファイルとしては一系列 (series) data=(データの値 (data values))—となる。季節調整を実行するにはスペック・コマンドの x11 が必要であり、もしそうでなければスペック命令である x11 回帰 (x11 regression), 期間変更の安定性 (slidingspans), 履歴 (history) (推定 (estimates) を含めて季節調整履歴を調べる) などの中で少なくとも一つが必要となる。その場合には X-12-ARIMA ではスペック・コマンド x11 が規定値として指定されていると扱うのでスペック・ファイルの中に—x11—と書かれていることと同等になっている。モデルの識別を実行するには識別 (identify) が必要であり、モデルの推定には通常はスペック命令アリマ (arima), 回帰 (regression), 推定値 (estimate) が必要となる。推定 (estimate) がなくてもスペック命令として外れ値 (outlier), 自動モデル (automdl), チェック (check), 予測 (forecast), x11, 期間変更の安定性 (slidingspans), 履歴 (history) 等が一つでもあればモデルの推定は行なわれる。その場合には X-12-ARIMA ではコマンド推定値 (estimate) があたかも存在するかのごとくに扱い、スペック・ファイルに—estimate—と書かれていることと同等となる。またスペック命令のアリマ (arima) がなければ推定は ARIMA(0 0 0) モデル (すなわちホワイト・ノイズ) の指定として推定

が実行されることになる。この場合はスペック・ファイルに—arima—と書かれていることと同等である。

一つの例外を除きスペック・ファイルに書かれている指示の順番とスペック命令のカッコ内の変数の順番は問題ではなく同一の結果をもたらす。ここでの例外は系列 (series) や集計 (composite) であるが、これらのコマンドは最初のスペック命令として書かれなければならないことである。スペック・ファイルの形式は自由であるが空白、タブ、空白行などを用いるとスペック・ファイルが読みやすくなる。コメントを付ける事もできる。入力表示にコメントやその他のルールの使い方については 5.3 節で議論する。

**重要事項** スペック・ファイルの最後には必ず改行を指定すべきである。そうでないと当該行が読まれないことになるが、これはフォートラン言語の制約である。

次に挙げる例 5.1 によりデフォルト X-11 を実行させる単純なスペック・ファイルを示しておく。このファイルを実行した出力ファイルのスペクトル診断によると曜日効果の存在を示唆しているが、この結果についてのメッセージが出力には書かれている。RegARIMA モデルを用いれば季節調整に先立ち曜日効果と予測により系列を延ばすことと同一の事を行なうことができる。例 5. 2 と例 5. 3 では最終の季節性と曜日効果を取り除く (例 5. 4) 前にモデルの ARIMA モデル部分を識別するスペック・ファイルを示している。ここでももちろん、X-11 のスペック・ファイルとともに 6 章で説明する X-11 の曜日効果の除去法を利用する事もできよう。

例 5.1 : デフォルト X-11 用の X-12-ARIMA スペック・ファイル

```
series{title = "Monthly Retail Sales of Household Appliance Stores"
data = ( 530 529 526 532 568 785 543 510 554 523 540 599
        574 619 619 600 652 877 597 540 594 572 592 590
        632 644 621 604 613 828 578 533 582 605 660 677
        682 684 700 705 747 1065 692 654 719 690 706 759
        769 730 740 765 791 1114 695 680 788 778 780 805
        852 823 831 836 913 1265 726 711 823 780 844 870
        865 915 920 935 1030 1361 859 852 954 895 993 1109
        1094 1173 1120 1159 1189 1539 1022 987 1024 1005 1054 1098
        1191 1191 1161 1201 1294 1782 1154 1059 1178 1126 1120 1233
        1260 1311 1302 1365 1395 1899 1123 1087 1210 1157 1159 1260
        1357 1265 1231 1287 1452 2186 1309 1242 1388 1400 1397 1527
        1654 1650 1555 1560 1836 2762 1541 1480 1619 1455 1510 1698
        1651 1749 1783 1863 2074 3051 1836 1690 1856 1796 1904 1927
        1978 2055 1976 2204 2423 3502 1977 1767 1935 1900 2073 2143
        2299 2247 2162 2274 2529 3731 2184 1901 2058 1974 2018 2091
        2239 2253 2157 2190 2397 3659 2170 2086 2297 2251 2311 2520)
start = 1972.jul}
x11{}
```

時系列データに対してモデル分析を行なうときには少なくとも X-12-ARIMA を 2 度実行していただくことを薦める。最初の実行はモデルの ARIMA モデル部分を識別することの為であり、二度目の実行は RegARIMA モデルを推定し、可能ならモデルを使って予測する為である。最初の実行の為のスペック・ファイルでは系列 (series) と識別 (identify) を使うと共に変換 (transform) と回帰 (regression) 等を用いる。二度目の実行ファイルで系列 (series), アリマ (arima), 推定 (estimate) 等を使い、さらに必要なら変換 (transform), 回帰 (regression), 外れ値 (outlier), チェック (check), 予測 (forecast) 等を用いる。プログラム X-12-ARIMA を二度実行させるときには二つの異なるスペック・ファイルが必要であるが、実際上は最初のスペック・ファイルを修正し

て二度目の実行ファイルを作れば良い。もし診断により推定されたモデルを修正することが必要であればスペック・ファイルをさらに訂正して X-12-ARIMA をもう一度実行すればよい。

#### 例 5.2 : RegARIMA モデル識別用の X-12-ARIMA スペック・ファイル

```
series{title = "Monthly Retail Sales of Household Appliance Stores"
data = ( 530 529 526 532 568 785 543 510 554 523 540 599
        574 619 619 600 652 877 597 540 594 572 592 590
        .
        .
        .
        2239 2253 2157 2190 2397 3659 2170 2086 2297 2251 2311 2520)
start = 1972.jul}
transform{function = log}
regression{variables = td}      # Comment: Series has trading-day effects
identify{diff=(0, 1) sdiff = (0, 1)}
```

モデルの識別を実行する為の典型的なスペック・ファイルは例 5.2 の形式のようになろう。この例では系列 (series), 変換 (transform), 回帰 (regression), 識別 (identify) などのスペック命令を使っている。まずデータが系列 (series) の中に定義され、系列の対数変換を行ない (変数変換 (transform) による)、系列に影響すると知られていたり可能性がある変数を回帰変数をコマンド回帰 (regression) により定める。ここで変数 (variables) = td は 6 個の曜日効果変数 (td6) を含むと共にうるう年効果を調整することにも言及しておこう。(3.3 節と 6 章の回帰 (regression) の説明を参照されたい。) コマンドの識別 (identify) により変換された変数の階差 (月の長さも調整して) に回帰変数 (6 個の曜日効果の変数) により回帰を行なう。ここで季節階差と非季節階差の最高次  $(1 - B)(1 - B^{12})$  を適用して回帰を行なっている。それからスペック・コマンドの識別 (identify) により原系列の回帰残差系列を計算し、季節階差と非季節階差のすべての組み合わせについて標本自己相関関数と標本偏自己相関関数の表をプロットして出力する。

最初の出力結果を検討し ARIMA モデル部分の識別 (例えば  $(0\ 1\ 1)(0\ 1\ 1)_{12}$  であったとすると) を行なう識別 (identify) の後にアリマ (arima) と推定 (estimate) をスペック・ファイルに付け加えることが考えられる。結果として得られるスペック・ファイルは例 5.3 で与えられる。(ただし、データのすべては載せていない。)

#### 例 5.3 : RegARIMA モデル推定用の X-12-ARIMA スペック・ファイル

```
series{title = "Monthly Retail Sales of Household Appliance Stores"
data = ( 530 529 526 532 568 785 543 510 554 523 540 599
        574 619 619 600 652 877 597 540 594 572 592 590
        .
        .
        .
        2239 2253 2157 2190 2397 3659 2170 2086 2297 2251 2311 2520)
start = 1972.jul}
transform{function = log}
regression{variables = td}      # Comment: Series has trading-day effects
# identify{diff=(0, 1) sdiff = (0, 1)}
arima{model = (0,1,1)(0,1,1)}
estimate{print = iterations}
```

このスペック・ファイルには命令として系列 (series), 変換 (transform), 回帰 (regression), アリマ (arima), and 推定 (estimate) が含まれている。ここでコマンドの回帰 (regression), ア



リマ (arima)、推定 (estimate) により次のモデルを指定し推定している：

$$(1 - B)(1 - B^{12})\left(y_t - \sum_{i=1}^6 \beta_i T_{it}\right) = (1 - \theta B)(1 - \Theta B^{12})a_t,$$

ただし  $T_{it}$  は 6 個の曜日回帰変数である。系列  $y_t$  は原データ系列からうる年効果を調整して対数変換を行なってからモデル分析が行なわれる。残差分析、外れ値の検出、あるいは予測などを行ないたいければそれに応じたスペック・コマンドを付け加えればよい。

求めた統計的モデルが適当である事を仮定するとスペック命令の `x11` と予測 (`forecast`) を入力スペック・ファイルに加える事により予測による外挿を利用する季節調整を行なう事ができる。そのようなスペック・ファイルとして例 5.4 を挙げておく。なおデータは一部分のみでは省略した。

#### 例 5.4 : 季節調整用の X-12-ARIMA スペック・ファイル

```
series{title = "Monthly Retail Sales of Household Appliance Stores"}
data = ( 530 529 526 532 568 785 543 510 554 523 540 599
        574 619 619 600 652 877 597 540 594 572 592 590
        .
        .
        .
        2239 2253 2157 2190 2397 3659 2170 2086 2297 2251 2311 2520)
start = 1972.jul}
transform{function = log}
regression{variables = td}      # Comment: Series has trading-day effects
# identify{diff=(0, 1) sdiff = (0, 1)}
arima{model = (0,1,1)(0,1,1)}
estimate{print = iterations}
forecast{maxlead = 60}
x11{seasonalma = s3x9}
```

このスペック・ファイルにより  $3 \times 9$  季節フィルター (`x11`) を曜日効果を調整した系列に適用して季節調整を行なうことができる。事前調整した系列は季節調整に先立ち計算した 60 期間の予測値 (予測 (`forecast`)) により外挿する。主要な出力ファイルにより得られた季節調整の良さについての診断ができる。また 6 章に説明する適当なスペック命令を含める事による追加的な診断結果を入れる事もできる。

## 5.1 プリントと保存

X-12-ARIMA からの出力の制御はスペック命令である印刷 (`print`) と保存 (`save`) の変数を使って行なう。印刷 (`print`) の変数は主要な出力ファイルへの出力を制御し、保存 (`save`) の変数はファイルに書き出す出力の表を制御する。説明の都合上、これら印刷 (`print`) と保存 (`save`) により制御される出力を表 (Tables) と呼ぶことにする。むろん、この種の表の中には (例えば ACF のプロットなど) 通常は表の形とはみなされない出力も含まれている。スペック命令である印刷 (`print`) と保存 (`save`) により制御されている表は印刷 (`print`) と保存 (`save`) の変数として説明されているように、指定されたプリント形式と拡張子とともに一覧となっている。保存 (`save`) を用いた出力表は精度が確保される形式で書かれているが、他のソフトウェアでさらに分析することができるようにラベル情報は最小限としている。保存された表は単一のタブをそれぞれ含む一貫した形式になっている。

スペック命令である印刷 (**print**) がなかったり (`—print=default—`)  
あるいは `—print=()` となっていればスペックであらかじめ指定している出力が主用な出力ファイルに書き出される。主たる出力ファイルへ出力することを止めたければ `print=none` とする必要がある。(ただし、このスペック `print=none` によっても幾つかのラベル情報などがスクリーンに書き出されることに注意しておく。) 他方、主な出力ファイルに書かれうるすべての利用可能な出力表とプロットを得るには `print=all` と設定することが必要である。また主な出力ファイルに書かれる (プロットを除く) 出力ファイルすべてを得るには `print=alltables` と指定すればよい。画面上に書き出される出力表の幾つかのみ得るには `print=brief` と指定すればよい。スペック・コマンドの `default`, `brief`, `none` などの出力水準に対して、さらに個々の表を加えるにはその名前を変数として付け加えればよい。これらのコマンドは必ずと言うわけではないが+の後にかかれる事が多い。例えば推定 (`estimate`) の中で `print = (+iterations + residuals)` は

```
print = (default + iterations + residuals)
```

と同等であるが、これはデフォルト `default` の出力に加えて、あらかじめ指定された繰り返し推定と推定された残差の結果の出力を意味している。コマンド `print=(none estimates)` を使うと出力は母数推定値のみとなる。デフォルトでは個々の表は出力されないが、出力変数名として- (すなわち `print=brief -acf` や `print=(all -iterations)`) とするとすべて出力される。利用者がどれかの出力表をファイルに保存したい時には適当なスペック・ファイルの中に保存 (`save`) 変数を書いておかなければならない。例えば推定値 (`estimate`) の中に `save=(mdl estimates)` とする等である。保存可能な表は印刷 (`print`) と保存 (`save`) など中にリストにある長い (`long`) 方の名前をスペック命令の記述に用いるか、保存する表の拡張子と同一の短い (`short`) 3文字名を用いることにより得る事ができる。例えば推定 (`estimate`) の中の `regcmatrix` は `rcm` と同一の指定である。ただし、既に出てきた鍵言葉である `none`, `all`, `alltables`, `default`, `brief` などは保存 (`save`) の中では使えない。また保存する表の名前は+か-を前につけてはならない。すべての表が保存できるとは限らないし、またいずれのスペック・ファイルでも表を保存する事ができるわけでもない。

保存 (`save`) により各ファイルに指定の表を書き込む事ができるわけであるが、保存されたファイルは出力と同一のディレクトリーに置かれ、主な出力ファイルと同一のファイル名と3文字の拡張子名を持つことになる。もし同じ名前のファイルが既に存在していれば新しく書き換えられる。ここで用いられる拡張子は印刷 (`print`) と保存 (`save`) の説明の中で一覧となってリストアップされている。例えば DOS 計算機上で入力ファイルとしてディレクトリー `C:\TSERIES` の中の `SALES.SPC` を使って `X-12-ARIMA` を実行させるとしてみよう。推定 (`estimate`) が `save = (mdl estimates)` を含んでいる時には結果として保存されるモデルの表と母数の推定値はファイル `C:\TSERIES\SALES.MDL` 及び `C:\TSERIES\SALES.EST` に書かれる。`X-12-ARIMA` によって使われる拡張子は明らかな矛盾を避ける (すなわち `.dat`, `.exe`, `.com`, `.for`, `.spc` など) が、利用者は `X-12-ARIMA` により保存されているファイルが意図しないうちに書き換えられてしまうことを注意深く避ける必要がある。保存したファイルのリストが出力の最初に書かれてるが、\*がついているファイルは存在していたファイルが書き換えられたことを意味する。またスペック・ファイルにエラーがあったり、その他の理由でプログラムが実行を停止した場合には幾つかあるいはすべての保存ファイルが書かれていない可能性もある。

## 5.2 日付

日付変数は幾つかのスペックに表われるが、その値は同一の形式に規定されている。月次データの日付は *year.month* であるが、この形式は他の周期（例えば *year.quarter*）等に一般化されている。年を表示するには4桁すべてを含めることが必要である。したがって67はAD(あるいはCE)67年でありAD1967年ではない。

月次データでは月は整数1から12、あるいは3文字の略月名 (*jan, feb, mar, apr, may, jun, jul, aug, sep, oct, nov, and dec*) を用いる。したがって1967.12と1967.*dec*は同一月を意味する。四半期データやその他の季節周期のデータでは整数のみが使われるので例えば1967.1や1967.4となる。

日付は系列の初期時点を定義したり時系列の一部を分析する際に使われる。また外れ値の回帰変数を指定する時にも利用される。例えば1978年4月に加法的な外れ値の為に回帰変数を指定し、1982年9月に水準の変化が始まることを指定するのは次のように書く。

```
regression { variables=(ao1978.apr ls1982.sep) }
```

日付けの季節性はスペック命令の系列 (*series*) の中のデータの季節性に一致している必要がある。すなわち、*ao1972.jan* は月次データには適当であるが、四半期データには使用が認められていない。

## 5.3 入力文での一般ルール

### 入力文字 (Allowable input characters)

引用で使われる文字を除き入力として認められる文字は英文字、数値、空白、タブ、新行 (*newline*) 文字、及びその他: `. , { } ( ) [ ] + - #` である。プログラムではスペック・ファイルにおける他のASCII文字は無視するが、それはフラッグや警告文が作られることでわかる。引用文の中では次のような文字がさらに使える: `! % ' * / : ; | ? @ \ _ | ^ `` 等である。二重の引用も複数の引用文で区切られた文の中では認められる。

### 大括弧 (Braces), 丸括弧 (parentheses), 角括弧 (brackets)

文字 `{ } , ( ) , [ ]` はそれぞれ異なる機能があり、互いに代替はできない。文字 `{ }` は変数を含むスペックに使われ、`( )` は変数への複数の数値リストを含むときに利用され、`[ ]` が利用されるのは (i) 例えば復活祭の回帰変数など特殊な変数を定義する値、すなわち回帰 (*regression*) `{ variables = (td Easter[14]) }` 等、あるいは (ii) 一部のラグの欠けたARIMAモデルに表われるラグ、たとえば `arima { model = (0 1 [1,3]) }` などで利用される。

### 活字の大きさ (Case sensitivity)

スペック名・変数・日付・鍵言葉 (例えば *none* と *all*) 等、事前に定められている回帰変数名 (*td* や *seasonal* など) は活字の大きさは問題とならない。したがってX-12-ARIMAではTDとtdは同一の変数であり共にスペック命令の回帰 (*regression*) やx11回帰 (*x11regression*) では変数 (*variables*) と認識される。

### コメント (Comments)

記号#の後に来る1行は#が引用文の中のない限りコメントとなる。スペックに一部分だけにコ

メントが必要なら、大括弧・丸括弧・角括弧で区切りをつけないなければならない。

### 等号 (Equals sign)

変数に値を指定するには等号=が用いられる。すなわち `print = none` あるいは `title = "Monthly Retail Sales of Household Appliance Stores"` などである。

### スペック・ファイルでの文長

スペック・ファイルでの行は132文字に制限されている。132文字を超える欄は無視される。したがって特にデータがスペック・ファイルの中で132文字を超える行を含んでいると入力データが切断されることになることに注意が必要である。ただし、1行132文字の制約は(スペック・ファイルでない)コマンドファイル(file)によって別のファイルから読み込まれるデータには適用されない。他のファイルの法はフォートラン(FORTRAN)での入力長の制約に依存しているのでシステムによって異なりうる。

### 重変数値

複数の数値からなる重変数値は例えば `variables=(td seasonal const)` などのように括弧でまとめなければならない。一変数は一つの値をとるか複数の数値を一つとしてとるが、一つの場合は括弧は付けても付けなくても良い。つまり `variables=td`, `variables=(td)`, `variables = (td seasonal)`, `start=1967.4`, `start=(1967.4)` などはずべて可能である。

### 空白リスト (Null list)

空白の変数、例えば `outlier{}` などを使う事ができる。空白の変数リストの変数はあらかじめ指定されたデフォルトの値をとることになる。

### 数値 (Numerical values)

数値は指数記号を含め自由な形式で使うことができる。例えば 400, 400.0, 400. と 4.e+2 はすべて同一な値である。

### 順番 (Ordering)

スペック・コマンドの順番についての唯一の制約は系列(series)あるいは集計(composite)が始めのスペックとなることである。スペックの中では変数の順番には制約がない。むろん、複数の値が指定される時には順番が依存することが明らかな場合もあり、例えばデータの変数、(系列(series), 変換(transform), 回帰(regression)など), モデル(model)の変数(アリマ(arima)など ARIMA モデルの特定化、期間(span)変数の日付(スペック系列(series)と外れ値(outlier))など)である。

### 区切り (Separators)

空白、タブ、空白行を文や行を区別するために使う事ができる。重変数値のリストの中ではカンマを `data=(0, 1, 2, 3, 4, 5)` などと用いる事ができる。欠落値をあらかじめ指定するデフォルト値で置き換える場合(数値が必要となる場合)には必ずカンマを用いなくてはいけない。例えばスペック命令の期間(span)では二つの値が要求される。コマンド文が仮に `span=(1967.4, )` となっている時には 1967.4 の後にカンマが存在するので期間(span)の二つ目の数値が欠けていることを意味する。この時にはあらかじめ定められた値(最後の観測の日付)が使われることになる。

### タイトルとファイル名

時系列名などのタイトルは少なくとも1つの許容される入力文字を含み、空白があっても('title'

or "title") などと書かれている必要がある。タイトルでは大文字と小文字の区別は維持される。タイトルに文字#が表われるときにはタイトルの一部分と見なされコメントとは見なされない。タイトルは1行79文字以内でなければいけない。またパス名を含みファイル名もタイトルについてのこれらの規則が適用される。

## 第6章 各スペック命令文の解説

---

次項から各スペック命令文をそれぞれ利用可能な変数とデフォルトで指定されている値を含め説明する。各スペック命令は幾つかの例を使って利用法を説明する。スペック・系列 (series) と変換 (transform) で説明している例は単にこれらのスペック命令の機能についての例示にすぎないのでスペックの機能をすべて説明しているわけではない。すなわち、X-12-ARIMA の入力ファイルとして例を使うと、有用な出力ファイルが得られるとは限らない。その他のスペック命令、すなわち(集計 (composite), x11, 識別 (identify), 回帰 (regression), アリマ (arima), 推定 (estimate), 外れ値 (outlier), チェック (check), 予測 (forecast), 期間変更の安定性 (slidingspans), 履歴 (history)) などでは例としては...により表されているデータそのもの (スペック系列 (series)、あるいは利用者が回帰 (regression) の中で利用する回帰変数など) 以外の部分はそのままでスペック・ファイルとして使う事ができる。読者は各スペック命令の解説で与えられている例は中身ばかりかファイルの形式も様々であることに気がつくであろう。これは意図的に利用者がスペック・ファイルの形式を定める時にかなり柔軟にできることを例示し、強調したことの結果である。

### [訳注]

より正確さを期すために第6章で説明している各スペック・ファイルに関連した出力やログ・ファイル診断の表の大部分は原文のままに掲載しておいた。これらを日本語に変換すると誤解が生じる可能性が高いと判断したが、あらかじめ設定してある回帰変数については表 3.1 に日本語に訳しておいたので参照されたい。

## アリマ (ARIMA)

---

### 解説

regARIMA モデルの ARIMA 部分を指定する。スペック命令 **regression** がないときは純粋な ARIMA モデルを定義することになる。モデルの ARIMA 部分には乗法的季節要素や、ゼロ係数を指定するラグ作用素も含まれる。繰り返し推定を行うにあたって、AR、MA パラメータの初期値が **ar** 変数と **ma** 変数により与えられる。また、あるパラメータを初期値のまま固定し、その他のパラメータの推定を行うこともできる。

### 使用法

```
arima \{ model = ([2 3] 1 1)(0 1 1)12
          title = "ARIMA Model"
          ar    = (0.3f, -0.14)
          ma    = (-0.7 0.85f) \}
```

### 引数

#### **ar**

非季節、季節自己回帰パラメータの初期値を、model 変数中に現れた順に与える。この変数があるときは、モデル中の全ての AR パラメータへ初期値を与えなければならない。パラメータの初期値を与えるには変数リストに数値を与えてもいいし、ゼロ係数を指定するラグ作用素であることを示してもいい。何も指定しないと欠損値としてデフォルトである 0.1 をとることになる。例えば、2つの AR パラメータがあるモデルにおいて、**ar**=(0.7, ) というのは **ar**=(0.7, 0.1) と同じことだが、**ar**=(0.7) とはできない。3つの AR パラメータがあるモデルについて、**ar**=(0.8, , -0.4) というのは **ar**=(0.8, 0.1, -0.4) と同じである。推定の際、パラメータをその初期値に固定しておきたいときは、**ar** リスト中、数値の後に '{f}' とつければよい。(e.g., **ar**=(0.7f, 0.1))

#### **ma**

自己回帰パラメータにおける **ar** と同様に、全ての移動平均パラメータに初期値を与える。

#### **model**

モデルの ARIMA 部分を定義する。形式は Box-Jenkins (1976) の標準的な表記にしたがう。この表記では、非季節 ARIMA モデルは  $(p\ d\ q)$  で表される。ここで、 $p$  は非季節的 AR 項の次数で  $d$  は非季節階差の次数、そして  $q$  は非季節的 MA 項の次数である。乗法的季節 ARIMA モデルは  $(p\ d\ q)(P\ D\ Q)$  により表される。ここで、 $p, d$  そして  $q$  は前と同様で、 $P$  は季節的 AR 項の次数、 $D$  は季節階差の次数で  $Q$  は季節的 MA 項の次数である。ここでは一番目の ARIMA 要素である  $(p\ d\ q)$  は非季節的 (i.e., その周期は 1) で、二番目の ARIMA 要素  $(P\ D\ Q)$  は季節周期が **series** スペックで与えられたものであるとみなされる。2つ以上の ARIMA 要素を与えることもできるし、ARIMA 要素にデフォルトと異なる季節周期を与えることもできる。詳しくは DETAILS を

参照のこと。

ARIMA 要素のオペレータの次数 ( $p d q$ ) は空白、またはコンマによって区切られ、例えば (0 1 1) と (0,1,1) は同じことを意味する。オペレータ中のラグに欠損値が含まれているときは、ラグのうちで存在するものを括弧でくくって表す。例えば、`model = ([2 3] 0 0)` は  $(1-\phi_2 B^2-\phi_3 B^3)z_t = a_t$  というモデルを意味する。

**print and save**

このスペック命令では表が出力されない。

**title**

ARIMA モデルのタイトルを引用符の中に表示する。使える文字数は 80 字以内である。タイトルは ARIMA モデルの記述と推定表の上に出力される。デフォルトでは ARIMA Model となる。

## 詳細事項

好きなだけ ARIMA 要素をモデル変数に含めることができるが、モデル中の AR、MA、階差係数は合計で 108 を超えてはならない。また、全ての AR、MA パラメータの最大ラグ数は 36 であり、全ての非季節、季節 ARIMA 要素の階差の最大値は 3 である。(最後の 2 つの制限は変更できる。—2.8 節を参照)

ARIMA 要素は通常、( $p d q$ )<sub>s</sub> という形式で表される。ここで、s はオペレータにおける季節周期である。従って、モデル変数に (0 1 1)<sub>6</sub> と書けば、 $1-B^6$  なる階差をとり、MA 項  $1-\theta_6 B^6$  を含むことになる。ARIMA 要素の後ろの季節周期 s が省略されているときは、次のデフォルトルールに従って s の値が決められる。1 番目の ARIMA 要素で季節周期が省略されれば、それは非季節モデル、つまり季節周期が 1 のモデルが仮定される。2 番目の ARIMA 要素で季節周期が省略されているときは、**series** スペックで決められた季節周期をもつようなモデルが仮定される。例えば、**series** スペックに `period = 12` とあるとき (または開始時が `year.month` の形で与えられ、周期が 12 であるとき)、`model = (0 1 1)(0 1 1)` と `model = (0 1 1)1(0 1 1)12` は同じものを意味する。さらに ARIMA 要素が追加されれば、季節周期が与えられない限り、それらは非季節的であることが仮定される。3 つの ARIMA 要素をもつモデルが例 6 に与えられているので参照のこと。また、季節周期が 1 のときは、季節周期が与えられていない ARIMA 要素は周期 1 をとることに注意。

MA 多項式の根が単位円の内部にあるように MA パラメータの初期値を与えてはならない。(Section 4.4. を参照。) そうしてしまうと、プログラムが止まり、初期値を与えなおして再度プログラムを動かすようにとのメッセージが出てくる。MA 多項式の根が単位円上に存在するような初期値は、反転不可能な多項式を推定しなくてよいつきのみ認められる。つまり、全く推定が行われないうち、もしくはその多項式中のパラメータ値が推定の間中、ずっと固定されているときのみそれが認められる。例えば、モデルが 1 次の季節 MA パラメータを含み、それが唯一の MA パラメータであるときに、`ma=(1.0f)` とすることは常に可能で、`ma=(1.0)` はそれが推定されないときのみ可能であるが、`ma=(1.1)` は常に認められない。

AR 多項式の尤度関数が厳密であれば (デフォルトである `exact=arma` に設定されていれば—**estimate** スペックを参照)、AR 多項式が非定常になる、つまり単位円の上、または内部に根が存



在するように AR パラメータの初期値を与えることはできない。そうしてしまうと、プログラムが止まり、初期値を与えなおして再度プログラムを動かすようにとのメッセージが出てくる。

ARIMA モデルの係数を固定すると、*AIC* や当てはまりの良さを確かめるための診断統計量、その他のモデル選択統計量が有効に働かなくなることもある。詳しくは `estimate` や `check` のセクションを参照されたい。

## 例

次に完全な形のスペック・ファイルの例を挙げる。

### 例 1

1 階の階差とラグ 1 の MA パラメータのある非季節 ARIMA モデル  $(1 - B)y_t = (1 - \theta B)a_t$  を推定する。

```
series { title = "Quarterly Grape Harvest" start = 1950.1
         period = 4
         data = (8997 9401 ... 11346) }
arima { model = (0 1 1)}
estimate { }
```

### 例 2

原系列の対数をとった  $y_t$  に対して季節 ARIMA モデル  $(1 - \phi_1 B - \phi_2 B^2)(1 - B)(1 - B^{12})y_t = (1 - \Theta_{12} B^{12})a_t$  を定義、推定する。スペック命令 `series` で開始時が月次の形で書かれているので、季節周期は 12 となることに注意。

```
series { title = "Monthly sales" start = 1976.jan
         data = (138 128 ... 297) }
transform {function = log}
arima { model = (2 1 0)(0 1 1)}
estimate { }
```

### 例 3

固定季節効果とトレンド定数を含み、回帰誤差が ARIMA (0 1 1) に従うような regARIMA モデルを定義、推定する。このとき、モデルは  $(1 - B)(y_t - \sum \beta_i M_{it} - c \cdot t) = (1 - \theta B)a_t$  なる形をしている。ここで  $M_{it}$  は固定季節効果を表す回帰変数である。

```
Series { Title = "Monthly Sales" Start = 1976.jan
        Data = (138 128 ... 297) }
Transform { Function = log }
Regression { Variables= (seasonal const)}
Arima { Model = (0 1 1)}
Estimate { }
```

### 例 4

1 階の階差と 1 つのゼロ係数を指定したラグ作用素 AR(2) を含むモデル、つまり  $(1 - \phi_2 B^2)(1 - B)y_t = a_t$  を定義し、推定する。

```
series{title = "Annual Olive Harvest" start = 1950
      data = (251 271 ... 240) }
arima{model = ([2] 1 0)}
estimate{ }
```

### 例 5

トレンド定数を含み、その回帰誤差  $z_t$  が 1 階の季節階差と 1 次の季節移動平均項をもち、非季節要素をもたない ARIMA モデル、 $(1-B^{12})z_t = (1-\Theta B^{12})a_t$  を定義、推定する。ここでは、ARIMA 要素が 1 つしかないので、**model** 変数に ARIMA 要素の季節周期を明記しなければならない。そうしないと、非季節モデルが与えられてしまうからである。

```
series { title = "Monthly sales" start = 1976.jan
         data = (138 128 ... 297) }
transform { function = log }
regression { variables = const }
arima { model = (0 1 1)12}
estimate { }
```

### 例 6

3 つの ARIMA 要素をもつモデルを定義、推定する。回帰誤差  $z_t$  は  $(1-\phi_1 B)(1-\phi_3 B^3)(1-B)z_t = (1-\Theta B^{12})a_t$  なる ARIMA モデルに従う。ここで、オペレータ  $1-\phi_3 B^3$  は、四半期データにおける自己相関を説明するために加えられる。なぜなら、それぞれの期は 3ヶ月よりなるからである。また、四半期要素しか与えてはならないことに注意。

```
series { title = "Monthly sales" start = 1976.jan
         data = (138 128 ... 297) }
transform { function = log }
regression { variables = (const seasonal)}
arima { model = (1 1 0)(1 0 0)3(0 0 1)}
estimate { }
```

### 例 7

回帰誤差  $z_t$  が “airline model”、ARIMA  $(0 1 1)(0 1 1)_{12}$  であるようなモデルを定義、推定する。ここで、季節 MA パラメータは 1.0 に固定されていて、 $z_t$  は  $(1-B)(1-B^{12})z_t = (1-\theta B)(1-1.0B^{12})a_t$  と表される。 $\theta$  の初期値が 0.1 なので、**ma** リストでは欠損値として扱われている。 $1-B^{12}$  によって過剰階差がとられているので、このモデルは例 3 のものと同一である。(過剰階差については Section 4.4 を参照。)

```
series { title = "Monthly sales" start = 1976.jan
         data = (138 128 ... 297) }
transform { function = log }
arima { model = (0 1 1)(0 1 1)12
        ma = ( ,1.0f)}
estimate { }
```

## 自動モデル選択 (automdl)

---

### 説明 (DESCRIPTION)

Reg-ARIMA モデルの ARIMA 部の特定には、X-11-ARIMA/88 (Dagum(1988) 参照) で用意されていたものと類似した、自動モデル選択のスペックが利用できる。これによりユーザーは時系列データにフィットしたモデルの選択が可能であり、また選択の基準となる境界値を任意に変更することができる。

### 使用法 (USAGE)

```
automdl{mode = both
  method = best
  file = "my.mdl"
  fcstlim = 25.0
  bcstlim = 25.0
  qlim = 15.0
  overdiff = 0.99
  identify = all
  outofsample = yes
  print = (none autochoice)
  savelog = automodel
}
```

### 引数 (ARGUMENTS)

#### **bcstlim**

mode=both のオプションにより逆予測が行われるように設定されている場合に、観測されているサンプルを用いた逆予測の誤差を利用した判定の境界値を設定するための引数。逆予測誤差の比率の絶対値を平均した数値が、設定された境界値と比較される。例えば、bcstlim=25 と設定すると、境界値が 25 この引数に代入される数値は、0 以上かつ 100 以下でなければならない。bcstlim のデフォルトの数値は 20 となっている。

#### **fcstlim**

観測されたサンプルに基づく予測誤差を用いた判定における境界値を設定する。データの最新の 3 年分の数値に関して、予測された値との誤差の比率の絶対値の平均が、ここで設定された境界値よりも小さくなければ、当該のモデルがこの自動選択スペックによって採択されることはない。例えば fcstlim=20 とすると、この境界値は 20 この引数に代入される数値は、0 以上かつ 100 以下でなければならない。fcstlim のデフォルトの値は 15 である。

#### **file**

この自動選択スペックで検討するモデルの候補を記載したファイルのパスとファイル名を指定する。モデルのタイプは arima スペックにおける model 引数と同じ記法を用いる。詳細は以下を参照のこと。この引数が指定されない場合には、プログラムは x12a.mdl に記載されたモデルの候

補を自動的に用いる。ただし x12a.mdl はカレント・ディレクトリに存在しなければならない。

#### **identify**

**outlier** スペックによる外れ値の自動識別および、**aicstestregression** スペック内の **aicstest** 引数による曜日効果の自動識別を、この自動モデル選択の中でどのように行うかを定める。仮に **identify = .all** とするならば、両者の自動判定がモデル候補の全てに対して行われる。また **identify = first** とするならば、モデル候補のうち、最初のものに対して外れ値と曜日効果の自動判定が行われる。そして最初のモデルに対して下された判定に従って、以下のモデルの診断がなされる。選択されたモデルが最初のモデルではなかった場合には、選択されたモデルに対して、自動選択の手順が再度実行される。デフォルトは **identify = first** である。

#### **method**

自動選択の手順において、選択基準を満たすもののうち、最初のを最終的に選択するか (**method = first** とする場合)、within-sample の予測誤差が最も小さいものを選択するか (**method = best** とする場合) を決定する。デフォルトは **method = first** である。

#### **mode**

ユーザーによって指定されたモデルの候補の中から、カナダ統計局の X-11-ARIMA のために作成され、Dagum(1988) に記載されているモデル選択の基準を満たすモデルを、プログラムが自動で発見できるようにするための引数。詳細は以下。もし **mode = fcst** とするならば、選択されたモデルにより一年分の予測値を生成し、**mode = both** ならば、予測値と逆予測値の両方を生成する。デフォルトは **mode = fcst** である。系列を前後に拡大するための逆予測値および予測値をくつがえすために、**forecast** スペックを用いることができる。モデル候補は、**file** 引数で指定したファイルから読み込まれるが、**file** 引数が使用されない場合には、x12a.mdl から読み込まれる。**arima** スペックと **automdl** スペックを同一のスペックファイルの中で同時に利用することは出来ない。

#### **outofsample**

いかなる種類の予測誤差を自動モデル選択の評価基準に用いるかを決定する。もし **outofsample=yes** とするならば、out-of-sample forecast errors が使用される。これは予測期間内のデータを推定と一年分の予測値の算出に用いるデータから除外することによって得られる。また **outofsample=no** とすれば、within-sample forecasts errors が用いられる。これは系列全体を用いたパラメーターの推定値を、予測値の算出に使用するものである。X-11-ARIMA との齟齬が生じないために、外れ値の調整は、既に外れ値として識別された予測値に対して行われる。デフォルトは **outofsample=no** である。

#### **overdiff**

この引数は過剰階差の診断において、MA パラメーターの推定値の和に対する閾値を設定する。プログラムは、少なくとも 1 階の季節階差を含むモデルに関しては、季節 MA パラメーターの推定値を、また少なくとも 1 階の階差を含むモデルに関しては、通常の MA パラメーターの推定値を計算する。もし通常の MA パラメーターの推定値の和が、ここで設定した閾値よりも大きければ、モデル自動選択のプロシーチャーは、そのモデルを過剰階差ということでリジェクト（採用しない）する。またもし季節 MA もパラメーターの推定値の和が、ここで定めた閾値よりも大きければ、モデル自動選択のプロシーチャーは、regression のスペックにおいて固定的な季節性を表す変数の使用を提案する内容の警告を出力するが、モデルをリジェクトはしない。この引数のデフォ

ルトの値は 0.9 である。設定すべき引数の値は、0.9 以下にすべきではなく、また 1 より大きくしてはならない。

### print

保存のためのオプションはこのスペックでは利用できない。出力のために利用できる表は、表 6-1 に挙げられている。

表 6-1: Automdl に関連した出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
autoheader	+	+	.	header for the automatic modelling output
automodels	+	+	.	output for each model used in the automatic model procedure
autochoice	+	+	.	model choice of automatic model procedure

*Name* gives the name of each table for use with the **print** and **save** arguments.

*Default* indicates which tables are printed (+) or not printed (·) by default.

*Brief* indicates which tables are printed (+) or not printed (·) when the brief print level is specified. See section 5.1 for more information on print levels.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the **print** and **save** arguments.

### qlim

この引数は、モデルの適合度を見るための Ljung-Box の Q 統計量の p 値の閾値を設定する。当てはめたモデルに関する Q 統計量の p 値が、ここで設定した値よりも大きくなければ、そのモデルは自動選択プロシージャによって受容されない。例えば、`qlim = 10` とすると、この閾値を 10 パーセントに指定したことになる。この引数となる値は、ゼロより大きくなくてはならず、100 以下でなければならない。デフォルトの値は 5 パーセントである。

### savelog

`savelog=automodel`、あるいは `savelog=amd` と設定することにより、モデルの自動選択の結果がセッション・ログ・ファイル (ログ・ファイルについては 2.6 章を参照) に出力される。

## 詳細事項 (DETAILS)

自動モデル選択プロシージャはデフォルトの設定では、対象となるモデルが以下の条件をパスするときに、そのモデルが受容可能であると判断する。

- (1) 最新の 3 年間について外挿された値の平均誤差率の絶対値が、15 パーセント以下である。
- (2) モデルの残差の無相関性をテストする Ljung-Box 統計量の p 値が、5 パーセントより大きい。
- (3) 過剰階差の兆候が見られない。

季節 MA ではない、通常の MA パラメーターの推定値の和が 0.9 よりも大きい場合に、過剰階差が存在するとみなされる。モデル・ファイル内のいずれのモデルも受容されない場合には、モデル選択は行われず。これらの閾値は、`fcstlim`、`qlim`、`overdiff` の引数を指定することにより変更可能である。

スペック・ファイルに `regression` スペックが含まれている場合には、指定された回帰の項は、自動選択プロシージャによって評価される全ての ARIMA モデルについて用いられることに注意されたい。また原系列は `transform` スペックによって変換される。

カナダ統計局によって開発された X-11-ARIMA では、モデル自動選択において以下のモデルが使用される。 $(0, 1, 1)(0, 1, 1)_s$ 、 $(0, 1, 2)(0, 1, 1)_s$ 、 $(2, 1, 0)(0, 1, 1)_s$ 、 $(0, 2, 2)(0, 1, 1)_s$  および  $(2, 1, 2)(0, 1, 1)_s$  である。ここで  $s$  は季節周期を表す (Dagum(1988) を参照)。これらのモデルは、プログラムと一緒に配布される `x12a.mdl` というファイル内で与えられるものであるが、もし `regression` によって固定的な季節変動を用いるように指定されているなら、これらのモデルを用いることは出来ない。

`file` 引数で示される (あるいは `x12a.mdl` によって示される) いずれのモデルも、一行に一つずつ書かれ、行末には、最後の行を除いて、"X" が記述されなければならない。

ユーザーは、行末の "X" のかわりにアスタリスク ("\*") を置いておくことによって、"デフォルト" となるモデルを指定することが出来る。これにより、自動モデル選択プロシージャによっていずれのモデルも選択されなかった場合に、プログラムは、`regression` スペックによって指定された回帰変数に基づいて、デフォルトの Reg-ARIMA モデルによって事前調整を行う。どのモデルも選択されなかった場合には、予測や逆予測は行われない。

X-11-ARIMA のデフォルトモデルの例は以下のようなものである。

```
(0 1 1)(0 1 1) *
(0 1 2)(0 1 1) X
(2 1 0)(0 1 1) X
(0 2 2)(0 1 1) X
(2 1 2)(0 1 1)
```

## 例

以下の例では完全なスペックファイルを示す。

### 例 1

自動 ARIMA モデル選択プロシージャを用いてモデルを選択し、そのモデル一年分の予測値によって系列を延長するために使用する。曜日効果と固定的な季節変動がモデルに組み込まれる。デフォルトの設定での季節調整が行われる。

```
series      { title = "Monthly sales"  start = 1976.jan
              data = (138 128 ... 297) }
regression  { variables = (td seasonal) }
automdl     { mode = fcst    file = "nosdiff.mdl" }
estimate    { }
x11         { }
```

`nosdiff.mdl` の内容は以下のとおり。

```
(1 1 0) X
(2 1 0) X
(0 1 1) *
(0 1 2) X
(2 1 2)
```

### 例 2

予測に関する閾値が 20 パーセントに、カイ二乗の閾値が 10 パーセントに、過剰階差の閾値が 0.99 にそれぞれ変更されている以外は、例 1 と同様である。また、最初に受容されるモデルが選択され、自動的な外れ値の検出が、`nosdiff.mdl` に示されている全てのモデルに対して行われる。

```

series      { title = "Monthly sales"  start = 1976.jan
             data = (138 128 ... 297) }
regression  { variables = td }
automdl     { mode = fcst   file = "nosdiff.mdl"
             method = first  fcstlim = 20  qlim = 10
             overdiff = 0.99 identify = all }
outlier     { }
estimate    { }
x11         { }

```

### 例 3

モデルの識別と選択の過程で、out-of-sample forecast errors が用いられる。それ以外は例 1 と同様である。

```

series      { title = "Monthly sales"  start = 1976.jan
             data = (138 128 ... 297) }
regression  { variables = td }
automdl     { mode = fcst   file = "nosdiff.mdl"
             outofsample=yes }
estimate    { }
x11         { }

```

## 集計 (composite)

---

### 説明 (DESCRIPTION)

このスペックは、集計された系列の、直接的または間接的な季節調整の手続きの一部を構成する。つまり、集計系列の調整値を得るために、メタ・ファイルから参照されるスペックのひとつである。メタ・ファイル内に指定されたスペック・ファイルでは、要素となる系列の指定と集計方法の指定をする必要がある (series スペックの `comptype` および `compwt` 引数を参照)。このスペックは `series` スペック内で使用される。

ユーザーは集計を含む調整の題名と、系列の名前を決めることができ、どの表をプリント又は保存するかを指定できる。また、間接的調整から得られるプロットのうち、どれを出力するかを決めることができる。

### 使用法 (USAGE)

```
composite{title = "Total one family housing starts"
           name = "hsift"
           decimals = 2
           modelspan = (1985.Jan,)
           spectrumstart = 1985.Jan
           print = (brief +indtest)
           save = (indseasonal)
           savelog = (indtest)}
```

### 引数 (ARGUMENTS)

#### **decimals**

アウトプット・ファイルの季節調整値の表において、何桁の少数を用いるかを指定する。この引数の値は 0 以上 5 以下の整数でなければならない。例えば "decimals=3" のように記述する。デフォルトでは 0 が指定される。

#### **modelspan**

Reg-ARIMA の係数の推定に用いる集計系列の期間を指定する。この引数は、ユーザーが初期のデータの値を予測に影響させたくない場合や、季節調整前の予備調整に使う回帰係数の推定に新しいデータの影響を与えたくない場合などに有効である。`modelspan` 引数は、データの始点と終点を表す 2 つの値を取る。空白の場合、分析しようとするデータの始点または終点が充てられる。例えば月次データについては、`modelspan=(1968.1, )` という指定をすると、他のスペックにおける Reg-ARIMA の推定は 1968 年 1 月からデータの末端までの数値に基づいて行われる。始点か終点のどちらかが空白の場合は、カンマが必要である。始点と終点の日付は、集計系列の範囲内であればならず、始点はデータの最後の日付よりも前でなければならない。

*O.per* という表現を用いた、その他の終点の日付の指定法は、`modelspan` で指定する終点を、



常に特定の月または四半期の最新の値にしたい場合に有効である。ここで *per* は月または四半期を指定する。もしデータ区間が12月以外の月で終わるとすると、`modelsspan=(,0.dec)` という指定により、モデル推定に利用されるデータ区間は、データの終点の年より一つ前の年までのデータが利用されることになる。

#### **name**

集計列の名前を指定する。名前は8文字以内クォーテーションで括られている必要がある。ここで指定された名前は、アウトプットのすべてのページにラベルとして記載される。

#### **print** 及び **save**

このスペックにより直接または間接的季節調整に関するオプションの表を出力できる。表 6-4 を参照。

#### **savelog**

ログ・ファイルに出力可能な事後診断結果が Table 6-5 に記載されている。

#### **spectrumstart**

集計系列と、直接および間接的に季節調整された系列および、修正された不規則変動系列のスペクトルを推定するのに使用するデータの始点の日付を指定する。この日付は `spectrumstart=year.seasonal period` という形式で記述されなければならない。これは例えば最新の10年間について、調整した系列に曜日効果や季節変動が残存していないかどうかを判定するのに利用できる。季節変動や曜日効果が增大している場合には、そのような残存が起こりうる。月時系列のスペクトル・プロットのデフォルトの開始時点は、系列が8年より長ければ、集計系列の終点から数えて8年目の時点である。8年より短い場合は、デフォルトはデータの開始時点となっている。四半期データについては、デフォルトは集計系列の開始時点である。例: `spectrumstart=1987.Jan.`

#### **title**

集計系列を表すタイトルを指定する。タイトルは79文字以内で、括弧（クォーテーション）で括られている必要がある。指定したタイトルは出力データの上に出力される。

### 稀にしか使わない引数 (RARELY USED ARGUMENTS)

#### **diffspectrum**

`diffspectrum=no` とした場合、(変換された) 原系列と季節調整済系列のスペクトルを計算する。デフォルト (`diffspectrum=yes`) では、階差を取った系列のスペクトルを算出する。

#### **saveprecision**

`save` 引数による別ファイルへ保存する表に関して、何桁までを扱うかを指定する。デフォルトの `saveprecision` は15である。例: `saveprecision=10`

#### **spectrumtype**

X-12-ARIMAによるスペクトル推定の方法を指定する。`spectrumtype = periodogram` とすると、系列のピリオドグラムが計算されプロットされる。デフォルト (`spectrumtype=arspec`) では自己回帰モデルに基づくスペクトルが計算される。

#### **yr2000**

表 6-2: Composite に関連する出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
compositesrs	+	+	cms	aggregated time series data, with associated dates
compositeplot	.	.	.	plot of the aggregate series
outlieradjcomposite	.	.	oac	aggregated time series data, adjusted for outliers.
header	+	+	.	header for indirect seasonal adjustment
indtest	+	+	.	test for adequacy of composite adjustment
indunmodsi	+	.	id8	final unmodified si-ratios (differences) for the indirect adjustment
indftestd8	+	.	.	F-test for stable and moving seasonality for the indirect seasonal adjustment
indreplacsi	+	.	.	final replacement values for extreme si-ratios (differences) for the indirect adjustment
indmovseasrat	+	.	.	moving seasonality ratios for the indirect seasonal adjustment
indseasonal	+	+	isf	final seasonal factors for the indirect adjustment
indseasonaldiff	+	+	isd	final seasonal difference for the indirect seasonal adjustment(only for pseudo-additive seasonal adjustment)
indseasonalplot	.	.	.	indirect seasonal factor plots, grouped by month or quarter
indseasadj	+	+	isa	final indirect seasonally adjusted series
indadjsatot	+	+	iaa	final indirect seasonally adjusted series, with yearly totals adjusted to match the original series
indsadjround	+	+	irn	rounded indirect final seasonally adjusted series
indresidualseasf	+	.	.	F-test for residual seasonality
indseasadjplot	.	.	.	plot of the final indirect seasonally adjusted series
indtrend	+	.	itn	final trend-cycle for the indirect adjustment
indtrendplot	.	.	.	plot of the final trend-cycle from the indirect seasonal adjustment
indirregular	+	.	iir	final irregular component for the indirect adjustment
indirregularplot	.	.	.	plot of the final irregular component from the indirect seasonal adjustment
orivsendadjplot	.	.	.	plot of the aggregate series with the indirect seasonally adjusted series
indmodoriginal	.	.	ie1	original series modified for extreme values from the indirect seasonal adjustment
indmodsadj	.	.	ie2	seasonally adjusted series modified for extreme values from the indirect seasonal adjustment
indmodirr	.	.	ie3	irregular component modified for extreme values from the indirect seasonal adjustment
indyrtotals	.	.	.	ratio of yearly totals of the original series and the indirect seasonally adjusted series

*Name* gives the name of each table for use with the `print` and `save` arguments.  
*Default* indicates which tables are printed (+) or not printed (·) by default.  
*Brief* indicates which tables are printed (+) or not printed (·) when the brief print level is specified. See section 5.1 for more information on print levels.  
*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the `print` and `save` arguments.

表 6-3: Composite に関連するログ・ファイル診断

<i>name</i>	<i>short</i>	<i>description of diagnostic</i>
indtest	itt	test for adequacy of composite adjustment
indm1	im1	M1 Quality Control Statistic from indirect adjustment
indm2	im2	M2 Quality Control Statistic from indirect adjustment
indm3	im3	M3 Quality Control Statistic from indirect adjustment
indm4	im4	M4 Quality Control Statistic from indirect adjustment
indm5	im5	M5 Quality Control Statistic from indirect adjustment
indm6	im6	M6 Quality Control Statistic from indirect adjustment
indm7	im7	M7 Quality Control Statistic from indirect adjustment
indm8	im8	M8 Quality Control Statistic from indirect adjustment
indm9	im9	M9 Quality Control Statistic from indirect adjustment
indm10	imt	M10 Quality Control Statistic from indirect adjustment
indm11	ime	M11 Quality Control Statistic from indirect adjustment
indq	iq	Overall index of the acceptability of the indirect seasonal adjustment
indq2	iq2	Indirect Q statistic computed without the M2 Quality Control statistic
indmovingseasratio	isr	Moving seasonality ratio from indirect adjustment
indicratio	iir	$\bar{I}/\bar{C}$ ratio from indirect adjustment
indfstabled8	id8	F-test for stable seasonality, performed on the final SI-ratios from indirect adjustment
indmovingseasf	isf	F-test for moving seasonality from indirect adjustment
indidseasonal	iid	Identifiable seasonality test result for indirect adjustment
peaks	spk	Visually significant peaks in spectra from indirect adjustment

*Name* gives the name of each diagnostic for use with the `savelog` argument.

*Short* gives a short names for the diagnostics in the `savelog` arguments.

yr2000=yes とすると、“X-11 フォーマット”のデータに関して、1945 年での“century cutoff”が行われる。すなわち、00 – 45 年については 20xx 年を表すと解釈され、46 – 99 年については 19xx 年を表すと解釈される。これがデフォルトの指定になっている。yr2000=no とした場合には、年代を表す 2 桁の数値は、すべて 20 世紀の年代を表すものと解釈され、4 桁の西暦に変換される。

注意: このオプションは (transform や x11regression など) 他のスペックでファイルが読み込まれる際には、プログラムの挙動に影響する。

## 詳細事項 (DETAILS)

composite スペックを伴ったスペック・ファイルは、集計系列の要素を構成する各系列に対応した複数のスペック・ファイル群を同時に用いる場合にのみ機能する。これらの複数のスペック・ファイルのファイル名は、メタ・ファイルの中にリストアップされている必要があり、そのメタ・ファイルの最後には、この composite スペックを伴ったスペック・ファイル自身のファイル名が記述されていなければならない。各要素系列の series スペック内の comtype 引数は、各要素系列が集計系列にどのようにして合成されるかを指定する。(X-12-ARIMA でメタ・ファイルをどのように実行するかについては 2 節を参照。)

このメタ・ファイルによる集計系列の調整の実行により、直接的な季節調整と同様に、集計系列の間接的な季節調整値が得られる。この間接的な調整は、各要素の comtype によって指定される調整の組み合わせであり、個々のスペック・ファイルの内容に応じてそれぞれ調整されたりされなかったりする。

直接的な調整は、各要素のスペック・ファイルの指示によって行われる。直接的な調整の出力をコントロールするには、`x11` スペックの `print` 引数と `save` 引数を利用すればよい。

集計された系列に対する間接的な季節調整において、要素の一つとして未調整の系列を含ませたい場合には、その要素の中の `x11` スペックにおいて `type = summary` のようにオプションを定めれば良い。

2.7章で言及したように、`-c` はメタ・ファイルを伴って実行される集計系列の季節調整に制約を課す場合にのみ利用される。集計の実行に際しては、`X-12-ARIMA` は通常は要素の組に対して季節調整を行う。`-c` が呼び出された場合には、季節調整およびモデリングに関してスペック・ファイルで各要素の系列に対して指定されたオプションは無視される。その場合は各要素の系列は、集計系列を作成することのみに用いられる。このオプションは集計系列に対して適切な `Reg-ARIMA` を特定しようとする時に有用である。

このスペックによって生成される季節調整値の診断に関する表は、独自のファイルとしては保存されないが、季節調整診断のオプションを、実行時に `-s` を付加することによって利用することが出来る。これによって集計系列の分析に関する情報を、診断結果を記載したファイル（拡張子が `.xdg` のファイル）の中に保存することが出来る。なお `savelog` 引数を指定することによって、特定の診断結果をログ・ファイル（`.log` を拡張子とするファイル）に書き込むことが出来る。詳しくは2節を参照のこと。

直接的または間接的季節調整について安定性分析（スライディング・スパン）が必要な場合は、安定性分析のオプションは各要素の系列ごとに指定されている必要がある。季節フィルタの長さが各要素によって異なるならば、`slidingspans` スペックの `length` 引数は、各要素系列が収納される期間が同一になるように設定されなければならない。

季節調整済系列の改訂について継続的にその履歴分析が行われる際には、直接的季節調整と間接的季節調整の両方についての改定値が生成される。この場合各要素系列それぞれについて、分析内容が特定されていなければならない。

調整済系列の階差系列および不規則成分については、そのスペクトルにおける季節周波数と曜日周波数でのピークが自動的に探索される。もしピークが発見されれば警告メッセージが出力される。`-n` フラグにより出力が省略される時には、スペクトルのプロットは出力されないのでプロットが必要ならプログラムを再度実行する必要があるとのメッセージが出力される。

## 例 (EXAMPLES)

以下の例は集計系列の季節調整のステップ全体を表している。またこれらは完全なスペック・ファイルである。

### Step 1

スペック・ファイルは各要素系列ごとにそれぞれ作成する必要がある。次の例では、各要素（北東部、中西部、南部、西部地域の住宅建築着工件数）をそれぞれ処理し、集計系列はこれらの単純な和としている。まず北東部の系列のスペック・ファイルの例（ファイル名は `cne1hs.spc`）を以下に示した。これは 3x9 季節フィルタを用いて季節調整が行われている。

```
series { title="NORTHEAST ONE-FAMILY Housing Starts"
         file="cne1hs.ori" name="CNE1HS" format="2R"
         comptype=add }
x11 { seasonalma=(s3x9)
      title=(
        "Component for Composite Adjustment"
        "of Total U.S. 1-Family Housing Starts") }
```

このスペック・ファイルによる CNE1HS の季節調整は、集計系列の間接的な季節調整の計算に加えられる。

```
series { title="West ONE-FAMILY Housing Starts"
         file="cwt1hs.ori" name="CWT1HS" format="2R"
         comptype=add }
x11 { type=summary }
```

このスペック・ファイルにより `cwt1hs.ori` に保存されている未調整の系列を、集計系列の間接的な季節調整に付け加えることになる。

### Step 2

四地域の合計からなる住宅建築着工件数系列の間接的な調整のためのスペック・ファイルを作成する。各系列の直接的調整は乗法的であり、また 3x9 移動平均季節フィルタを使うものとする。直接的な調整による季節成分と、間接的な調整による内包的な季節成分の、両方の系列が保存されることになる。スペック・ファイル (`clfths.spc` というファイル名のファイル) は以下の通りである。

```
composite { title="TOTAL ONE-FAMILY Housing Starts"
            name="C1FTHS" save=(indseasonal) }
x11 { seasonalma=(s3x9)
      title="Composite adj. of 1-Family housing starts"
      save=(D10) }
```

### step 3

各要素と集計系列に関するメタ・ファイルを作成する。このメタ・ファイルは `hsiftot.mta` として保存されるが、内容は以下の通りである。

```
cne1hs
cmw1hs
cso1hs
cwt1hs
clfths
```

集計系列に関するスペック・ファイルは最後に提示するので確認されたい。

**Step 4**

この例を X-12-ARIMA に実行させるには次のように入力すればよい。

**x12a -m hs1ftot**

そしてリターンキーを押せばよい。

## チェック (check)

---

### 解説

残差項からモデル診断を行うための統計量を指定する。標本 ACF, PACF, Ljung-Box Q 統計量とその p 値, 残差の要約統計量, 正規性の検定, 残差のスペクトルとヒストグラムを指定できる。

### 使用法

```
<check> {<maxlag> = 24
          <print> = (none +histogram +acf)
          <save>  = (acf)
          <savelog> = normalitytest
        }
```

### 引数

#### maxlag

表示する標本 ACF, PACF のラグの数。デフォルトでは月次データについて 36, 四半期データについて 12 である。

#### print 及び save

表 6-2 にこのスペックで利用可能な出力をまとめておく。

#### savelog

ログファイル (2.6 節を参照) に対する診断は表 6-3 の通りである。

### 詳細

スペック **check** は推定したモデルから得られた残差を用いる。スペック **estimate** がない場合には **check** スペックで強制的に推定する。

帰無仮説のもとで Ljung-Box Q 統計量は漸近的に  $\chi^2$  分布に従う。推定した AR, MA 項の数以下の自由度について表示される。自由度 0 は無視される。

X-12-ARIMA は Reg-ARIMA モデルの残差の正規性からの乖離を検定する為の二つの統計量を計算する。これらは共に尖度を示す量である。1つは Geary (1936) および Gastwirth and Owens (1977) で与えられた、Geary の a 統計量である。

$$a = \frac{\frac{1}{n} \sum_{i=1}^n |X_i - \bar{X}|}{\sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}}$$

ここで  $\bar{X}$  は標本平均である。そしてもう 1つは標本尖度である。

表 6-4: Check に関連する出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
acf	+	·	acf	autocorrelation function of residuals with standard errors, and Ljung-Box Q-statistics computed through each lag
acfplot	+	·	·	plot of residual autocorrelation function with $\pm 2$ standard error limits
pacf	·	·	pcf	partial autocorrelation function of residuals with standard errors
pacfplot	·	·	·	plot of residual partial autocorrelation function with $\pm 2$ standard error limits
acfsquared	+	·	ac2	autocorrelation function of squared residuals with standard errors, and Ljung-Box Q-statistics computed through each lag
acfsquaredplot	+	·	·	plot of squared residual autocorrelation function with $\pm 2$ standard error limits
normalitytest	+	·	·	Geary's a and Kurtosis statistic tests for the normality of the model residuals
specresidual	·	·	spr	spectral plot of the Reg-ARIMA model residuals
histogram	+	·	·	histogram of standardized residuals and the following summary statistics of the residuals: minimum, maximum, median, standard deviation, and robust estimate of residual standard deviation ( $1.48 \times$ the median absolute deviation)

*Name* gives the name of each table for use with the `print` and `save` arguments.

*Default* indicates which tables are printed (+) or not printed (·) by default.

*Brief* indicates which tables are printed (+) or not printed (·) when the brief print level is specified. See section 5.1 for more information on print levels.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the `print` and `save` arguments.

表 6-5: Check に関連するログ・ファイル診断

<i>name</i>	<i>short</i>	<i>description of diagnostic</i>
normalitytest	nrm	Test results from the normality tests on the Reg-ARIMA model residuals (Kurtosis and Geary's a statistics)
ljungboxq	lbq	Significant lags for the Ljung-Box Q statistic

*Name* gives the name of each diagnostic for use with the `savelog` argument.

*Short* gives a short names for the diagnostics in the `savelog` arguments.



$$b_2 = \frac{m_4}{m_2^2} = \frac{n \sum_{i=1}^n (X_i - \bar{X})^4}{(\sum_{i=1}^n (X_i - \bar{X})^2)^2}$$

これらの性質は Snedecor and Cochran (1980) の Section 5.14 で議論されている。

この統計量のうち一方が極端な値であれば、標準化された残差は標準正規分布に従っていないことになる。X-12-ARIMA は、Pearson(1938) および Pearson and Hartley (1954) で与えられた表によって、有意水準 1% の検定を行う。もし Reg-ARIMA モデルがデータによく適合しているならば、このような正規性の欠如は通常は問題とならない。

しかしながら、データの特定の変動がモデルによって良く捕捉されていないければ、統計量は極端な値を取り得る。そのような変動はしばしば追加的な回帰変数や異なる回帰変数 (曜日効果、休日効果、外れ値等) を用いることによって表現できる。従って統計量の極端な値は、用いる回帰変数の選択を再考するための指標として利用することができる。

統計量が極端な値となるような、その他の重要な効果も存在する。例えば係数の確率的変動ないし時変条件付分散などといったものであり、これは Reg-ARIMA モデルでは表現できないので、X-12-ARIMA の検定統計量や予測の信頼区間の信頼性を低下させる。こうした効果の存在は、二乗残差についての Ljung-Box Q-統計量が極端な値を取ることによって示されることがしばしばある。

残差の 2 乗に関する ACF のラグの数は季節の数に等しく設定される (月次データでは 12, 四半期データでは 4)。この値は `maxlag` では変更できない。

ARIMA モデルで係数を固定すると `check` スペックの Box Ljung Q-統計量の自由度と  $P$  値は無効になる。これは先にモデルの当てはめを行って推定した係数を固定していた場合に起きる。

$P$  値は統計的に有意でない係数が 0 を値に持っているときに (近似的に) 有効であると期待される。

## 例

以下のようなスペック・ファイルを例として挙げておく。

### 例 1

指定したモデルの残差についてすべての診断結果を表示する。標本自己相関、標本偏自己相関はラグ 36 (月次データのデフォルトの値) まで計算する。`estimate` スペックがなくても `check` スペックがモデル推定を強制的に実行する。

```
series { title = "Monthly Retail Sales" start = 1964.jan
         file = "sales1.dat" }
regression { variables = (td ao1967.jun
                        ls1971.jun easter[14]) }
arma { model = (0 1 1)(0 1 1) }
check { print = (all) }
```

### 例 2

例 1 と同じ系列について標本偏自己相関のテーブルとプロットを除いたモデル診断を行う。残差の自己相関はラグ 24 まで計算する。

```
Series { Title = "Monthly Retail Sales"   Start = 1964.jan
        File = "SALES1.DAT" }
Regression { Variables = (TD A01967.jun
                          LS1971.jun Easter[14]) }
Arima { Model = (0 1 1)(0 1 1) }
Check { Print = (All -PACF -PACFplot) Maxlag = 24 }
```

## 推定 (ESTIMATE)

---

### 解説

スペック命令 `regression` や `arima` で定義された Reg-ARIMA モデルを推定する。推定の際には多くのオプションが認められていて、推定結果には AR、MA、回帰パラメータの点推定値と標準誤差が出力される。その中には、分散  $\sigma^2$  の最尤推定量；個々の回帰パラメータに対しての t-統計量；特定の回帰効果（もしモデルに含まれていれば）に関して、パラメータが全体として有意かを調べるための  $\chi^2$ -統計量；そして、（もし、厳密な尤度関数が用いられていれば）尤度に基づくモデル選択統計量などが含まれる。 $\chi^2$ -統計量によって調べることでできる回帰効果には季節効果、曜日効果、そしてユーザーによって定義された回帰効果などがある。

### 使用法

```
estimate{tol = 1.0e-5
maxiter = \[200]
exact = arma
outofsample = yes
print = (none +model +estimates +lkstats)
save = (model)
saveolog = (aic bic)
```

### 引数

#### **exact**

推定、尤度の評価、予測をするにあたって厳密な尤度を用いるか、条件付尤度を用いるかを定める。デフォルトは `exact=arma` であり AR、MA パラメータ両方に対して厳密な尤度関数が用いられる。その他のオプションとしては次のものが挙げられる。`exact = ma` とすると MA 部分には厳密な、AR パラメータには条件付の尤度関数が用いられる。`exact =none` とすると AR、MA パラメータの両方に対して条件付尤度が用いられる。

#### **maxiter**

ARMA 推定（AR、MA パラメータに対して行う非線形繰り返し推定）における繰り返し回数の許容最大数である。回帰係数のあるモデルでは、この制限は全ての IGLS 繰り返し推定において、ARMA 繰り返し推定の回数の合計数に適用される。回帰係数のないモデルではこの制限は 1 回の ARMA 繰り返し推定における繰り返しの最大数のことである。デフォルトは `maxiter = 200` となっている。

#### **outofsample**

直近 3 年分の予測誤差の平均的大きさを計算するときに、どの種類の予測誤差が用いられるかを決定する。これは診断統計量として使われる。もし `outofsample=yes` となっているならサンプル外予測誤差が用いられる。；これは、データセットから予測期間のデータを除いてから推定し、(直近

の3年分のデータそれぞれに対して)1年分の予測を行うことで得られる。もし `outofsample=no` となっているならサンプル内予測誤差が用いられる。つまり、全データからモデルのパラメータ推定を行い、そのモデルを使って、直近3年分のそれぞれの予測を行う。デフォルトは `outofsample=no` である。

### print 及び save

表 6.6 にこのスペック命令で利用可能な出力結果が与えられている。

表 6-6: Estimate に関連する出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
options	+	.	.	header for the estimation options
iterations	.	.	itr	detailed output for estimation iterations, including log-likelihood values and parameters, and counts of function evaluations and iterations
iterationerrors	.	.	.	error messages for estimation iterations, including failure to converge
model	+	+	mdl	if used with the <code>print</code> argument, this controls printing of a short description of the model; if used with the <code>save</code> argument, this creates a file containing regression and arima specs corresponding to the model, with the estimation results used to specify initial values for the ARMA parameters
regcmatrix	.	.	rcm	correlation matrix of regression parameter estimates if used with the <code>print</code> argument; covariance matrix of same if used with the <code>save</code> argument
armacmatrix	.	.	acm	correlation matrix of ARMA parameter estimates if used with the <code>print</code> argument; covariance matrix of same if used with the <code>save</code> argument
estimates	+	+	est	regression and ARMA parameter estimates, with standard errors
averagefcsterr	+	.	.	average magnitude of forecast errors over each of the last three years of data.
lkstats	+	+	lks	log-likelihood at final parameter estimates and, if <code>exact = arma</code> is used (default option), corresponding model selection criteria (AIC, AICC, Hannan-Quinn, BIC)
lformulas	.	.	.	formulas for computing the log-likelihood and model selection criteria
roots	.	.	rts	roots of the autoregressive and moving average operators in the estimated model
regressioneffects	.	.	ref	$X\hat{\beta}$ , matrix of regression variables multiplied by the vector of estimated regression coefficients
residuals	.	.	rsd	model residuals with associated dates or observation numbers

*Name* gives the name of each table for use with the `print` and `save` arguments.

*Default* indicates which tables are printed (+) or not printed (.) by default.

*Brief* indicates which tables are printed (+) or not printed (.) when the brief print level is specified. See section 5.1 for more information on print levels.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the `print` and `save` arguments.

### savelog

ログ・ファイル (2.6 節参照) への出力の際に利用可能な診断統計量が表 6.7 にまとめられている。

表 6-7: Estimate に関連するログ・ファイル診断

<i>name</i>	<i>short</i>	<i>description of diagnostic</i>
aic	aic	Akaike's Information Criterion (AIC)
aicc	acc	Akaike's Information Criterion (AIC) adjusted for the length of the series
bic	bic	Baysean Information Criterion (BIC)
hannanquinn	hq	Hannan-Quinn Information Criterion
averagefcsterr	afc	Average forecast error over the last three years of data

*Name* gives the name of each diagnostic for use with the `savelog` argument.  
*Short* gives a short names for the diagnostics in the `savelog` arguments.

### tol

非線形推定における収束許容差である。繰り返し推定が収束することを調べるために、対数尤度の変化量が `tol` と比較されることとなる。回帰変数のあるモデルでは、`tol` は IGLS 繰り返し推定 (新たな AR、MA パラメータに対して回帰パラメータが再推定される) の収束性を調べるために用いられる。回帰変数のないモデルでは、IGLS 繰り返し推定は行われない。そして `tol` は AR、MA パラメータを推定するのに使用される非線形繰り返し推定の収束性を調べるのに用いられる。デフォルトでは `tol = 1.0e-5` となっている。

### 稀にしか使われない引数

#### file

以前に X-12-ARIMA を実行したときのモデルを含むファイルの名前である。そうしたファイルは `save=model` または `save=mdl` と設定することで作成される。ファイル名は引用符付で示されなければならない。もしファイルがカレント・ディレクトリになれば、そのパスを与えておく必要がある。`file` 変数を使うときは、`arima` スペックにおける `model`、`ma`、`ar` の各変数、そして `regression` スペックにおける `variables`、`user`、`b` の各変数を使うことはできない。

#### fix

`file` 変数で指定されたファイル中のモデルの係数を固定するか、もしくは更なる推定のための初期値として使うかを決める。`fix = all` とするなら回帰パラメータ、ARMA パラメータ両方の推定値がモデルファイルにおける値に固定される。`fix = arma` とするなら ARMA パラメータの推定値のみがモデル・ファイルの値に固定される。`fix = none` とするなら全てのパラメータ推定値が固定されない。デフォルトは `fix = nochange` であり、これによりモデル・ファイルで固定するように示された係数値はそのまま保存され、その他の係数は再推定可能となる。

### 詳細事項

X-12-ARIMA による推定結果は”標準的な”仮定の下で漸近的に妥当なものとなる。(つまり、十分長い期間の時系列に対しては近似的に正しいとみなされるということ)、3.5 節を参照されたい。厳密な尤度関数が使われているときのみ、尤度に基づくモデル選択統計量が導出される。モデル選択統計量の使用に関してのコメントについては 4.5 節を参照されたい。

繰り返し推定が収束すれば、X-12-ARIMA はそのことを示すメッセージとその推定結果を出力す

る。繰り返し推定が収束しなければ、X-12-ARIMA はそのことを示すメッセージと繰り返しの最終段階におけるパラメータの値を出力する。この値をパラメータ推定値として使うべきではない。その代わりに、繰り返しが終了したときに得られたパラメータ値を初期値としてプログラムを再実行すべきである。収束問題に関しての潜在的な原因とその対処法が4節で論じられている。

tol は大きすぎても小さすぎてもいけない。tol を大きく設定してしまうと真の最尤推定値から大きく乖離した推定値が得られてしまう。また、tol を小さく設定してしまうと、繰り返しの回数が不必要に多くなってしまったり、結果の正しさについて謝った印象を与えてしまいかねない。tol に対しての大きすぎる、もしくは小さすぎる値とは何なのかは考えている問題に依存する。そこで、デフォルトである  $\text{tol} = 10^{-5}$  がほどよい妥協点として与えられている。倍精度のコンピュータ精度 (PC と Sun4 コンピュータではだいたい  $10^{-14}$ ) を下回るような値に tol を設定すると、エラーが出力されてしまう。しかし、コンピュータ精度の限度に近いというのは非常に小さな値である。

回帰変数を含むモデルでは、それぞれの IGLS 繰り返し推定中の ARMA 繰り返し推定における収束性を調べるために2つ目の収束許容差が必要となる。プログラムでは、最初の2回の IGLS 繰り返し推定では  $100 \times \text{tol}$  に設定され、その後、tol に再設定される。(最初のほうの IGLS 繰り返し推定では回帰パラメータがかなり大きく変化するので、最初のうちは ARMA パラメータが tol の範囲内に収まっているかということは問題にならない。) 従って、tol がデフォルト値である  $10^{-5}$  に設定されていれば、ARMA の収束許容性は最初の2回の IGLS 繰り返し推定では  $10^{-3}$  となり、その後  $10^{-5} (= \text{tol})$  となる。また、回帰変数のあるモデルではそれぞれの IGLS 繰り返し推定に対して、ARMA 繰り返し推定の回数の最大値が必要とされる。このリミットは40に設定されている。

もし、ある特定の IGLS 繰り返し推定における ARMA 繰り返し推定が収束しなかったとしても、それは普通問題にならない。プログラムは次の IGLS 繰り返しに進む。そうすると、そのときの ARMA 繰り返し推定はうまく収束するかもしれない。実際、全体の収束性に関して必要なのは、最後の IGLS 繰り返し推定での ARMA 繰り返し推定が収束することで、しかも、ARMA 繰り返し推定の総数が `maxiter` 以内でありながら、IGLS 繰り返し推定が収束許容性である tol の範囲内で収束することである。

`print=roots` と設定すると、推定されたモデルにおける全ての AR、MA オペレータの根の表が作成される。その根に加え、表にはそれぞれの根の絶対値や周波数 (`[-.5, .5]` 上に) も出力される。絶対値が1より大きい根は単位円外に存在し、それは定常な AR オペレータ、反転可能な MA オペレータに対応している。(4.4節を参照のこと。) AR パラメータに対しての尤度関数が条件付であるとき (`exact = ma` または `exact = none` と設定されているとき) のみ、AR 部分の根が単位円上、もしくは単位円内 (絶対値  $\leq 1$ ) に存在し得る。MA 部分の根は単位円内 (絶対値  $\leq 1$ ) には存在しない。なぜなら、反転可能性が満たされないと推定できなくなっているからである。単位円上 (絶対値 = 1) にある MA 部分の根はその値を丸めたとき、もしくは推定の際に MA オペレータの全てのパラメータが固定されているときに起こり得る。

表6.8に示されたサンプル結果では、非季節的 AR(2) 多項式が、 $z = x \pm y$ ,  $x = -.6784$ ,  $y = .8817$  でその絶対値が  $r = (x^2 + y^2)^{1/2} = 1.1125$  であるような1組の共役複素根をもつことがわかる。この数値は単位根 (1.000) に近いので、根における非負のほうの周波数を調べてみるとよい。つま

り、系列が決定的な周期要素を含んでいるかどうかを見るために、 $z = re^{\pm i2\pi\lambda}$  における  $\lambda \geq 0$  の数値を調べるのである。このようにする理由は次の通り。モデル化された時系列が、周期  $1/\lambda$  の周期的要素  $f(t)$ , i.e.  $f(t+1/\lambda) = f(t)$  をもつならば、推定された高次の AR 多項式は  $e^{\pm i2\pi\lambda}$  に近い根を持つことが多い。(ただし、階差オペレータが  $e^{\pm i2\pi\lambda}$  を根として持たなければ) なぜこういったことが起こるのか説明できる理論的結果もあるのだが、直感的な説明は次の通りである。周期  $1/\lambda$  の最も簡単な関数、

$$f(t) = A \cos(2\pi\lambda t + c)$$

において、その根が  $e^{\pm i2\pi\lambda}$  である AR(2) 多項式  $\phi(B) = (1 - e^{-i2\pi\lambda}B)(1 - e^{i2\pi\lambda}B)$  に

$$\phi(B)f(t) = 0$$

なる性質があったとする。従って、この AR(2) 要素は  $f(t-1)$  と  $f(t-2)$  から  $f(t)$  を完全に予測することができる。2次以上の AR オペレータをもつモデルをあてはめるとき、AR パラメータは  $\phi(B)f(t) = 0$  を満たすような値をとりがちになる。(AR(1) 多項式に関しては  $e^{i2\pi\lambda}$  が定数、つまり  $\lambda = 0, 1/2$  のときで十分である。) それゆえ、絶対値  $r = 1$  なる AR 根があれば、時系列中におおまかな周期要素が含まれることが多い。

月次系列では、季節効果の周波数は  $\lambda = 1/12, 2/12, 3/12, \dots, 6/12$  (それぞれ 0.0833, 0.1666, 0.2500,  $\dots$ , 0.5000 と等しい) である。周波数  $\lambda = 0$  はトレンドの動きを表し、周波数  $\lambda = 0.3482$  は曜日効果を表している。上の表では非季節的 AR 根が 0.3544 だったが、これは曜日効果を示す周波数にとっても近いことに注意しておく。実際、表中の時系列には強い曜日要素が存在し、自動モデル化を行うプロシージャはそれを説明しようとして、AR(2) 要素を取り入れたのである。というのも、regression スペックにおいて曜日回帰が行われていなかったからである。

MA 多項式において、季節周波数またはトレンド周波数をもつ単位根に近い根が存在するとき、MA 多項式は階差もしくは季節階差多項式と共通の根をもっていることが多い。そうした共通因子  $\kappa(B)$  の存在は時系列に決定的トレンドや季節要素が含まれていることを示す。より明確に言うと、3節の方程式 (3) (ただし回帰変数は省略) の表記において、 $\kappa(B)f(t) = 0$  を満たす関数  $f(t)$  が存在して、時系列  $y_t$  が

$$\phi(B)\Phi(B^s) \left\{ \frac{(1-B)^d(1-B^s)}{\kappa(B)^D} \right\} y_t = f(t) + \left\{ \frac{\theta(B)\Theta(B^s)}{\kappa(B)} \right\} a_t$$

のようにモデル化されることを意味する。

上の例の表ではモデルの季節的移動平均多項式は  $\Theta(B^{12}) = 1 - \Theta B^{12}$ ,  $\Theta = 0.6417$  で、その根 ( $\Theta(z) = 1 - \Theta z$  の根) は  $1/\Theta = 1.5583$  である。経験的に  $1/\Theta$  が 1.10 以下であるときは、固定化された季節効果のみをもつモデル ( $D = 0$  で、regression スペックにおいて variables=seasonals とされたモデル) と取り替えるのがよいとされる。

非季節的 MA 多項式が 1 に近い根 (絶対値が 1 に近く、周波数が 0 に近い根) をもつとき、時として過剰階差がとられていることがある。つまり、階差次数  $d$  と非季節的 MA 次数  $q$  の両方を 1 つずつ小さくした代替モデルを考えてみるべきである。そして、t-統計量が有意となれば、その代替モデルにトレンド定数 (上の式で  $f(t) = C$  とおかれたもので、regression スペックで variables=const としたもの。) を含めてみるべきである。

ARIMA モデルや、**regression** スペック、**x11regression** スペックによる回帰モデルにおいて係数を固定すると、出力されたモデル選択統計量 *AIC*, *AICC*, *Hannan Quinn*, *BIC* は無効となってしまう。これは、以前にモデルを当てはめたときに得られた推定値に値を固定したときに起こる。P-値は統計的に有意でない係数が 0 に固定されたときに、(近似的に) 期待通りの妥当性をもつであろう。

## 例

次に完全なスペックファイルの例を挙げておく。

The following examples show complete spec files.

### 例 1

$(1 - B)(1 - B^{12})y_t = (1 - \theta B)(1 - \Theta B^{12})a_t$  なるモデルの回帰係数を一般化最小二乗法で推定する。ただし、 $M_{it}$  は月次の固定された季節効果を表す回帰変数である。MA パラメータ  $\theta$  は 0.25 に固定されている。残差がスペックファイルと同じ名前でカレントディレクトリに保存されている。ただし、拡張子は *.rsd* である。

```
series { title = "Monthly Sales" start = 1976.1
         data = (138 128 ... 297) }
regression { variables = seasonal }
arima { model = (0,1,1) ma = (0.25f) }
estimate { save = residuals }
```

### 例 2

$(1 - \phi B)(1 - B)(1 - B^{12})z_t = (1 - \Theta B^{12})a_t$  なる季節モデルを推定する。ここで、**tol** は  $10^{-4}$  に設定されている。これはデフォルトと比べて緩い収束許容差である。さらに、繰り返し最大数を 100 に減らしてある。モデルに回帰パラメータが存在しないので、**tol** と **maxiter** は、 $\phi$  と  $\Theta$  を推定するための、1 回の非線形 ARMA 繰り返し推定に適用される。パラメータ推定に用いられる尤度関数は MA に対しては厳密、AR に対しては条件付のものである。**print** 変数により尤度とパラメータ値が各々の繰り返しの段階で出力されている。そして、繰り返しの最終段階のものに続いて、推定された AR、MA オペレータの根も出力されている。**save** 変数により、最後に得られた regARIMA モデルがファイルに保存される。

```
series { title = "Monthly Inventory" start = 1978.12
         data = (1209 834 ... 1002) }
transform { function = log }
regression { variables = (td ao1999.01) }
arima { model = (1,1,0)(0,1,1)}
estimate { tol = 1e-4 maxiter = 100 exact = ma save = mdl
          print = (iterations roots) }
```

### 例 3

例 2 で保存された regARIMA モデルの推定値が、**file** 変数により今回の実行でも用いられている以外は例 2 と同様である。全てのパラメータ推定値がモデル・ファイル中の値に固定されている。

```
series { title = "Monthly Inventory" start = 1978.12
         data = (1209 834 ... 1002) }
```



```
transform { function = log }
estimate { file = "Inven.mdl"
          fix = all }
```

#### 例 4

3つのデータ値を追加した以外は例3と同様である。また、追加したデータが異常値かどうかを判定しようとしている。例2と3ではデータ期間の最終年月は1999年12月である。regARIMAモデルのパラメータはこのデータ期間より得られた値に固定されているが、これは例2で保存されたものである。

```
series { title = "Monthly Inventory" start = 1978.12
         data = (1209 834 ... 1002 1425 901 1375) }
transform { function = log }
estimate { file = "Inven.mdl"
          fix = all }
outlier { span=(2000.01,) }
```

## 予測 (forecast)

---

### 解説

推定したモデルを用いて `series` で指定した時系列について予測と逆予測を行う。変換された系列については点予測とその標準誤差が出力され、原系列については点予測と予測区間が出力される。

### USAGE

```
forecast(maxlead = 12
         maxback = 12
         probability = .95
         exclude = 10
         print = (none +original +transformed)
         save = (variances)
```

### 引数

#### **exclude**

予測の前に系列の直近部分を（またはスベック `series` の引数 `span` で指定したスパンから）どれだけ除くかを指定する。デフォルトでは系列の最後（あるいは最後のスパン）から予測する。すなわち `exclude = 0` である。

#### **maxback**

逆予測の数を指定する。デフォルトは 0 で最大値は 60（60 の限界は変更できる。2.8 節を参照。）  
注意：逆予測は 15 年よりも長い時系列データについては計算されない。また、スベック `series` の引数 `modelspan` で指定した開始時が、スベック `series` の引数 `span` で指定したスパンと異なる場合には、逆予測は計算されない。

#### **maxlead**

予測の数を指定する。デフォルトは 12 で最大値は 60。（60 の限界は変更できる。2.8 節を参照。）

#### **print and save**

オプションで利用できる出力は表 6-8 にまとめてある。

#### **probability**

正規性を仮定したときの予測区間に入る確率を指定する。デフォルトでは `probability=.95` などで変換されたデータの予測区間は点予測  $\pm 1.96 \times$  予測標準誤差となる。

### 詳細

予測は推定された（あるいは評価された）モデルについて実行される。もし `estimate` スペックがなければ `forecast` スペックが強制的に予測前に推定を行う（デフォルトの設定）。予測に用いられるモデルは `regression` と `arima` で指定されたモデルである。もし `outlier` スペックがあ

表 6-8: Forecast に関連する出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
transformed	+	·	ftt	point forecasts on the transformed scale, with corresponding forecast standard errors
variances	·	·	fvr	forecast error variances on the transformed scale, showing the contributions of the error assuming the model is completely known (stochastic variance) and the error due to estimating any regression parameters (error in estimating AR and MA parameters is ignored)
forecasts	+	·	fmt	point forecasts on the original scale, along with upper and lower prediction interval limits

*Name* gives the name of each table for use with the `print` and `save` arguments.

*Default* indicates which tables are printed (+) or not printed (·) by default.

*Brief* indicates which tables are printed (+) or not printed (·) when the brief print level is specified. See section 5.1 for more information on print levels.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the `print` and `save` arguments.

ればモデルは認識された外れ値について説明変数を追加する。外れ値として想定された値は推定したパラメータを通じて間接的に、また系列の最後の方で外れ値が検出された場合には直接的に、予測に影響を与える。

モデルに移動平均項がある場合には予測は、推定されたモデルから得られる残差に依存する。`estimate` スペックの引数 `exact` により正確な尤度関数から計算されたのか、それとも条件付き尤度関数から計算されたのか判断される。

予測の標準誤差はモデルの回帰パラメータの推定からの誤差が含まれるが、AR,MA パラメータの推定誤差は含まれない。

モデルにユーザー定義の回帰変数が含まれている場合には、それらが予測区間全ての点について適用される。

変換されたデータについての予測区間は  $K$  を正規分布のパーセント点として

$$\text{点予測} \pm K \times \text{予測標準偏差}$$

と計算される。原系列の点予測と予測区間は変換 (Box-Cox 変換, ロジスティック変換, 前処理 (スペック `regression` で `variables = td` を指定したときには月の長さや四半期の長さにより調整する)) の逆変換で計算される。`transform` スペックにユーザー定義の前処理が含まれているときには予測期間を通じて逆変換がなされる。予測期間の中で前処理が定義されていなければそのままである。この場合ユーザー定義の前処理は効果を持たない。

`exclude > 0` とする理由はデータが得られている期間について予測を実行してモデルの性能を評価するためである。X-12-ARIMA では関係するデータがある場合には予測誤差 (観測値 - 予測値) が表示される。`exclude = 0` とすると *within-sample* 比較となる。スペック `series` の引数 `span` を使うと *out-of-sample* 比較となる。(例 4 参照。)

予測と逆予測が 1 回の X-12-ARIMA で季節調整とともに実行されるとき、原系列に加えてそれらにも季節調整方法が適用される。季節調整で Reg-ARIMA モデルが指定され、`forecast` スペックがない場合には、モデルから 1 年先の予測が作成される。予測なしに季節調整を指定するには

maxlead = 0 とする。

Reg-ARIMA モデルで曜日効果、外れ値、休日効果、ユーザー定義の回帰変数を推定する前に行う調整が原系列になされると、予測、逆予測についても調整される。

警告：x11 スペックで季節調整が指定されていると exclude は用いられない。その場合には exclude は 0 に設定され、警告メッセージが表示される。

## 例

次に完全なスペック・ファイルの例を挙げておく。

### 例 1

月次データで最後の月から 12 カ月先までの予測を 95 パーセントの予測区間（点予測 ± 1.96 × 標準誤差）で行う。これらはすべてデフォルトオプションである。estimate スペックはないが forecast スペックがあるのでデフォルトのオプションで推定が強制的に実行される。変換されたデータの点予測と予測区間の限界は指数関数で元に戻され、regression で variables = td を指定したために変換された原系列を  $m_t/\bar{m}$  をかけて元に戻している。

```
SERIES { TITLE = "Monthly sales"  START = 1976.JAN
          DATA = (138 128 ... 297) }
TRANSFORM { FUNCTION = LOG}
REGRESSION { VARIABLES = TD }
ARIMA { MODEL = (0 1 1)(0 1 1)12 }
FORECAST { }
```

### 例 2

例 1 で用いたデータについて 24 カ月先までの予測を行う。outlier スペックがあるので、予測に用いるモデルには AO,LS が含まれる。さらに regression スペックで指定された曜日効果の変数が含まれる。

```
Series { Title = "Monthly Sales"  Start = 1976.jan
         Data = (138 128 ... 297) }
Transform { Function = Log}
Regression { Variables = Td }
Arima { Model = (0 1 1)(0 1 1)12 }
Estimate { }
Outlier { }
Forecast { Maxlead = 24 }
```

### 例 3

10 個のデータを取り除いた上で 15 期先までの予測を行う。データを取り除く前に推定を行い、そのあとで 10 個のデータを除いて予測を行う。それら 10 個のデータについて within-sample 予測誤差が表示される。それぞれの予測には点予測に 1.645 × 予測標準誤差を加減した予測区間限界（90 パーセント）が表示される。

```
series { title = "Monthly sales"  start = 1976.jan
         data = (138 128 ... 297) }
transform { function = log }
regression { variables = td }
arima { model = (0 1 1)(0 1 1)12 }
```

```

estimate { }
forecast { maxlead = 15
           probability = .90
           exclude = 10 }

```

#### 例 4

系列は1992年3月が最後であるが、引数 `span` によって最後の24個のデータがモデル推定に用いるデータから取り除かれる。推定されたパラメータを用いて1990年3月から最後まで24個のデータが予測される。この最後の24個のデータに対する *out-of-sample* 誤差が表示される。(例3と比較。)

```

series { title = "Monthly sales" start = 1976.jan
         data = (138 128 ... 297)
         span = ( ,1990.mar) }
transform { function = log }
regression { variables = td }
arma { model = (0 1 1)(0 1 1)12 }
estimate { }
forecast { maxlead = 24 }

```

#### 例 5

月次データの最後から12ヵ月先までの予測を95パーセントの予測区間について実行する。(点予測  $\pm 1.96 \times$  標準誤差。) これらは全てデフォルトの設定である。12個の逆予測も行う。またデフォルトの乗法型季節調整を逆予測系列、曜日効果を修正した原系列、予測系列に対して行う。

```

series { title = "monthly sales" start = 1980.jan
         file = "ussales.dat" }
transform { function = log }
regression { variables = td }
arma { model = (0 1 1)(0 1 1)12 }
forecast { maxback=12 }
x11{ }

```

## 履歴 (history)

---

### 解説

一連の時系列データから一連のコマンドの実行をリクエストするのに最適なスペック。(i) 重複する季節調整や大規模な季節調整の調整結果 (ii) 将来時点の予測誤差 (iii) 尤度統計量の履歴を作成するために便利である。ユーザーは履歴を作成する最初の日と (i)-(iii) の種類を指定できる。(iii) 予測誤差を指定した場合はユーザーは予測に用いるベクトルを指定できる。注意: 分析の履歴を作成するとプログラムの実行時間を大幅に増加させる。

### 使用法

```
history(estimates = (sadj fcst trend)
        sadjlags = (1,2,3,12)
        trendlags = (1,2,3)
        target = final
        start = 1975.jan
        fstep = ( 1 2 )
        fixmdl = no
        fixreg = outlier
        endtable = 1990.Jan
        print = ( all -revvalsa )
        save = ( sar trr fcsterrors )
        savelog = ( aveabsrevesa aveabsrevtrend )
)
```

### 引数

#### endtable

季節調整とトレンドの推定と期間ごとのトレンドの変化を要約する表を作成するための最後の日付を指定する。そうすることで統計量の要約が最後の (あるいはほとんど最後の) 季節調整やトレンドの推定結果に基づいていることを確認できる。endtable で値を指定しなければ系列の最後、または、sadjlags や trendlags で指定された最大のラグよりも 1 だけ大きい値に設定される。このオプションは予測や尤度の推定の履歴分析には影響をもたらさない。例: endtable= 1980.jun

#### estimates

Reg-ARIMA モデルまたは X-11 季節調整のどの推定結果を履歴分析に用いるかを決定する。例: estimates = (sadj aic)。デフォルトは季節調整系列 sadj。表 6-29 に推定量の説明をまとめた。

#### fixmdl

Reg-ARIMA モデルを再推定するかを決める。fixmdl=yes とすれば history の実行中に ARIMA モデルのパラメーターと Reg-ARIMA モデルの回帰係数は全系列から推定される値に固定される。(モデル・スパン (期間) に関しては modelspan により指定する。) fixmdl=no とすれば Reg-

表 6-9: 引数 estimates の選択

名前	オプションの内容
<b>sadj</b>	最終的な季節調整系列 (および合成季節調整系列を計算していれば間接季節調整系列)。
<b>sadjchng</b>	最終的な季節調整系列の月ごとの (四半期ごとの) 変化
<b>trend</b>	最終的なヘンダーソン (Henderson) トレンド要素
<b>trendchng</b>	最終的なヘンダーソン (Henderson) トレンド要素の月ごとの (四半期ごとの) 変化
<b>seasonal</b>	確定したプロジェクトの季節要素
<b>aic</b>	Reg-ARIMA モデルの AICC や最大化された対数尤度
<b>fcst</b>	Reg-ARIMA モデルから計算される予測値と平均自乗予測誤差 model. <b>警告:</b> 予測がなされた時にだけ用いられる。forecast を参照のこと spec.

ARIMA モデルのパラメーターは最後のデータが更新されるたびに再計算される。デフォルトでは `fixmdl=no` である。この引数は適切な Reg-ARIMA モデルが存在しない場合には無視される。

#### fixreg

Reg-ARIMA モデルまたはその他の不規則要素の回帰モデル (irregular component regression) において説明変数の係数を固定するかどうか、指定する。それらの係数は series スペックや composite スペックで示されたモデル・スパンから得られる値に固定される。ほかの係数はそれぞれの history span に関して再推定される。曜日 (**td**) 休日 (**holiday**) 異常値 (**outlier**) やその他のユーザー定義の回帰 (**user**) などの効果は固定される。この引数は適切な Reg-ARIMA モデルや他の irregular component regression があてはまらない時や `fixmdl=yes` の場合には無視される。

#### fstep

history を実行して分析する予測誤差のステップ数 (同時に分析できるのは 4 個の異なるステップまで) をベクトルで指定する。例えば `fstep=(1 2 12)` は 1 ステップ、2 ステップ、12 ステップの予測誤差を分析する。デフォルトでは月次データに関して (1 12) であり、四半期データに関して (1 4) である。

**注意:** このベクトルで与えられる値は forecast スペックで `maxlead` を指定した時の値を超えない値で 1 以上の値を指定する。

#### print 及び save

出力として利用できる表は表 6-10 にまとめた。

#### sadjlags

ラグ付きの季節調整の調整結果の分析に用いたいラグ (最大で 5 個のラグまで全て 0 以上) をベクトルで指定する。それぞれのラグについて計算された結果は調整時点を越えた多くの観測値を用いて計算された結果である。すなわち、`sadjlags` で与えられるそれぞれの `revisionlagi` の値に対して、 $t$  時点の調整結果を計算する時には、 $t + \text{revisionlag}_i$  時点までのデータを用いられる詳しくは詳細を参照のこと。

このオプションは引数 `estimates` において季節調整系列の調整履歴や月ごと (四半期ごと) の変化を指定している時に限り有効である。デフォルトではラグ付きの季節調整の調整結果は表示されない。

#### start

履歴調整結果の開始日を指定する。この引数を用いなければ季節調整が実行される限りデフォルト

表 6-10: History に関連する出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	内容
header	+	+	.	履歴分析のヘッダー
outlierhistory	+	+	rot	改訂履歴に関して除去された外れ値と残された外れ値の記録 (自動外れ値識別が用いられた場合だけ表示される。)
sfilterhistory	.	.	sfh	改訂履歴の中のそれぞれの観測値に対して選択された季節調整フィルターの記録 (自動季節調整フィルター選択がなされた場合だけ表示される。)
sarevisions	+	.	sar	季節調整データの同時点から最近の推定までの改定
sasummary	+	+	.	季節調整改訂結果の統計量の要約
saestimates	.	.	.	季節調整データの同時点と最近の推定結果
chngrevisions	+	.	chr	同時点から最近までの月ごとの (四半期ごとの) 季節調整系列の改訂結果
chnghsummary	+	+	chs	同時点から最近までの月ごとの (四半期ごとの) 季節調整系列の統計量の要約
chnghestimates	.	.	.	同時点から最近までの月ごとの (四半期ごとの) 季節調整系列の推定値の変化
indsarevisions	+	.	iar	同時点から最近までの 間接季節調整系列の推定値の変化
indsasummary	+	+	.	同時点から最近までの間接季節調整系列の統計量の要約
indsaestimates	.	.	.	同時点から最近までの 間接季節調整系列の推定値
trendrevisions	+	.	trr	同時点から最近までのトレンド要素の改訂結果
trendsummary	+	+	.	同時点から最近までのトレンド要素の統計量の要約
trendestimates	.	.	.	同時点から最近までのトレンド要素の推定値
trendchngrevisions	+	.	trc	同時点から最近までの月ごとの (四半期ごとの) トレンド要素の改訂結果
trendchnghsummary	+	+	.	同時点から最近までの月ごとの (四半期ごとの) トレンド要素の統計量の要約
trendchnghestimates	.	.	.	同時点から最近までの月ごとの (四半期ごとの) トレンド要素の変化

*name* それぞれ表の名前で、引数 `print` や `save` で用いられる。

*default* で表示されるのは (+)、表示されないのは (.)

*Brief* では `brief` プリントで表示されるのは (+)、表示されないのは (.)

プリントレベルは指定する。詳しくは 5.1 節を参照のこと

ファイルを保存した時に *Ext* はファイルの拡張子を示す。・は表が保存できないことを表す。

ファイルの拡張子は `print` や `save` での略称を与える。



表 6-10: History に関連する出力 (つづき)

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
<code>sfrevisions</code>	+	·	<code>sfr</code>	revision from concurrent to most recent estimate of the seasonal factors, as well as the revision from concurrent to most recent estimate of the projected seasonal factors
<code>sfsummary</code>	+	+	·	summary statistics for seasonal factors
<code>sfestimates</code>	·	·	·	concurrent, projected, and most recent estimate of the seasonal factors
<code>lkhhistory</code>	+	+	<code>lkh</code>	Log-Likelihood and AICC Values for each observation in the revisions history.
<code>fcerrors</code>	+	+	<code>fce</code>	accumulating sums of squared forecast errors and evolving mean square forecast errors for each observation in the revisions history.
<code>fcsthory</code>	·	·	<code>fch</code>	concurrent forecasts and forecast errors for each observation in the revisions history.

*Name* gives the name of each table for use with the `print` and `save` arguments.

*Default* indicates which tables are printed (+) or not printed (·) by default.

*Brief* indicates which tables are printed (+) or not printed (·) when the `brief` print level is specified. See section 5.1 for more information on print levels.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the `print` and `save` arguments.

表 6-11: History に関連するログ・ファイル診断

<i>name</i>	<i>short</i>	診断の内容
<code>aveabsrevesa</code>	<code>asa</code>	季節調整系列の改訂結果の絶対値の平均
<code>aveabsrevchnng</code>	<code>ach</code>	季節調整系列の月ごとの (四半期ごとの) 変化の改訂結果
<code>aveabsrevindsa</code>	<code>iaa</code>	間接季節調整系列の改訂結果の絶対値の平均
<code>aveabsrevtrend</code>	<code>atr</code>	最終トレンド要素の改訂結果の絶対値の平均
<code>aveabsrevtrendchnng</code>	<code>atc</code>	トレンド要素の月ごとの (四半期ごとの) 変化の改訂結果の絶対値の平均
<code>aveabsrevsf</code>	<code>asf</code>	最終季節項の改訂結果の絶対値の平均
<code>aveabsrevsfproj</code>	<code>asp</code>	プロジェクトの季節項の改訂結果の絶対値の平均

*name* は引数 `savelog` で用いる時のそれぞれの診断の名前を意味する。

*Short* は引数 `savelog` で用いる時の略称を意味する。

トでは最大の季節フィルターの長さを用いられる。デフォルトの開始日は  $3 \times 3$  かまたは、定常フィルターを用いる場合には系列の開始から 6 年後で、 $3 \times 9$  フィルターを用いる場合には系列の開始から 12 年後になる。季節調整がなされなければ、系列の開始日から 8 年後になる。例：  
`start=1990.jun.`

#### `target`

同時推定からの残差かまたは季節調整の調整結果を定義する最後の推定からの残差かを指定し、トレンドに関して `sadjlags` で指定したラグに関して計算するかそれとも、`trendlags` で指定したラグに関して計算するかを指定する。デフォルトでは `target=final` である。変更する場合には `target=concurrent` とする。

#### `trendlags`

`sadjlags` と同様、この引数はラグつきのトレンド要素に関して履歴分析の中でどのラグを用いるかを指定する。最大で 5 つまでのラグ (値は 0 以上) を指定できる。このオプションは引数 `estimates` において最終的なトレンド要素の調整履歴や月ごと (四半期ごと) の変化を指定している時に限

り有効である。デフォルトではラグつきのトレンド推定の調整結果は表示されない。

## 稀にしか使わない引数

### **fixx11reg**

履歴分析の中で不規則要素の回帰モデルを再推定するか否かを指定する。fixx11reg=yes のとき、不規則要素の回帰モデルで回帰係数は分析の中で全系列に関して推定された値に固定される。fixx11reg=no のとき、不規則要素の回帰モデルのパラメーターは履歴期間の最後の時点が先の時点に指定されたら、その度に再推定される。デフォルトでは fixx11reg=no である。この引数は不規則要素の回帰モデルが指定されない限り無視される。

### **outlier**

履歴分析の中で Reg-ARIMA モデルが再推定されたら自動的に外れ値を探すか指定する。outlier スペックが用いられていなければ無効である。outlier=keep ならば系列全体を用いて自動的に外れ値を識別する。デフォルトではこの設定であり、外れ値の影響を考えて、係数は他の Reg-ARIMA モデルのパラメーターが推定されると再推定される。outlier=remove ならば、その外れ値と見られる回帰変数の値は履歴分析の間取り除かれる。そのため、その影響は推定されず、系列から除外される。このオプションにより、外れ値の識別をしないとどうなるかを調べることができる。outlier=auto ならば、自動的に識別された外れ値の中で履歴分析の開始日より前で outlierwin までの観測値だけは Reg-ARIMA モデルに含められる。それぞれの原系列から一部を取り除いて行う推定の中で、最後の outlierwin + 1 個の観測値に関してだけは自動的に外れ値の識別が行われる。識別された外れ値はその実行結果の中では用いられるが、分析に用いられるデータの最後の時点から outlierwin で指定した時点までにその外れ値が入っている場合にはその一連の履歴分析の中でしか用いられない。

### **outlierwin**

履歴分析の中で、どれだけの観測値を外れ値の分析に用いるかを指定する。デフォルトでは月次データで 12、四半期データで 4 である。この引数は outlier スペックが用いられ、history スペックの中で outlier=auto が用いられる場合に有効である。

### **refresh**

Reg-ARIMA モデルのパラメーターの推定でどの 2 つの初期値が用いられるかを指定する。refresh=yes ならば、最後のモデルの評価から推定されるパラメーターが次の Reg-ARIMA モデル推定の初期値として用いられる。refresh=no ならば、Reg-ARIMA モデルのパラメーターの初期値が系列全体から導かれる推定値として指定される。デフォルトでは refresh=no である。

### **x11outlier**

履歴分析の中で全ての不規則要素の回帰モデルに対して AO 外れ値を識別するかを指定する。x11outlier=yes ならば、AO 外れ値の識別はそれぞれの履歴分析に対して実行される。その AO 外れ値は、系列全体に対して AO 外れ値の識別がなされる時には不規則要素の回帰モデルに加えられるが、履歴分析の前には不規則要素の回帰モデルから取り除かれる。x11outlier=no ならば、AO 外れ値は自動的に認識されて履歴分析の中でも用いられる。データの全スパンから探された外れ値の日づけがある履歴分析のデータスパンに入っていなければ、外れ値はその分析のモデル

から取り除かれる。この AO 外れ値の影響を考慮して、係数は他の Reg-ARIMA モデルのパラメーターが推定されると再推定される。しかし、その他の AO 外れ値は自動的に識別されず、推定もされない。デフォルトでは `x11outlier=yes` である。

## 詳細

このオプションが呼び出されると、X-12-ARIMA は季節調整の初期の推定から最後の推定までの一連の改訂結果を表示する (表 6-12)。また、X-12-ARIMA は外れ値の予測誤差の履歴を作成して、Reg-ARIMA モデル推定から得られる尤度統計量を計算する。これらの改訂結果と履歴は以下のように得られる。

与えられた系列  $y_t$  ( $t = 1, \dots, T$ ) に対し系列  $y_1, y_2, \dots, y_n$  ( $t \leq n \leq T$ ) から計算される季節調整値を  $A_{t|n}$  とする。 $t$  時点の観測値の同時季節調整値は  $A_{t|t}$  で表され、最近の、すなわち直近の調整値は  $A_{t|T}$  で表される。季節調整値の改訂の割合は

$$R_t = \frac{A_{t|T} - A_{t|t}}{A_{t|t}}$$

で定義されこれがプログラムで計算される。乗法的季節調整や対数線形季節調整から得られるトレンド要素と季節調整項の改訂も割合で計算される。

加法的季節調整では、全ての  $A_{t|t}$  が同じ符号を持っている時には  $R_t$  は同じように計算される。(これはトレンドについても同じことが言える。) そうでない時には、差が計算される:

$$R_t = A_{t|T} - A_{t|t}$$

加法的季節調整では季節項の改訂は常に差として計算される。 $S_{t|T} - S_{t|t}$  すなわち、推定された季節項を用いて  $S_{t|T} - S_{t|t^*}$  と計算される。ここで  $t^*$  は  $t$  月に関して推定するのに用いた最後の月となる。

$C_{t|n}$  を系列  $y_1, y_2, \dots, y_n$  から計算された、 $t$  月に関する月ごとの (四半期ごとの) 季節調整済み系列の変化とする。すなわち、

$$C_{t|n} = \frac{A_{t|n} - A_{t-1|n}}{A_{t-1|n}}$$

その変化の割合は

$$R_t = C_{t|T} - C_{t|t}$$

で定義される。月ごとのトレンド項に関する改訂も同様に計算される。

引数 `sadjlags` と `trendlags` は同じ系列に対して異なるラグをとって改訂を分析する。その計算は引数 `target` の値に依存する。表 6-12 では異なる `target` の値に対してどのように計算されるかを示している。

1年と2年のラグ (月次データに対しては 12 や 24。四半期データに対しては 4 や 8) が `sadjlags` で指定された場合には  $t$  時点より 1年後と 2年後の季節調整値の間の改訂率が計算される: 月次データに対しては

$$RY_t = \frac{A_{t|t+24} - A_{t|t+12}}{A_{t|t+12}}$$

表 6-12: 改訂分析で得られる尺度

推定値 (Estimate)	Concurrent Target	Final Target
季節調整系列	$(A_{t t+lag_i} - A_{t t})/A_{t t}$	$(A_{t T} - A_{t t+lag_i})/A_{t t+lag_i}$
最終トレンド要素	$(T_{t t+lag_i} - T_{t t})/T_{t t}$	$(T_{t T} - T_{t t+lag_i})/T_{t t+lag_i}$
季節調整系列 (トレンド) の変化	$C_{t t+lag_i} - C_{t t}$	$C_{t T} - C_{t t+lag_i}$

*Estimate* は季節調整による推定値を意味する。  
*Concurrent Target* はターゲットが同時推定の時のラグつきの改訂履歴の計算式を意味する。  
*Final Target* はターゲットが最終の推定の時のラグつきの改訂履歴の計算式を意味する。  
 $A_{t|i}$  は  $t$  時点の  $i$  時点までの系列から計算された季節調整値を表す。  
 $T_{t|i}$  は  $t$  時点の  $i$  時点までの系列から計算されたトレンド項を表す。  
 $C_{t|i}$  は  $t$  時点の  $i$  時点までの系列から計算された季節調整値の変化を表す。

これは季節調整済み系列と季節調整系列の月ごとの変化 (四半期ごとの変化) についてのみ計算される。

ラグ付きの改訂結果は改訂の様子を良くあらわしている。同時推定の結果はさらにデータを付け加えるとその調整値がどれだけ変化するかを示している。最終推定の結果はその推定値が最終的な値にどのくらいの速さで収束計算されたかを示している。`sadjlags` と `trendlags` を用いる別の理由としては、同時推定はしばしば現時点の予備データに基づく推定になっていることが挙げられる。確定したデータが2ヶ月後にならないと利用できない場合それぞれの月に関して確定値に基づく調整の改訂を知るためには、`sadjlags = 2` とすれば良い。トレンドに関しては、トレンドの同時推定はしばしば不安定になることがしばしば挙げられる。そのため、分析者は X-12-ARIMA トレンドを推定する前に、数ヶ月の間、データが利用できるようになるまで待つことが望ましい。3ヶ月待つならば、`trendlags = 3` からトレンド推定の改訂が計算できる。季節調整の合成のため系列の改訂履歴が指定される時には、季節調整系列が直接的、間接的という2種類計算される。それぞれの系列要素に関して (たとえその系列が季節調整されていなくとも) 指定されなければならない。`composite` スペックを説明する例を参照のこと。直接的季節調整系列の改訂履歴 (表 6-12 の `sadj`) は間接的季節調整にだけ便利である。

自動季節調整フィルターのオプションを利用する場合にはプログラムがデータ・スパン (期間) が改訂履歴分析の中で変更されるごとに季節調整フィルターの選択がなされる。季節調整フィルターが分析に変更を与えた場合には警告文が表示され、それぞれのデータ・スパンに関して選択された季節調整フィルターの表が表示される。

予測の改訂結果の開始日は `fstep` の値に依存する。 $n$ -ステップ先予測誤差の履歴の開始日は履歴分析の開始日から  $n$  期先の日からである。例えば: `fstep = (1 12)` と `start = 1992.jan` を指定すると1-ステップと12-ステップ先予測の履歴が1992年2月からと、1993年1月から計算される。

ある状況では、プログラムは自動的に固定したモデルの係数を履歴分析に用いる。それは、Reg-ARIMA モデル推定において引数 `start` で指定できる履歴改訂分析の開始時点から設定されるデータ数が、データの打ち切りのために60以下になってしまった時や、系列の長さのためや、`series` や `composite` スペックの中の引数 `span` や `modelspan` の値のために起こりうる。この場合履歴分析の中で ARIMA と Reg-ARIMA モデルの回帰の係数は全ての系列 (または、指定されているならばモデルスパン) から推定される値に固定される。係数は引数 `modelspan` で指定された最後の日付よりも遅い日付のデータをデータ・スパンの中に含んでいる時にも固定される。特に極端な

場合としては、履歴分析の開始日よりモデル・スパンの最後の日が早い場合には Reg-ARIMA モデルの係数は履歴分析の間固定される。x11regression スペックの回帰モデルでも同様に扱われる。x11regression の引数 span で指定されたためにデータ・スパンが 60 以下になると、係数は固定される。

## 例

次に完全なスペック・ファイルを挙げておく。

### 例 1

乗法的季節調整では Reg-ARIMA モデル、予測、後ろ向き予測、回帰の時の外れ値の調整といった場合を除いて 3x9 の季節移動平均が全ての月に対して実行される。季節調整系列の改訂履歴は (デフォルトの設定では) 全てのデータに対して 12 年間の期間後 (3x9 季節項が用いられているため) 実行される。加えて、同時観測値の 2 期間後の推定値も計算される。

```
Series { Title = "Sales Of Livestock" Start = 1967.1
        File = "cattle.ori" }
X11 { SeasonalMA = S3X9 }
History { sadjlags = 2 }
```

### 例 2

外れ値とレベルシフトをあらかじめ調整しておいて季節 ARIMA モデルを実行する。回帰変数は定数項と取引日効果と 2 つのレベルシフト項である。(1 つは 1972 年 5 月でもう 1 つは 1976 年 9 月) 季節 ARIMA は (0,1,2)(1,1,0)12 を指定する。1 ステップ先推定誤差の履歴を作成する。分析の開始時点は 1975 年 1 月で、これは 1 ステップ先推定誤差は 1975 年 2 月に関して計算されることを示している。

```
series { title = "Exports of Leather goods"
        start = 1969.jul file = "expleth.dat" }
regression { variables = (const td ls1972.may ls1976.oct) }
arima { model = (0 1 2)(1 1 0) }
estimate { }
history { estimates = fcst fstep = 1 start=1975.jan }
```

### 例 3

例 2 と同じ Reg-ARIMA モデルとデータを用いて 1 ステップ予測と 12 ステップ予測の予測誤差を計算する履歴の開始時点は 1975 年の 1 月。履歴はファイルに保存する。そのファイルには、12 ステップ予測誤差が定義されないところ (1975 年の 2 月から 12 月まで) には 0 が表示される。

```
series { title = "Exports of Leather goods"
        start = 1969.jul
        file = "expleth.dat" }
regression { variables = (const td ls1972.may ls1976.oct) }
arima { model = (0 1 2)(1 1 0) }
estimate { }
history { estimates = fcst save = r6 start = 1975.jan }
```

### 例 4

3×3 の季節移動平均と Reg-ARIMA モデルを用いて乗法的月次季節調整を実行する。Reg-ARIMA モデルは最後の 12 月のデータまで含めて計算される。Reg-ARIMA モデルの係数は全

での12月について計算され、季節調整系列の改訂履歴とトレンド要素は6年後から計算される。また、系列の季節調整の改訂履歴はその系列を含む合成系列の間接的季節調整の計算に統合される。(メタファイルの中の合成系列のスペックファイルは適切な history スペックの中に含まれなければならない。例5を参照。)

```
series { title = "Housing Starts in the Mid-West"
         start = 1967.1
         file = "hsmwtot.ori"
         modelspan = (,0.Dec)
         comptype=add
}
regression { variables = td }
arima { model = (0 1 2)(0 1 1) }
x11 { seasonalMA = S3X3 }
history { estimates = (sadj trend) }
```

#### 例 5

全ての月について Reg-ARIMA モデルの予測系列から拡張する形で合成月次季節調整を 3x3 季節移動平均で実行する。Reg-ARIMA モデルは最後の12月までのデータを用いて計算される。直接・間接季節調整系列と直接季節調整のトレンド項の改訂履歴は、Reg-ARIMA モデルのパラメーターを12月ごとに再計算して計算する。それぞれの推定値の改訂率が別のファイルに保存される。

```
composite{ title = "Total Housing Starts in the US"
           modelspan = (,0.Dec)
}
regression { variables = td }
arima { model = (0 1 1)(0 1 1) }
x11 { seasonalMA = S3X3 }
history { estimates = (sadj trend)
         save = (sar iar trr) }
```

## 識別 (identify)

---

### 解説

Reg-ARIMA モデルの ARIMA パートを同定するための表と、標本自己相関関数 (ACF) と標本偏自己相関関数 (PACF) のプロットの出力を設定する。ACF と PACF は **diff** 引数と **sdiff** 引数によって規定される非季節階差と季節階差の階数のうちの全ての組み合わせについて出力される。スペック・回帰 **regression** があると、ACF と PACF は回帰の残差の特定の階差系列に関して計算される。スペック・回帰 **regression** がなければ、ACF と PACF は原系列の特定の階差系列に関して計算される。

### 利用法

```
identify(diff = (0, 1)
         sdiff = (0, 1)
         maxlag = 36
         print = (none +acf +acfplot +pacf +pacfplot])
}
```

### 引数

#### **diff**

非季節階差の次数を設定する。この値が 0 であることは階差を取らないことを意味し、1 は 1 階の階差  $(1 - B)$ 、2 は 2 階の階差  $(1 - B)^2$  を表す。ACF と PACF は **sdiff** で設定される季節階差との全ての組み合わせについて計算される。デフォルトの値は **diff**=(0) である。

#### **maxlag**

ACF と PACF の両方の表とプロットについて、計算されるラグの大きさを設定する。デフォルトの値は月次データで 36、四半期データで 12 である。

#### **print and save**

出力に表示できる表については表 6-13 を参照。

#### **sdiff**

季節階差の階数を設定する。0 は季節階差を取らないことを意味し、1 ならば 1 階の季節階差  $(1 - B^s)$  等となる。ACF と PACF は **diff** で設定される非季節階差との全ての組み合わせについて計算される。デフォルトの値は **sdiff**=(0) である。

### 詳細事項

**regression** スペックが存在するならば、プログラムは (**transform** スペックを実行した後に) 系列と回帰変数に対して、**diff** 引数と **sdiff** 引数によって設定された最大の階差を取る。次に階差

系列は階差を取られた回帰変数に対して回帰される。次にその結果である回帰係数 ( $\hat{\beta}_i$ ) は、階差を取らない回帰効果 ( $\sum_i \hat{\beta}_i x_{it}$ ) の計算に利用される。そして原系列 ( $y_t$ ) から回帰効果を引くことによって階差を取らない回帰誤差 ( $\tilde{z}_t = y_t - \sum_i \hat{\beta}_i x_{it}$ ) が求められる。この回帰誤差系列と **diff** と **sdiff** によって設定されたその階差は、次に ACF と PACF を生成するのに利用される。

このようなプロセスにはひとつの例外がある。もし **regression** スペックにおいて定数項が設定されているならば (`variables = (const ...)`)、定数項は回帰が行われる時点ではモデルに含まれているが、元の系列から回帰効果を引き去る段階では含まれない。詳細は 3.4 節を参照されたい。

ACF と PACF は **diff** と **sdiff** によって設定される非季節階差と季節階差の全ての組み合わせに対して計算される。例えば、`diff = (0, 1)` および `sdiff = 1` としたならば、その場合には ACF と PACF は  $(1 - B^s)\tilde{z}_t$  と  $(1 - B)(1 - B^s)\tilde{z}_t$  について計算される。ただしここで  $\tilde{z}_t$  はすでに述べたように系列の回帰誤差とし、 $s$  は **series** スペック内で設定される季節周期であるとする。またもし `diff = (0, 1, 2)` および `sdiff = (0, 1)` としたならば、ACF と PACF は次の 6 系列について計算されることになる。すなわち  $\tilde{z}_t$ 、 $(1 - B)\tilde{z}_t$ 、 $(1 - B)^2\tilde{z}_t$ 、 $(1 - B^s)\tilde{z}_t$ 、 $(1 - B)(1 - B^s)\tilde{z}_t$  および  $(1 - B)^2(1 - B^s)\tilde{z}_t$  である。

スペック **identify** とスペック **estimate** の両方が存在する場合、スペック **identify** が先に実行される。スペック **identify** はスペック **regression** が存在するならばその情報を利用するが、スペック **arima** については無視をするという点に注意されたい。

ユーザーは階差操作によって、ユーザー定義の回帰変数も含めて、回帰の説明変数が非正則とならないように注意する必要がある。さもないと非正則性は致命的なエラーをもたらす場合がある。そのようなことが生じるひとつのケースは、**sdiff** を正の値 (例えば 1) としている一方で、**regression** スペックにおいて `variables = (seasonal)` としているような場合である。

もし ACF と PACF に対して設定したラグの次数が、原系列 (あるいは階差系列) の長さをよりも大きいならば、ACF と PACF は可能な最大のラグまでについてのみ計算される。

## 例

以下に完全なスペック・ファイルの例を示す。

### 例 1

$y_t = \log(Y_t)$  となる月次系列に関して必要な階差の次数を識別するために有用な ACF の表を生成する。ただし  $Y_t$  は **series** スペックで設定された原系列であるとする。ここで ACF は  $y_t$ 、 $(1 - B)y_t$ 、 $(1 - B^{12})y_t$  および  $(1 - B)(1 - B^{12})y_t$  に対して計算される。**regression** スペックは存在しないので、回帰効果としては何も除去されない。ACF はデフォルトであるラグ 36 個分について計算される。

```
series { title = "Monthly Sales"  start = 1976.jan
         data = (138 128 ... 297) }
transform { function = log }
identify { diff = (0, 1)
          sdiff = (0, 1)
          print = (none +acf) }
```



## 例 2

ACF と PACF を求める前に固定的な季節効果が除去される。regression スペックは固定的な季節変数と同様に、トレンド要素を含んでいる。identify スペックは、 $(1-B)y_t$  を階差回帰変数  $(1-B)x_{it}$  に回帰し、階差を取らない回帰効果を  $\tilde{z}_t = y_t - \sum_{i=2}^{12} \tilde{\beta}_i x_{it}$  によって除去することで、固定的な季節効果を取り除く（トレンドの定数項  $\tilde{\beta}_1 x_{1t}$  は引き去らない）。次に  $\tilde{z}_t$  and  $(1-B)\tilde{z}_t$  の ACF と PACF が計算される。定数項は  $(1-B)y_t$  がゼロでない平均を持つことを許すような、線形な定数トレンドである。例えば  $x_{1t} = t$ 。

```
SERIES      { TITLE = "MONTHLY SALES"  START = 1976.JAN
              DATA = (138 128 ... 297) }
REGRESSION  { VARIABLES = (CONST SEASONAL) }
IDENTIFY    { DIFF = (0,1) }
```

## 例 3

Reg-ARIMA モデルの AR パートと MA パートを識別するための ACF と PACF を生成する。ACF と PCF の表は出力しない。 $Y_t$  を例 1 と同じ系列とし、1 階の非季節階差と 1 階の季節階差が選択され、曜日効果と復活祭効果を含むものとする。regression スペックが存在するので、identify スペックはまず最初に  $(1-B)(1-B^{12})y_t$  を  $(1-B)(1-B^{12})x_{it}$  に対して回帰する。ここで  $x_{it}$  は曜日効果と復活祭効果を表す回帰変数であるとする。また  $y_t$  は、原系列  $Y_t$  を月の長さ効果 (length-of-month effects) を調整の上で対数を取った系列であるとする。(regression スペックの td の解説を参照されたい。)  $\tilde{\beta}_i$  を推定された回帰係数を表すとすると、この identify スペックは、回帰残差系列  $(1-B)(1-B^{12})(y_t - \sum_i \tilde{\beta}_i x_{it})$  に関する ACF と PACF のプロットを生成する。ACF と PACF は 30 のラグに渡って計算される。

```
Series { Title = "Monthly Sales"  Start = 1976.Jan
        Data = (138 128 ... 297) }
Transform { Function = Log }
Regression { Variables = (Td Easter[14])}
Identify { Diff = (1)  Sdiff = (1)  Maxlag = 30
          Print = (None +ACFplot +PACFplot) }
```

## 例 4

モデルの識別のために (16 のラグに渡って) ACF と PACF を生成し、四半期系列に対する暫定的なモデルの推定を行う。ここでは 1971 年第一四半期において既知のレベルシフトが存在する。この効果は  $(1-B)(1-B^4)y_t$  を階差を取ったレベル・シフト変数に回帰することによって推定される。階差を取らない回帰残差を求める為に、回帰効果は  $\tilde{z}_t = y_t - \tilde{\beta}LS71.1_t$  によって除去され、ACF と PACF は  $\tilde{z}_t$ 、 $(1-B)\tilde{z}_t$ 、 $(1-B^4)\tilde{z}_t$  および  $(1-B)(1-B^4)\tilde{z}_t$  に対して計算される。スペック identify はスペック arima の情報を無視する。

以下のスペックファイルでは、 $(1-B)(1-B^4)(y_t - \beta LS71.1_t) = (1-\theta B)(1-\Theta B^4)a_t$  という Reg-ARIMA モデルについての推定と標準的な事後診断が行われる。このような暫定的なモデルの推定をモデル識別のための ACF と PACF の計算と同時に行うことは、適切な ARIMA モデルに関する事前の情報がある場合には、しばしば有用である。これは、データに対するモデルの当てはめが過去に行われており、その後系列に新しいデータが付け加わって延長された場合などに相当するであろう。もし事後診断の結果、暫定的なモデルが不適切だと思われるならば、ユーザー

は事後診断結果と `identify` スペックの出力の両方の情報を、新しいモデルの選択に利用することになるだろう。

```
series { title = "Quarterly Sales" start = 1963.1 period = 4
         data = (56.7 57.7 ... 68.0) }
regression { variables = (ls1971.1) }
arma { model = (0 1 1)(0 1 1) }
identify { diff = (0, 1) sdiff = (0, 1) maxlag = 16 }
estimate { }
check { }
```

## 外れ値 (outlier)

---

### 解説

加法的な外れ値 (additive outliers)、一時的変化の外れ値 (temporary change outliers)、レベルシフト、またはこれら3つの組み合わせによって表される異常値の自動的な検出を行う。外れ値が検出された後は、適切な回帰変数が自動的に検出された外れ値としてモデルに組み込まれて、モデルの再推定が行われる。この過程は外れ値が1つも検出されなくなるまで繰り返される。もし2つかそれ以上の個数のレベルシフトが検出されたなら (またはスペック regression によってレベル・シフト変数が設定されているならば)、2つ以上の連続したレベル・シフトの効果により、一時的なレベル・シフトを打ち消してしまうという帰無仮説を検定するために、 $t$ 統計量を計算することができる。

### 利用法

```
outlier{types = all
  critical = 3.75
  method = addall
  span = (1973.may, 1992.sep)
  lsrn = 0
  print = (none +header)
  save = (tests)
}
```

### 引数

外れ値検出の境界値となる  $t$ -statistic の絶対値を設定する。デフォルトの値は Reg-ARIMA モデルの推定に用いられるモデル期間 (スパン) に応じて決定される。これは Ljung(1993) の定式化に基づいている。

この引数に対してただひとつの値が設定された場合 (`critical = 3.5`)、その値が全てのタイプの外れ値に対する境界値として適用される。もし3つの値のリストが引数として与えられるならば (`critical = (3.5, 4.0, 4.0)`)、リストの最初の数値は加法的な外れ値 (additive outliers) の境界値 (この例では 3.5) として、二番目の数値はレベル・シフトの境界値 (この例では 4.0) として、三番目の数値は一時的変化の外れ値 (temporary change outliers) の境界値 (この例では 4.0) としてそれぞれ用いられる。リストに `critical = (3.25, , 3.25)` のように欠落が含まれる場合には、その部分はモデル期間 (スパン) に対応したデフォルトの値が用いられる。境界値の値を大きくすることは、外れ値に対する敏感性を小さくすることになり、外れ値として処理される観測データの個数が小さくなる可能性がある。

#### lsrn

一時的変化の外れ値 (temporary change outliers) を作るために、2回かそれ以上の個別の連続

的なレベル・シフトが取り消されるという帰無仮説を検定するための  $t$  統計量が計算される。この  $t$  統計量は、外れ値のパラメーターをそれぞれの標準誤差で割った値の合計として算出される。(Otto and Bell (1993) を参照。) これは自動的に検出されるレベル・シフト変数と、regression スペックで指定されたレベル・シフト変数の両方に対して適用される。Lsrun は 0 から 5 の値を取り得る。1 では一時的変化の外れ値 (temporary change outliers) についての  $t$  統計量が計算されない。もし lsrn に指定された数値が外れ値の自動検出で見つかるレベルシフトの総数を超過していたならば、lsrun はそのその数を新たな値としてリセットされる。lsrun のデフォルトの値は 0 である。この設定では一時的変化の外れ値 (temporary change outliers) の  $t$  統計量は計算されない。

#### method

検出された外れ値をモデルにどのような形で連続的に付加していくかを設定する。選択肢は method = addone か method = addall のいずれかである。この 2 つの方法の説明については詳細事項を参照されたい。デフォルトの値は method = addone である。

#### print and save

出力される内容のリストは表 6-15 に記載されている。

表 6-13: Outlier に関連する出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
header	+	.	.	options specified for outlier detection including critical value, outlier span, and types of outliers searched for
iterations	.	.	oit	detailed results for each iteration of outlier detection including outliers detected, outliers deleted, model parameter estimates, and robust and non-robust estimates of the residual standard deviation
tests	.	.	.	$t$ -statistics for every time point and outlier type on each outlier detection iteration
temporaryls	+	+	.	summary of $t$ -statistics for temporary level shift tests
finaltests	.	.	fts	$t$ -statistics for every time point and outlier type generated during the final outlier detection iteration

*Name* gives the name of each table for use with the print and save arguments.  
*Default* indicates which tables are printed (+) or not printed (.) by default.  
*Brief* indicates which tables are printed (+) or not printed (.) when the brief print level is specified. See section 5.1 for more information on print levels.  
*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the print and save arguments.

#### span

外れ値を探索する期間の始期と終期を指定する。始期と終期は両方とも時系列に含まれている必要があり、また series スペックの modelspan 引数においてモデルスパンが指定されているならば、その範囲にも含まれている必要がある。また始期は終期よりも先でなければならない。span = (1976. jan, ) のように引数の値が欠落している場合は、時系列の始期と終期がデフォルトして設定される。(もし series スペックの modelspan 引数が指定されているならば、上記のように、始期と終期は modelspan 引数での指定に従う。)

### types

検出する外れ値の種類を指定する。選択肢は `types = ao`、加法的な外れ値 (additive outliers) のみ; `types = ls`、レベルシフトのみ; `types = tc`、一時的変化の外れ値 (temporary change outliers) のみ; `types = all`、加法的な外れ値、一時的変化の外れ値およびレベルシフトを同時に; `types = none`、外れ値検出のオフ。またデフォルトの値は `types = (ao ls)` である。

### 稀にしか使われない引数

#### tcrate

一時的変化の外れ値のための説明変数の影響が減衰していく割合を定義する。この引数の値は 0 以上 1 以下の数値でなければならない。デフォルトの値は `tcrate=0.7 ** (12/period)` であり、`period` は一年間に含まれる観測値の個数を表す。これにより `tcrate` のデフォルトの値で、異なる周期の系列に対しても一年間で同じ割合で減衰していくことが保証される。`regression` スペックでこの引数が設定された場合、ここで設定をする必要はない。

### 詳細事項

データの最初の時点におけるレベルシフトの推定は、それ以前の系列が不明であるため、不可能である。従って、データの始点における LS の検定統計量は計算されない。またどのようにデータの最後の時点におけるレベルシフトは、その時点での AO と区別することができず、データの 2 番目の時点における LS はデータの始点における AO と区別することができない。よって AO が同時に検出されている場合には、2 番目の時点と終点におけるレベルシフトに関する検定統計量は計算されない。計算されないレベルシフトに関する検定統計量は 0 として出力される。

同様に、終点での一時的変化の外れ値 (temporary change outliers) はその時点での AP と区別できないため、AO が同時に検出されている場合には、終点での TC の検定統計量は計算されない。計算されない TC の検定統計量は 0 として出力される。

`addone` は以下のような形で働く。プログラムは外れ値検出の機能しうる全ての時点において、指定された外れ値のタイプのそれぞれについて  $t$  統計量を計算する。もし  $t$  統計量の絶対値の最大値が境界値を越えていたならば、外れ値として検出され、適切な回帰変数がモデルに付加される。プログラムはそのようにして得られた新しいモデルを推定し、さらなる外れ値を探索する。この過程は外れ値が見つからなくなるまで繰り返される。この点に関して、明確でない外れ値 ( $t$  統計量が境界値を超過しなくなったもの) をモデルから削除する手順が踏まれる。これは 1 時点について最も明確でない外れ値に対して行われ、モデルに残っている全ての外れ値が明確に有意であるとみなされるまで続けられる。この削除のプロセスでは、通常の (頑健でない) 残差分散の推定量が使われるが、これは外れ値検出の過程で得られるものとは幾分異なった  $t$  統計量をもたらすことがある。

`addall` は全体としては `addone` と同様のステップを踏むが、`addall` では絶対値で  $t$  統計量が境界値を越えた全ての外れ値が組み込まれるという点で異なる。典型的には、この方法によりモデルに付加された外れ値変数は、新しいモデルの推定が行われた場合に明確に有意ではなくなることもありうる。従って `addall` は検出の過程で付加される外れ値のうち、不必要なものを削除する

過程に大きく依存している。

`addone` と `addall` の違いは、最終的に異なる外れ値の組み合わせを生成し得る点である。この二つの方法の実用上の相違点は重要である。まず第一に、`addone` は `addall` に較べて計算時間がより長くかかるということである。第二に `addall` は、モデルにおいて許される回帰変数の個数を上回る、多数の外れ値を付加してしまう可能性がある。この場合にはプログラムはエラーメッセージを出力し停止する。対応策としては、より少数の外れ値しか検出されないように境界値を上げるか、`addone` に切り替えることであるが、このような事態はそれほど頻繁には起こらないと思われる。

両方の方法について、検出された全ての外れ値の  $t$  統計量は `print=tests` を指定することによって出力することができる。このオプションでは相当量の結果を出力する。

境界値を選択するには判断力と経験が必要である。低次の ARIMA モデルから発生させた 200 個の長さの系列によるシミュレーションによると、Chang, Tiao, and Chen (1988) は AO の検出に関して、高い感応度には 3、中間の感応度には 3.5、低感応度には 4 という数値を用いることを奨励している。

外れ値検出は `regression` スペックと `arima` スペックで指定されたモデルと推定されたパラメータの値から出発する。もし `estimate` が欠けていたならば、`outlier` スペックは、デフォルトの推定オプションによって、外れ値の検出に先立って強制的にモデルの推定を実行する。もし特定の既知の時点での外れ値の存在が予め疑われるならば、適当な AO, TC または LS の回帰変数を、スペック `regression` 内で指定しておくことができる。

外れ値検出の結果は指定された Reg-ARIMA モデルに非常に大きく依存する。ある観測値は、その他の観測値に較べてモデルに対するフィットの度合いが小さい場合に外れ値として分類される。従って非常に不適切に Reg-ARIMA モデルを利用すると、不適切な外れ値の調整につながってしまうことになる。

## 例

次に完全なスペック・ファイルの例を挙げておこう。

### 例 1

AO と LS のデータ区間全体での同時探索を行う。`addone` が用いられ、境界値はデフォルトに従いデータ長に依存して決定される。もしこのモデルでの外れ値検出によって確認されるレベル・シフトが 2 つ以上存在するならば、TC を形成するために 2 から 5 回の連続的なレベル・シフトを取り消すかどうかを検定するための  $t$  統計量が計算される。スペック `estimate` がないので、スペック `outlier` ではデフォルトとして設定された方法で強制的にモデルの推定を行う。

```
series { title = "Monthly sales" start = 1976.jan
         data = (138 128 ... 297) }
arima { model = (0 1 1)(0 1 1)12 }
outlier {lsrum = 5 types=(ao ls) }
```

### 例 2

AO のみを `addall` を用いて境界値を  $t = 4.0$  として探索する。スペック `series` において `span` 引数が用いられているので、そこでの区間 (1980 年 1 月から 1992 年 12 月) のみがモデルの推定と

外れ値の検出に使用される。lsrun はデフォルトの値である 0 を取るので、regression スペックで指定されている 2 つのレベル・シフトは、一時的な LS として相殺されるてしまうか否かの検定は行われない。

```
Series { Title = "Monthly Sales" Start = 1976.Jan
        Data = (138 128 ... 297)
        Span = (1980.Jan, 1992.Dec) }
Regression { Variables = (LS1981.Jun LS1990.Nov) }
Arima { Model = (0 1 1)(0 1 1)12 }
Estimate { }
Outlier { Types = A0 Method = Addall Critical = 4.0 }
```

### 例 3

例 2 と同一の区間を使用してモデルの推定が行われるが、1987 年と 1988 年においてのみ LS の有無が探索される。デフォルトである addone が使用されるが、境界値は  $t = 3.0$  である。2 つの LS の全ての可能な組み合わせについて、一時的な LS としてキャンセルアウトするかどうかの検定が行われる。

```
series { title = "Monthly sales" start = 1976.jan
        data = (138 128 ... 297)
        span = (1980.jan, 1992.dec) }
arima { model = (0 1 1)(0 1 1)12 }
estimate { }
outlier { types = ls
        critical = 3.0
        lsrn = 2
        span = (1987.jan, 1988.dec) }
```

### 例 4

例 2,3 と同様の区間を用いてモデル推定が行われるが、AO,TC および LS について探索される。デフォルトの addone が利用されるが、境界値に関しては AO は  $t_{AO} = 3.0$ 、LS は  $t_{LS} = 4.5$ 、TC は  $t_{TC} = 4.0$  が用いられる。

```
series { title = "Monthly sales" start = 1976.jan
        data = (138 128 ... 297)
        span = (1980.jan, 1992.dec) }
arima { model = (0 1 1)(0 1 1)12 }
estimate { }
outlier { critical = (3.0, 4.5, 4.0) }
```

## 回帰 (regression)

---

### 解説

Reg-ARIMA モデルに含まれる回帰変数を設定する。また ARIMA モデルの同定を補助するスペック `identify` によって除去すべき回帰効果を設定する。あらかじめ定義されている回帰変数は、`variables` 引数によって選択できる。利用できる回帰変数は、定数、固定的季節性、曜日変動、休日変動、加法的外れ値 (additive outliers)、レベル・シフト、一時的変化である。状態の変化 (change-of-regime) に対する回帰変数は、季節性と曜日効果の回帰変数に対して設定することができる。`user` 引数によって、ユーザーが定義した回帰変数をこれらに付け加えることができる。ユーザー定義変数に対しては、`data` 引数によるデータの指定か、`file` 引数によるファイルの指定のうちどちらかが必ず必要である (両方ではない)。`regression` スペックは予め定義された回帰変数とユーザー定義変数の両方を利用することができる。

### 使用法

```
regression{variables = (const)
    seasonal or sincos[1, 2, 3]
    td or tdnolpyear or tdstock[31] or
    tdccoef or tdnolpyear
    lom or loq
    lpyear
    easter[8]
    labor[8]
    thank[1]
    ao[1967.apr]
    ls[1972.sep]
    tc[1979.sep]
    rp[1965.nov-1968.may]
    print = (none)
    save = (rmx)
    savelog = aictest
    user = (temp precip)
    usertype = holiday
    start = 1955.jan
    data = (25 0.1 ...) or file = "weather.dat"
                                format = "(2f5.1)"
    aictest = ( easter user
               td or tdnolpyear or tdstock or
               tdccoef or tdnolpyear
    )
}
```

### 引数

#### `aictest`

Reg-ARIMA モデルに回帰変数の組を導入するかどうかを決定するために、AIC に基づく選



択を行うかを設定する。この引数に対して利用可能な回帰変数は `td`、`tdnolpyear`、`tdstock`、`tdlcoef`、`tdlnoyear`、`easter` および `user` である。例えば、もし曜日変動を含むモデルが選択されたならば、AIC（系列の長さに応じて調整をし、以後は AICC と記述する）は、曜日変動の回帰変数を含む場合と含まない場合の両方について計算される。デフォルトではより小さい AICC を与えるモデルが、予測値の生成や外れ値の検出等に用いられる。1 種類以上の回帰変数が設定された場合、AIC テストは、(a) 曜日変動の回帰変数、(b) 復活祭の回帰変数、(c) ユーザー定義変数、という順で行われる。同じ種類の変数が複数ある場合には（例えば複数の曜日効果変数）、`aicctest` のプロシージャはそれらをひとつのグループとして処理する。従って最終的なモデルには、そのタイプの変数が全て含まれるか、1 つも含まれないかのどちらかとなる。判定の手順についてのさらなる情報は詳細事項を参照されたい。このオプションが設定されない場合、AIC に基づくモデル選択は行われない。

#### **data**

ユーザー定義回帰変数となるデータを指定する。このデータの期間は調整しようとする系列の期間（あるいは `series` スペックの `span` 引数によって範囲が指定されているならばその範囲）を含むものでなければならない。また `forecast` スペックにおいて指定された予測と逆予測の範囲をも含んでいる必要がある。データはフリーフォーマットで読み込まれる。この引数によって与えられる数値は、`user` 引数によって名前を指定されたユーザー定義変数に割り当てられる。この割り当ては、第 1 時点での全てのユーザー定義変数を通して進められ、また第 2 時点での全ての変数を通して進められる。`data` 引数が設定されている場合、`file` 引数を同時に用いることはできない。

#### **file**

全てのユーザー定義変数の値を記述したファイルのファイル名を指定する。ファイル名はクォーテーションマークで括られている必要がある。ファイルがカレントディレクトリ内に無いならば、ファイルのパスが指定される必要がある。`data` におけるのと同様、データ長は調整しようとする系列全体と、予測と逆予測の範囲の両方をカバーしていなければならない。`file` 引数が設定されているならば、`data` 引数は使用することができない。

#### **format**

`file` 引数で指定された回帰変数のデータに使われているフォーマットを指定する。引数の値としては 4 つの可能性がある。

- (a) FORTRAN 形式。これはクォーテーションで括られている必要があり、括弧が必要である。  
(例: `format='(6f12.0)'`);
- (b) date-value 形式。これは年と月または四半期と、それに応じたユーザー定義変数の値が、この順序でフリーフォーマットの各行に記載されたファイルである。従って 3 つの回帰変数を含むデータファイルでは、1991 年 7 月の値が 0,0,1 ならば、1991 7 0 0 1 となる。全てのユーザー定義変数は同一のファイル上に、`user` 引数に指定された順序と同じ順序で記載される必要がある。(例: `format='{datevalue}'`);
- (c) x12save 形式。これは X-12-ARIMA が表を保存するために用いる形式である。この形式では、前回の X-12-ARIMA の実行で保存されたファイルを読むことができる。(例: `format='{x12save}'`)

もし `format` 引数が設定されないならば、データはフリーフォーマットで読み込まれる。`free`

*format*では、1つの行に記述されている全ての数字が、次の行が始まるまで読み込まれ、個々の値は1つ以上の(カンマやタブではなく)空白によって区切られる必要がある。**Format**は**data**引数と同時に使用することはできず、**file**とのみ用いられる。

#### print and save

出力に関するオプションは表を参照されたい。

表 6-14: Regression に関連する出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
regressionmatrix	.	.	rmx	values of regression variables with associated dates
aictest	+	+	.	output from AIC-based test(s) for trading day, easter, and user-defined regression variables
outlier	+	+	otl	combined regARIMA outlier factors
aoutlier	+	+	ao	regARIMA additive (or point) outlier factors
levelshift	+	+	ls	regARIMA level change and ramp outlier factors
temporarychange	+	+	tc	regARIMA temporary change outlier factors
tradingday	+	+	td	regARIMA trading day factors
holiday	+	+	hol	regARIMA holiday factors
regseasonal	+	+	a10	regARIMA user-defined seasonal factors
userdef	+	+	usr	factors from user-defined regression variables

*Name* gives the name of each table for use with the **print** and **save** arguments.

*Default* indicates which tables are printed (+) or not printed (.) by default.

*Brief* indicates which tables are printed (+) or not printed (.) when the **brief print** level is specified. See section 5.1 for more information on print levels.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the **print** and **save** arguments.

#### savelog

**savelog=aictest** または **savelog=ats** という設定により、**aictest** 引数によって指定された AIC に基づくモデル選択の結果をログファイルに出力することができる。(ログファイルにかんして詳しくは 2.6 節を参照されたい。)

#### start

ユーザー定義変数のデータの始まりの日付を指定する。デフォルトはデータの始点となっている。適切な値はデータの開始時点までの(あるいは **series** スペックの **span** 引数で指定された開始時点までの)任意の日付となる。

#### user

ユーザー定義変数の名前を設定する。名前は、モデルに組み込まれるユーザー定義変数の全てに対して必要である。与えられた名前はプログラムの出力において、推定された係数のラベルとして使用される。ユーザー定義変数の数値は、**data** 引数か **file** 引数のどちらか一方を用いて与えられなければならない。ユーザー定義変数の個数の最大値は 52 である。(この最大個数は変更可能である。詳しくは 2.8 節を参照。)

#### usertype

モデルによって推定された回帰効果の種類を、個々のユーザー定義変数に割り当てる。これにより変数とその推定された効果が利用可能となり、同じタイプの予め定義された回帰変数と同様に出力される。このオプションはプログラムによって与えられた回帰効果に対する代替の変数を試してみる場合に有用である。

ユーザー定義変数の効果の種類は、定数 (**constant**)、季節性 (**seasonal** or **userseasonal**)、

曜日効果 (**td**)、ストック曜日効果 (**tdstock**)、月の長さ (**length-of-month**) (**lom**)、1 四半期の長さ (**length-of-quarter**) (**loq**)、閏年 (**lpyear**)、休日ないし復活祭 (**holiday**, **easter** または **US holidays thanks** および **labor**)、外れ値 (**ao**, **ls**, **rp** および **tc**) およびその他のユーザー定義変数 (**user**) である。ひとつの効果の種類は、**regression** スペック (**usertype=td**) で指定された全てのユーザー定義変数について設定することができる (**usertype=td**)。または個々のユーザー定義変数それぞれについて設定することができる (**usertype=(td td td td td td holiday user)**)。ユーザー定義変数に対して一度タイプの割り当てが行われると、**aicctest** や **noapply** といったその他の引数に関する追加の指定は、**user** ではなくここで指定されたタイプを使用しなければならない。このオプションが利用されない場合は、全てのユーザー定義変数は **user** 引数のタイプとなる。ユーザー定義変数の指定に関するさらなる情報については詳細事項を参照されたい。

#### **variables**

モデルに組み込む予め定義された回帰変数のリストを設定する。これらの変数に対応するデータはプログラムによって計算され、多くの場合はカレンダーの関数となる。利用可能な回帰変数の表は詳細事項を参照されたい。また追加的情報や実際に利用される回帰変数を示した表に関しては、3.3 節を参照されたい。

### 稀にしか使われない引数

#### **aicdiff**

回帰変数付きのモデルが選択されるために、**aicctest** で指定された回帰変数を含むモデルの AIC 値が回帰変数を含まないモデルの AIC 値をどれだけ下回らなければならないかを指定する。デフォルトの値は **aicdiff=0.0** である。このオプションが **aicctest** 引数との関係でどのように使われるべきかについては、詳細事項を参照されたい。

#### **b**

**variables** 引数と **user** 引数で回帰変数が指定された時に、回帰係数の初期値を設定する。**b** 引数は **Reg-ARIMA** モデルにおけるすべての回帰係数に対して初期値を割り当てなければならないが、**regression** スペックにおいて **variables** 引数と **user** 引数よりも後で設定される必要がある。初期値は、引数のリスト内の値を指定するか、または以下の例のように欠落値を示すことによって、パラメーターに対して割り当てられる。欠落値はデフォルトの値である 0.1 となる。例えば、2 つの回帰変数を含むモデルでは、**b=(0.7, )** という設定は **b=(0.7,0.1)** と同等であるが、**b=(0.7)** という記述は許されない。3 つの回帰変数を含むモデルについては、**b=(0.8,, -0.4)** という設定は **b=(0.8,0.1,-0.4)** と同等である。パラメーターを特定の値で固定するには、**b** 引数のリストにおいて {f} を用いればよい。例えば **b=(0.7f, 0.1)** とすればよい。

#### **centeruser**

(標本) 平均または季節の平均をユーザー定義変数から取り除くための設定を行う。**centeruser=mean** とすると、ユーザー定義変数の平均が回帰変数から差し引かれる。また **centeruser=seasonal** とすると、各月 (もしくは各四半期) のそれぞれの平均がユーザー定義変数から差し引かれる。このオプションが設定されない場合は、ユーザー定義変数はすでに適切に中央に配置されていると見なされ、調整されない。

### eastermeans

easter[w] 変数によって導入される復活祭変数の季節性を除去するために、表 3-1 の脚注 5 に示されているように、長期 (400 年) の月平均を使用する (eastermeans=yes) か、またはその代わりに復活祭変数の係数の計算に用いられるデータ期間から月毎の平均を算出して用いる (eastermeans=no) かを指定する。デフォルトは eastermeans=yes である。この引数は、回帰変数に復活祭変数が含まれていない場合か、または復活祭変数だけが sceaster[w] となっている場合には無視される (詳細事項を参照)。

### noapply

スペック x11 で設定された季節調整が適用される前の事前調整において、regression スペックで指定された回帰変数のうち、その係数推定値に基づく回帰効果を原型列から除去しない変数のリストを指定する。

適用可能なタイプは、曜日効果 (td)、復活祭、レイバーデー、感謝祭からクリスマス休暇 (holiday)、加法的外れ値 (ao)、レベル・シフト (ls)、一時的変化 (tc) およびユーザー定義変数 (user) である。

### tcrate

一時的変化に対応する回帰変数が減衰する割合を指定する。この値は 0 以上 1 以下の数値でなければならない。デフォルトの値は tcrate=0.7 \*\* (12 / period) である。ここで period は一年分の観測値の個数である (四半期系列ならば 4 である)。デフォルトがこの形式となっていることにより、デフォルトでの tcrate では、異なる周期の系列に対しても、回帰変数は一年間で同率の減衰をすることになる。

## 詳細事項

予測が行われる際には、X-12-ARIMA は予め定義されてる回帰変数のうち、選択されたものについて、その値は予測期間全体に対して生成される。もしユーザー定義変数が使用されているならば、変数の値はやはり予測期間全体に対して与えられなければならない (逆予測についても同様)。ユーザー定義変数に関する 52 個という制限に加えて、モデルに含まれる回帰変数は全体で 80 個という制限がある。(この制限は変更可能である。2.8 章を参照。) 後者の制限は、予め定義された回帰変数とユーザー定義変数と outlier スペックによって自動的に付加される回帰変数の合計数に関する制限である。ユーザー定義変数の、予測期間を含まない最大長は 600 である。(この制限もまた変更可能である。2.8 章を参照。)

もし variables 引数において const が設定されているならば、結果的に得られる回帰変数は、ARIMA モデルにおける階差操作によって得られる系列に定数項が存在するのを考慮に入れるということになる。もし ARIMA モデルが和分を含まないならば、その場合には const はゼロでない平均に対応した通常の定数からなる回帰変数である。もし ARIMA モデルが和分を含んでいるならば、この回帰変数はトレンド定数 (trend constant) と呼ばれる。後者のケースでは、実際に用いられる回帰変数は、階差操作を行った後に 1 が並んだ列になるような系列として定義される。これについての議論は 3.3 章を参照されたい。

月次のフロー系列に曜日効果が存在すると思われる時には、概して variables 引数に対して td

表 6-15: (指定された) 回帰変数

Variable	Description
<b>const</b>	Trend constant regression variable to allow for a nonzero overall mean for the differenced data.
<b>seasonal</b>	Fixed seasonal effects parameterized via $s - 1$ seasonal contrast variables ( $s =$ seasonal period). The resulting variables allow for month-to-month (or quarter-to-quarter, etc.) differences in level, but have no net effect on overall level. Seasonal cannot be used with <b>sincos</b> and also not in models with seasonal differencing except as a partial change of regime variable (see DETAILS).
<b>sincos[ ]</b>	Fixed seasonal effects (for $s =$ seasonal period) parameterized via trigonometric regression variables of the form $\sin(\omega_j t)$ and $\cos(\omega_j t)$ at seasonal frequencies $\omega_j = (2\pi j/s)$ for $1 \leq j \leq s/2$ (dropping $\sin(\omega_j t) \equiv 0$ for $j = s/2$ for $s$ even). Each frequency to be included must be specified, i.e., for monthly series <b>sincos</b> [1, 2, 3, 4, 5, 6] includes all seasonal frequencies while <b>sincos</b> [1, 2, 3] includes only the first three. <b>Sincos</b> [ ] cannot be used with <b>seasonal</b> or in models with seasonal differencing.
<b>td</b>	Estimate monthly (or quarterly) flow trading-day effects by including the <b>tdnolpyear</b> variables (see below) in the model, and by handling leap-year effects either by rescaling (for transformed series) or by including the <b>lyear</b> regression variable (for untransformed series). <b>Td</b> can only be used for monthly or quarterly series, and cannot be used with <b>tdnolpyear</b> , <b>td1coef</b> , <b>tdnolpyear</b> , <b>lyear</b> , <b>lom</b> , <b>loq</b> , or <b>tdstock</b> []. If <b>td</b> is specified, do not specify <b>adjust = lyear</b> or <b>adjust = lom</b> ( <b>adjust = loq</b> ) in the transform spec.
<b>tdnolpyear</b>	Include the six day-of-week contrast variables (monthly and quarterly flow series only): (no. of <i>Mondays</i> ) - (no. of <i>Sundays</i> ), ..., (no. of <i>Saturdays</i> ) - (no. of <i>Sundays</i> ). <b>Tdnolpyear</b> cannot be used with <b>td</b> , <b>td1coef</b> , <b>tdnolpyear</b> , or <b>tdstock</b> .
<b>td1coef</b>	Estimate monthly (or quarterly) flow trading-day effects by including the <b>tdnolpyear</b> variable (see below) in the model, and by handling leap-year effects either by rescaling (for transformed series) or by including the <b>lyear</b> regression variable (for untransformed series). <b>Td1coef</b> can only be used for monthly or quarterly series, and cannot be used with <b>td</b> , <b>tdnolpyear</b> , <b>tdnolpyear</b> , <b>lyear</b> , <b>lom</b> , <b>loq</b> , or <b>tdstock</b> []. If <b>td1coef</b> is specified, do not specify <b>adjust = lyear</b> or <b>adjust = lom</b> ( <b>adjust = loq</b> ) in the transform spec.
<b>tdnolpyear</b>	Include the weekday-weekend contrast variable (monthly and quarterly flow series only): (no. of <i>weekdays</i> ) - $\frac{5}{2}$ (no. of <i>Saturdays and Sundays</i> ). <b>Tdnolpyear</b> cannot be used with <b>td</b> , <b>tdnolpyear</b> , <b>td1coef</b> , or <b>tdstock</b> .
<b>lyear</b>	Include a contrast variable for leap-year (monthly and quarterly flow series only): 0.75 for leap-year Februaries (first quarters), -0.25 for non-leap year Februaries (first quarters), 0.0 otherwise. <b>Lyear</b> cannot be used with <b>td</b> , <b>td1coef</b> , or <b>tdstock</b> .
<b>lom</b>	Include length-of-month as a regression variable. If <b>lom</b> is requested for a quarterly series, X-12-ARIMA uses <b>loq</b> instead. Requesting <b>lom</b> when $s \neq 12$ or 4 results in an error. <b>Lom</b> cannot be used with <b>td</b> , <b>td1coef</b> , or <b>tdstock</b> .
<b>loq</b>	Include length-of-quarter as a regression variable. If <b>loq</b> is requested for a monthly series, X-12-ARIMA uses <b>lom</b> instead. The same restrictions that apply to <b>lom</b> apply to <b>loq</b> .
<b>tdstock[w]</b>	Estimate day-of-week effects for inventories or other stocks reported for the $w^{\text{th}}$ day of each month. The value $w$ must be supplied and can range from 1 to 31. For any month of length less than the specified $w$ , the <b>tdstock</b> variables are measured as of the end of the month. Use <b>tdstock</b> [31] for end-of-month stock series. <b>Tdstock</b> can be used only with monthly series and cannot be used with <b>td</b> , <b>tdnolpyear</b> , <b>td1coef</b> , <b>tdnolpyear</b> , <b>lom</b> , or <b>loq</b> .
<b>easter[w]</b>	Easter holiday regression variable (monthly or quarterly flow data only) which assumes the level of daily activity changes on the $w$ -th day before Easter and remains at the new level until the day before Easter. The value $w$ must be supplied and can range from 1 to 25. To estimate complex effects, several of these variables, differing in their choices of $w$ , can be specified.
<b>labor[w]</b>	Labor Day holiday regression variable (monthly flow data only) that assumes the level of daily activity changes on the $w$ -th day before Labor Day and remains at the new level until the day before Labor Day. The value $w$ must be supplied and can range from 1 to 25.
<b>thank[w]</b>	Thanksgiving holiday regression variable (monthly flow data only) that assumes the level of daily activity changes on the $w$ -th day before or after Thanksgiving and remains at the new level until December 24. The value $w$ must be supplied and can range from -8 to 17. Values of $w < 0$ indicate a number of days after Thanksgiving; values of $w > 0$ indicate a number of days before Thanksgiving.

表 6-15 (指定された) 回帰変数 (つづき)

Variable	Description
sceaster[w]	Statistics Canada Easter holiday regression variable (monthly or quarterly flow data only) assumes that the level of daily activity changes on the $(w - 1)$ -th day and remains at the new level through Easter day. The value $w$ must be supplied and can range from 1 to 24. To estimate complex effects, several of these variables, differing in their choices of $w$ , can be specified.
ao date	Additive (point) outlier variable, AO, for the given date or observation number. For series with associated dates, AOs are specified as <i>ao date</i> . For monthly series this is <i>ao year.month</i> (e.g., <i>ao1985.jul</i> or <i>ao1985.7</i> ), for quarterly series this is <i>ao year.quarter</i> (e.g., <i>ao1985.1</i> for an AO in the first quarter of 1985), and for annual series this is <i>ao year</i> (e.g., <i>ao1922</i> ). For series without associated dates, AOs are specified as <i>ao observation number</i> , e.g., <i>ao50</i> for an AO at observation 50. More than one AO may be specified. All specified outlier dates must occur within the series. (AOs with dates within the series but outside the span specified by the <i>span</i> argument of the series <i>spec</i> are ignored.)
ls date	Regression variable for a constant level shift (in the transformed series) beginning on the given date, e.g., <i>ls1990.oct</i> for a level shift beginning in October 1990. More than one level shift may be specified. Dates are specified as for AOs and the same restrictions apply with one addition: level shifts cannot be specified to occur on the start date of the series (or of the span specified by the <i>span</i> argument of the series <i>spec</i> ).
tc date	Regression variable for a temporary level change (in the transformed series) beginning on the given date, e.g., <i>tc1990.oct</i> for a temporary change beginning in October 1990. More than one temporary change may be specified. Dates are specified as for AOs, and the same restrictions apply.
rp date-date	Ramp effect which begins and ends on the given dates, e.g., <i>rp1988.apr-1990.oct</i> . More than one ramp effect may be specified. All dates of the ramps must occur within the series. (Ramps specified within the series but with both start and end dates outside the span specified by the <i>span</i> argument of the series <i>spec</i> are ignored.) Ramps can overlap other ramps, AOs, and level shifts.

を設定することが推奨される。フローの時系列とは、日次の系列が月次集計されることによって作られている系列のことである。この場合には、プログラムが閏(うるう)年効果をどのように処理するかは、スペック **transform** の情報に依存する。系列が変換(ボックス=コックス変換ないしロジスティック変換)されているならば、閏年効果は事前調整によって除去される。すなわち、系列は変換の前に  $lp_t$  という成分によって除算されている。ただし  $lp_t$  は時点  $t$  が閏年の2月ならば  $lp_t = 28.25/29$  であり、閏年でない年の2月ならば  $lp_t = 28.25/28$  で、それ以外の月ならば  $lp_t = 1.00$  であるとする。

もし系列が変換されていないならば、閏年変数 **lpyear** がモデルに組み込まれる。 $LP_t$  と表されるこの変数の値は、時点  $t$  が閏年の2月ならば  $LP_t = 29 - 28.25$  となり、閏年でない年の2月ならば  $LP_t = 28 - 28.25$  となり、それ以外の月では  $LP_t = 0.00$  となる。両方のケースにおいて、回帰変数 **tdnolpyear** と、(no. of Mondays) - (no. of Sundays), ..., (no. of Saturdays) - (no. of Sundays) という変数がモデルに組み込まれる。

閏年変数は、月(または四半期)の長さの変数 (length-of-month (quarter)) に対しては、季節性を含まない要素である。もし変換された系列に対するモデルに、**lom** 変数を用いて月の長さの変数を組み入りたいならば、これは **lom** 変数と **tdnolpyear** 変数を指定することによって可能となる。例えば **variables = (lom tdnolpyear ...)** とする。もし変換されていない系列に月の長さ効果に関する事前調整を行いたいならば、これは **regression** スペックにおいて **variables = (tdnolpyear ...)** とし、**transform** スペックにおいて **adjust = lom** 指定することによって可能となる。

スペック `transform` において `adjust=1om` が指定されているならば、`variables` 引数に対するリストで `td` か `1om` のどちらかが指定されると、矛盾が発生することになる。この矛盾は、二つの設定が月の長さで除算された系列に対してなされたか、または月の長さの調節と `1om` の両方が使用されたかの、どちらかの理由によって生じる（こうしたことは一般に回帰係数に対応する方程式系に非正則性をもたらすことになる）。この場合は、(i) `transform` スペックから `adjust=1om` という設定を除くか、(ii) `variables` 引数に対するリストにおいて `td` を `tdnolpyear` に置き換えるか `1om` を落とすかの、どちらかの対応がなされるべきである。

四半期のフローの時系列に対しても同様の曜日効果のオプションが利用可能である。その場合には上記の説明の中で、`1om` を `1oq` に置き換えて読めばよい。

`1om` と `1oq` の値は同一である。もしどちらかが設定されたならば、`series` スペックで指定された季節周期に従って、どちらが使われるかが決定される。従って `period = 12` なら `1om` が、`period = 4` なら `1oq` がそれぞれ導かれる。また `1om` や `1oq` は `tdnolpyear` が指定されなくても設定され得る点に注意されたい。これは、一週間の曜日の効果 (day-of-week effects) を含まない系列に対して、月の長さ（または四半期の長さ）効果によって生じる固定的な季節変動を説明するために可能になっている。予め定義された期間の長さ (length-of-period) 変数については、月次または四半期のフローの系列に対してのみ使うことができる。

在庫などのようなストック系列に関しては、月次系列についてのみ曜日効果を推定することができる。`Tdstock[w]` は、6日分の曜日を7番目の曜日に対比させることによって6つの回帰変数を生成する。ここで `w` は1から31までの値を取る。詳細は3.3章を参照。`w` の値は指定されなければならない。この数値は、一ヶ月の中でストックの値が報告される日か、または月の最後の日のうち、小さいほうを表す。従って `tdstock[31]` は月の最終日 (end-of-month) のストックに対して使用される。

休日効果の回帰変数 (復活祭、レイバーデー、感謝祭) はフローの系列である。復活祭変数は月次系列と四半期系列の両方について利用可能である。レイバーデーと感謝祭の変数は月次系列に対してのみ利用可能である。

レジームの変化の変数は、季節性 (seasonal)、三角関数による季節性 (sincos)、曜日効果 (`td`, `tdnolpyear`, or `tdstock`)、閏年効果 (`lpyear`)、月の長さ (`1om`)、四半期の長さ (`1oq`) を示す回帰変数に対して設定することができる。レジームの変化の変数は2種類が利用可能である。それは全体に対するものと部分に対するものである。

表6-16に示されているように、レジーム変化の変数は、`regression` スペックの `variables` 引数において、各変数の名前に1つか2つのスラッシュで囲って変化の日付を付加することによって設定できる。レジームの変化として指定された日付によって、モデリングされる系列は2つに分割される。早いほうの区間は変化点以前のデータを含み、新しい方の区間は変化点以後のデータを含む。部分的なレジーム変化の変数は二つの区間のうち一方のみに制限され、他方の区間ではゼロとなる。全体のレジーム変化の変数は、関心のある基本的な回帰と、前期における部分的なレジームの変化に関する回帰の両方を推定する。例えば `variables = (td/1990.jan/)` という全体のレジーム変化は、`variables = (td td/1990.jan//)` という指定と同等である。この設定ではプログラムは曜日効果に関して、`td` と `td/1990.jan//` の両方の係数推定値を出力する。

全体のレジーム変化の変数を利用して得られる係数推定値には有用な解釈がある。基本的な回

表 6-16: 回帰変数タイプの変更

Type	Syntax	Example
Full change of regime regressor	reg/date/	td/1990.jan/
Partial change of regime regressor, zero before change date	reg//date/	td//1990.jan/
Partial change of regime regressor, zero on and after change date	reg/date//	" td/1990.jan//

帰変数を  $X_{jt}$ 、変化点を  $t_0$  とすると、前期での部分的なレジーム変化の変数は

$$X_{jt}^E = \begin{cases} X_{jt} & t < t_0 \\ 0 & \text{for } t \geq t_0 \end{cases}$$

となり、後期での変化は  $X_{jt}^L = X_{jt} - X_{jt}^E$  と表される。**transform** スペックの指定どおりに変換されたデータに対しては、全体のレジーム変化変数は

$$\sum_j a_j X_{jt} + \sum_j b_j X_{jt}^E = \sum_j a_j X_{jt}^L + \sum_j (a_j + b_j) X_{jt}^E$$

という形式になる。この式の右辺から、基本の回帰変数  $X_{jt}$  に対する係数  $a_j$  は後期の変数  $X_{jt}^L$  の係数であると解釈でき、 $X_{jt}^E$  に対する係数  $b_j$  は、前期に対する回帰係数の値を得るのに必要な後期における変化の程度を計測していると解釈することができる。従って、統計的に有意な  $b_j$  の値は、レジーム変化の性質を示していることになる。

部分的なレジーム変化変数の利用について、その他の2つの自然な例について述べる。まず第一に、**variables = (td//1990.jan/)** という指定は、1990年までは統計的に有意な曜日効果が存在しないが、それ以降では有意な曜日効果が存在する可能性のあるような系列において、曜日効果を推定するために利用することができる。第二に、**arima** スペックにおいてARIMAモデル季節差分が指定された場合、あるいは**automdl** スペックによって推定されたモデルにおいて、**variables = (seasonal//1990.jan/)** という指定によって、1990年1月に発生した季節パターンの固定的変化を推定すること可能となり、またその変化の統計的有意性を検定することが可能となる。

**aicctest** 引数によって、**variables** 引数のリストから回帰変数を削除したり、リストに変数を付け加えることができる。**aicdiff** にゼロでない正の数値を与えることで、**aicctest** のプロシージャが検定される回帰変数をモデルに組み込むことが起こりにくくなる。**aicdiff** 引数に対応した数値を  $\Delta AICC$  書くとする。この数値のデフォルトの値は0である。また  $AICC^{with}$  ( $AICC^{without}$ ) は、**aicctest** 引数によって特定された回帰変数を含む(含まない)モデルに関するAICCの値を表すものとする。もしこれらの回帰変数が**variables** 引数のリストにのせられていないならば、

$$AICC^{with} + \Delta AIC < AICC^{without}$$

が成り立つ場合に、その回帰変数がモデルに加えられる。もしこれらの回帰変数が**variables** 引数のリストにのせられているならば、それらはこの式の不等号が成り立つ場合にのみ、Reg-ARIMAモデルの中に残される。

二番目のケースにおいて、**aicctest = (tdstock)** とされているならば、**tdstock[31]** によって指定される、月の終わり(end-of-month)のストック変数がモデルに加えられる。なぜなら **tdstock[w]** における  $w$  の値のデフォルトが31だからである。



`aicctest = (easter)` とした上で、`variables` 引数のリストに復活祭効果の回帰変数が指定されていないならば、さらなる可能性が生じる。その場合3つの付加的なモデルが考慮される。それは `easter[w]` 変数のうち  $w = 1, 8, 15$  となる場合を含む Reg-ARIMA モデルである。最も小さい AICC 値を与える復活祭の変数が、少なくとも  $\Delta AIC = 0$  とした時に復活祭変数なしのモデルよりも小さい AICC を与えるならば、その変数がモデルに残される。

$w$  の値の差が7以下になるような `easter[w]` 変数の組については、AICC は高い信頼性を持ってモデルを識別することができないということを、シミュレーション実験の結果は示している。`history` スペックによって生成されるアウトオブサンプルの予測の診断結果から、そのような回帰変数に関して、正確な予測を与えるモデルを見つけることによって、モデルを識別することができる場合がある。また従って、そのような変数はおそらくデータに含まれる復活祭効果をより良く表現できるだろう。

曜日効果変数が `aicctest` 引数と `variables` 引数の両方で指定される場合、指定される回帰変数のタイプは同一でなければならない。ストック曜日効果変数のサンプル日とレジーム変化の変数で指定された変化点の日付は、`aicctest` 引数に含まれてはならない。これらは `variables` 引数の中で指定されると想定される。例えば、もし `variables=(tdstock[15] ao1995.jan)` としたならば、`aicctest` は `tdstock` としなければならない。

ユーザー定義変数は(それらが季節性の回帰変数でなければ)、季節性を含まない形でプログラムに投入されるべきである。`x11regression` の詳細事項で述べられている加法的分解に対する季節性除去の方法は適切であると思われる。なぜなら回帰変数は `regARIMA` モデルの加法的要素であるからだ。もし季節性除去がなされなければ、季節的成分は推定された全ての季節成分を含むわけではないことになる。その他の問題は、季節性を含む回帰変数には係数推定値と推定された効果が存在するが、これらは互いに相関があり、従って解釈が困難となる。

`usertype` 引数によってユーザー定義変数に対してタイプが割り当てられた場合、そのようなタイプのユーザー定義変数から導かれる要素は、`regression` スペックにおいて指定された同一のタイプの変数と結合する。`noapply` 引数にタイプ名が指定されない限り、結果的に得られる要素は、`x11` スペックによって決定されている季節調整要素の計算のために、系列から調整されるだろう。

`usertype=seasonal` と設定することにより、そのユーザー定義変数は `variables` 引数で指定された季節回帰変数と全く同様に扱われることが可能となる。それらの回帰変数から推定された季節効果は、季節調整に先立って調整されるということはない。対照的に、`usertype=userseasonal` と設定すると、`x11` スペックで設定された季節調整が実行される前に、原系列を調整するユーザー定義変数から季節成分が生成される。組み合わせられた季節成分は X-11 および回帰の成分から生成される。

もし `format = "datevalue"` か `format = "x12save"` が指定されているならば、ユーザー定義変数の開始時点は自動的にデータファイルから読み取られることに注意されたい。従って `regression` スペックの `start` 引数によってデータ開始時点を指定する必要はない。

`regression` か `x11regression` のうちどちらかのスペックによって推定された効果が系列の調整に使用されないということを指定するために `noapply` オプションが利用されないならば、曜日効果と休日効果は `regression` と `x11regression` において同時に設定されないかもしれない。

`eastermeans` 引数についての2通りの選択肢は顕著に異なる休日要素をもたらす。しかしこの

選択は (使用される regARIMA モデルが季節階差や固定的季節回帰変数を含むならば) 予測には全く影響を与えず、また通常は結合された季節成分と休日成分に対してもごくわずかな影響しか与えない。なぜならこの選択の違いを埋め合わせるために季節成分が変化するためである。

`eastermeans=yes` の設定の下で季節性を取り除いた復活祭変数を得るために使用される季節ごとの平均は、グレゴリオ暦における最初の 400 年間である 1582 年から 1982 年までの復活祭に関する日付の頻度から生成される。こうした頻度は Montes (1997) で与えられている。グレゴリオ暦での復活祭の日付を計算するアルゴリズムは Duffet-Smith (1981) で与えられている。

季節性の無い時系列について、このスペックを用いて推定される曜日効果と休日効果を調整するには、`x11` スペックにおいて `type=trend` と指定すればよい。

定数の係数に対して `b=()` 引数が用いられる場合は、*AIC* やその他の選択のための統計量は役に立たないことがある。`estimate` スペックの詳細事項を参照されたい。

## 例

次に完全なスペック・ファイルの例を挙げておく。

### 例 1

ARIMA (0 1 1) モデルにしたがう誤差項、固定的季節性およびトレンドを持つモデルの推定。

```
SERIES      { TITLE = "Monthly sales"  START = 1976.JAN
              DATA = (138 128 ... 297) }
REGRESSION { VARIABLES = (CONST SEASONAL) }
ARIMA { MODEL = (0 1 1) }
ESTIMATE { }
```

### 例 2

最終的な不規則成分に対して 4 番目と 5 番目の季節周波数についてのサインとコサインの回帰変数を最小二乗法であてはめて、その季節周波数における”視覚的に明白な”スペクトルが統計的に有意であるかどうかを検定する。

```
series { title = "Irregular Component of Monthly Sales"
         start = 1976.jan
         file = "sales.d13"
         format = "x12save"
       }
regression { variables = (const sincos[4,5]) }
estimate { }
```

### 例 3

月次系列に対して曜日効果、復活祭、レイバーデー、感謝祭の回帰変数を指定する。各休日効果の継続期間が指定される。`td` が指定され、系列が対数変換されているので、変換前の原系列は閏年成分によって除算され、回帰変数 `tdnolpyear` が変換後の系列に対して当てはめられる。`identify` スペックに従い、(変換と月の長さ調整の後に) 階差操作を最大に適用された系列を、同様の階差操作を適用された回帰係数に対して回帰することを通し、て回帰係数が推定される。`identify` スペックは、回帰の誤差に対する ARIMA モデルの特定に利用するために、様々な(回帰残差に関する)ACF と PACF を生成する。

```
Series { Title = "Monthly Sales" Start = 1976.Jan
        Data = (138 128 ... 297) }
Transform { Function = Log }
Regression { Variables = (TD Easter[8]
                        Labor[10] Thank[3]) }
Identify { Diff = (0 1) SDiff = (0 1) }
```

#### 例 4

ARIMA モデル  $(0\ 1\ 1)(0\ 1\ 1)_{12}$  を回帰誤差に利用する。例 3 と同様の回帰変数と、`td` 引数が与える標準的な閏年効果による時系列の分割の代わりに `lom` 変数を含むモデルの推定を行う。`(td` を `tdnolpyear` で置き換えることにより標準的な閏年効果による分割を回避する。) 曜日効果と復活祭の変数の有意性を (AICC を用いて) テストする。回帰の誤差項には  $(0\ 1\ 1)(0\ 1\ 1)_{12}$  の ARIMA モデルが使われる。

```
series      { title = "Monthly sales" start = 1976.jan
              data = (138 128 ... 297) }
transform { function = log }
regression { variables = (tdnolpyear lom easter[8]
                        labor[10] thank[3])
            aicctest = (tdnolpyear easter) }
arima { model = (0 1 1)(0 1 1) }
estimate { }
```

#### 例 5

四半期時系列に対して、曜日効果と 2 つの加法的外れ値、2 つのレベルシフトを含むモデルを推定する。これらの効果ともに、変換された系列は ARIMA  $(0\ 1\ 1)(0\ 1\ 1)_4$  モデルに従う。

```
Series {Title = "Quarterly Sales" Start = 1963.1 Period = 4
        Data = (1039 1241 ... 2210) }
Transform { Function = Log }
Regression { Variables = (A01967.1 LS1985.3 LS1987.2 A01978.1 TD) }
Arima { Model = (0 1 1)(0 1 1) }
Estimate { }
```

#### 例 6

1985 年第 3 四半期から 1987 年第 1 四半期までの一時的なレベルシフトのためのユーザー定義変数を推定する。この一時的なレベルシフトの効果は `identify` スペックによって行われる回帰を通して、ARIMA の同定のための ACF と PACF の計算に先立って、除去される。

```
series {title = "Quarterly sales" start = 1981.1
        data = (301 294 ... 391) period = 4 }
regression {user = t1s
            data = (0 0 0 0 0 0 0 0 0 0 0 0 ...
                  0 0 1 1 1 1 1 1 1 1 1 1 0 0 0 0 ... 0) }
identify { diff = (0 1) sdiff = (0 1) }
```

#### 例 7

定数と固定的な季節変動と 2 つのユーザー定義変数を含むモデルを推定する。2 つのユーザー定義変数のデータはカレントディレクトリの `weather.dat` というファイルに保存されている。このファイルはこのモデルでは使用されない他の変数のデータも含んでいる。この 2 つのユーザー定義変数のためのデータは、FORTRAN 形式のこのファイルから最初の 16 列を飛ばすことによ

て得られる。ユーザー定義変数に関するデータセットがモデル化される時系列のためのデータの前に始まるので、開始時点が指定されている。

```
series {title = "Monthly Riverflow" start = 1970.1
      data = (8.234 8.209 ... 8.104) period = 12 }
regression { variables = (seasonal const)
            user = (temp precip)
            file = "weather.dat"
            format = "(t17,2f8.2)"
            start = 1960.1 }
arima { model = (3 0 0)(0 0 0) }
estimate { }
```

#### 例 8

月末のストック曜日効果と加法的外れ値を含む小売在庫の月次系列のモデルを推定する。回帰効果を差し引いた変換後の系列は ARIMA (0 1 0)(0 1 1)<sub>12</sub> モデルに従う。ストック曜日効果をモデルに残すべきかどうかを (AICC によって) 決定する。

```
series {title = "Retail Inventory - Family Apparel"
      start = 1967.1 period = 12
      data = (1893 1932 ... 3201) }
transform { function = log }
regression { variables = (tdstock[31] ao1980.jul)
            aictd=tdstock }
arima { model = (0 1 0)(0 1 1) }
estimate { }
```

#### 例 9

固定的な季節性と曜日効果を含む月次の売上系列に対するモデルを推定する。1985年12月におけるレジーム変化の変数が、2つの回帰変数の両方に対して設定される。回帰効果を差し引いた変換後の系列は ARIMA (0 1 1) モデルに従う。

```
series {title = "Retail Sales - Televisions"
      start = 1976.1 period = 12
      file = 'tvsales.ori' }
transform { function = log }
regression { variables = (td/1985.dec/ seasonal/1985.dec/) }
arima { model = (0 1 1) }
estimate { }
```

#### 例 10

例 9 と同様に、部分的なレジーム変化変数が季節性と曜日効果の 2つの回帰変数と組み合わせて用いられる。付加的な回帰変数は 1985年12月以前では 0 になるように設定される。

```
series {title = "Retail Sales - Televisions"
      start = 1976.1 period = 12
      file = 'tvsales.ori' }
transform { function = log }
regression { variables = (td td//1985.dec/
                        seasonal seasonal//1985.dec/) }
arima { model = (0 1 1) }
estimate { }
```

### 例 11

回帰変数として固定的トレンドと曜日効果が指定される。ARIMA モデルの選択に自動選択のプロシージャが適用される。曜日効果に対する事前調整の後に加法的季節調整が行われる。

```
series      { title = "Exports of pasta products"
              start = 1980.jan  data = "pasta.dat" }
regression { variables = (const td)                }
automdl     { mode = both                          }
x11        { mode = add                            }
```

### Example 12

選択された回帰効果は、季節ごとの平均、定数、いくつかの外れ値、曜日効果および復活祭効果である。1988年、1989年、1990年の特別なセールプロモーションがユーザー定義変数として使われるが、このデータは promo.dat のファイル名で 3f12.0 フォーマットで保存されている。モデルの ARIMA 部分は (2,1,0) である。全ての回帰効果についての事前調整がなされた後に季節調整が行われる。最終的な季節調整済系列からは、復活祭効果と曜日効果は除去される。予測値は 24 期に渡って生成される。

```
series{ title = "Retail sales of children's apparel"
        file = "capprl.dat"  start = 1975.1          }
transform{ function = log    }
regression{
  variables = (const td ao1976.oct ls1991.dec easter[8]
              seasonal)
  user = (sale88 sale89 sale90)
  start = 1975.1  file = "promo.dat"  format = "(3f12.0)"
}
arima{ model = (2 1 0)  }
forecast{ maxlead = 24  }
x11{ save=seasonal appendfcst=yes }
```

### 例 13

ユーザー定義変数が、事前調整、最終調整、ファイルの出力ファイルの保存に関して加法的外れ値と同様に処理されること以外は、例 12 と同様である。

```
series{ title = "Retail sales of children's apparel"
        file = "capprl.dat"  start = 1975.1          }
transform{ function = log    }
regression{
  variables = (const td ao1976.oct ls1991.dec easter[8]
              seasonal)
  user = (sale88 sale89 sale90)
  start = 1975.1  file = "promo.dat"  format = "(3f12.0)"
  usertype = ao
}
arima{ model = (2 1 0)  }
forecast{ maxlead = 24  }
x11{ save=seasonal appendfcst=yes }
```

### 例 14

曜日効果と外れ値を含む Reg-ARIMA モデルを指定する。回帰係数の最初の値を指定し、外れ値

変数の係数をそれらの値に固定する。このモデルを用いて 12 期分の予測値を生成する。曜日効果と外れ値に関する事前調整の後、デフォルトである乗法的季節調整が施される。

```

series{
  period=12          format='2L'
  title='MidWest Total Starts'
  file='mwtoths.dat' name='MWTOT '
}
transform{ function=log }
arima{ model=(0 1 2 )(0 1 1 ) }
estimate{ save=mdl }
forecast{ maxlead=12 }
regression{
  variables=(ao1977.jan ls1979.jan ls1979.mar ls1980.jan td)
  b=( -0.7946F -0.8739F 0.6773F -0.6850F 0.0209
      0.0107 -0.0022 0.0018 0.0088 -0.0074 )
}
x11{ }

```

### 例 15

X-11 フォーマットのファイルからデータが読み込まれる。データの開始時点はこのファイル内の情報によって決定されることに注意されたい。従って開始時点を指定する必要はない。曜日効果と休日効果を含む Reg-ARIMA モデルを指定する。自動外れ値検出を行い、モデル診断の結果を出力する。このモデルを用いて 12 期分の予測値を生成する。曜日効果と外れ値に関する事前調整の後に、3x3 の季節移動平均を用いた乗法的季節調整が施される。最終的な季節調整済系列からは休日効果と曜日効果は取り除かれる。曜日効果と休日効果については個別の出力ファイルに保存される。

```

Series { Format='1L' File='bdptrs.dat' Name='BDPTRS'
         Title='Department Store Sales' }
Transform { Function=Log }
Regression { Variables=( Td Easter[8] )
             Save = ( Td Holiday ) }
Arima { Model=(0 1 1)(0 1 1) }
Outlier { }
Estimate { }
Check { }
Forecast { }
X11 { Mode = Mult Seasonalma = S3X3
      Title=( 'Department Store Retail Sales Adjusted For'
              'Outlier, Trading Day, And Holiday Effects' )
}

```

## 系列 (series)

---

### 説明 (DESCRIPTION)

時系列データに対して指定すべき項目は、系列の名前、系列の開始時点の日付、季節周期 (月次データに対して 12、四半期データに対して 4)、および系列の中で分析に用いる期間などである。データ系列は、スペック `series` の中に `data` を用いて直接挿入することができるが、`file` を用いてデータが記述された他のファイルを指定することもできる。X-12-ARIMA がデータのメタ・ファイルを使って実行される場合には、スペック `series` 中でデータを指定してはならない点に注意する必要がある。なぜならデータ系列はデータのメタ・ファイルの中で指定されるからである (詳細は 2.5.2 節を参照のこと)。

### 使用法 (USAGE)

```
series{title = "Example Series"
      start = 1967.1
      period = 12
      span = (1970.1,)
      modelspan = (1985.Jan, 0.Dec)
      name = "tstsr"
      data = (480 ... 1386) or file = "example.dat"
              format = "2r"

      decimals = 2
      precision = 1
      comptype = add
      compwt = 1.0
      spectrumstart = 1975.Jan
      print = (none +header)
      save = (spn)
}
```

### 変数 (ARGUMENTS)

#### `comptype`

この引数は合成された (もしくは集計された) 系列について、その各要素がどのように統合されるかを指定する。各要素となる系列はその他の系列に対して、足し合わせるか、引き去るか、掛け合わせるか、除するかのをいずれかを選ぶことができる。デフォルトのままでは合成は行われない。

### compwt

合成系列の各要素について、合成する前にウェイトとして定数倍するかどうかを指定する。この定数はゼロよりも大きくなければならない(例えば `compwt=0.5` とする)。この引数は必ず `comptype` と同時に使用される。デフォルトではウェイトは 1 になっている。

### data

時系列データを記述したベクトルを指定する。データは以下のようなフォーマットで記述されたものについて、行に沿って順に読み込まれる。各データは最低一つのスペースかカンマ、もしくは改行によって分割されている必要がある。観測値の個数は、与えられるデータベクトルの長さと同じ数値に自動的に決定される。なお `data` 引数を使用した場合には、`file` 引数を同時に使用することはできない。

### decimals

季節調整の結果が記述される出力ファイル内で桁の少数を用いるかを指定する。この引数の値は 0 以上 5 以下の整数でなければならない。例えば `decimals=5` のように記述する。デフォルトでは 0 が指定される。

### file

時系列データを記述したファイルの名前を指定する。ファイル名はクォーテーション・マークでくくられていなければならない。データ・ファイルがカレント・ディレクトリ以外の場所にあるならば、完全なパス名によって指定する必要がある。パス名やファイル名は使用する OS (オペレーティング・システム) に依存する。`file` 引数を使用する場合は、`data` 引数を同時に使用することはできない。

### format

この引数はデータファイルを参照する際に、そのファイルがフリー・フォーマット (通常のテキスト・ファイル形式) ではない場合に使用する。使用可能なファイル形式は 5 種類ある。

- (a) FORTRAN フォーマットのファイル。指定の際には丸括弧でくくったものをさらにクォーテーションでくくる必要がある。例えば `format="(6f12.0)"` などとする。;
- (b) X-11 および X-11-ARIMA で使用するフォーマット。各バージョンに応じて 2 文字のコードを指定する。例えば `format="1r"` などとする。;
- (c) "date-value" フォーマットのファイル。これは年次、月次もしくは四半期などの区分が、各観測値に並べてフリー・フォーマット形式で記述されたファイル形式である。従って 1991 年の 7 月に 32531 という値をとっているならば、その行は `1991 7 32531` と記述されることになる。例えば `format="{datevalue}"` とする。;
- (d) X-12-ARIMA フォーマットのファイル。これは表を保存するためのものである。この形式により、以前に実行した X-12-ARIMA のプログラムの結果から得られたファイルを読み込むことができる。例えば `format="{x12save}"` とする。;
- (e) TRAMO 形式および SEATS 形式のファイル。この形式を指定することにより、モデリングのためのプログラムである TRAMO や、季節調整プログラム SEATS のために作成されたデータ・ファイルを X-12-ARIMA で使用することができる。例えば `format="{tramo}"` とする。



(b)で触れた X-11 形式においては、データは各年ごとのラベルと合わせて 6 文字か 12 文字で保存される。使用できるフォーマットの完全なリストと使用法については DETAILS を参照されたい。この format 引数が与えられない場合には、データはフリー・フォーマットとして読み込まれる。フリー・フォーマットにおいては各行の数値が全て順に読み込まれていく。また数値は一つ以上のスペースで区切られていなければならない。(区切り文字はカンマやタブではいけない)

Format 引数は data 引数とは同時には使用できず、file 引数とのみ併用される。

#### **modelsspan**

これはデータの内 Reg-ARIMA モデルの係数の推定に使用する区間を指定するためのものである。この引数は例えば、データの初期部分が予測に影響することを避けたり、逆に最新の部分のデータが季節調整の前段階である事前調整における回帰モデル推定に影響を及ぼさないようにするといった目的のために使用される。既に述べた **span** 引数と同様に、**modelsspan** 引数についても開始時点と終了時点の二つの値を指定する。開始時点か終了時点のどちらかの値が省略されていても良い。例えば月次データについては、**modelsspan**=(1968.1, ) という記述によって、Reg-ARIMA モデルの形がどのようになるとしても、モデルの推定は 1968 年 1 月から最後までデータに基づいて行われることになる。開始時点か終了時点のどちらかが省略される場合にはカンマを入れておく必要がある。開始時点と終了時点の日付は **series** の中で読み込まれるデータ区間に含まれていなければならない。開始時点は終了時点よりも先でなければならない。

#### **name**

分析しようとする時系列データの名前を指定する。名前は 64 文字以内で、クォーテーションでくくられている必要がある。最初の 16 文字が全てのページにラベルとして記載されることになる。**format** 引数が予め設定してあるならば、この名前の中の最初の六文字 (または、**format**="{cs}" が与えられているならば八文字) が、プログラムが適切な系列を読み込んでいるかどうかの確認に使用される。また、系列が格納してある場所から特定の系列を見つけ出すためにも使われる。

#### **period**

系列の持つ季節性の周期を指定する。季節調整の際には、月次データのための 12 と、四半期データのための 4 の、どちらかの値だけが有効である。それ以外の季節周期は 12 によって指定することができる。(この制限は解除することができる。詳細は 2.8 節を参照のこと) デフォルトでは **period** 引数は 12 になっている。

#### **precision**

時系列データから読み込む数字の桁数を指定する。このオプションは必ず **format** 引数と同時に使用される。値は 0 から 5 までの整数値をとる。デフォルトでは 0 になっている。**precision** 引数が **series** の内部で使用される場合に、**format** で指定されるはずのデータ・フォーマットが用いられるならば、**precision** 引数は無視される。

#### **print and save**

表 6 から 19 はスペック **series** 内で出力として使用可能な表を示している。

#### **savelog**

スペクトルのプロットが視覚的に明らかなピークを有しているときに、診断のために使うことの出来る結果をログ・ファイルに保存する。**savelog**=peaks か **savelog**=spk とすることで、スペクトルに関する情報がログ・ファイルに保存される。

### **span**

計算および分析に使用するデータの区間を指定する。**span** 引数は、開始時点と終了時点の2つの値を代入する。開始時点か終了時点かのどちらかが省略されていても良い。例えば月次データの場合は、**span=(1968.1, )** という記述により、データの開始時点は1968年1月となり、終了時点は **data** か **file** 引数で指定されるデータの最後尾のデータの日付になる。開始時点か終了時点の値を省略する場合には、必ずカンマが必要になる。開始時点と終了時点は、使用するデータ全体の中に含まれていなければならない、開始時点は終了時点より先でなければならない。

### **spectrumstart**

原系列、季節調整済系列、修正済みの不規則成分について、スペクトル推定に使用するデータの開始時点指定する。この日付は **spectrumstart=year.seasonal period** に含まれていなければならない。これは曜日効果や季節性が調整済み系列の中に残存していないかどうかを最新の7年の調整値から判断するために使うことも出来る。季節性や曜日効果の変動している場合は、残存することがある。デフォルトでは、月次データのスペクトルのプロットには、分析に用いるデータが8年以上あるならば、系列の最後から最新の8年分が使われる。その際にはデータの最初の値の日付が開始時点とされる。また四半期データについては、デフォルトの開始時点はデータの最初の値の日付となる。例えば **spectrumstart=year.seasonal period** とする。

### **start**

**start=year.seasonal period** という形式内での時系列データの開始時点指定する。(5.2節および以下の例を参照) **start** のデフォルトの値は **-1.1-** である。(DETAILSを参照)

### **title**

時系列データの表題を指定する。表題はクォーテーションでくくられた79文字以内の文字列とする。これはアウトプットの各ページに記載される。(ただし **-p** オプションが使われる場合を除く。2.7節を参照)

## 稀にしか使わない引数 (RARELY USED ARGUMENTS)

### **diffspectrum**

これによって **diffspectrum=no** と記述された場合、(変換された)元系列か、季節調整済み系列のスペクトルが計算される。デフォルト (**diffspectrum=yes**) では、差分をとった系列のスペクトルが計算される。

### **missingcode**

データ系列のうちプログラムが欠損値として処理する数値を指定する。このオプションは **Reg-ARIMA** モデルの推定やモデルの自動選択を指定するスペックファイルの中でのみ使用される。デフォルトの値は **-99999** である。例えば **missingcode=0.0** などとする。

### **saveprecision**

**save** で引数として指定されたファイルに表を保存する際、何桁の少数までを保存するかを指定する。デフォルトの値は **15** である。例えば **saveprecision=10** などとする。

### **spectrumtype**

**X-12-ARIMA** によって作成されるスペクトルのプロットで用いるスペクトルのタイプを選択する。

spectrumtype = periodogram と記述した場合、時系列のピリオドグラムが計算されプロットされる。デフォルト (spectrumtype=arspec) では自己回帰モデルを当てはめた場合の推定値に基づいて計算されたスペクトルが与えられる。

#### trimzero

trimzero=no とされた場合、file の引数を指定することにより読み込まれるデータに含まれる、先頭か最後のゼロはデータ系列の値として扱われる。デフォルト (trimzero=yes) では、データの先頭、もしくは最後に存在するゼロは無視される。format が datevalue, x12save, または tramo を値として使用された場合は、trimzero の内容に関わらず、全てのデータは通常の値として解釈される。

#### yr2000

yr2000=yes とした場合、“century cutoff”が1945年を境にして適用される。すなわち“X-11 formats”で記述されているデータ系列について、年次が00から45と表記されているものは2000年から2045年、46から99と表記されるものは1946年から1999年を、それぞれ表すものと解釈される。これはデフォルトの設定になっている。もしyr2000=noとしたならば、二桁の数字で表された年次は全て20世紀の年代を示すものとして解釈される。

### 詳細事項 (DETAILS)

データ系列の個数と終了時点の日付はデータを読み込んだ後にプログラムによって識別される。X-12-ARIMA は600の最大値を受け入れる。(この制限は変更可能であるが、詳しくは2.8節を参照。)

もしスペック・ファイルがあるディレクトリから他のディレクトリにコピーされていたり、他のコンピューターからコピーされたものだった場合には、file 引数で指定するファイルのパスや名前が適切なものであるかどうかを確認する必要がある。

スペック series はスペック composite を含むスペック・ファイルの中で用いることはできない。スペック composite は合成されたデータ系列の季節調整を行うためのものである。

以下の表は format 引数に関して、有効な X-11 フォーマットの二文字のコードの各々に対応したデフォルトのフォーマットを表したものである。FORTRAN フォーマットについても同様である。

表 6.17: X-11 フォーマット・コードのデフォルト設定

Code	FORTRAN Format for Monthly Data	FORTRAN Format for Quarterly Data	Description
1r	(12f6.0,I2,A6)	(4(12x,f6.0),I2,A6)	Year and identifier on the right, data in 6-digit fields
2r	(6f12.0,/,6f12.0,I2,A6)	(4f12.0,24x,I2,A6)	Year and identifier on the right of the second line, data in 12-digit fields
1l	(A6,I2,12f6.0)	(A6,I2,4(12x,f6.0))	Year and identifier on the left, data in 6-digit fields
2l	(A6,I2,6f12.0,/,8x,6f12.0)	(A6,I2,4f12.0)	Year and identifier on the left of the first line, data in 12-digit fields
cs	(A8,I2,10X,12E16.10,18X)	(A8,I2,10X,12E16.10,18X)	Data in CANSIM data base utility format, data in 16-digit fields

これらのフォーマットは decimals によって修正することができる。もし decimals が X-11

フォーマットのコードを使用しないスペック `series` の内部で用いられたなら、`decimals` は無視される。

もし X-11 フォーマットのコードが指定されたなら (あるいは `format = "datevalue"` か `format = "x12save"` とされたなら)、データ系列の開始時点はデータ・ファイルから自動的に読み込まれることに注意する必要がある。これにより、`start` によってデータの開始時点を指定しなくても良いことになる。

もし一つのスペック・ファイルによって複数のデータ・ファイルを処理するためにデータ・メタ・ファイルを使用するならば、X-11 フォーマットを用いることは避けるべきである。これらのフォーマットは該当するデータがファイルにふくまれていることを確認するために、(`name` で指定される) データ系列の名前を必要とする。従ってデータ・メタ・ファイル内で指定される全てのデータ・ファイルは、同一の名前である必要があることになる。これは望ましくないことが多い。

フォーマットされたデータの読み込みの際には、X-12-ARIMA はデータの末端のゼロを切り捨てる。この規則は、例 3 で示されるような形式の入力データを用いる際に適用される。もしデータの末端のゼロが本当にデータの値であるなら、`trimzero=no` とすることでそのような扱いが可能になる。しかしながら、データの先頭にあるゼロが本当のデータの値で、末端のゼロがそうでないならば、そのデータ・ファイルはフリー・フォーマットとして処理が可能になるように整形されなければならない。下の例 4 がこの規則について触れている。

`span` 引数と `modelspace` 引数は、データの末端部を除去することにより、`forecast` スペックと同時に使用して、*out-of-sample* の予測値に関して比較対象となる系列を生成することができる。この 2 つの引数が両方とも用いられる場合、モデルの推定には指定された範囲のデータのみが用いられる。

`modelspace` 引数で指定されたデータ範囲の始点が、`span` 引数で指定されたデータの始点と異なる場合、スペック `forecast` 内の `maxback` 引数で設定された値は無視され、逆予測は実行されないということに注意されたい。

原系列の中で `missingcode` の値と同じ値が発見されると、その時点に加法的な外れ値を表す回帰変数が挿入され、欠損値がモデルの推定において適当と思われる大きさの数値に置き換えられる。Reg-ARIMA モデルが推定された後は、このような欠損値に相当する回帰係数から生成される要素を利用して、原型列の調整が行われる。調整された値は欠損値の推定値である。

## 例 (EXAMPLES)

以下に示す例は、Example 8 を除き、意味のあるアウトプットが得られないという点で、完全なスペックファイルではないということに注意されたい。意味のある結果を得るには、`x11`、`arma`、`estimate` といった他のスペックを付け加える必要がある。

### 例 1

`data` 引数を利用した系列の指定。

```
series{
  title="A Simple Example"
  start=1967.jan          # period defaults to 12
  data=( 480 467 514 505 534 546 539 541 551 537 584 854
        522 506 558 538 605 583 607 624 570 609 675 861
```

```
1684 1582 1512 1508 1574 2303 1425 1386) }
```

### 例 2

1940年から1993年にかけてのデータに関して、先端部と末端部のデータを除外した例。最初の6年間は第二次世界大戦後の分析のために除外されている。このデータは1991年から93年のデータを out-of-sample の予測のパフォーマンスの評価に利用することができる。

```
series { data = (879 899 985 ...) # There are 216 data values
         start = 1940.1          # ending in 1993.4
         period = 4              # Quarterly series
         span = (1946.1, 1990.4) }
```

### 例 3

この例は X-12-ARIMA が、季節調整プログラム X-11-ARIMA のフォーマットで保存されたファイルからどのようにデータを読み込むかを示している。ここでは1970年7月から1993年2月のデータが、c:\data\sales1.dat というパス名で、以下のような内容で保存されているとする。

```
146.4 109.2 132.1 144.8 116.1 100.370SALES
142.9 158.8 196.2 244.0 251.6 245.5 244.2 213.8 188.9 197.2 181.2 161.371SALES1
.
.
.
148.8 177.2 0 0 0 0 0 0 0 0 0 093SALES1
```

データは—(12f6.1,i2,a6)—の形式であり、最後の8文字は最後の年と系列のIDを表している。

```
SERIES{ TITLE = "Monthly data in an X-11 format"
        PERIOD = 12
        FILE = "C:\DATA\SALES1.DAT" # a DOS path and file
        PRECISION = 1
        FORMAT = "1x" }
```

フォートランのフォーマットでは空白はゼロと判断されるので、入力される系列は最初に6つのゼロを含むことになる。また末端には10個のゼロを含む。詳細事項で述べたように、X-12-ARIMA はデフォルトでは先端と末端の連続するゼロは破棄するようになっているので、実際のデータ系列は正確な月が割り当てられる(146.4 to July, 1970, etc.)。また、それぞれの行で一年が表されるので、ユーザーは start 引数を用いる必要はない。

### 例 4

この例は X-12-ARIMA で利用する為には修正する必要があるデータ・ファイルの例を示している。オリジナルのデータ・ファイルは1980年2月から1990年11月のもので、(6f4.0,1x,i4) フォーマットで以下のようにになっているとする。

```
0 342-256 491 0 0001
-234 922-111 2 0 199 0002
.
.
```

581-987-423 10 0 0022

このファイルは、いくつかの数値が並んで入力されており、26列から29列に行数が(0001, 0002, ...)という形で挿入されているために、フリー・フォーマットのファイルとして扱うことができない。フリー・フォーマットにおいては、行数もデータとして扱われることになる。このファイルは1980年2月を始点とする(6f4.0)フォーマットのファイルとしても扱われない。なぜならtrimzero = yesをデフォルト設定としているX-12-ARIMAでは、先端と末端のゼロを落としてしまい、誤って342を1980年2月のデータとして割り当ててしまうからである。trimzero = noと指定することにより、ゼロがデータ系列に付け加えられ、先端と末端の空白がゼロとして読み込まれる。解決策は、ファイルがフリー・フォーマットとして読み込まれるように、データ・ファイルを修正することである。これには行数の記録を除去し、数値と数値の間に空白を作る必要がある。修正されたファイルがexample4.newであり、以下のようになっている。

```
      0 342 -256 491  0
-234 922 -111  2  0 199
.
.
.
581 -987 -423 10  0
```

これにより次のようなスペック series によりデータをデータ・ファイルから正確に読み込めるようになる。example4.new

```
series {title = "Data read correctly in with trimzero = no"
        start = 1980.2  period = 12
        file = "example4.new" } # file is in current directory
```

#### 例 5

この例はX-12-ARIMAが“date-value”フォーマットのデータをどのように読み込むかを示している。このデータは1970年7月から1993年2月までで、次のような内容でc:\data\sales1.edtというパス名で保存されている。

```
1970  7  14624
1970  8  10952
1970  9  13251
1970 10  14408
.
.
.
1993  1  14838
1993  2  17762
```

それぞれのデータは年と月と数値を表している。

```
SERIES{ TITLE = "Monthly data in a date-value format format"
        PERIOD = 12
        FILE = "C:\DATA\SALES1.EDT" # a DOS path and file
        FORMAT = "DATEVALUE" }
```

例 3 で示したように、X-11-ARIMA 同様、データの開始時点はデータ・ファイルから直接読み込まれる。従ってユーザーは start 引数を指定する必要は無いことに注意されたい。

#### 例 6

この例は example 5 と同じであるが、この系列は合成系列の一要素として扱われている。DECIMALS は 2 とされ、モデル推定に用いられるデータ期間は、開始時点から 1992 年 12 月である。

```
SERIES{ TITLE = "Monthly data in a date-value format format"
        PERIOD = 12
        FILE = "C:\DATA\SALES1.EDT"    # a DOS path and file
        FORMAT = "DATEVALUE"
        COMPTYPE = ADD
        DECIMALS = 2
        MODELSPAN = (,1992.DEC)
}
```

#### 例 7

この例は X-12-ARIMA が欠損値をどのように処理するかを示している。データは前の二つの例と同じものを用いているが、欠損値を示すコードが 1990 年 1 月に挿入されている点だけが異なる。

```
1970  7  14624
1970  8  10952
1970  9  13251
1970 10  14408
.
.
.
1990  1  -99999.
.
.
.
1993  1  14838
1993  2  17762
```

以下のスペックでは 1990 年 1 月の欠損値のコードは、スペック・ファイルにおいて推定されることになる Reg-ARIMA モデルを仮定したうえで、外れ値として適当な大きさの数値に置き換えられる。

```
SERIES{ TITLE = "Monthly data in a date-value format format"
        PERIOD = 12
        FILE = "C:\DATA\SALES1.EDT"    # a DOS path and file
        FORMAT = "DATEVALUE"
}
```

#### 例 8

この例は、データが季節性や曜日効果を含んでいるかどうかを判定するために、対数変換したデータの 1 階差分を取った系列のスペクトルをプロットする方法を示している。これは完全なスペック・ファイルである。

```
series{ title = "Spectrum analysis of Building Permits Series"
        start = 1967.Jan
        file = "permits.dat"
```

```
format = "(12f6.0)"
spectrumstart = 1987.Jan
print = (none +specori)
}
transform[ function = log ]
```

#### 例 9

この例は、X-12-ARIMAによって保存されたデータを、X-12-ARIMAでどのように読み込むかを示している。以前のプログラムの実行により、外れ値が調整されたデータ系列がc:\data\sales1.a11として保存されているとする。

```
SERIES{ TITLE = "Monthly data in a file saved by X-12-ARIMA"
PERIOD = 12
FILE = "C:\DATA\SALES1.A11" # a DOS path and file
FORMAT = X12SAVE }
```

例 3 で示した X-11-ARIMA フォーマットのファイルや、例 5 で示した "date-value" フォーマットのファイルと同じように、データの開始時点はデータ・ファイルから直接読み込まれるので、start 引数を指定する必要が無いことに注意されたい。



## 期間変更の安定性 (slidingspans)

---

### 解説

オプションのスペック・コマンドとして期間変更の安定性分析がある。このコマンドにより時系列データから重なりのあるいくつかの期間を選び、それぞれの期間における季節調整系列の性質を比較することができる。安定性分析 (スライディング・スパン) の初期時点 (`start`)、期間の長さ (`length`)、スライディング・スパン統計量 (`cutseas`, `cutchn`, `cuttd`) を定める閾値、分析を実行する際にどのように RegARIMA モデルの母数推定値を得るか (`fixmdl`)、さらに RegARIMA モデルの外れ値の自動識別機能を利用するか (`outlier`) などを指定することができる。

### 利用法

```
slidingspans{start = 1975.jan
             length = 132
             numspans = 3
             cutseas = 3.0
             cutchn = 3.0
             cuttd = 2.0
             outlier = yes
             fixmdl = no
             fixreg = outlier
             print = (long -ssheader)
             save = (sfspan)
             savelog = (percent)
             }
```

### 変数

#### `cutchn`

季節調整系列における月次・四半期・年の変化率についての閾値。ある月 (四半期) において、その月を含む複数のデータ期間から計算した季節調整系列の変化率の最大値がこの閾値を超えると、その月 (四半期) の推定値は信頼が置けない値と認識される。この値は 0 以上でなければならない。デフォルトの値は 3.0 である。例: `cutchn=5.0`

#### `cutseas`

季節項と季節調整系列の閾値。ある月 (四半期) において、季節項と季節調整系列の変化率が絶対値でその閾値を超えている場合には、その月 (四半期) の推定値は信頼が置けない値として認識される。この値は 0 以上でなければならない。デフォルトの値は 3.0 である。例: `cutseas=5.0`

#### `cuttd`

曜日効果の項の閾値。ある月 (四半期) において、曜日効果の項の変化率が絶対値でその閾値を超えている場合には、その月 (四半期) の推定値は信頼が置けない値として認識される。この値は 0

以上でなければならない。デフォルトの値は 2.0 である。例：cuttd=1.0

#### **fixmdl**

スライディングスパンを季節調整する前に、regARIMA モデルの母数推定値の初期値をどのように設定するかを指定する。この引数は regARIMA モデルを当てはめない時には無視される。

**fixmdl=yes** では、もとの regARIMA モデルの母数の推定値をそれぞれのスパンの regARIMA モデルの初期値とする。それらの母数は固定され、推定し直されることはない。これがデフォルトの設定になっている。

**fixmdl=no** では全系列に対して行った regARIMA モデルの推定を記憶しておく。estimate スペックで固定されていると、その値と同じ値に設定されたままになる。**fixmdl=clear** では全ての母数の初期値としてデフォルトの 0.1 が用いられる。そして全ての母数が推定しなおされる。

#### **fixreg**

Reg-ARIMA モデル、irregular component regression のいずれかの回帰係数の組を固定するか、指定する。それらの係数は series スペックや composite スペックにおけるモデルスパンから得られた値に固定される。その他の回帰係数はそれぞれのスライディングスパンで推定しなおされる。曜日効果 (td)、休日 (holiday)、はずれ値 (outlier)、ユーザーが定義した回帰効果 (user) は固定される。この引数は regARIMA モデル、irregular component regression が系列に当てはめられない時と **fixmdl=yes** と時には無視される。

#### **length**

標本期間に対応して月・四半期で数えられた時系列データ期間の長さで計算結果を比較する為に指定する。利用者が指定する場合には 3 年以上 17 年以下でなければならない。期間の長さの指定がなければプログラムは利用者が選んだ季節フィルターにもとづく長さ（あるいは利用者が季節フィルターを指定していなければあらかじめ指定されている長さ）が選択される。月次系列の場合には例えば **length=96** とする。

#### **numspans**

比較のために作成する出力結果のスライディングスパンの数を指定する。2 から 4 の間の数を指定する。この引数を指定しないと、ユーザーが **length** で指定した数に応じてスパンの最大の値が自動的に選択される。**length** を指定しないとプログラムが自動的に選択する。例：**numspans=4**

#### **outlier**

Reg-ARIMA モデルでそれぞれの期間について推定しなおす際に外れ値検出を行うかどうか指定する。この引数は **outlier** スペックを利用しないと効果がない。この引数を指定しないと、それぞれのスライディングスパンでの Reg-ARIMA モデルの再推定で外れ値検出がなされない。**regression** スペックで外れ値について指定していればそれが適用される。**outlier=yes** とするとそれぞれの期間で Reg-ARIMA モデルが推定されるごとに外れ値検出を行う。**outlier=keep** とすると全データに対する Reg-ARIMA モデルの推定で認識された外れ値が適用される。全データについての推定で認識された外れ値が当該スライディングスパンの期間外にあった場合には、その期間におけるそのモデルの外れ値とは見なされない。

#### **print** と **save**

出力に用いられるテーブルは表 6-18 の通り。

表 6-18: Slidingspans に関連する出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
header	+	+	.	header text for the sliding spans analysis
ssftest	+	.	.	F-tests for stable and moving seasonality estimated over each of the sliding spans
factormeans	+	.	.	range analysis for each of the sliding spans
percent	+	+	.	table showing the percent of observations flagged as unstable for the seasonal and/or trading day factors, final seasonally adjusted series (if necessary), and the month-to-month (or quarter-to-quarter) changes
yypercent	.	.	.	additional entry for the percent of observations flagged as unstable for the year-to-year changes
summary	+	.	.	tables, histograms and hinge values summarizing the percentage of observations flagged for unstable seasonal and/or trading day factors, final seasonally adjusted series (if necessary), and month-to-month (or quarter-to-quarter) changes
yysummary	.	.	.	additional tables, histograms and hinge values summarizing the percentage of observations flagged for the year-to-year changes
sfs	.	.	sfs	seasonal factors from all sliding spans
chngs	.	.	chs	month-to-month (or quarter-to-quarter) changes from all sliding spans
sas	.	.	sas	seasonally adjusted series from all sliding spans
ychngs	.	.	ycs	year-to-year changes from all sliding spans
tds	.	.	tds	trading day factors from all sliding spans"
indfactormeans	+	+	.	range analysis for the implicit adjustment factors of the indirectly seasonally adjusted series
indpercent	+	+	.	tables of the percent of observations flagged as unstable for the seasonal factors and month-to-month (or quarter-to-quarter) changes of the indirect seasonal adjustment
indypercent	.	.	.	additional entry for the percent of observations flagged as unstable for the year-to-year (or quarter-to-quarter) changes of the indirect seasonal adjustment
indsummary	+	.	.	tables, histograms and hinge values summarizing the percentage of observations flagged for unstable seasonal factors, month-to-month (or quarter-to-quarter) and year-to-year changes for the indirect adjustment
indysummary	.	.	.	additional tables, histograms and hinge values summarizing the percentage of observations flagged for the year-to-year changes of the indirect seasonal adjustment

*Name* gives the name of each table for use with the `print` and `save` arguments.

*Default* indicates which tables are printed (+) or not printed (.) by default.

*Brief* indicates which tables are printed (+) or not printed (.) when the brief print level is specified. See section 5.1 for more information on print levels.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the `print` and `save` arguments.

Table 表 6-18: Slidingspans に関連する出力 (つづく)

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
indsfspan	.	.	sis	indirect seasonal factors from all sliding spans
indchngspan	.	.	cis	indirect month-to-month (or quarter-to-quarter) changes from all sliding spans
indsaspan	.	.	ais	indirect seasonally adjusted series from all sliding spans
indycngspan	.	.	yis	indirect year-to-year changes from all sliding spans

*Name* gives the name of each table for use with the print and save arguments.

*Default* indicates which tables are printed (+) or not printed (·) by default.

*Brief* indicates which tables are printed (+) or not printed (·) when the brief print level is specified. See section 5.1 for more information on print levels.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the print and save arguments.

### ssavelog

ログファイルに対する診断として利用可能なのは (2.6 節参照) 季節調整値と観測値の比についてスライディングスパン分析により不安定かどうか判定することだけである。savelog=percent か、savelog=pct と指定すると、ログファイルにその情報を保存する。

### start

スライディング・スパンの開始時点を指定する。デフォルトでは 2 つ目のスパンの開始月である。

例: start=1990.jan

### 稀にしか使われない引数

#### additivesa

スライディング・スパン分析で加法型季節調整を季節調整系列の最大階差 (additivesa = difference) から計算するか、それとも原系列と季節調整系列の調整比から (additivesa = percent) 計算するかを指定する。このオプションは階差 (additivesa = difference)、またはパーセント変化 (additivesa = percent) を月次、四半期、年次のいずれの分析から作成するかも指定する。もしいづれかの期間の中で季節調整系列が 0 以下の値を取っているならばスライディング・スパン分析は階差について実行される。

#### fixx11reg

irregular component 回帰モデルをスライディング・スパン分析の中で再推定するかどうか指定する。fixx11reg=yes とすると、回帰係数は全系列から推定された値に固定される。一方 fixx11reg=no とすると、パラメータはそれぞれのスパンについて再推定される。デフォルトは fixx11reg=no である。

#### x11outlier

A0 外れ値検出を irregular component 回帰モデルの中で実行するかどうか指定する。x11outlier=yes では A0 外れ値検出はそれぞれの期間で実行される。全期間に対して A0 外れ値の自動識別が実行されたときには A0 外れ値の説明変数は irregular component 回帰モデルに追加されているがスライディング・スパン分析の前に説明変数からはずされる。x11outlier=no では全期間に対して

実行された AO 外れ値検出の結果がそれぞれの期間についても適用される。全系列に対して実行された AO 外れ値検出の結果、外れ値と認識されたデータが当該期間に含まれていない場合には、その値はその期間において外れ値とは見なされない。

irregular component 回帰モデルのパラメータが推定される時には必ず AO 外れ値に対する係数も再推定される。しかし AO 外れ値の候補となる変数が追加されることはない。デフォルトは x11outlier=yes である。

## 詳細事項

本節は 2 つの小節に別れている。前半は診断に関する説明であり、後半は引数についての追加的な情報である。

### 主な診断法とその解釈

安定性の診断法は、Findley, Monsell, Shulman and Pugh (1990) および Findley and Monsell (1986) において詳細に説明され、その他の有力な診断法との比較が行われている。ここではより簡略な説明を与える。基本的な診断法は、季節調整の計算に用いられるデータ期間がシステムティックに変化していく時に、季節調整とその月から月への変化がどのように変動するかを表すための記述統計的方法である。システムティックな変化とは、任意の隣り合った二つの区間が、常に一方が他方に対して一年遅れて始まって終わるということである。区間の長さは、調整に用いられた季節フィルターの長さに従って決定される。最後の区間の終了時点は、通常は時系列の最新の日付となる。十分なデータが利用可能ならば、4 つの区間が用いられる。最も早く開始する区間に対して  $j = 1$  という指数が割り振られ、その次に始まる区間に対して  $j = 2$  が割り振られる。それ以降も同様である。

季節調整値が全て正の値を取る系列に対しては、二つの最も重要な安定性に関する統計量である  $A(\%)$  と  $MM(\%)$  が次のようにして計算される。第  $t$  月が少なくとも 2 つの区間に含まれているとして、その区間のうちの 1 つが第  $j$  区間であるとする。このとき  $A_t^j$  が、全データに基づいて計算された季節(および曜日、休日)調整値を表すとし、また  $A_t^j$  が、診断の対象となる季節調整法(オプションによって指定された調整法)が第  $j$  区間のみに対して適用された場合に得られる調整値を表すとする。この季節調整法は

$$\frac{\max_j A_t^j - \min_j A_t^j}{\min_j A_t^j} > .03. \quad (1)$$

となるるとき、不安定であると言う。さらに  $t$  と  $t-1$  の両方が少なくとも 2 つの区間に含まれているような第  $t$  月に対して、"季節調整値前期比変化率 (seasonally adjusted month-to-month percent change)"  $100 \times (A_t - A_{t-1})/A_{t-1}$  が、

$$\max_j \frac{A_t^j}{A_{t-1}^j} - \min_j \frac{A_t^j}{A_{t-1}^j} > .03. \quad (2)$$

となるときに、不安定であると言う。

(1) では  $j$  は第  $t$  月を含む全ての区間について変化するものとし、(2) では第  $j$  区間は第  $t$  月と第  $t-1$  月の両方を含んでいなければならない。

$A(\%)$  は、(1) の左辺が定義されるような月の数 (少なくとも 2 つの期間に共通して含まれる月の数) に関して、不安定な調整値が与えられる月の割合を表すために使われる。(2) に関する同様の量は  $MM(\%)$  と表される。以下で議論するような特殊な状況を除いては、選択した手順によって実行される季節調整が、 $A(\%) > 25.0$  ( $> 15.0$  は問題があると思われる) または  $MM(\%) > 40.0$  という結果をもたらすようならば、その季節調整を用いることは推奨できない。

季節調整された系列において、年から年への変化 (year-to-year percent changes)  $100 \times (A_t - A_{t-12}) / A_{t-12}$  に対して、同じように定義される統計量  $YY(\%)$  が存在する。これに対して月から月への変化 (month-to-month changes) の場合と同じ境界値によって不安定性が定義される。境界値は (1) と (2) のように、通常は .03 である。調整系列におけるこうした年から年への変化は、第  $t$  月と第  $t-12$  月の間の時点で変化が起こっている場合には、トレンドの方向性に関して誤った指標となることがあり得るため、調整系列においては調整値自体と月から月への変化の方がより重要である。従って統計量  $YY(\%)$  はその他の統計量よりも重要度は低いが、データによっては月から月への変化に関心が持たれる場合があるので、X-12-ARIMA の出力には含まれている。出力されるテキストには  $YY(\%)$  が極端な値として 10.0 よりも大きい場合に記載されるが、我々の実験では、そのような  $YY(\%)$  の値を取る系列は通常は  $A(\%)$  や  $MM(\%)$  もまた極端に大きな値を取るようになるので、過大な  $YY(\%)$  についての情報は多くの場合不要となる。どのようなケースにおいても、ある季節調整法が  $YY(\%)$  の値のみから棄却されるということはないであろう。

$A(\%)$  や  $MM(\%)$  が過大な値を取る理由が判明し、それほど問題とならない場合が時々ある。例えばこうしたことは、不安定な調整値や変化のある月が、現在の年から数年前の既知の問題が存在する区間や、非常に問題があるだろうと予測され得るような特定の月に集中している場合に発生し得る。特定の月とは、冬の天候の違いに非常に左右されやすいことがあらかじめ知られているような系列における冬の月などである。安定性分析の出力結果を用いることにより、そうした集中は識別しやすくなる。

境界値が 0.3 を越えたところで "軽く" 増加することによって、 $A(\%)$  と  $MM(\%)$  の値が極めて適切なレベルにまで劇的に減少する場合は、出力結果によって示され得る。(1) と (2) の左辺が .03 から .05 の間の値になっている月の大半は、非常に季節変動の大きい、ユーザーがより一層の不確実性を許容するであろうと思われる月であって、またこれらの統計量が実質的に .05 より大きくなる月は少なかったという理由から、我々は、境界値を 0.5 まで増加させることが適切と思われるいくつかの系列を発見した。

我々はこの経験に刺激されて、様々なセンサス局の時系列に関して、適切な境界値と季節成分の大きさとの間の統計的關係、すなわち季節変動の大きさに応じて境界値を増加または減少させるために利用できるような関係、を確立することを目的とした探索的研究を行った。しかしながら、考察した時系列においては、適切な境界値と季節変動の大きさとの間に相関関係は発見されなかった。例えば、.03 という境界値によって良い  $A(\%)$  と  $MM(\%)$  が得られるような、非常に大きい季節変動を含む時系列は比較的多く存在し、また季節変動が程々の大きさでありながら、.05 という境界値によっても許容できない大きさの  $A(\%)$  と  $MM(\%)$  をもたらす時系列も数多く存在する。実際、固定的季節効果 (全ての 1 月が同じ季節要素をもっている、等) を含む系列に関しては、季節調整値は不規則成分の変動に対して極めて敏感である一方、季節変動の大きさに対して

は全く敏感でないということが、シミュレーション実験によって簡単に示される。

ある系列に対する季節調整オプションの選択によって、安定性分析による診断では不適切であると判断されるような調整値がもたらされた場合、季節フィルタの長さ、曜日効果調整の方法、予測に関する設定等で、異なるオプションを選択することにより、適切と判断される調整結果が得られることはしばしばある。オプションのいかなる選択によっても適切な調整値が得られないならば、問題は、その系列が何らかの意味で"季節的"であるかどうかではなく、その系列の季節的変動が十分に反復的であるかどうか、あるいは利用可能な時系列データの中で十分明白に現れているかどうかということである。ここで考察したどのようなオプションをとる場合にも X-12-ARIMA は十分に信頼できる推定結果をもたらすのである。

#### 引数に関する追加的情報

スライディング・スパン分析ではそれぞれの季節調整に応じて異なる調整値を比較し、曜日効果の有無を調べる。乗法型や対数加法型の季節調整モデルについて季節項、月次、年次の季節調整系列の変化を分析する。乗法型や対数加法型の曜日効果についても、曜日効果の項と季節調整系列を分析する。曜日効果なしの加法型季節調整モデルについては季節調整系列とその月次、年次変化を分析する。曜日効果の分析がなされれば季節調整系列、曜日効果ありの調整系列について分析する。

注意：加法型調整モデルでは比較的小さい値や負の値がスライディング・スパン統計量のほとんど全てに少なからず影響を与える。そのため通常は引数 `print` や `save` の中で `saspan` を利用することで分析主目的となる期間について行き過ぎた不安定性がないかどうか調べる。加法型季節調整モデルについてスライディング・スパン分析を行うには、より深い研究が必要である。

スライディング・スパン分析についてより詳しくは Findley, Monsell, Shulman, and Pugh(1990) を参照。

`arima` スペックで ARIMA モデルを自動選択するオプションを指定すると、選択されたモデルについてすべてのスライディング・スパンが適用される。自動選択の過程でどのモデルも選択されなかった場合にはスライディング・スパン分析ではモデル推定が行われない。

このスペックの多くのテーブルは個別ファイルに保存されないが、コマンド実行時に季節調整診断要約のオプション `-s` を指定すると、スライディング・スパン分析の結果が診断要約ファイル(ファイルの拡張子は `.xdg`) に保存される。さらに引数 `savelog` はログ・ファイル(拡張子は `.log`) に選択した診断を保存するようになれる。より詳しくは 2 節を参照。

合成した季節調整系列のスライディング・スパン分析を行うにはそれぞれの系列でオプション指定しなければならない。それぞれの系列で季節調整フィルタの長さが異なる場合にはそれぞれのオプション指定の前に `length` で長さを指定して、分析する期間が同じ長さになるようにしなければならない。

自動季節調整フィルタ選択のオプションを指定した場合には季節フィルタは原型列の季節調整を作成するのに用いたフィルタになる。当該期間においてユーザーの指定した外れ値が実際には外れ値とは見なされない場合にはその変数はモデルから取り除かれ、未来の期間について定義されている場合には `Reg-ARIMA` モデルが再定義される。ユーザー定義の回帰ではそれぞれの期間で定数(全ての説明変数が 0)になっていないかどうか調べている。

## 例

次に完全なスペック・ファイルの例を示しておく。

### 例 1

月次系列の乗法型季節調整で全月に 3x9 季節因子を利用する。スライディング・スパン分析のすべてのオプションについてデフォルトで実行する。

```
SERIES { FILE = "TOURIST.DAT"  START = 1976.1  }
X11 { SEASONALMA = S3X9  }
SLIDINGSPANS {  }
```

### 例 2

四半期データの対数加法型モデルで 3x9 季節調整フィルターを最初の 2 期に対して、3x5 季節調整フィルターを最後の 2 期間について指定し、7 次のヘンダーソンのトレンド・フィルターを指定する。スライディング・スパン分析は閾値 5.0 として実行する。

```
Series {
  File = "qstocks.dat"
  Start = 1967.1
  Title = "Quarterly stock prices on NASDAC"
  Freq = 4
}
X11 {
  Seasonalma = ( S3x9 S3x9 S3x5 S3x5 )
  Trendma = 7
  Mode = Logadd
}
Slidingspans {
  Cutseas = 5.0
  Cutchng = 5.0
}
```

### 例 3

曜日効果と自動外れ値検出の項を回帰変数に入れた季節 ARIMA モデル。回帰変数は定数項、曜日効果項、1982 年 5 月と 1982 年 9 月の接続項である。ARIMA モデルは (0,1,2)(1,1,0)<sub>12</sub> である。加法型季節調整を外れ値と取引効果項について行う。スライディング・スパン分析は自動識別の過程で見つかったすべての外れ値についてそれぞれの期間ごとに Reg-ARIMA モデルを再推定する。

```
series { title = "Number of employed machinists"
         start = 1980.jan file = "machine.emp"
}
regression { variables = (const td rp82.may-82.oct) }
arma { model = (0 1 2)(1 1 0) }
outlier { }
estimate { }
check { }
forecast { }
x11 { mode = add }
slidingspans { outlier = keep }
```

### 例 4

既存の回帰項は定数、曜日効果、固定季節項とする。ARIMA モデルは (3,1,0) とする。60 期先まで予



測を行う。季節調整は、調整前の曜日効果を用いて実行する。それぞれの期間について Reg-ARIMA モデルを再推定する。

```
series { title = "Cheese Sales in Wisconsin"
         file = "cheez.fil"  start = 1975.1  }
transform { function = log }
regression { variables = (const seasonal tdnolpyear) }
arima { model = (3 1 0) }
forecast { maxlead = 60 }
x11 { save = seasonal  appendfcst = yes }
slidingspans { fixmdl = no }
```

#### 例 5

スライディング・スパン分析を乗法型モデルに適用する。3つの40期からなるスライディング・スパンを用いて実行する。これによりたとえば系列が3x9季節調整フィルター(44四半期)を用いるとくに適当なスライディング・スパン分析としては期間が短い、季節調整の安定性を分析することができる。

```
Series      {
  File = "qstocks.dat"
  Start = 1987.1
  Title = "Quarterly stock prices on NASDAC"
  Freq = 4
}
X11         {
  Seasonalma = S3x9
}
Slidingspans {
  Length = 40
  Numspans = 3
}
```

## 変換 (transform)

---

### 解説

Reg-ARIMA モデルの推定のために予測系列の変換や調整を特定する。この目的のために系列は Box-Cox (power) 変換, もしくは対数変換してすることができる。月の長さが調節され, ユーザーによって定義された事前調整因子により分割される。ユーザーによって定義された事前調整因子は `data` 引数もしくは特定のファイルの `file` 引数のどちらかで与えておかなければならない。季節調整では永続的に除去されるファクターを特定することに加え, 季節調整ファクターが計算されるまで除去される一時的なファクターも特定することができる。

### 使用法

```
transform{function = log or power = 0.0
  adjust = lom
  title = "prior adjustment factors"
  start = 1975.jan
  data = (1.25 ... 1.90) or file = "prioradj.dat"
  format = "(6f12.3)"
  name = "Adjfac"
  mode = ratio
  print = (none)
  save = (prioradj)
  savelog = atr
}
```

### 引数

#### adjust

月次データに月の区間調整をする (`adjust = する` (`adjust = log`), もしくは月次データや四半期データに閏年調整をする (`adjust = lpyear`). (DETAILS を見よ.), もしくは何もしない `adjust = none`. デフォルトはなにもしない。

コマンド `regression` や `x11regression` の `variables` 引数において `td` や `td1coef` が特定された場合, もしくは加法, 擬加法季節調整が特定された場合, `adjust` 引数は使ってはならない。閏年調整が可能なのは長期の対数変換が `power` もしくは `function` 引数で特定された場合である。

事前調整では `adjust` 引数で特定された因子を用い曜日効果などの全ての回帰変数により調整する。 `Regadjust = all` は全ての回帰変数を調整する。 `Regadjust = td` は曜日日数の回帰変数だけを調整する。 `Regadjust = none` は回帰変数の調整を全く行わない。 デフォルトは `Regadjust`

#### data

予め調節された一つもしくは二つの系列は `mode=diff` (下を見よ.) でない限り正の値を取らなければならない。なぜなら代入された時系列の値に対応するように分割する目的があるからだ。デフォルトの値は 1 のベクトルである。(先行調整なし)。 `data` (or `file`) が使われたときは調整因

子は系列における全ての観測値に供給されなければならない。(series スペックの span 引数が特定された場合はその期間)。一般的には調整因子は必要な予測(事後)にもそれぞれまた供給しなければならない。(DETAILS を見よ。) 調整因子はフリー・フォーマットで読み込む。出発時点が調整因子に与えられたら、系列の最初よりも前に開始しても良い。data 項目が適応されたら、file 項目を適応できない。mode = diff の場合は data の値は系列から除かれる。それらは負でも良い。

二つの系列は data を通じて代入することができ、このとき永続的、一時的な先行調整因子の両方が type の集合で特定される。より詳しい情報は DETAILS を見よ。

#### file

ユーザー定義の先行調整因子を含むファイルの名前。ファイルの名前はクォートで囲う必要がある。ファイルがカレントディレクトリにない場合にはパスを与える必要がある。file 項目が使われた場合は data 項目を使うことはできない。ファイルのデータがフリーフォーマットでない限り format 項目を使う必要がある。

永久的、一時的の両方の先行調整因子は type 集合で特定される。因子は単一もしくは二つのファイルから代入することができる。詳しくは DETAILS を見よ。

#### format

ファイルから事前調整因子を読むために使われるフォーマットについて述べる。5つのタイプの入力採用される。

- (a) 適当な FORTRAN フォーマット、クォートで囲う必要がありまた最初と最後の挿入を含む必要がある。(例: `format="(6f12.0)";`)
- (b) 二つの文字コードが前のバージョン、X-11 や X-11-ARIMA で使われていたデータフォーマットからの選択に対応している。(例: `format="1r";`)
- (c) 年、月、四半期といった"データ値"のフォーマットがあり、それぞれの観測値の値はこの順にフリーフォーマットでデータファイルに列挙されている。(例: `format="{datevalue}";`)
- (d) フォーマット X-12-ARIMA はテーブルの保存に使う。前の X-12-ARIMA から保存されるファイルで読み込むこともできる。(例: `format="{ x12save}";`)
- (e) TRAMO や SEATS のプログラムを系列の中やその記述を読むのに使う場合のフォーマット。これは X-12-ARIMA で TRAMO モデルのプログラムや SEATS 季節調整プログラムのためにフォーマットされたデータファイルを X-12-ARIMA で読み込むことを可能にする。

(b) で述べられた予め定義された X-11 データ・フォーマットで、データは年ごとに各年のデータに対応した系列ラベルとともに 6 ないし 12 の文字領域に保存される。これらの系列の完全なリストは DETAIL セクションのスペック series を参照にせよ。フォーマット項目が与えられときデータは次の行を読む前にデータはフリー・フォーマットで読み込まれることになる。free format では一つのラインのすべての数が次のラインを読む前に読み込まれる。一つ以上のスペースで数は分けられなくてはならない。(コンマやタブによってではない)。Format は file 項目と共に使われ data とは使われない。

永久的, 一時的な事前調整因子が2つの異なるファイルからそれぞれ異なるフォーマットで代入された場合は, 上のフォーマットが特定される. 詳しくは DETAILS を見よ.

#### function

系列  $Y_t$  を変換して **series** スペックに代入する. 変換は **log**, 平方根, 逆数, ロジスティック変換を用いる. 代替的な方法として AIC に基づいて **log** 変換と変換なしの選択 (**function=auto**) を行うこともできる. この際, **regression** と **arima** スペックで特定された **regARIMA** か, **automdl** で特定された **automatic model file** の最初のモデルを使う. (DETAILS を見よ). デフォルトでは何も変換しない (**function = none**). **function** と **power** の項目を両方入れてはならない. **注意:** 季節調整のための事前因子が **Reg-ARIMA** モデルから発生した場合, これらの項目で使われる値には制限が加わる; 詳細事項を見よ.

#### mode

時系列にどのようにユーザー定義の事前調整因子を当てはめるか特定する. 系列ごとの事前調整因子がパーセントではなく (e.g., (100 100 50 ...)), 割合で与えられる場合は (e.g., (1.0 1.0 .5 ...)), **mode=ratio** とおく. 事前調整が原系列から引かれる場合は **mode=diff** とおく. 季節調整モードが **x11** スペックで乗法可能か対数変換で加法可能であれば **mode=diff** を使う. 因子は **log** でスケール変換されているものとする. 原系列と同じにするためには因子を指数変換しておけばよい. この引数が特定されない場合は事前調整因子はパーセントであると想定される.

永続的な事前調整因子と, 一時的なものの両方が **type** によって特定された場合は, 両立するならば2つの値まで特定可能である. (e.g., **diff** は割合やパーセントで特定できない.) 詳しくは DETAILS を見よ.

#### name

事前調整因子の名前. 名前は64文字まででクォートで囲わなければならない. 最初の16文字まで事前調整因子のラベルとしてプリントすることができる. **format** の **X-11** が特定されたとき最初の6文字 (**format="{cs}"**なら8文字) の名前は定義済みのフォーマットと共に使われ, プログラムが正しい系列を読み込んでいるかの確認, 多くの因子系列が保存されているファイルから特定の系列の検索に利用される.

永続的な事前調整因子と一時的なものの両方が **type** によって特定された場合は, ユーザーは両方の事前調整因子の集合の系列名を特定するか, 名前を全く打ちこまないことのどちらかを実行できる. 詳しくは DETAILS を見よ.

#### power

入力系列  $Y_t$  の変換は **Box-Cox** 変換を用いる.

$$Y_t \rightarrow y_t = \begin{cases} \log(Y_t) & \lambda = 0; \\ \lambda^2 + (Y_t^\lambda - 1)/\lambda & \lambda \neq 0. \end{cases}$$

この **Box-Cox power** 変換の定式化は  $\lambda$  が1に近いとき  $Y_t$  に近く,  $\lambda$  が0に近いとき  $\log Y_t$  に近くなるように構成されている. 更に変換された値は  $Y_t$  が1より大きい場合は正の値をとる.

$\lambda$  は必ず与えなければならない (e.g., **power = .33**). デフォルトは変換なしである ( $\lambda = 1$ ), i.e., **power = 1.log** 変換, 平方根変換, 逆変換を **function** を使うことで代替的に利用できる. 同一のスペックファイルで **power** と **function** を両方用いてはならない.

**注意:** 季節調整の先行調整因子が **Reg-ARIMA** モデルから生成される場合, これらの引数で

表 6-19: Transform に関連する出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
priorfactors	+	+	a2	prior-adjustment factors, with associated dates
prioradjusted	.	.	a3	prior-adjusted series, with associated dates
transformed	.	.	trn	prior-adjusted and transformed data, with associated dates

*Name* gives the name of each table for use with the **print** and **save** arguments.  
*Default* indicates which tables are printed (+) or not printed (.) by default.  
*Brief* indicates which tables are printed (+) or not printed (.) when the brief print level is specified. See section 5.1 for more information on print levels.  
*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the **print** and **save** arguments.

使われる値に制約がある。詳細事項を見よ。

**precision**

事前調整因子のファイルから読み込む 10 進数の数。このオプションは **format** の定義済みのフォーマットと共にのみ使用可能。この値は 0 から 5 の値を取り、(for example, **precision=5**) を含める。デフォルトは 0。定義済みのフォーマットのどれも用いない **transform** スペックに **precision** が用いられる場合、この引数は無視される。

永続的な事前調整因子と、一時的なものの両方が **type** によって特定された場合は、2 つの値まで特定可能である。詳しくは **DETAILS** を見よ。

**print and save**

以下のオプションな出力テーブルが利用可能。テーブル 6 - 23 に可能な出力がリストアップされている。

**savelog**

**savelog=autotransform** や **savelog=atr** を指定すると log ファイルへの出力の自動変換選択法が適用される。log ファイルの詳細は 2.6 節を見よ。

**start**

ユーザー定義の事前調整因子の開始日。デフォルトは系列の開始日である。妥当な値は 1 番新しい系列の開始日である。(可能ならば **series** スペックの **span** によって特定された期間の開始日)。永続的な事前調整因子と、一時的なものの両方が **type** によって特定された場合は、2 つの事前調整因子の 2 つの開始日まで特定可能である。詳しくは **DETAILS** を見よ。

**title**

ユーザー定義の事前調整因子。タイトルは引用符 (クォート) で囲む必要があり、79 文字まで入れることができる。

**type**

ユーザー定義の事前調整因子が永続的なもの (原形列と最後の季節調整系列からは除かれたもの) か一時的なもの (季節因子と生む原形列からは除かれるが最後の季節調整系列からは除かない) かを特定する。

この引数で値が与えられた場合 (**type = temporary**) にのみ、ユーザー定義の事前調整因子が一つだけと予想される。ユーザー定義の事前調整因子の両方のタイプが与えられた場合、(**type = (temporary permanent)**)、2 つの事前調整因子が期待される。より詳しい情報は詳細事項を

見よ。デフォルトは `type = permanent` である。

### 稀にしか使われない引数

#### **aicdiff**

自動変換選択のオプション (`function=auto`) が呼び出されたら変換しないように AICC に差が定義される。デフォルトの値は `aicdiff = -2.0`。このオプションの変換選択での使用方法に関してのより詳しい状況は DETAILS を見よ。

#### **trimzero**

`trimzero=no` の場合、つまりファイルへ書き込まれた時系列の最初もしくは最後が 0 ならば、それは系列の値として取り扱われる。デフォルト (`trimzero=yes`) は最初と最後の 0 を無視する。`format` 引数で `datevalue`, `x12save`, もしくは `tramo` が用意されたなら、`trimzero` の値に関わらず全ての入力値は系列の値として取り扱われることに注意する。

### 詳細事項

仮に Box-Cox もしくはロジスティック変換が月数 (閏年) 調整やユーザー定義の事前調整因子の関連で特定されたならば、時系列は開始月数や事前因子の調整をし、そして Box-Cox もしくはロジスティック変換する。月数とロジスティック変換の両方が特定されたら組み合わせた調整因子の予測が用いられる。 (`length-of-month × prior adjustment`)。四半期の長さや閏年の調整も同様である。

仮に `adjust` の `lom` や `loq` が特定された場合は、正しい調整因子が `series` スペックで特定される `period` により決定される。月次系列  $Y_t$  ではまず観測値がそのつき ( $m_t$ ) の日数で割られその後、平均の月の長さ (30.4375) をかけて  $(30.4375 \times Y_t)/m_t$  が得られる。四半期の期間調整も同じように行われ、 $q_t$  を  $t$  四半期の日数とすると結果は  $(91.3125 \times Y_t)/q_t$  になる。変換や月次調整されたデータの予測値は出力時に原形列に戻される。(スペック `forecast` の文章を見よ。)

月間、四半期データは閏年と閏年でない年の一月に関して `adjust = lpyear` を特定することで調整可能である。月次データは一月の観測値に  $28.25/m_t$  をかける。 $m_t$  はその年の一月の日数。他の月は影響されない。

X-12-ARIMA はユーザー定義の事前調整因子 (一時的もしくは永続的) 600 の最大値を許容する。ただし予測期間を含まない。(この限界は変更可能---2.8 節を見よ)

調整因子が予測期間に供給された場合、事前調整系列の予測値はこれらの因子によって逆変換 (もしくはかける、`mode = diff` の場合はたす。) される。調整因子が予測期間に供給されない場合は予測の逆変換はただ月次ないし四半期データの Box-Cox 変換や `log` 変換を行うことを指す。これはユーザー定義の事前調整因子が予測期間を通して 1 (`mode = diff` ならば 0) であると強く仮定する。

季節調整が `x11` スペックと共に呼び出される場合、`power` や `function` のどの値も Reg-ARIMA モデルの系列を予測する目的で使用可能。あてはまらないのは原形列や最終的に季節調整された系列を調整するために使用された回帰係数により因子が求められた場合である。この場合、可能な変換は `log` 変換 (乗法可能、`log` 変換して加法可能)、無変換のみである。無変換は出力 1 で特定可

能 (加法季節調整). 季節調整が行われない場合は出力変換は使用されていない.

`function=auto` で系列が正の値を持つ場合, プログラムは無変換から `log` 変換までの変換を適当に選択し `Reg-ARIMA` 変換, 無変換の系列に当てはめる. `log` 変換が選ばれないのは

$$AICC_{nolog} - AICC_{log} < \Delta_{AICC} \text{ or } AICC_{log} + \Delta_{AICC} > AICC_{nolog}$$

の場合である. ここで  $AICC_{log}$  は変換された系列に `Reg-ARIMA` モデルをあてはめた `AICC` の値であり,  $AICC_{nolog}$  は無変換の系列に `Reg-ARIMA` モデルをあてはめた `AICC` の値でもの,  $\Delta_{AICC}$  は, デフォルト値が `-2` の `aicdiff` の入力値である. 負の値の  $\Delta_{AICC}$  は `log` 変換の選択にバイアスがあることを意味する. 系列が `0` や負の値をとるならば変換を行わない.

`Reg-ARIMA` モデルが `regression` と/や `arima` で特定されたならばこのモデルから `AICC` が計算され検定に使われる. モデルが特定されないならば自動モデル選択が `automdl` スペックから特定され, プログラムはまず自動モデルファイルからモデルを読み込みそのモデルから `AICC` を計算する.

仮に `automdl` スペックの `file` から自動モデルが特定されない場合は, プログラムはカレントディレクトリの `x12a.mdl` のファイルを用いる. このファイルの文法構造の詳細は `automdl` スペックの文章を見よ. `x11` や `x11regression` スペックで季節調整が特定された場合, プログラムは選択された変換に適した季節調整モードになる. (`log` 変換に関し乗法可能で無変換では加法可能). プログラムは現在自動変換選択オプションにおけるユーザー定義の事前調整因子の使用を許していない.

一時的もしくは永続的なユーザー定義の事前調整因子はスペック `transform` で付加されるある特別な特徴を利用する必要がある. 例えばデータ入力に関してはユーザーは事前調整因子の各タイプを特定することが可能である. プログラムにおいて `type` 引数によって所与のなかで, どのタイプの事前因子が参照されたかわかる. 例えば以下のようにすればよい.

```
transform{
  type=(temporary permanent)
  file=("temp.fil" "perm.fil")
  format=("(6F12.5)" "(F15.3)")
  start=(1980.jan 1975.jul)
  mode=(ratio percent)
}
```

一時的な事前調整因子は `temp.fil` から読み込まれ (`6F12.5`) のフォーマットが使われている. これらの因子は 1980 年の 1 月から始まる. 永続的な事前調整因子は 1975 年に始まり `perm.fil` から読み込まれ (`F15.3`) フォーマットを用いる.

2 つの記入が `file` で与えられているがそのうち一つだけが `format`, `start`, `mode` もしくは `precision` のどれかの場合には両方の因子に適応される. `name` の値の数は `type` から示唆される事前調整因子の数に対応しなければならない.

`data` が事前調整因子の 2 つの入力に使われたとき, データは 2 列の行列であるとされる. 列データのタイプは `type` によって示唆される. 下の例では

```
transform{
  type=(temporary permanent)
  data=( 1.055 1.000
```

```

      .
      .
      .
      1.033  1.000 )
start=1980.jan
mode=ratio
}

```

最初の列は一時的な調整因子 (値は 1.055, 0.990, と 1.025) で第 2 列は永続的な先行調整因子と解釈される. 同様な仮定は 2 つの調整タイプが一つのデータファイルにあるときも成り立つ. 下の例参照.

```

transform{
  type=(temporary permanent)
  file="both.fil"
  start=1980.jan
  mode=ratio
}

```

## 例

注意: 以下の例は "完全" なスペック・ファイルではない. 有用な出力には追加的なスペックが付け加えなければならない. (e.g., `x11`, `identify`, `arima`, や `estimate`)

### 例 1

1967 年のストライキのユーザー定義の事前調整と月数調整を特定化する.

```

series { data = (879 899 462 670 985 973 ...)
  start = 1967.jan }
transform { data = (1 1 .5 .75 1 1 ...)
  mode = ratio
  adjust = lom }

```

### 例 2

現在のドル (名目) をドル (実) に変える物価デフレータの log 変換とユーザー定義の事前調整を特定する. デフレータ系列の開始日は時系列モデル化される前なので特定される.

```

series {title = "Total U.S. Retail Sales --- Current Dollars"
  file = "retail.dat"
  start = 1980.jan}
transform {function = log
  title = "Consumer Price Index"
  start = 1970.jan # adj. factors start January, 1970
  file = "cpi.dat"
  format = "(12f6.3)"}

```

### 例 3

ユーザー定義の調整因子の読み込みに既存のフォーマットが使われ, 因子が一時的な事前調整因子に適用されるだけで後は例 2 と同様である.

```

series {title = "Total U.S. Retail Sales --- Current Dollars"
  file = "retail.dat"
  start = 1980.jan}

```



```

transform {function = log
          title = "Consumer Price Index"
          start = 1970.jan # adj. factors start January, 1970
          file = "cpi.dat"
          format = "1R"
          precision = 3
          name = "cpi"
          type = temporary
        }

```

#### 例 4

四半期時系列データの変化の安定化に立方根変換が特定される。

```

SERIES {TITLE="Annual Rainfall"
        FILE="RAIN.DAT"
        PERIOD=4
        START=1901.1}
TRANSFORM {POWER=.3333}

```

#### 例 5

この例は 2 つのユーザー定義の事前調整系列因子の例：消費者物価指数で最終的な季節調整データから除去し、(1 つ目) 現在価値 (永続的な先行効果) と (2 つ目) 一時的な影響 (一時的な先行効果) に系列の値を分ける。これらの因子はそれぞれのファイルから読み込まれる。ファイルは同様のフォーマットなので `format` と `precision` には一価の値が投入される。

```

series {title = "Retail Sales of computers --- Current Dollars"
        file = "rscomp.dat"
        start = 1980.jan
      }
transform {function = log
          title = "Consumer Price Index & Strike Effect"
          type = (permanent temporary)
          start = 1970.jan # adj. factors start January, 1970
          file = ("cpi.dat" "strike.dat")
          format = "1R"
          precision = 3
          name = ("cpi" "strike")
        }

```

#### 例 6

自動変換選択法を系列変換に `log` 変換を用いるかどうかの決定に使う。 `model` スペックが与えられていないので、プログラムは、自動モデル選択のプロシージャで検定に必要な AICC を計算するのに使用される `x12a.mdl` ファイルの中から、最初のモデルを読み込むことになる。AICC の検定の誤差は 0 に初期化し、モデル推定で同様の AICC の値を出すような変換をプログラムに選択可能にしている。

```

series {title = "Total U.K. Retail Sales"
        file = "ukretail.dat"
        start = 1978.jan
      }
transform {function = auto
          aicdiff = 0.0
        }

```

## X11

---

### 解説

センサス局の改訂されたプログラム X-11 と X-11Q により季節調整を行いたいユーザーは次のようなコマンドを用いて実行できる。すなわち季節性の分解形式 (**mode**)、季節あるいはトレンド移動平均 (**seasonalma** と **trendma**)、季節調整における外れ値の調整 (**sigmalim**) などである。またユーザーは 1970 年代に X-11 のために作られた Bateman と Mayes による方法で計算された復活祭の調整 (**x11easter**) を指定するオプションを選べる。 **print** と **save** によって指定される出力オプションとして最終的なテーブルと X-11 の季節調整方法の分析を含めることができる。 X-12-ARIMA では、データの分析や調整問題に追加的なスペックを使うことができ、事前回帰調整の補正を行ったり、系列を前後に拡張する。このような操作は系列を補正し、 X-11 はこの系列からより良い季節調整を与える。 X-11 の季節調整診断について詳しくは Shiskin, Young, Musgrave (1967) と Lothian and Morry (1978) を見よ。 曜日効果の調整と他の休日調整は **x11regression** スペックから得ることができる。

### 使用法

```
x11{mode = pseudoadd
  seasonalma = s3x9
  trendma = 13
  sigmalim =(1.25 2.75)
  title ="3x9 moving average, mad"
  appendfcst = yes
  x11easter = yes
  force = totals
  keepholiday = yes
  final = user
  print = (brief +b2 )
  save = ( d10 d11 )
  savelog = ( m7 q )
}
```

### 引数

#### **appendfcst**

セーブオプションで保存と選択された X-11 tables に予測値を含めるかどうかきめる。 予測値は以下のテーブルに保存される。 a16, b1, d10, d16, h1 の x11 スペック, a6, a7, a8, a8.tc, a9, a10 の回帰スペック, c16 と c18 の x11 の回帰スペックである。 **appendfcst=no** ならば、予測値は保存されない。 デフォルトは予測値を含めない。

#### **final**

事前調整のタイプのリストは **regression** と **outlier** のスペックから得ることができる。 それらは最終的な季節調整系列からは除去されるものである加法的な外れ値 (**final=ao**)、水準変化と

ramp outliers(final=ls), 一時的変化 (final=tc), ユーザー定義の回帰変数から出てくるもの (final=user) は除去することができる。この選択が特定されない場合, 最終的な季節調整系列はこれらの効果を含む。

#### force

季節調整系列を以下のように修正するよう指定する。(a) 年毎の季節調整系列と原形列の合計を等しくする (force=totals). (b) 季節調整系列を各カレンダー年の値に調整し, どの年も rounded seasonally adjusted series の合計が the rounded annual total の合計と等しくなるようにする (force=round). (c) 年毎の合計を合わせそれから調整する (force=both). force=totals の場合は, 一年の合計の違いが季節調整の値としてでてくる。この際, 月毎 (四半期毎) に原形列と連動させる。より詳しくは Huot (1975) と Cholette (1978) をみよ。季節調整パターンがかわっている, 曜日効果の調整 (これについては詳細事項をみよ) が行われているならば, この強制的方法は推奨できない。

#### mode

実行される季節調整分解のモードを決定する。四つの選択がある。: 乗法型 (mode=mult), 加算 (mode=add), 偽加算 (mode=pseudoadd), log 加算 (mode=logadd)。デフォルトモードは mult で, 自動変換選択手法が transform スペックで起動されない限りこうなる。後者の場合は系列に選択された変換にあうモードが当てはめられる。(mult が log 変換で add が無変換)。

#### print and save

表 6-20 はデフォルトで利用可能な表の一覧である。表 6-21 はこの引数を指定することによって印刷や保存が可能な表を示している。表 6-22 は print 引数で指定できるプロットを表している。

#### savelog

log ファイル (see section 2.6) への出力で利用可能な診断結果は表 6-27 にリストアップされている。

#### seasonalma

季節因子の推定に季節移動平均 (季節"フィルター"とも呼ばれる。)を用いるよう指定。これらの季節移動平均は  $n \times m$  moving averages で, 連続する  $m$  項の平均の系列に対し  $n$  項の平均がとられることを意味する。Table 6-27 の季節調整フィルターは系列全体, あるいは特定の月, 四半期に選択される。同様の移動平均が全てのカレンダー月や四半期に用いられる場合, 単一の値だけを記述すればよい。同じカレンダー月や四半期で異なる季節移動平均が望まれる場合は両方記述して, 各月, 四半期において望まれる移動平均を指定しなければならない。四半期系列のれいは以下のようなものである。seasonalma=(s3x3 s3x9 s3x9 s3x9)。季節移動平均が指定されない場合は, プログラムは最後の季節フィルターを自動的に選択する。この選択は seasonalma=msr とおけば呼び出される。これは X-11-ARIMA/88 の moving seasonality ratio procedure を使って実行される。詳しくは詳細事項。これは 前のヴァージョン X-11 と X-11-ARIMA から変更したことで前は季節移動平均が指定されない。3x3 移動平均が各繰り返しにおける季節因子の初期値を計算するのに用いられる。3x5 の移動平均は最後の季節因子を計算するのに使われる。この季節フィルターの系列は seasonalma=xX11default と入力することで指定できる。

#### sigmalim

表 6-20: X11 のデフォルト出力

<i>name</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
x11easter	+	h1	X-11 Easter adjustment factors
combholiday	+	chl	combined holiday prior adjustment factors, A16 table
ftesthol	+	.	holiday F-test, H1 table
adjoriginal	+	b1	original series, adjusted for prior effects and forecast extended
irrwtd	.	c17	final weights for the irregular component
unmodsi	.	d8	final unmodified si-ratios (differences)
ftstd8	+	.	F-tests for stable and moving seasonality, D8
replaci	.	d9	final replacement values for extreme si-ratios (differences), D iteration
movseasrat	.	.	moving seasonality ratios for each period
seasonal	+	d10	final seasonal factors
seasonaldiff	+	fsd	final seasonal difference (only for pseudo-additive seasonal adjustment)
seasadj	+	d11	final seasonally adjusted series
seasadjtot	+	saa	final seasonally adjusted series with constrained yearly totals (if <i>force</i> = <i>round</i> or <i>force</i> = <i>both</i> )
saround	+	rnd	rounded final seasonally adjusted series (if <i>force</i> = <i>round</i> ) or the rounded final seasonally adjusted series with constrained yearly totals (if <i>force</i> = <i>both</i> )
residualseasf	.	.	F-test for residual seasonality
trend	.	d12	final trend-cycle
irregular	.	d13	final irregular component
adjustfac	+	d16	combined seasonal and trading day factors
adjustdiff	+	fad	final adjustment difference (only for pseudo-additive seasonal adjustment)
calendar	+	d18	combined holiday and trading day factors
yrtotals	.	e4	ratio of yearly totals of original and seasonally adjusted series
origchanges	.	e5	percent changes (differences) in original series
sachanges	.	e6	percent changes (differences) in seasonally adjusted series
revsachanges	.	e6a	percent changes (differences) in seasonally adjusted series with revised yearly totals
rndsachanges	.	e6r	percent changes (differences) in rounded seasonally adjusted series
trendchanges	.	e7	percent changes (differences) in final trend component series
x11diag	+	.	summary of seasonal adjustment diagnostics
qstat	+	.	quality control statistics
tdaytype	+	.	trading day factors printed by type of month
specsa	+	sp1	spectral plot of differenced, seasonally adjusted series (or of the logged seasonally adjusted series if <i>mode</i> = <i>logadd</i> or <i>mode</i> = <i>mult</i> )
specirr	+	sp2	spectral plot of outlier-modified irregular series

*Name* gives the name of each table for use with the *print* and *save* arguments.

*Brief* indicates which tables are printed (+) or not printed (.) when the brief print level is specified. See section 5.1 for more information on print levels.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the *print* and *save* arguments.

表 6-21: X11 に関連するその他の出力

name	ext	description of table
ftestb1	.	F-test for stable seasonality, B1 table
trendb2	b2	preliminary trend-cycle, B iteration
sib3	b3	preliminary unmodified si-ratios (differences)
replacsib4	.	preliminary replacement values for extreme si-ratios (differences), B iteration
seasonalb5	b5	preliminary seasonal factors, B iteration
seasadjb6	b6	preliminary seasonally adjusted series, B iteration
trendb7	b7	preliminary trend-cycle, B iteration
sib8	b8	unmodified si-ratios (differences)
replacsib9	.	replacement values for extreme si-ratios (differences), B iteration
seasonalb10	b10	seasonal factors, B iteration
seasadjb11	b11	seasonally adjusted series, B iteration
irregularb	b13	irregular component, B iteration
irrwtb	b17	preliminary weights for the irregular component
tdadjorigb	b19	original series adjusted for preliminary trading day
extremeb	b20	extreme values, B iteration
adjoriginalc	c1	original series modified for outliers, trading day and prior factors, C iteration
trendc2	c2	preliminary trend-cycle, C iteration
modsic4	c4	modified si-ratios (differences), C iteration
seasonalc5	c5	preliminary seasonal factors, C iteration
seasadjc6	c6	preliminary seasonally adjusted series, C iteration
trendc7	c7	preliminary trend-cycle, C iteration
replacsic9	c9	modified si-ratios (differences), C iteration
seasonalc10	c10	preliminary seasonal factors, C iteration
seasadjc11	c11	seasonally adjusted series, C iteration
irregularc	c13	irregular component, C iteration
tdadjorig	c19	original series adjusted for final trading day
extremec	c20	extreme values, C iteration
adjoriginald	d1	original series modified for outliers, trading day and prior factors, D iteration
trendd2	d2	preliminary trend-cycle, D iteration
modsid4	d4	modified si-ratios (differences), C iteration
seasonald5	d5	preliminary seasonal factors, D iteration
seasadjd6	d6	preliminary seasonally adjusted series, D iteration
trendd7	d7	preliminary trend-cycle, D iteration
unmodsioux	d8b	final unmodified SI ratios, with labels for outliers and extreme values
autosf	.	automatic seasonal factor selection
trendadjls	tal	final trend-cycle adjusted for level shift outliers
irregularadjao	iao	final irregular component adjusted for point outliers
modoriginal	e1	original series modified for zero-weighted extreme values
modseasadj	e2	seasonally adjusted series modified for zero-weighted extreme values
modirregular	e3	irregular component modified for zero-weighted extreme values
robustsa	e11	robust final seasonally adjusted series
mcdmovavg	f1	MCD moving average of the final seasonally adjusted series

*Name* gives the name of each table for use with the `print` and `save` arguments.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved.

The file extensions given can also be used as short names for the tables in the `print` and `save` arguments.

季節調整の反復で極端に異常な値を下方修正するための重みの上限と下限のシグマを指定する。引数 **sigmalim** として下限と上限のシグマの 2 値が入力可能。可能な値は 0 より大きな実数で上限より下限を小さくとる。(例: **sigmalim**=(1.8 2.8))。値を入れないデフォルトの場合は下限 1.5, 上限 2.5 となる。例えば **sigmalim**=(,3.0) と書けば上限 3.0 で下限が 1.5 となる。どちらかの値を代入しない場合はコンマが必要。X-12-ARIMA が極値の調整のためにこれらシグマをどう取り扱っているかの説明は詳細事項を見よ。

#### **title**

括弧 (クォーテーション) 中の季節調整のタイトルはユーザーの便宜のためである。これは 8 つのタイトルのリストから複数選択可能である。例はタイトルが二つの場合。リストが決定されたら、それぞれのタイトルはスペック・ファイルの異なる行に書かなければならない。このリストはタイトルページの系列のタイトルの下にプリントされる。デフォルトの季節調整タイトルは存在しない。

#### **trendma**

最後のトレンド・サイクルを推定するのにどの Henderson 移動平均を用いるか指定する。3 から 101 までの奇数を指定できる。例: **trendma**=23。何も選択されない場合はプログラムはデータの統計的特徴に依存してある移動平均を選ぶ。月次データは 9-, 13- または 23-term Henderson 移動平均が選択される。四半期データはプログラムは 5- または 7- の Henderson 移動平均を選択する。

#### **type**

**type**=summary ならば、プログラムはトレンドサイクルや異常、残差季節因子に関連する診断の推定を行う。更に既に季節調整されたもしくは季節性がない残差からの入力系列から残差の営業日と休日を推定する。これらの因子は除去できない。最後の季節調整系列の出力系列 (表 D11) は原系列 (表 A1) に等しい。

**type**=trend では、プログラムは季節要素を推定することなしに最後のトレンドサイクルと不規則な要素を推定してしまう。入力系列は既に季節調整されたものか季節性がないものと仮定する。このオプションでは推定された営業日や休日効果や永久的な予測調整が調整された系列 (table D11) や計算されたトレンド (table D12) から除去される。composite spec とともにメタファイルが集計の間接的な調整に使われるなら、これらのオプションが季節調整されていない集計要素に使われる。デフォルトでは **type**=sa で、プログラムは系列の季節分解を計算する。全ての **type** の 3 つの値で、最後の季節調整 (printed in the D 11 table of the main output file) が要素の間接的な季節調整の形式に使われる。

#### **x11easter**

Bateman and Mayes のノンパラメトリック手法を用い復活祭効果を推定するかどうかをの指定。**x11easter**=yes ならば予め季節調整がなされるれ、この調整から生ずる 3 月、4 月の変動は 4 つにグループ分けされる。これらのグループ毎の平均から復活祭効果の推定を行う。推定値は季節調整される系列からは除去される。これらの選択は月次系列の複数調整が行われる場合にのみ可能。X-11 の復活祭効果のより詳しいことは Monsell (1989) か Chen and Findley (1998) を参照せよ。デフォルトは **x11easter**=no。

## 稀にしか使われない引数

### **calendarsigma**

極端な値の探索や調整の為に標準誤差が使用されるが、各カレンダー月 (quarter) 毎あるいは2つの補完的なカレンダー月の集合 (quarters) 毎に別々に計算するかどうかの指定. `calendarsigma=all` なら各カレンダー月毎 (quarter). `calendarsigma=signif` ならコクランの仮説検定でカレンダー毎の変動に分散不均一性が認められた場合にのみ各カレンダー月毎 (quarter) に標準誤差を計算する. `calendarsigma=select` なら月の集合 (quarters) は二つに分けられてそれぞれのグループで標準誤差が計算される. `select` の選択では `sigmavec` を2つの月 (quarters) のグループの一つを指定に使う必要がある. `calendarsigma` が指定されない場合は標準誤差は5年毎の変動から Dagum (1988) の方法で計算される.

### **forcestart**

`force` の関連で、この選択は季節調整のトータルが計算される年の開始の指標を指定する. `forcestart` がしてされない限りは、カレンダー一年が想定されていて `force=totals` でも暗に指定される. `forcestart` で月 (月の名前か Section 5.2 の略称) や四半期 (`q1` 第一四半期, `q2` 第二四半期など) を指定すれば代替的な年の開始 (例えば会計年度) が選べる. 例えば九月もしくは十月に会計年度を指定する場合は `forcestart=october` や `forcestart=oct` と指定する. 第三四半期に会計年度を指定する場合は `forcestart=q3` とおく.

### **itrendma**

トレンド・サイクルの初期推定の移動平均モデルを指定. `itrendma = centered1yr` なら年毎の移動平均, `itrendma = cholette2yr` なら2年毎の移動平均, Pierre Cholette of Statistics Canada が開発した修正された Leser フィルター, が使われる. 詳しくは Leser (1963) and Cholette (1979) を見よ. デフォルトは1年毎の移動平均で短い周期的で変動調整された系列やトレンドの急な変化がある系列では2年のフィルターが使われる.

### **keepholiday**

最終的な季節調整で休日効果を推定するプログラムを使うかどうか決める. デフォルトでは `keepholiday=no` でプログラムからの休日効果因子は最終的な季節調整から除去される. `keepholiday=yes` ならばプログラムからの休日効果因子は最終的に季節調整まで持ち越される. デフォルトは季節ならびに休日効果両方の系列調整をするのに使われる.

### **print1stpass**

`print1stpass=yes` ならば, X-11 の復活祭や変動の回帰調整に使われる変動要素を発生させるのに必要な季節調整からの出力をプリントアウトする. `print1stpass=no` ならばこの出力は抑制され, X-11 の復活祭や変動の回帰調整に関連するテーブルのみの出力となる. デフォルトは `print1stpass=no` である. `print1stpass=yes` ならば `print` でさらにどのテーブルがプリントできるか指定する.

### **sfshort**

系列が5年より長い場合は季節調整を得る際にどの季節フィルターを使うか指定. デフォルトは `sfshort=no` で安定した季節調整フィルターが使われる. 引数 `seasonalma` には関係しない. `sfshort=yes` では X-12-ARIMA は可能な限り `seasonalma` の季節フィルターを使う.

### **sigmavec**

2つの月(四半期の)のうちの1つを指定し `calendarsigma=select` のオプションのもとでその不規則変動の標準誤差が計算される。ユーザーは月(月のフルネームか Section 5.2 の略語)が四半期(`q1` が第一四半期 `q2` が第二四半期 etc)かを入力する

**Warning:** `calendarsigma=select` のときだけこの引数は指定される。

**spectrumaxis**

`spectrumaxis=same` では異なる原形列と端点を修正した異なる季節調整がなされた系列を同じ軸にスペクトルプロットする。`spectrumaxis=diff` ならばそうしない。デフォルトは `spectrumaxis=diff` である。

**trendic**

変動とトレンド分散の比 (irregular-to-trend variance ratio) が Henderson の移動平均の端点でのウェイトを決めるのに使われるがその指定。Doherty (1991) からの手法。変数が指定されない場合は `trendic` の値は Henderson のトレンドフィルターの長さに依存する。デフォルト値は5番目の Henderson のトレンドフィルターの端点でのウェイトに近い値をとる。これは X-11 と X-11-ARIMA で最初に行われた。

**true7term**

7期の Henderson で利用される端点でのウェイトの指定。`true7term = yes` ならば主な Henderson フィルターが利用できない場合、系列の最後の観測値に7期の Henderson フィルターの非対称端点ウェイトが利用される。`true7term = no` ならば5期のものが使われ、これは Statistics Canada によってリリースされた X-11-ARIMA 前の版と同様。デフォルトは `true7term = no`。

**詳細事項**

**季節調整の形式 (Modes of seasonal adjustment):** どの X-12-ARIMA の季節調整でも原型列 (0) は3つの基本的な要素に分解される。

**トレンドと循環項 (Trend-Cycle) (C):** 中長期の系列の動き、結果としての変化点を含む。

**季節要素 (Seasonal) (S):** 年毎の同じ月、四半期で全く同じように起こるトレンドの年内変動。

**不規則変動 (Irregular) (I):** 季節因子やトレンドが系列から除去された残りの残差要素。(特定されたら営業日、休日効果も除く)。これはとても短い期間の変動として特徴づけられる。これらはストライキなどや短い期間の経済変動などによってとても大きくなることもある。

与えられた季節調整の特徴に主に依存しながらいくつかの異なったモデルが使われる。例えば C, S, I が組み合わさって原型列を構成する。X-12-ARIMA は4つの異なる分解モデルに適切な季節調整モードを与える。下のテーブルは原型列 (0) と季節調整済系列 (SA) の両方の場合に変数 `mode` の4つの値と対応するモデルを与える。

Table 6-22: 季節調整の形式とモデル

Entry for	Name for	Model for	Model for
mult	Multiplicative	$O = C \times S \times I$	$SA = C \times I$
add	Additive	$O = C + S + I$	$SA = C + I$
pseudoadd	Pseudo-Additive	$O = C \times [S + I - 1]$	$SA = C \times I$
logadd	Log-Additive	$Log(O) = C + S + I$	$SA = exp(C + I)$



デフォルトの季節調整モードは乗法型である。多くの季節経済時系列は系列の増減の水準に応じて季節変動の増減の大きさもきまる。この種の季節性は乗法的季節性 (**multiplicative seasonality**) と呼ばれる。乗法型な要素を推定するためにプログラムは Shiskin, Young, and Musgrave (1967), Dagum (1988), and Baxter (1994) で詳しく与えられている ratio-to-moving average method が使われる。擬加法モデルがある月の値が極端に小さく (休暇や天候による) 残りの月が乗法型な季節性がある場合に考えられる。系列の水準に季節変動の大きさが影響を受けない場合は加法モデルが適当である。対数加法モデルは別の乗法型分解を与え、これは時系列モデルが当てはまるある計量分析において有益である。対数加法季節調整では、トレンド要素は対数系列の加法分解から計算される。原型列と同じ単位のトレンドを作るには加法トレンドを指数変換する必要がある。これはトレンド推定の加法バイアスを生む: X-12-ARIMA では Thomson and Ozaki (1992) のバイアス補正を使い調整している。乗法型、擬加法型、対数加法型季節調整では季節要素と不規則な要素は割合で表される。多くの出力ではそれらは 100% で表される。加法季節調整では季節要素と不規則な要素は原型列と同じ単位で、中心が 0 である。季節多重度が大きくない限り、多くの乗法型季節性を持つ系列の乗法型、擬加法季節調整はとてよく似た結果を生む。もっとも小さな因子は 0.7 かそれより小さい場合、乗法型季節性と擬加法型季節性の間に大きな違いはない。最も小さな季節因子が 0.5 かそれより小さい場合はこの違いが重要になる。乗法型季節性では季節因子が小さい月 (四半期) では大きい月に比べて多くのより極端な値 (Table C17 の 100 より小さい値、特に 0) を生む。よって擬加法型の季節調整をするのが好ましい。いつ擬加法型季節調整を使うかの詳細は Baxter (1994) を見よ。単純化のためにこの議論は営業日や休日効果では無視される。これらが推定されたら、分解に追加的な因子として加えられ、定義にも依るが調整された系列の年の総和は原型列の年の総和と大きく異なり得る。

#### **Downweighting of extreme irregulars:**

不規則な要素の平均を  $\mu_I$  とする。(1.0 は乗法型季節調整、0.0 は加法的季節調整)  $\sigma_{X11}$  は月や四半期の不規則な要素の標準誤差を推定する。 $I_t - \mu_I$  の絶対値は  $\sigma_{X11}$  をかけたより低いシグマ極限で、不規則変動  $I_t$  はフルウェイトである。 $I_t - \mu_I$  の絶対値は  $\sigma_{X11}$  をかけた乗法シグマ極限より大きい、不規則値は 0 で、季節因子の計算で  $I_t$  は  $\mu_t$  に変わることを意味する。その他は  $I_t$  は下方ウェイトがかけられている。

#### **Automatic seasonal filter selection:**

この手順は X-11-ARIMA/88 からのもので Dagum(1988) を見よ。最初の二つの季節調整の繰り返しは  $3 \times 3$  の移動平均が初期季節因子の計算に使われ、 $3 \times 5$  の移動平均が最後の季節調整因子に使われる。三回目と最後の反復では  $3 \times 3$  移動平均が初期季節因子を計算するのに使われる。しかし最後の反復ではプログラムは移動季節比 ( $\bar{I}/\bar{S}$ , global MSR と呼ばれる)。プログラムは global MSR の大きさをもとに  $3 \times 3, 3 \times 5, 3 \times 9$  の移動平均選択する。季節移動平均比のより詳しい情報は Lothian (1984) を見よ。

#### **Forecast extension:**

導入で述べたように、Reg-ARIMA モデルは予測 (もしくは逆にたどる) で系列を拡張し最新の観測値の季節調整を改善する。X-12-ARIMA は Reg-ARIMA モデルがどのような forecast を指定しようとも季節調整の 1 年予測プライヤーで系列を拡張する。予測を用いた拡張なしに季節調整を行う場合は forecast で maxlead = 0 とおく。

表 6-23 : 設定環境  
SI-ratios in Table D 8.B Assumed  
Affected by a Level Shift

Percent Change in Level( $\Delta_L$ )	Length of Henderson Filter				
	23	13	9	7	5
$\Delta_L \leq 1.1$	0	0	0	0	0
$1.1 < \Delta_L \leq 1.2$	1	1	0	0	0
$1.2 < \Delta_L \leq 1.3$	1	1	1	0	0
$1.3 < \Delta_L \leq 1.5$	2	1	1	0	0
$1.5 < \Delta_L \leq 1.8$	2	1	1	1	0
$1.8 < \Delta_L \leq 1.9$	2	2	1	1	0
$1.9 < \Delta_L \leq 2.0$	3	2	1	1	0
$2.0 < \Delta_L \leq 2.6$	3	2	1	1	1
$2.6 < \Delta_L \leq 2.9$	3	2	2	1	1
$2.9 < \Delta_L \leq 3.6$	4	2	2	1	1
$3.6 < \Delta_L \leq 5.5$	4	3	2	1	1
$5.5 < \Delta_L$	5	3	2	1	1

**Residual seasonal and trading day effects in the adjusted series:**

季節、営業日のピークのスペクトルの探索するルーチン。視認できるピークは見つかりと警告を出しピークを見つけたプロットは印刷される。制限された (the -n flag) 選択が使われた場合はプロットはメインでは出力されない。しかし-n flag なしでプログラムを再度走らせるよう示唆する。

**Level shifts and the final Henderson trend:**

水準シフトが推定され季節調整の事前系列が除去された場合はそれらは最終的な Henderson のトレンド・サイクル (Table D12) で戻される。よってこの要素は観測された水準を保つ。水準が調整された時系列のトレンド・サイクルの数表は print = trendadjls とすれば得られる。

**Easter adjustment:**

このスペックにより復活祭の調整での選択は Reg-ARIMA モデルでの休日効果が regression により指定された、あるいは復活祭調整がスペック x11regression の中で指定された場合には使用できない。

**Annual Totals:**

季節調整の総和を原系列の総和と合わせるのは季節調整の質を、特に季節パターンが変化しつつある場合に、下げる。曜日の調整を行うのは集計された曜日効果に変化しゼロと緩やかに異なるから自然ではない。

**Table of SI values with labels for extreme values:**

Table D 8.B は未修正の X-11 季節調整プログラムによって生み出された季節不規則値 (トレンドなし、したがって SI 値と呼ばれる) のうちどれが極値調整で修正されなければならないか (Table C 17 にある不規則ウェイトによって示される) もしくはどれが regARIMA はずれ値によって影響を受けたであろうかの情報を与える。(ユーザーもしくは outlier スペックで指定される。) X-11 の極値として認識された各 SI 値が "\*" の次にプリントされる。単一の regARIMASI はずれ値がモデルにあった場合は SI 値が "#" の次にプリントされる。少なくひとつの regARIMA はずれ値と関連した極端な SI 値は "&" の次に印刷される。複数の regARIMA はずれ値の SI は "@" の次に印刷される。全てのはずれ値の塊の初期と最終の間の点のはずれ値としてマークされる。乗法型季

節調整で、水準変化によって最も影響を受けやすい水準変化はずれ値の前後の SI は文字 " " が値の隣にマークされる。このようにして記し付けられた観測数はレベルシフトの大きさと Table 6-30 で述べているような SI 比を生むトレンドに用いられる Henderson フィルターの長さに依存する。

#### Treatment of nonseasonal series:

非季節系列は type=trend 選択を用いてトレンドサイクルと不規則要素に分解することが可能。この分解は X-11 の季節調整分解の簡略化から得られる。この分解は Henderson とトレンドと極値探索だけ保持する。

### 例

#### 例 1

全てがデフォルトオプションの乗法型季節調整 (プログラムは移動季節比を使って季節フィルターの長さを選択) 1976 の 1 月に月次系列が始まりカレントディレクトリのファイル klaatu.dat にフリーフォーマットで保存される。

```
Series { File="klaatu.dat" Start = 1976.1 }
X11 { }
```

#### 例 2

乗法型月次季節調整、全ての月で 3×9 季節因子、23-期の Henderson 移動平均をトレンドサイクルに使う。不規則要素の回帰において営業日の回帰の優位性の検定を行う。(系列の長さには AIC の版を使う。)

```
Series { File="klaatu.dat" Start = 1976.1 }
X11 { SeasonalMA = s3x9 TrendMA = 23 }
X11regression { variables = td aicctest=td }
```

#### 例 3.

四半期季節調整、3×3 季節因子を最初の 2 四半期に用い、3×5 季節因子を残り 2 四半期に使う。7 期の Henderson トrend 移動平均

```
series {
    file="qhstarts.dat"
    start = 1967.1
    period=4
}
x11 {
    seasonalma = (s3x3 s3x3 s3x5 s3x5)
    trendma = 7
}
```

#### 例 4.

乗法型月次季節調整は、あるデフォルトモデルのタイプが自動モデル選択において指定された場合、全ての月で 12 月間の長さの ARIMA 予測拡張を用い 3×9 季節移動平均で行われる。季節調整系列の会計年度の総和が原系列の 9 月に始まる会計年度の総和に等しくなるようにする。

```
SERIES { TITLE="EXPORTS OF TRUCK PARTS" START =1967.1
          FILE = "X21109.ORI" }
AUTOMDL{ MODE=FCST }
X11 { SEASONALMA = S3X9 FORCE=TOTALS FORCESTART = OCT }
```

### 例 5.

月次データの事前調整を得るのに季節 ARIMA モデルが使われる。例では予測拡張を行わない。回帰変数を定数に指定、曜日効果、2つの水準変化、1つは1972年の5月、2つめは1976年の9月。モデルの ARIMA 部分は  $(0,1,2)(1,1,0)_{12}$  である。はずれ値の調整後の加法的季節調整系列、水準変化、Reg-ARIMA モデルを用いた曜日効果。季節分解の不規則変動ではずれ値を探すためのシグマ極限は2.0と3.5である。alltables プリントレベルを季節調整出力に用いる。

```
SERIES{ TITLE = "EXPORTS OF LEATHER GOODS"  START = 1969.JUL
        DATA = (815 866 926 ... 942) }
REGRESSION{ VARIABLES = (CONST TD LS1972.MAY LS1976.OCT) }
ARIMA{ MODEL=(0 1 2)(1 1 0) }
ESTIMATE{ }
FORECAST{ MAXLEAD=0 }
X11{ MODE = ADD PRINT = ALLTABLES SIGMALIM = (2.0 3.5) }
```

### 例 6.

使われる最初に定義される回帰効果は曜日変数と定数である。ユーザー定義の回帰変数は1988年と1990年の特別な曜日効果を捕らえる為に使われる。ファイル special.dat からこれらの変数が読み込まれる。モデルの ARIMA 部分は  $(3,1,0)(0,1,1)_{12}$  です。季節区間の12はこれがデフォルトなので指定できない。12の予測、12の後ろ向きの予測の拡張と、ユーザー定義の回帰効果と Reg-ARIMA 曜日で乗法型季節調整を行う。2つの季節調整のラインリストが指定される。

```
series { title = "Unit Auto Sales"  file = "autosal.dat"
        start = 1985.1 }
transform { function = log }
regression { variables = (const td)  user = (sale88 sale90)
            file = "special.dat"  format = "(2f12.2)" }
arma { model = (3 1 0)(0 1 1)12 }
forecast { maxlead=12  maxback=12 }
x11 { title = ("Unit Auto Sales"
             "Adjusted for special sales in 1988, 1990")
     }
```

### 例 7.

最初に定義された X-11 データのフォーマットでファイルからデータを読む。データファイルから供給される情報を初期値に用いるので指定する必要はない。あるはずれ値を取る対数変換されたデータに Reg-ARIMA モデルを用いる。これを5年の予測に用いる。全ての月の3x9の季節移動平均を使い乗法型季節調整を行う。

```
series { title="NORTHEAST ONE FAMILY Housing Starts"
        file="cne1hs.ori"  name="CNE1HS"  format="2R" }
transform { function=log }
regression {
    variables=(ao1976.feb ao1978.feb ls1980.feb
              ls1982.nov ao1984.feb)
}
arma { model=(0 1 2)(0 1 1) }
forecast { maxlead=60 }
x11 { seasonalma=(s3x9)
      title="Adjustment of 1 family housing starts"
    }
```

例 8.

あらかじめ定義した回帰効果は定数である。1980, 1985, 1991 のストライキのユーザー定義の回帰変数はファイル `strikes.dat` に含まれている。モデルの ARIMA の部分は  $(0,1,1)(0,1,1)12$  である。モデルはスペックで指定されているのでデフォルトで一年の予測を行う。ユーザー定義の回帰係数の事前予測後の季節調整。X-11 曜日効果と休日効果そして Bateman-Mayes 法での推定による復活祭効果での系列の調整。

```
series{ title="Automobile Sales"
        file = "carsales.dat"
        start = 1975.1 }
transform{ function = log }
regression{ variables = ( const )
            user = (strike80 strike85 strike90)
            file = "strike.dat" format = "(3f12.0)"
            }
arima{ model = (0 1 1)(0 1 1)12 }
x11{ title = ("Car Sales in US - Adjust for strikes in 80, 85, 90")
    save=seasonal appendfcst=yes
    x11easter=yes
    }
x11regression { variables = td }
```

## X11 回帰 (X11-REGRESSION)

---

### 解説

欠測値を伴わない時系列に関する x11 スペックの結合に用いるの任意のスペック。このスペックは、あらかじめ定義されているかまたは使用者が決めた標準的ではない説明変数による回帰モデルによって曜日効果を評価する。使用者は引数によりあらかじめ決められた回帰係数を選ぶことができる。あらかじめ決められた変数は曜日変数 (取引日や休日) と加法的外れ値のことである。状態変化オプションは取引日を説明変数におくことができる。使用者が決めた曜日効果を表す説明変数は引数を通じて含めることができる。任意の使用者の定義した変数のデータは file データ引数または名前がファイル引数で付けられたファイルを用いなければならない。(両方を用いることはできない。) 指定された回帰モデルはあらかじめ決められた回帰変数と使用者が定義した回帰変数両方を含めることができる。

### 使用法

```
x11regression{variables = (td or tdstock[31]
    easter[8]
    labor[8]
    thank[1]
    ao1967.apr )
    user = (temperature precip)
    start = 1955.jan
    data = (25 0.1 ...)
    or
    file = "weather.dat"
    format = "(2f5.1)"
    tdprior = ( 1.4 1.4 1.4 1.4 1.4
        0.0 0.0 )
    aictest = ( easter user
        td or tdstock)
    span = (1980.jan,1995.dec)
    sigma = 2.75
    or
    critical = 3.5
    outliermethod = addone
    outlierspan = (1973.may, 1992.sep)
    usertype = holiday
    print = ( brief +b15 )
    save = ( c16 c18 )
    savelog = aictest
}
```

### 引数

#### aictest

もし指定された回帰変数が使用者の regARIMA が含まれているのなら、AIC を基礎においた比

較を用いて決定することで指定される。この変数として記載できるのは **td**、**tdstock**、**tdlcoef**、**easter** と **user** である。例えば、もし曜日のモデルの選択が指定できたら AIC の値（時系列の長さによる修正があれば AICC の値を参照して）は指定された曜日変数を含むものおよび含まないモデルについて得ることができる。初期設定では予測値および外れ値の特定するには小さな AICC を持つモデルが用いられる。もし 1 種類以上の説明変数が指定されたら、AIC 検定が順次以下の手順で実行される：(a) 曜日説明変数、(b) **easter** 説明変数、(c) 使用者の定義した説明変数。同じ種類のいくつかの変数が存在したら（例えばいくつかの **td** 説明変数）、**aicctest** 操作がそれらを一まとめりとして活用される。それは、全てのこの種類の変数は最終的モデルを含むか何も含まないかである。検定操作についての更なる情報は DETAILS を見よ。もしこのオプションが指定できなければ自動的に AIC-を基礎にした選択を行なうことはできない。

#### **critical**

棄却限界（しきい値）の定め、加法的な外れ値（極端な外れ値）を見つけるのに外れ値  $t$  統計量の絶対値と比較する。この引数は引数 **sigma** を用いないか説明変数が取引日のみのものであるかではない限り用いられる。割り当てられた値は 0 より大きな実数値である。例：**critical=4.0** 初期棄却限界は外れ値を求める区間内の観測数によって決められる。（下にある引数 **outlierspan** を見よ。）表 6-15 は外れ値を探す期間の長さ数の初期棄却限界を与えている。大きな（小さな）棄却限界は不規則なものを少しだけ（たくさん）外れ値とみなす。極端な不規則値に対するプロテクションを用いたくなくれば、大きな **critical** 値を用いなければならない。

#### **data**

使用者の定義した回帰変数に割り付けた値。あらかじめ決められた時間の枠はその枠の系列（または **series** スペックの引数 **span** によって指定されたスパン）を覆っていないといけない。あらかじめ決められた時間の枠は **forecast** スペックで要請された予測および逆予測の時間の枠も覆っていないといけない。データの値は自由なフォーマットで読み込むことができる。この引数で与えられた数値は、引数 **user** で名付けられた使用者が定義した変数の順番通りに与えられなければならない。この割り当ては、はじめの時点における使用者の定義した全ての変数を通じて行なわれるべきであり、続いて次の時点の全ての変数等として行なわれる。もし、引数 **data** が使われたら、引数 **file** 使うことができない。

#### **file**

ファイルに入ったデータの名前は使用者の定義した全ての回帰変数を評価する。ファイル名括弧（クォーテーション）で囲まれないといけない。もしファイルがカレント・ディレクトリになければパスを与えないといけない。引数 **data** のようにあらかじめ与えられた時間の枠は系列と任意の予測逆予測の時間の枠を含まなければならない。もし引数 **file** が使われたら、引数 **data** が使われないといけない。

#### **format**

引数 **file** の中で名付けられたファイルの中の回帰変数の値を読み込む時に使われたフォーマットを表す。3 種類の input が選ばれうる。

- (a) コーテーション記号で囲まれかつ前後の丸括弧で囲まれた正当な FORTRAN のフォーマット  
(例: `format='(6f12.0)'`);
- (b) 年次、月次または四半期の “date-value” フォーマットと与えられた観測に関する使用者の定

義した各々の回帰変数の値がその順序で与えられ、データファイルの中の各行のフォーマットは自由である。よって、1991年7月に0,0,1の値を3つの回帰変数に持つ行は1991 7 0 0 1のように書き表される。全ての使用者の定義した回帰変数はいつも同じ順番でないといけなく、その順序は引数 **user** の順序通りでないといけなく。(例: `format='{x12save}'`)

- (c) "x12save" フォーマット X-12-ARIMA は表を保存するのに用いられる。これにより前回使用者が X-12-ARIMA を実行して保存したファイルを読み込むことができる。(例: `format='{x12save}'`)

もしデータに引数 **format** が与えられていなければ自由なフォーマットから読み込む。自由な形式のフォーマットでは次の行に移る前にその行の全ての数字が読み込まれる。各数字は一文字以上のスペースを空けないといけなく。(コンマやタブを用いれない) **format** は引数 **data** といっしょに用いることはできず、**format** のみで使われないといけなく。

#### **outliermethod**

モデルにおいて検出された外れ値を順次加えるプログラムを決定する。とりうる選択肢は **method** = **addone** または **method** = **addall** である。これらの二つの方法の解説については DETAILS の節中の **outlier** スペックにおいて見てみよう。初期設定では **method** = **addone** になっている。この引数は引数 **sigma** が使われていたら使用することができない。

#### **outlierspan**

外れ値が探された不規則なコンポーネントの期間のはじめと終わりの日にちを特定する。はじめの日と最後の日には系列の中に必ず含まれていないとならず、はじめの日には必ず最後の日より前にないといけなく。欠損値、例えば `outlierspan = (1976.jan, )`、は最初の日や最後の日は適切にはずさねばならない。(もし、引数 **span** が **series** スペックにあれば、**series** スペックで与えられた期間の最初の日と最後の日からそれらに移さなければならない。) この引数は引数 **sigma** といっしょに用いることができない。

#### **print and save**

表 6-24 はスペック **x11regression** の出力が可能なものの一覧である。

`savelog=aictest` または `savelog=ats` とセットすると引数 **aictest** を指定することによって AIC を基準とした選択手続きの結果がログファイルに出力する。(ログファイルについての詳細は 2.6 節を見よ。)

#### **sigma**

曜日回帰が行なわれる前に極端な値をとる不規則な要素外すシグマ極限。不規則値と平均の差がここで入力した値と標準偏差の積より大きいときに極端な値として外される。各不規則値は月次(四半期)型によって決められる標準誤差を持つ。月次型は月の長さとの始まる週の日によって決められる。この引数は曜日の列がモデルの中にあるか、引数 **critical** がある場合以外に用いることができない。割り当てられた実数は 0 より大きく、初期設定は 2.5 である。(曜日系列変数が推定される唯一の回帰変数である場合にのみ実施される) 例: `sigma=3.0`。

#### **span**

回帰モデルの係数の推定に用いられた不規則な要素の期間(データ間隔)を指定する。この引数は、使用者が回帰の推定に影響を及ぼす初めの方のデータを季節調整前の事前調整で用いることをしたくない場合に用いることができる。**series** スペックのところでは詳細が述べられている **modelspace**



表 6-24: X11regression に関連する出力

<i>name</i>	<i>default</i>	<i>brief</i>	<i>ext</i>	<i>description of table</i>
prior <sub>td</sub>	+	+	a4	prior trading day weights and factors
extreme <sub>valb</sub>	.	.	b14	irregulars excluded from the irregular regression, B iteration
x11 <sub>regb</sub>	.	.	.	preliminary irregular regression coefficients and diagnostics
trading <sub>dayb</sub>	.	.	b16	preliminary trading day factors and weights
comb <sub>tradingdayb</sub>	.	.	b18	preliminary trading day factors from combined daily weights
holiday <sub>b</sub>	.	.	bxh	preliminary holiday factors
calendar <sub>b</sub>	.	.	bxo	preliminary calendar factors
comb <sub>calendarb</sub>	.	.	bcc	preliminary calendar factors from combined daily weights
extreme <sub>val</sub>	+	.	c14	irregulars excluded from the irregular regression, C iteration
x11 <sub>reg</sub>	+	.	.	final irregular regression coefficients and diagnostics
trading <sub>day</sub>	+	+	c16	final trading day factors and weights
comb <sub>tradingday</sub>	+	+	c18	final trading day factors from combined daily weights
holiday	+	+	xh1	final holiday factors
calendar	+	+	xca	final calendar factors (trading day and holiday)"
comb <sub>calendar</sub>	+	+	xcc	final calendar factors from combined daily weights
xrout <sub>lierhdr</sub>	+	.	.	options specified for outlier detection including critical value and outlier span
xrout <sub>lieriter</sub>	.	.	xoi	detailed results for each iteration of outlier detection including outliers detected, outliers deleted, model parameter estimates, and robust and non-robust estimates of the residual standard deviation
xrout <sub>liertests</sub>	.	.	.	t-statistics for every time point of each outlier detection iteration
xr <sub>egressionmatrix</sub>	.	.	xrm	values of irregular regression variables with associated dates
xr <sub>egressioncmatrix</sub>	.	.	xrc	correlation matrix of irregular regression parameter estimates if used with the <code>print</code> argument; covariance matrix of same if used with the <code>save</code> argument
xaic <sub>test</sub>	+	+	.	output from AIC-based tests for trading day and holiday

*Name* gives the name of each table for use with the `print` and `save` arguments.

*Default* indicates which tables are printed (+) or not printed (.) by default.

*Brief* indicates which tables are printed (+) or not printed (.) when the `brief` print level is specified. See section 5.1 for more information on print levels.

*Ext* gives the file extension used if the table is saved. A dot indicates that the table cannot be saved. The file extensions given can also be used as short names for the tables in the `print` and `save` arguments.

スペックのように、引数 **span** は二つの値、興味のある期間のはじめの日にちと最後の日にち、をとる。欠測値は用いられた系列の最初の日か最後の日から外される。例えば、月次データでは宣言 **span**=(1968.1, ) は 1968 年から始まって解析期間の最後の日で終る時系列データから推定されることを指定する。はじめの日か最後の日が抜けている場合はコンマが必要である。モデルの期間の初日と最後の日は必ず両方とも **series** スペックによる解析のデータの期間中を指定せねばならない。またはじめの日は最後の日より先に置かれなければならない。

最終日の指定においてはさらに、*0.per* の形で、**span** 最後の日が、データが解析された期間中の暦上でもっとも最近の当該月 (四半期データの場合は四半期) になるようにすることができる。ここで、*per* は暦の月 (四半期) を表す。よって、もし 12 月以外の月で終るある期間のデータを考えると、**span**=(,0.dec) によって期間中の最後の前の年の 12 月でデータを取り終えるように回帰係数をとどめることができる。

#### **start**

使用者の定める回帰変数の値のはじめの日のこと。初期設定では系列のはじめの日になる。有効な値は系列のはじめの日以前の任意の日にち (またはもし存在するのならば **series** スペックの引数 **span** 指定された以前の日にち) である。

#### **tdprior**

月曜日から始まる使用者が挿入する 7 日分の重み付けのリストで、季節調整に先立って行ないたい X-11 取引日調整を指定する。これらの重みはプログラムにより和が 7 になるように調整される。このオプションは乗法的かつ対数加法的季節調整にのみ使われることができる。値は 0 に等しいかそれより大きな実数値でなければならない。例: **tdprior**=(0.7 0.7 0.7 1.05 1.4 1.4 1.05).

#### **user**

使用者の定める回帰変数名前のリストを指定する。係数が推定される使用者が定義した各変数では名前が必要である。与えられた名前はプログラムの出力において推定された係数のラベルとして用いられる。使用者が定義した変数の値は引数 **data** または引数 **file** を用いて与えられなければならない。使用者が定義する回帰変数の最大値は 52 である。(この最大値は変更することができる— 2.8 節を見よ)

#### **usertype**

使用者が定める回帰変数の型を割り当てる。使用者の定める回帰効果は曜日変数 (**td**)、曜日ストック変数 (**tdstock**)、休日変数 (**holiday**, イースター変数 **easter**, および合衆国の休日 **thanks**, **labor**)、加法的外れ値 (**ao**)、または使用者の定義した (**user**) 回帰効果である。一種類の効果の型は **x11regression** スペック (**usertype**=**td**) で定められた使用者の定義した全ての回帰変数が指定されるか各使用者の定義した回帰変数はそれ自体の型 (**usertype**=(**td td td td td td td holiday user**)) で与えられるかである。使用者の定義した回帰変数の割り当てられた型については **DETAILS** を見よ。

#### **variables**

モデルに含まれたあらかじめ定められた回帰変数のリスト。これらの変数の値はプログラムによって計算され、ほとんどの場合暦の関数になっている。使用できるあらかじめ定められた変数に関する議論と表は **DETAILS** を見よ。

## 稀にしか使われない引数

### **aicdiff**

引数 **aicdiff** で指定された回帰変数が受容されるのに必要な AICC との差異を定める。初期設定では **aicdiff=0.0** である。

### **b**

引数 **variables** および引数 **user** の順番で表れる回帰パラメータの不規則な要素の初期値または固定値を指定する。もし引数 **b** が存在すれば、**regARIMA** モデル中の全ての回帰係数を初期値に割り当てなければならない。初期値は引数のリストまたは下の例のように欠損していることを明示的に指示することによりパラメータに割り当てられる。欠損値は初期設定値 0.1 をとる。例えば二つの回帰係数を含むモデルにおいては **b=(0.7, )** は **b=(0.7, 0.1)** と同値である。ただし **b=(0.7)** と書くことはできない。3変数の回帰変数を含むモデル **b=(0.8, , -0.4)** は **b=(0.8, 0.1, -0.4)** と同値である。パラメータが指定された値のまま固定するには、**b** リスト値の直後に **f** を付ける。つまり **b=(0.7f, 0.1)** とする。

### **centeruser**

使用者の定義した回帰変数から(標本)平均または季節平均の除去を指定する。もし **centeruser=mean** とすると各回帰変数の平均が使用者の定義した回帰変数から取り除かれる。**centeruser=seasonal** は各使用者の定義した回帰変数から暦上の月内(四半期)の平均が取り除かれる。もしこのオプションが指定されなければ使用者が定義した適切に中央に配置されていると仮定されており、修正は行なわれない。

### **eastermeans**

長期の月次平均(400年分)が表3-1の脚注5で説明されたように **easter[w]** 変数を用いて生成されたイースターのための回帰変数を非季節化するか、または、イースターのために用いられる回帰係数 (**eastermeans=no**) を計算するのに用いられるデータの期間から平均を計算する。初期設定は (**eastermeans=no**) である。この引数は回帰モデルの中にイースターのための回帰変数が入っていないか、イースターのための回帰変数が **sceaster[w]** (Detailsを見よ) のみの場合には、無視される。

カレンダー調整の要素が休日の積(または、加法的季節調整が用いられたら和)が同じ値になることを阻止する (**forcecal=yes**) か阻止しない (**forcecal=no**) かを指定する。この引数は休日および取引日回帰係数がこのスペックの引数 (**variable**) の中で指定される時のみ作動する。

### **noapply**

原系列または最終的な調整系列の外で調整されるべきではないモデルにおける推定値 **x11regression** スペック内に定義されている回帰の効果の種類のリスト。利用できる効果にはモデルの曜日効果 (**td**) とイースター、勤労感謝の日感謝祭ークリスマス休日効果 (**holiday**) が含まれる。

### **umdata**

使用者の定義した回帰変数を含むモデルが推定される前に不規則な系列  $I_t$  (または  $\text{Log } I_t$ ) から取り除かれる平均調整値の入力配列。この引数 **umfile** は DETAILS に述べられているあらかじめ定義された平均関数が使用者が定義した回帰変数を含むモデルにとって不相応な場合用いられる。平均調整関数は調整の方法に依存した関数である。更なる情報は DETAILS を見よ。

これらの値の時間の枠は系列の時間の枠(または、もしあれば **series** スペックの引数 **span** で指

定された期間)を必ず含まないといけない。これらの時間の枠は **forecast** スペックにおいて要請された予測と逆予測の時間の枠も含まないといけない。データの値は好きなフォーマットで読みこまれる。もし引数 **umdata** が使われたら引数 **umfile** は使うことができない。

#### **umfile**

使用者の定義した回帰変数を含むモデルが推定される前に不規則な系列  $I_t$  (または  $\text{Log } I_t$ ) から取り除かれる平均調整値を含んだファイルの名前。これは DETAILS に解説があるように、あらかじめ定義された回帰変数が使われたときのみ  $I_t$  から取り除かれた平均関数を戻す。ファイル名はコーテーション・マークで囲まれなければならない。もしファイルがカレント・ディレクトリになれば、パスが与えられなければならない。引数 **umdata** のようにデータの時間の枠の値は系列と任意の予測と逆予測を含まなければならない。もし引数 **file** が使われたら、引数 **umdata** を使うことはできない。

#### **umformat**

引数 **umfile** の中に名付けられたファイルから回帰変数のデータを読む時に使われるフォーマットを示す。5種類の入力が許されている。

- (a) コーテーション・マークで囲まれて前後に括弧でくくられた有効な FORTRAN フォーマット  
(例: `umformat='(6f12.0)'`);
- (b) "date-value" フォーマット、この中には年次、月次、または四半期の別、およびそれに対応した与えられた観測に対する付随した平均調整の値がこの順に並んでいて、データ・ファイル中の各列は自由なフォーマットで入っている。よって、1991年の7月の1.01を平均調整したデータの列は 1991 7 1.01 と書かれる。(例: `umformat='{datevalue}'`);
- (c) "x12save" フォーマットで、X-12-ARIMA が表を保存するときに用いられる。これを用いることにより使用者はまえに走らせた X-12-ARIMA が保存したファイルを読むことができる。  
(例: `umformat='{x12save}'`)
- (d) 前のバージョンである X-11 および X-11-ARIMA で使われたデータ・フォーマットの集合と一致した2文字のコード。(例: `umformat='{tramo}'`)
- (e) TRAMO および SEATS プログラムが系列およびその説明により読み込むのに使うフォーマット。これにより X-12-ARIMA が TRAMO モデリング・プログラム および季節調整プログラム SEATS のためにフォーマットしたデータ・ファイルを読み込むことができる。

(d) でみられたあらかじめ定義された X-11 データフォーマット内で、各年のデータに年と系列のラベルを加えた6から12文字の文字群が保存されている。これらのフォーマットの完全なリストは **series** スペックの DETAILS の節を見よ。もし引数 **umformat** が与えられていなければデータは自由なフォーマットで読み込むことができる。**Umformat** は引数 **umdata** といっしょに使うことはできない、**umfile** といっしょにのみ使うことができる。

#### **umname**

**umfile** におけるファイル名において保存された系列の値の名前。名前は必ずコーテーションで囲まれていなければならない。64文字まで含ませることができる。初めの16文字までが平均調整値の使用

者の定めた平均のラベルとしてプリントされる。引数 **umformat** のあらかじめ定めたフォーマットを指定すると、この名前の初めの6文字 (もし **umformat**=' '{cs}'"ならば8文字はプログラムが正しい系列を読んでいるかチェックしたり多くの系列が格納されたファイルからある系列を見つけ出すためのあらかじめ定められたフォーマットに用いる。

#### **umprecision**

使用者が定義した平均から読み込まれる小数点以下の桁数。このオプションはあらかじめ定められたフォーマットの引数 **umformat** とのみいっしょに用いることができる。この値は0から5までの整数を用いなければならない。(例えば、**umprecision**=5) 初期設定は0である。

**umdata** または **umfile** で指定された平均調整値の初日。初期設定は系列の初日である。適切な値は系列の初日までの任意の日である。(もしあれば **series** スペックの引数 **span** で指定された期間の初日までの日)

#### **umtrimzero**

もし **umtrimzero=no** であれば、引数 **umfile** を通じて入る使用者の平均時系列の初めか終わりの0を系列の値としてみなす。初期設定 (**umtrimzero=yes**) は0を無視する原因になる。引数 **format** を **datevalue** か **x12save** または **tramo** とすれば **umtrimzero** の値にかかわらず、全ての入力値は系列の値とみなされる。

### 詳細事項

このスペックはカレンダー効果や他の効果を効果を調整しない予備季節調整の不規則要素  $I_t$  から推定するのに用いる。最小二乗法を効果の回帰モデルに適用することで推定が行われる。最も簡単な場合は以下でモデルは

$$I_t - 1.0 = \beta' X_t + e_t,$$

となる。ここで  $X_t$  は興味のある基本的な効果を表す変数である。他の場合はより複雑な  $I_t$  の線型変換がモデルの右辺に現れる。全ての場合において誤差項  $e_t$  は独立の定数として  $t$  統計量とカイ2乗統計量とAICが計算される。残念なことに  $I_t$  を用いたフィルタリング操作は  $e_t$  の複数の過程を保証するものではない。上述の統計量からの推定された効果の統計的有意性の決定は **regression** を用いた **regARIMA** からの推定された効果の決定より信頼性が低い。つまり **x11regression's** 統計量は **regression's** よりも実際はそうでない時に効果は有意と示しがちである.. 効果が本当にある場合は **regression** と **x11regression** スペックからの推定はほとんど同じである。それらがはっきり違うと分かる時は **regression** は **x11regression** よりも大きい。( **history** のスペックの予測診断は十分な長さの系列の推定された効果を比べるのに用いられる。 Findley=Monsell=Bell=Otto=Chen (1998)、及び Soukup=Findley (2000) を見よ。 ) **x11regression** はユーザーが興味のあるデータ期間によく当てはまる **regARIMA** を見つけられなかったときまで使用しない。

Table 6-25: X11regression の (指定された) 回帰変数

Variable	Description
td	Estimates monthly (or quarterly) flow trading-day effects by adding the <code>tdnolpyear</code> variables (see Table 6-17) to the model. The derivations of February from the average length of 28.25 are handled either by rescaling (for multiplicative adjustments) or by including the <code>lpyear</code> regression variable (for additive and log-additive adjustments). <code>Td</code> cannot be used with <code>tdstock[]</code> or <code>td1coef</code> .
td1coef	Estimate monthly (or quarterly) flow trading-day effects by including the <code>tdnolpyear</code> variable (see below) in the model, and by handling leap-year effects either by rescaling (for transformed series) or by including the <code>lpyear</code> regression variable (for untransformed series). <code>Td1coef</code> can only be used for monthly or quarterly series, and cannot be used with <code>td</code> or <code>tdstock[]</code> .
tdstock[ <i>w</i> ]	Adds 6 stock trading-day variables to model the effect of the day of the week on a stock series estimated for the $w^{\text{th}}$ day of each month. The value $w$ must be supplied and can range from 1 to 31. For any month of length less than the specified $w$ , the <code>tdstock</code> variables are measured as of the end of the month. Use <code>tdstock[31]</code> for end-of-month stock series. <code>Tdstock</code> can be used only with monthly series and cannot be used with <code>td</code> or <code>td1coef</code> .
easter[ <i>w</i> ]	Easter holiday regression variable (monthly or quarterly flow data only) which assumes the level of daily activity changes on the $w$ -th day before Easter and remains at the new level until the day before Easter. The value $w$ must be supplied and can range from 1 to 25. To estimate complex effects, several of these variables, differing in their choices of $w$ , can be specified.
labor[ <i>w</i> ]	Labor Day holiday regression variable (monthly flow data only) that assumes the level of daily activity changes on the $w$ -th day before Labor Day and remains at the new level until the day before Labor Day. The value $w$ must be supplied and can range from 1 to 25.
thank[ <i>w</i> ]	Thanksgiving holiday regression variable (monthly flow data only) that assumes the level of daily activity changes on the $w$ -th day before or after Thanksgiving and remains at the new level until December 24. The value $w$ must be supplied and can range from -8 to 17. Values of $w < 0$ indicate a number of days after Thanksgiving; values of $w > 0$ indicate a number of days before Thanksgiving.
sceaster[ <i>w</i> ]	Statistics Canada Easter holiday regression variable (monthly or quarterly flow data only) assumes that the level of daily activity changes on the $(w - 1)$ -th day and remains at the new level through Easter day. The value $w$ must be supplied and can range from 1 to 24. To estimate complex effects, several of these variables, differing in their choices of $w$ , can be specified.
ao <i>date</i>	Adds an additive (point) outlier variable, AO, for the given date or observation number. For series with associated dates, AOs are specified as <code>ao<i>date</i></code> . For monthly series the form is <code>ao<i>year.month</i></code> (e.g., <code>ao1985.jul</code> or <code>ao1985.7</code> ), for quarterly series it is <code>ao<i>year.quarter</i></code> (e.g., <code>ao1985.1</code> for an AO in the first quarter of 1985). More than one AO may be specified. All specified outlier dates must occur within the series. (AOs with dates within the series but outside the span specified by the <code>span</code> argument of the series <code>spec</code> are ignored.)

### 乗法的分解の為の不規則成分回帰モデル

不規則要素には季節性やトレンドは(乗法分解において定数の水準が1.0を超えるようには)存在しないと仮定される。したがって、不規則変動の回帰モデルの回帰変数は通常季節もしくはトレンド要素を持たないはずである。この理由から regression で使用される曜日(表3-1と表6-15を見よ。)と復活祭の回帰変数は x11regression のスペック(表6-25を見よ。)では修正されている。季節調整分解のさまざまなタイプは Findley, Monsell, Bell, Otto and Chen (1998) の1.4章からのものである。

この修正の性質を以下で述べるの月次の営業日と休日を組み合わせた形式で示す。

$$\gamma_0 m_t + \sum_{j=1}^6 \gamma_j (d_{j,t} - d_{7,t}) + \delta' \mathbf{H}_t, \quad (1)$$

ここで  $d_{j,t}$  は月  $t$  の  $j$  のタイプの週の数で ( $j = 1, \dots, 7$  は月曜, ... は土曜をそれぞれ表す。)  $m_t = \sum_{j=1}^7 d_{j,t}$  ( $t$  月の長さ) と  $\mathbf{H}_t$  は休日変数の列ベクトルを表す。カレンダーの定義からこれら興味のある変数のほとんどの時間間隔は周期的で  $m_{t+48} = m_t$ ,  $d_{j,t+336} = d_{j,t}$ ,  $\mathbf{H}_{t+P} = \mathbf{H}_t$  で  $P$  は  $\mathbf{H}_t$  に含まれる休日変数による。(グレゴリー暦の全ての提案されるコレクションが使われた場合、復活祭カレンダーは 38,000 年 = 456,000 月の期間がある。この理由から近似された周期性が  $\mathbf{H}_{t+P} \approx \mathbf{H}_t$  を持つように  $P$  を選ぶことがより実用的である。)  $f_t$  が期間  $12p$  月の近似関数  $f_{t+12p} \approx f_t$  ならば、その(近似)季節と水準要素はそのカレンダー一月の平均、

$$f_t^* = \frac{1}{p} \sum_{j=1}^p f_{t+12j},$$

で与えられ、近似期間は12ヶ月の周期で  $f_{t+12}^* \approx f_t^*$ 。季節と水準効果が分割により  $f_t$  から除かれたら、結果としての非季節化、 $f_t$  の水準中立の要素は  $f_t/f_t^*$ 。これらのアイデアを上関数(1)に当てはめるためには  $p$  が多重度 28 ならば  $d_{j,t}^* = d_{7,t}^*$ ,  $1 \leq j \leq 6$  で結果としてのこのカレンダー効果関数の季節と水準要素は以下であることに注意する必要がある。

$$\gamma_0 m_t^* + \delta' \mathbf{H}_t^*,$$

ここで

$$m_t^* = \begin{cases} m_t & , m_t = 30, 31 \\ 28.25 & , m_t = 28, 29 \end{cases}$$

となる。それゆえ時系列が(1)の形の曜日と休日効果を含むとき、乗法非季節化と非トレンド化からの不規則要素は曜日と休日効果を持つことが期待できる。

$$\frac{\gamma_0 m_t + \sum_{j=1}^6 \gamma_j (d_{j,t} - d_{7,t}) + \delta' \mathbf{H}_t}{\gamma_0 m_t^* + \delta' \mathbf{H}_t^*} = \frac{\frac{m_t}{m_t^*} + \sum_{j=1}^6 \alpha_j ((d_{j,t} - d_{7,t})/m_t^*) + \beta' \frac{\mathbf{H}_t}{m_t^*}}{1 + \beta' \frac{\mathbf{H}_t^*}{m_t^*}} \quad (2)$$

右辺の表現は  $\alpha_j = \gamma_j/\gamma_0$  と  $\beta = \delta/\gamma_0$ 。しかし、曜日効果と休日効果は通常 2.3% の範囲であるから、近似

$$\left(1 + \beta' \frac{\mathbf{H}_t^*}{m_t^*}\right)^{-1} \approx 1 - \beta' \frac{\mathbf{H}_t^*}{m_t^*}$$

は (2) に当てはめることができる。このファクターによって (2) の右辺分子をかければ係数の積を含む項は一般的に小さくなり無視できる。これは (2) の線形近似を生む。

$$\frac{m_t}{m_t^*} + \sum_{j=1}^6 \alpha_j \left( \frac{d_{j,t} - d_{7,t}}{m_t^*} \right) + \beta' \left( \frac{\mathbf{H}_t - \mathbf{H}_t^*}{m_t^*} \right). \quad (3)$$

近似を得るために、これも使う、

$$\frac{m_t}{m_t^*} = 1 + \frac{1}{28.25} (m_t - m_t^*),$$

また閏年変数を含む項を取り扱う  $LY_t = m_t - m_t^*$  その積は  $\beta' (\mathbf{H}_t - \mathbf{H}_t^*) / m_t^*$  は無視できる。式 (3) は以下の形式の線形回帰モデルの不規則変動  $I_t$  を示唆する。

$$I_t - \frac{m_t}{m_t^*} = \sum_{j=1}^6 \alpha_j \left( \frac{d_{j,t} - d_{7,t}}{m_t^*} \right) + \beta' \left( \frac{\mathbf{H}_t - \mathbf{H}_t^*}{m_t^*} \right) + \kappa' \mathbf{AO}_t + e_t,$$

ここで  $\mathbf{AO}_t$  は加法的はずれ値の変数を含む回帰ベクトルを意味する。このモデルを使う代わりに X-12-ARIMA で X-11 と X-11-ARIMA の曜日効果の回帰モデルに従って以下のスケール変換されたモデルの最小二乗法からの (3) の係数を得る。

$$m_t^* I_t - m_t = \sum_{j=1}^6 \alpha_j (d_{j,t} - d_{7,t}) + \beta' (\mathbf{H}_t - \mathbf{H}_t^*) + \kappa' \mathbf{AO}_t + \varepsilon_t \quad (4)$$

ここで `td` は `x11regression` の引数 `variables` で指定され、複数の休日効果を `easter[w]`, `labor[w]`, `thank[w]` で指定する。Table 3-1 の下に説明されているようにこれらの休日変数の関連する回帰変数は非季節化された形式  $\mathbf{H}_t - \mathbf{H}_t^*$  でこれらが回帰スベックから推定される。これはモデルの季節部分からと分解の季節要素のなかでのみ季節効果が起こるようにしたものである。X-11-ARIMA/88 に従って `sceaster[w]` に関連する回帰は決して非季節化されない。実際、`sceaster[w]` の回帰変数で指定される (4) の  $\mathbf{H}_t^*$  の入力は 0 である。

曜日要素と休日要素

カレンダー要素 (3) は休日要素と曜日要素の積として近似的に分解される。

$$1 + \beta' \left( \frac{\mathbf{H}_t - \mathbf{H}_t^*}{m_t^*} \right) \quad (5)$$

$$\frac{m_t}{m_t^*} + \sum_{j=1}^6 \alpha_j \left( \frac{d_{j,t} - d_{7,t}}{m_t^*} \right) = \frac{\sum_{j=1}^7 (1 + \alpha_j) d_{j,t}}{m_t^*}, \quad (6)$$

(ここで  $\alpha_7 = -\sum_{j=1}^6 \alpha_j$  である)。数  $1 + \alpha_j$  は *daily weights* と呼ばれる。曜日要素 (6) は以下のよう  
に書ける。

$$\frac{m_t}{m_t^*} + \frac{\sum_j^{(5)} \alpha_j}{m_t^*},$$

ここで  $\sum_j^{(5)}$  は  $j$  の和  $d_{jt} = 5$  を表す。この形式は月の長さ効果から別に営業日効果は 1 月に 5 回起こる曜日効果だけによることを示す。曜日効果だけが推定された場合は形式は  $\beta = 0$ 。で応用される。



もし1つかそれ以上の曜日の”ウェイト” $1 + \alpha_j$ が負の値をとり、かつ **reweight** オプションが設定されているならば、曜日要素の計算において  $\alpha_j < -1$  を満たす全ての  $\alpha_j$  は  $\alpha_j = -1$  で置き換えられ、また  $\alpha_j \geq -1$  を満たす全ての  $\alpha_j$  は  $\alpha_j = (1 + \alpha_j)w - 1$  で置き換えられる。ここで

$$w = 7 \left\{ \sum_{\alpha_i \geq -1} (1 + \alpha_i) \right\}^{-1}$$

であるとし、 $\alpha_j > -1$  となるいずれの  $\alpha_j$  に対しても、**b** 引数によって固定された値が割り当てられていないとする。もし固定された値があるならば、 $\alpha_j > -1$  を満たす  $\alpha_j$  のうち固定されていないもののみが置き換えられ、その場合の  $w$  は全ての  $\alpha_i > -1$  に対して

$$w = \left\{ 7 - \sum_{\alpha_i \text{ fixed}} (1 + \alpha_i) \right\} \left\{ \sum_{\alpha_i \text{ not fixed}} (1 + \alpha_i) \right\}^{-1}$$

と定義される。

休日効果またはストックの曜日効果のみの推定

**x11regression** スペックの **variables** 引数において休日効果か曜日効果のみが指定されているならば、**X-12-ARIMA** はこれらの効果を

$$I_t - 1 = \beta' (\mathbf{H}_t - \mathbf{H}_t^*) + \kappa' \mathbf{A}\mathbf{O}_t + e_t \quad (7a)$$

および

$$I_t - 1 = \sum_{j=1}^6 \alpha_j D_{j,t} + \kappa' \mathbf{A}\mathbf{O}_t + e_t \quad (7b)$$

という形式のモデルを適用して OLS により推定する。ここで  $D_{j,t}$  は表 3-1 の **tdstock[w]** に関する回帰変数を表す。これらのモデルはそれぞれ

$$1 + \beta' (\mathbf{H}_t - \mathbf{H}_t^*) \quad (8a)$$

および

$$1 + \sum_{j=1}^6 \alpha_j D_{j,t} = 1 - \alpha_j(t) \quad (8b)$$

という形式のカレンダー効果調整の成分を導く。ここで  $\alpha_j(t)$  は 第  $t$  月の  $w$  番目の日の係数を表す。

ユーザー定義のフローの曜日効果と休日効果

(4) の回帰モデルは、係数推定値と独立に知られている不規則成分  $I_t$  の平均関数を表す要素として  $m_t/m_t^*$  をもたらす。これはユーザー定義変数が用いられている場合の、既知の要素に関するデフォルトの設定でもある。このデフォルト設定が適切ならば、少なくとも1つのユーザー定義の曜日変数ないし休日変数を含む回帰モデルは、

$$m_t^* I_t - m_t = \alpha' \mathbf{T}\mathbf{D}_t + \beta' \tilde{\mathbf{H}}_t + \kappa' \mathbf{A}\mathbf{O}_t + \varepsilon_t, \quad (9)$$

という形式になる。ここで  $\mathbf{TD}_t$  と  $\tilde{\mathbf{H}}_t$  は指定された営業日変数と休日変数のベクトルを表す。ユーザー定義変数は `file` 引き数か `data` 引き数によって導入される。プログラムはユーザー定義変数の季節性除去は行わない。これらは適切に季節性除去がなされた上でプログラムに導入されるべきである。X-12-ARIMA は、(5) と (6) と同様にして休日成分と曜日成分に近似的に分解されるカレンダー成分

$$\frac{m_t}{m_t^*} + \alpha' \frac{\mathbf{TD}_t}{m_t^*} + \beta' \frac{\tilde{\mathbf{H}}_t}{m_t^*}$$

を計算する。休日効果のみか、またはストックの曜日効果のみが推定される場合は、デフォルトの既知の平均関数は定数 1.0 となり、モデルと結果的に得られる休日成分は、(7) と (8) の類似形となる。

上述したデフォルトの既知の平均関数が適切でなければ、ユーザーは平均関数  $\mu_t$  を、`umfile` 引き数か `umdata` 引き数によって導入することができる。この場合は推定される回帰モデルは

$$I_t - \mu_t = \alpha' \mathbf{TD}_t + \beta' \tilde{\mathbf{H}}_t + \kappa' \mathbf{AO}_t + e_t \quad (10)$$

となり、カレンダー成分

$$\mu_t + \alpha' \mathbf{TD}_t + \beta' \tilde{\mathbf{H}}_t$$

のみが得られる。(10) から推定される係数  $\alpha$ 、 $\beta$  は、(9) から得られる係数とは異なるスケールに基づいており、大まかに言って

$$\frac{1}{48} \sum_{j=0}^{47} \frac{1}{m_t^*} \simeq .03288$$

だけ小さくなっている。スケールの違いに関する同様の概算が、`x11regression` の代わりに `regression` から計算される、あるいは (4) の代わりに (7) から計算されるカレンダー効果の係数についても成り立つ。

#### その他の分解法に対する不規則成分の回帰モデル

以下では、加法的、擬加法的および対数加法的な分解において、予め定義された回帰変数によるフローの曜日効果と休日効果の推定に用いられるモデルを提示する。ユーザー定義変数、フローの曜日効果ないし休日効果の回帰変数を含む場合のモデルの適切な修正は、乗法的分解に関する上述の説明で示された修正と同様である。

#### 加法的分解

もし `x11` スペックで `mode=add` となっているならば、カレンダー効果は

$$I_t = \alpha_0 LY_t + \sum_{j=1}^6 \alpha_j (d_{j,t} - d_{7,t}) + \beta' (\mathbf{H}_t - \mathbf{H}_t^*) + \kappa' \mathbf{AO}_t + e_t$$

という形のモデルにより OLS を用いて推定される。従ってこのカレンダー効果は正確に曜日効果  $\alpha_0 LY_t + \sum_{j=1}^6 \alpha_j (d_{j,t} - d_{7,t})$  と休日効果  $\beta' (\mathbf{H}_t - \mathbf{H}_t^*)$  との和になる。

#### 擬加法的分解

もし `x11` スペックで `mode=pseudoadd` となっているならば、 $\bar{m} = 30.4375$  および  $LY_t = m_t - m_t^*$

とした上で、カレンダー効果は

$$\bar{m}(I_t - 1) - LY_t = \sum_{j=1}^6 \alpha_j (d_{j,t} - d_{7,t}) + \beta' (\mathbf{H}_t - \mathbf{H}_t^*) + \kappa' \mathbf{A} \mathbf{O}_t + \varepsilon_t$$

という形式のモデルにより OLS を用いて推定される。カレンダー効果

$$1 + \frac{1}{\bar{m}} LY_t + \sum_{j=1}^6 \alpha_j \left( \frac{d_{j,t} - d_{7,t}}{\bar{m}} \right) + \beta' \left( \frac{\mathbf{H}_t - \mathbf{H}_t^*}{\bar{m}} \right)$$

は、曜日効果と休日効果の成分を得るために、

$$\left\{ 1 + \frac{1}{\bar{m}} LY_t + \sum_{j=1}^6 \alpha_j \left( \frac{d_{j,t} - d_{7,t}}{\bar{m}} \right) \right\} \left\{ 1 + \beta' \left( \frac{\mathbf{H}_t - \mathbf{H}_t^*}{\bar{m}} \right) \right\}$$

のように近似的に分解することができる。

対数加法的分解

もし x11 スペックで mode=logadd となっているならば、カレンダー効果は

$$m_t^* (\log I_t + 1) - m_t = \sum_{j=1}^6 \alpha_j (d_{j,t} - d_{7,t}) + \beta' (\mathbf{H}_t - \mathbf{H}_t^*) + \kappa' \mathbf{A} \mathbf{O}_t + \varepsilon_t. \quad (11)$$

という形式のモデルにより OLS を用いて推定される。これらは

$$\exp \left\{ -1 + \frac{m_t}{m_t^*} + \sum_{j=1}^6 \alpha_j \left( \frac{d_{j,t} - d_{7,t}}{m_t^*} \right) \right\} \exp \left\{ \beta' \left( \frac{\mathbf{H}_t - \mathbf{H}_t^*}{m_t^*} \right) \right\}$$

のようにして曜日効果と休日効果に完全に分解される。曜日効果に関するその他の二つの有用な形式

$$\begin{aligned} & \exp \left\{ -1 + \frac{m_t}{m_t^*} \right\} \exp \left\{ \sum_{j=1}^6 \alpha_j \left( \frac{d_{j,t} - d_{7,t}}{m_t^*} \right) \right\} \\ & \exp \left\{ -1 + \sum_{j=1}^6 (1 + \alpha_j) \frac{d_{j,t}}{m_t^*} \right\}, \end{aligned} \quad (12)$$

があり、閏年効果  $\exp \{-1 + m_t/m_t^*\} \simeq m_t/m_t^*$  と日にち別のウェイト  $(1 + \alpha_j)$  を強調する形になっている。

変数 `tdprior` が使用される場合

上記のモデルのいずれの係数に対しても、`b` 引数を適切に指定することによって、定数値を割り当てることができる。乗法的分解における (6) や対数加法的分解における (12) で表される曜日成分に対応した7日のウェイトがわかるような事前情報がある場合がしばしばある。“事前の”ウェイト  $1 + \alpha_j^{(p)}$ ,  $1 \leq j \leq 7$  が `tdprior` 引数によって割り当てられている場合には、その系列は、x11 スペックで mode=mult 引数が指定されているならば

$$\frac{\sum_{j=1}^7 (1 + \alpha_j^{(p)}) d_{j,t}}{m_t^*} = \frac{m_t}{m_t^*} \left\{ 1 + \sum_{j=1}^6 \alpha_j^{(p)} \left( \frac{d_{j,t} - d_{7,t}}{m_t} \right) \right\} \quad (13)$$

によって、また `x11` スペックで `mode=logadd` が指定されているならば

$$\exp \left\{ -1 + \frac{m_t}{m_t^*} \right\} \exp \left\{ \sum_{j=1}^6 \alpha_j^{(p)} \left( \frac{d_{j,t} - d_{7,t}}{m_t^*} \right) \right\} \quad (14)$$

によって事前調整が施される。`b` の代わりに `tdprior` を用いることの1つの利点は、事前調整済の系列の不規則成分に明確な曜日効果が残存しているかどうかを自動的にテストし、残存するあらゆる影響を除去する目的で調整成分を計算するために、`aictest` を使用することができるということである。しかしながら、(13) または (14) による事前調整が  $m_t/m_t^*$  と  $\exp \{-1 + m_t/m_t^*\}$  で表される閏年効果をそれぞれ除去するため、`variables` 引数で `td` が指定された場合には、残存する効果を推定するために `x11regression` で用いられる (4) と (12) のモデルは修正される必要がある。`mode=mult` の場合には、(4) の代わりに

$$m_t l_t - m_t = \sum_{j=1}^6 \alpha_j (d_{j,t} - d_{7,t}) + \beta' (\mathbf{H}_t - \mathbf{H}_t^*) + \kappa' \mathbf{A} \mathbf{O}_t + \varepsilon_t,$$

とし、また `mode=logadd` の場合には (12) の代わりに

$$m_t^* \log l_t = \sum_{j=1}^6 \alpha_j (d_{j,t} - d_{7,t}) + \beta' (\mathbf{H}_t - \mathbf{H}_t^*) + \kappa' \mathbf{A} \mathbf{O}_t + \varepsilon_t.$$

とする必要がある。はじめのモデルからはカレンダー効果の成分

$$1 + \sum_{j=1}^6 \alpha_j \left( \frac{d_{j,t} - d_{7,t}}{m_t} \right) + \beta' \left( \frac{\mathbf{H}_t - \mathbf{H}_t^*}{m_t} \right)$$

が得られ、(13) との乗算により合成されたカレンダー成分が得られる。この結果は

$$\frac{m_t}{m_t^*} \left\{ \sum_{j=1}^7 (1 + \alpha_j^{(p)} + \alpha_j) \frac{d_{j,t}}{m_t} + \beta' \left( \frac{\mathbf{H}_t - \mathbf{H}_t^*}{m_t} \right) \right\} \quad (15)$$

と近似される。二番目のモデルからは

$$\exp \left\{ \sum_{j=1}^6 \alpha_j \left( \frac{d_{j,t} - d_{7,t}}{m_t^*} \right) + \beta' \left( \frac{\mathbf{H}_t - \mathbf{H}_t^*}{m_t^*} \right) \right\}$$

というカレンダー成分が導かれ、(14) との乗算により合成された成分

$$\exp \left\{ -1 + \sum_{j=1}^6 (1 + \alpha_j^{(p)} + \alpha_j) \frac{d_{j,t}}{m_t^*} \right\} \quad (16)$$

が得られる。(15) と (16) の式は、係数  $\alpha_j$  に対する統計的に有意な  $t$  値が、事前のウェイト  $1 + \alpha_j^{(p)}$  に明らかに修正の必要があることを表すと解釈できることを示している。

これで不規則成分の回帰モデルとその成分に関する議論が完結する。

`x11regression` によって指定される回帰モデルは `x11` スペックでの B と C による反復計算の不規則変動の系列から推定される。スペック・ファイルは `arima` と `automdl` と `regression` 含

み、最初に `x11regression` を用いて効果を推定し、それを他のスペックで指定された予測、逆予測の拡張や推定に用いられるデータから除去する。Bateman-Mayes の復活祭効果の調整を呼び出した場合、この調整は `x11regression` によって推定された効果の事前調整を持つ系列から計算される。予測がなされた場合は、`X-12-ARIMA` は全ての予測期間で予め選択された回帰のためにデータを発生させる。ユーザー定義の回帰変数があるときは全ての予測期間においてデータ値が供給されなければならない。極限 52 ユーザー定義回帰係数に加えてモデルでは極限 80 回帰変数がある。(これらの極限は変わるが詳しくは 2.8 節を見よ。) 後者の極限は事前はずれ値探索から発生した回帰係数の数と事前定義されたユーザー定義の回帰係数との数の和である。予測期間を含まないユーザー定義の回帰変数の系列の長さの最大値は 600。(この極限は可変であるが詳しくは 2.8 節を見よ。) 曜日調整及び (あるいは) 休日調整は `Reg-ARIMA` か不規則変動モデルのどちらか一方から得なければならない。これらの効果は `regression` と `x11regression` の両方から推定されなければならない。 `noapply` オプションを調整に 1 つの要素しか用いていないことを保証するために使う必要がある。引数 `aicctest` の効果は引数 `variables` で名付けられた回帰変数を削除したり `variables` で指定されたモデルに回帰変数を加えてしまう可能性がある。 `aiccdiff` の正の値の効果はモデルにテストされた変数を含める `aicctest` 手法をより難しくする  $\Delta_{AIC}$  を `aiccdiff` と関連付けるものとする。それはデフォルトでは 0 である。  $AICC^{with}$  (と  $AICC^{without}$ ) を `aicctest` で指定した回帰変数 (を用いない) のモデルの AICC 値とする。 `variables` リストに名前がないときは以下の場合に回帰モデルに加えらる。

$$AICC^{with} + \Delta_{AIC} < AICC^{without}.$$

変数リストに名前があるときはこの不等式を満たすときに限り `Reg-ARIMA` モデルそのまま入れておく。2 番目の場合は `aicctest = (tdstock)` なら `tdstock[31]` に指定される月末在庫変数は新たに加えられた変数である。なぜなら `tdstock[w]` の中の  $w$  はデフォルトでは 31 となるからである。 `variables` で復活祭効果に変数になく `aicctest = (easter)` であればより多くの可能性がある。3 つの追加的なモデルが考えられる。つまりそれぞれ  $w = 1, 8, 15$  の回帰変数 `easter[w]` の `Reg-ARIMA` モデルを指定することで得られるモデルのことである。もっとも小さい AICC の復活祭の回帰変数のモデルは復活祭の回帰変数がないモデルより少なくとも  $\Delta_{AIC} = 0$  だけ AICC が小さいならばそれを使う。他の場合は復活祭の回帰変数を用いない。 `aicctest` と `variables` の両方で曜日効果の回帰変数があるときは指定される回帰変数は一致しなければならない。ストックの曜日効果のサンプル日と範囲変化 (`change-of-regime`) の回帰変数で指定されるデータは引数 `aicctest` に含めなければならない。例えば `variables=(tdstock[15] ao1995.jan)` ならば `aicctest` の入力 `tdstock` でなければならない。営業日 (`td` or `tdstock`) や休日もしくは特定のアメリカの休日 (`easter`, `thanks`, `labor`) では回帰変数のタイプは引数 `usertype` のユーザー回帰変数のなかで指定される。引数 `usertype` は `x11regression` で指定されるのと同様の変数からの回帰要素と組み合わされるユーザー定義の回帰変数から来る要素である。結果としての要素は引数 `noapply` にタイプ名が現れない限り `x11` スペックで決定される季節要素の調整計算のための系列とは別に調整される。 `x11regression` がスペック・ファイルの中でスペック `x11` なしに使われる場合には、回帰に使われる不規則要素はデフォルト指定の `x11{}` から得られる。ここで引数 `eastermeans` の 2 つの選択子ではかなり異なる休日効果を生じることになる。通常、休日効果と季節要素を組み合

わせたものに選択は少ししか影響しない。というのは季節要素が選択による違いを補償してくれるからである。eastermeans=yesのもとで季節性を除いた復活祭回帰変数を得るために月次平均はGregorian calendar, 1583-1982.の最初の400年間の復活祭のデータの頻度から決められる。頻度はMontes (1997)に与えられている;復活祭のデータをグレゴリオ暦で計算するアルゴリズムはDuffet-Smith (1981)に与えられている。季節性を持たない時系列では、このスペックの方法で推定された曜日と休日の調整はスペック x11においてtype=trendの設定することで得ることができる。引数b=( )が係数を固定するのに用いられた場合は、AICや他のモデル選択統計量が不適当になるかもしれない。詳しくはestimateを参照せよ。

## 例

以下に完全なスペック・ファイルの例を挙げる。

### 例 1

デフォルト・オプションの乗法季節調整(季節フィルターの長さを選択する移動平均比を使うプログラム)。月時系列は1976の1月に始まりフリー・フォーマットでカレント・ディレクトリーのファイルklaatu.datに保存される。曜日効果調整は不規則要素の回帰変数を使って行われる。

```
Series { File = "westus.dat"
         Start = 1976.1
       }
X11 { }
X11Regression { Variables = td
               }
```

### 例 2

例1と同様に不規則要素の回帰で曜日効果や復活祭効果の回帰変数を入れるかどうかはAICを基にした検定だけによって行われる。

```
Series { File = "westus.dat"
         Start = 1976.1
       }
X11 { }
X11Regression { Variables = td
                Aictest = (td easter)
               }
```

### 例 3

復活祭前後の期間のユーザー定義の回帰変数は曜日回帰変数に沿った不規則回帰変数の中に含める。AOはずれ値の識別は不規則回帰変数の手順の中で行われる。

```
series {
  file = "ukclothes.dat"
  start = 1985.Jan
}
x11 { }
x11regression{
  variables = td          critical = 4.0
  user = (easter1 easter2) file = "ukeaster.dat"
  usertype = holiday     start = 1980.Jan
}
```

#### 例 4

デフォルトの季節調整を行う。スペック `x11regression` の曜日回帰変数は初期の値に固定される；復活祭回帰変数は推定される。

```
series{
  format = '2R'
  title = 'MIDWEST ONE FAMILY Housing Starts'
  name = 'CMW1HS'
  file = 'cmw1hs.ori'
  span = (1964.01,1989.03)
}
x11{ }
x11regression{
  variables = (td easter[8])
  b = ( 0.4453f 0.8550f -0.3012f 0.2717f
        -0.1705f 0.0983f -0.0082)
}
```

#### 例 5

このスペック・ファイルから事前に曜日のウェイトが与えられる。曜日のウェイトは不規則回帰変数を組み合わせた曜日要素にこれらのウェイトと組み合わせて計算される。

```
series {
  file = "nzstarts.dat" start = 1980.Jan
}
x11 { }
x11regression{
  variables = td
  tdprior = (1.4 1.4 1.4 1.4 1.4 0.0 0.0)
}
```

#### 例 6

曜日効果の推定のために不規則要素の回帰変数を使う。(範囲の変更が1990の1月にあった.)。Bateman-Mayesの手法により復活祭の推定が行われる。これらの効果が最初に推定され、`x11` スペックで指定された季節調整が曜日や休日調整が行われた系列に施される。

```
series{
  title = 'Motor Home Sales'
  start = 1967.1
  span = (1972.1, )
  name = 'SB0562'
  file = 'C:\final.x12\TOB05601.TXT'
  format = '2L'
}
X11REGRESSION {
  variables = (td/1990.1/ labor[10] thank[10] )
}
x11{
  seasonalma = x11default
  signalim = (1.8 2.8)
  appendfcst = YES
  save = (D11 D16)
  x11easter = yes
}
```

### 例 7

予め定義された回帰効果は曜日変数と定数である。ユーザー定義の回帰変数は 1980,1985,1991 年でファイル `strikes.dat` 中にある。ARIMA 部分は  $(0,1,1)(0,1,1)_{12}$  である。モデルはスペックで指定されているので、年の予測はデフォルトでなされる。季節期間の 12 はデフォルトなので特に指定しない。ユーザー定義の回帰効果の事前調整ごの季節調整系列、X-11 の曜日 (何かで推定されたユーザー事前因子) と復活祭効果の系列調整と最終的な季節調整系列から休日効果を除去する。2つの季節調整タイトルが指定される。

```
series{ title = "Automobile Sales"
        file = "carsales.dat"
        start = 1975.Jan }
transform{ function = log }
regression{ variables = (const)
            user = (strike80 strike85 strike90)
                file = "strike.dat"
                format = "(3f12.0)" }
arima{ model = (0 1 1)(0 1 1)12 }
x11{ title = ("Car Sales in US"
             "Adjust for strikes in 80, 85, 90")
     save = seasonal appendfcst = yes
     x11easter = yes
     final = holiday      }
x11regression{ variables = td }
```



## 参考文献

---

- Abraham, B., and G.E.P. Box (1978), "Deterministic and Forecast-Adaptive Time-Dependent Models," *Applied Statistics*, 27, 120-130.
- Abraham, B., and J. Ledolter (1983), *Statistical Methods for Forecasting*, New York: John Wiley and Sons.
- Akaike, H. (1973), "Information Theory and an Extension of the Likelihood Principle," in the *Second International Symposium on Information Theory*, eds. B.N. Petrov and F. Czaki, Budapest: Akademia Kiado, 267-287.
- Baxter, M.A. (1994) "A Guide to Seasonal Adjustment of Monthly Data With X-11 (Third Edition)," Central Statistical Office, United Kingdom.
- Bell, W.R. (1983), "A Computer Program for Detecting Outliers in Time Series," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 634-639.
- Bell, W.R. (1987), "A Note on Overdifferencing and the Equivalence of Seasonal Time Series Models With Monthly Means and Models With (0,1,1)<sub>12</sub> Seasonal Parts When  $\Theta = 1$ ," *Journal of Business and Economic Statistics*, 5, 383-387.
- Bell, W.R. (1999), "An Overview of *regarima* Modeling," forthcoming Research Report, Statistical Research Division, U. S. Census Bureau.
- Bell, W.R., and S.C. Hillmer (1983), "Modeling Time Series with Calendar Variation," *Journal of the American Statistical Association*, 78, 526-534.
- Bobbitt, L. and Otto, M. C. (1990), "Effects of Forecasts on the Revisions of Seasonally Adjusted Values Using the X-11 Seasonal Adjustment Procedure," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 449-453.
- Box, G.E.P., and D.R. Cox (1964), "An Analysis of Transformations," *Journal of the Royal Statistical Society, B*, 26, 211-252.
- Box, G.E.P., and G.M. Jenkins (1976), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden Day.
- Box, G.E.P., and G.C. Tiao (1975), "Intervention Analysis with Applications to Economic and Environmental Problems," *Journal of the American Statistical Association*, 70, 70-79.

- Brockwell, P.J., and R.A. Davis (1991), *Time Series: Theory and Methods*, New York: Springer-Verlag.
- Burman, J.P., M.C. Otto, and W. R. Bell (1987), "An Iterative GLS Approach to Maximum Likelihood Estimation of Regression Models with ARIMA Errors," Research Report No. 87/34, U. S. Census Bureau.
- Chang, I., and G.C. Tiao (1983), "Estimation of Time Series Parameters in the Presence of Outliers," Technical Report No. 8, Statistics Research Center, University of Chicago.
- Chang, I., G.C. Tiao, and C. Chen (1988), "Estimation of Time Series Parameters in the Presence of Outliers," *Technometrics*, 30, 193-204.
- Chen, B., and D.F. Findley (1998) "Comparison of X-11 and *regarima* Easter Holiday Adjustments," Statistical Research Division, U. S. Census Bureau, manuscript in preparation.
- Cholette, P.A. (1978) "A Comparison and Assessment of Various Adjustment Methods of Sub-Annual Series to Yearly Benchmarks," Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.
- Cholette, P.A. (1979) "A Comparison of Various Trend Cycle Estimators," Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.
- Dagum, E.B. (1988) "X-11-ARIMA/88 Seasonal Adjustment Method - Foundations and Users Manual," Statistics Canada.
- Doherty, M. (1991) "Surrogate Henderson Filters in X-11," Technical Report, Statistics New Zealand, Wellington, New Zealand.
- Duffet-Smith, P (1981) "Practical Astronomy With Your Calculator, 2nd Edition," Cambridge University Press.
- Findley, D.F. (1985) "On the Unbiasedness Property of AIC for Exact or Approximating Linear Stochastic Time Series Models," *Journal of Time Series Analysis*, 6, 229-252.
- Findley, D.F. (1999) "Akaike's Information Criterion II" in the *Encyclopedia of Statistical Science, Update Volume 3*, eds. S. Kotz, C. B. Read, and D. L. Banks. New York: John Wiley and Sons, 2-6.
- Findley, D.F. and C.C. Hood (1999). "X-12-ARIMA and Its Application to Some Italian Indicator Series," to appear in *Seasonal Adjustment Procedures - Experiences and Perspectives*, Istituto Nazionale di Statistica (ISTAT), Rome, 231-251.
- Findley, D.F., B.C. Monsell, H.B. Shulman, and M.G. Pugh (1990), "Sliding Spans Diagnostics for Seasonal and Related Adjustments," *Journal of the American Statistical Association*, 85, 345-355.
- Findley, D.F., B.C. Monsell, W.R. Bell, M.C. Otto, B.C. Chen (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program," *Journal of Business and Eco-*

- nomics Statistics*, 16, 127-176 (with Discussion).
- Fuller, W.A. (1976), *Introduction to Statistical Time Series*, New York: John Wiley.
- Geweke, J. (1978), "Revision of Seasonally Adjusted Time Series," SSRI Report No. 7822, University of Wisconsin, Department of Statistics.
- Gomez, V., A. Maravall, and D. Peña (1999). "Missing Observations in ARIMA Models: Skipping Approach versus Additive Outlier Approach," *Journal of Econometrics*, 88, 341-363.
- Griswold, R.E., and M.T. Griswold (1997) *The ICON Programming Language, 3rd Edition*, San Jose: Peer-to-Peer Communications.
- Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel (1986), *Robust Statistics: The Approach Based on Influence Functions*, New York: John Wiley.
- Hannan, E.J., and B.G. Quinn (1979), "The Determination of the Order of an Autoregression," *Journal of the Royal Statistical Society*, B, 41, 190-195.
- Hillmer, S.C., and G.C. Tiao (1979), "Likelihood Function of Stationary Multiple Autoregressive Moving Average Models," *Journal of the American Statistical Association*, 74, 652-660.
- Hood, C.C. (1998a) "X-12-Graph: A SAS/GRAPH Program for X-12-ARIMA Output, User's Guide for X-12-Graph Interactive for PC/Windows, Version 1.0", U. S. Census Bureau, Washington, DC.
- Hood, C.C. (1998b) "X-12-Graph: A SAS/GRAPH Program for X-12-ARIMA Output, User's Guide for X-12-Graph Batch for PC/Windows, Version 1.0", U. S. Census Bureau, Washington, DC.
- Huot, G. (1975) "Quadratic Minimization Adjustment of Monthly or Quarterly Series to Annual Totals," Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.
- Hurvich, C.M., and C. Tsai (1989), "Regression and Time Series Model Selection in Small Samples," *Biometrika*, 76, 297-307.
- Leser, C.E.V. (1963), "Estimation of Quasi-Linear Trend and Seasonal Variation," *Journal of the American Statistical Association*, 58, 1033-1043.
- Ljung, G.M. (1993), "On Outlier Detection in Time Series," *Journal of the Royal Statistical Society*, B, 55, 559-567.
- Ljung, G.M., and G.E.P. Box (1978), "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65, 297-304.
- Ljung, G.M., and G.E.P. Box (1979), "The Likelihood Function of Stationary Autoregressive-Moving Average Models," *Biometrika*, 66, 265-270.
- Lothian, J. and M. Morry (1978), "A Test of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Program," Research Paper, Seasonal Adjustment and Time Series Staff,

Statistics Canada.

Monsell, B.C. (1989) "Supplement to Technical Paper No. 15 - The Uses and Features of X-11.2 and X-11Q.2," Statistical Research Division, U. S. Census Bureau.

Montes, M. J. (1997), "Frequency of the Date of Easter over one 400 year Gregorian Cycle" [Online]. Available: <http://www.smart.net/|mmontes/freq2.html> [1999, December 8].

Montes, M. J. (1998), "Calculation of the Ecclesiastical Calendar" [Online]. Available: <http://www.smart.net/|mmontes/ec-cal.html> [1999, December 8].

More, J.J., B.S. Garbow, and K.E. Hillstrom (1980), User Guide for MINPACK-1, Report ANL-80-74, Argonne National Laboratory, Argonne, Illinois.

Otto, M.C., and W.R. Bell (1990), "Two Issues in Time Series Outlier Detection Using Indicator Variables," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 182-187.

Otto, M.C., and W.R. Bell (1993), "Detecting Temporary Changes in Level in Time Series," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 170-174.

Pierce, D.A. (1971), "Least Squares Estimation in the Regression Model With Autoregressive-Moving Average Errors," *Biometrika*, 58, 299-312.

SAS Institute Inc. (1990), *SAS/GRAPH Software: Reference, Version 6, First Edition, Volume 1*, Cary, NC: SAS Institute.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *Annals of Statistics*, 6, 461-464.

Shiskin, J., A.H. Young, and J.C. Musgrave (1967) "The X-11 Variant of the Census Method II Seasonal Adjustment Program," Technical Paper No. 15, U.S. Department of Commerce, U. S. Census Bureau.

Soukup, R.J. and Findley, D. F. (2000) "Detection and Modeling of Trading Day Effects," to appear as an SRD Technical Report, U. S. Census Bureau.

Thomson, P., and T. Ozaki (1992) "Transformation and Seasonal Adjustment," submitted for publication.

Vandaele, W. (1983), *Applied Time Series and Box-Jenkins Models*, New York: Academic Press.

Wilson, G.T. (1983), "The Estimation for Time Series Models, Part I. Yet Another Algorithm for the Exact Likelihood of ARMA Models," Technical Report No. 2528, Mathematics Research Center, University of Wisconsin-Madison.

## 解題

---

この付録ではプログラム X-12-ARIMA を実際に利用する上で特に注意する必要があると思われる二つの基本的事柄について簡単に解説する。

### A. X-12-ARIMA(2002) とは？

米国センサス局においては 1980 年代末頃から同局時系列研究グループにより新たな季節調整法の開発が X-12 開発計画として進行していた。1996 年になり時系列研究グループはその頃から一般にも利用が広がりつつあったインターネットによりプログラムの公開を開始した。ここで公開と云う意味は実はセンサス局の公式なソフトウェアとして配布したということではなく、あくまで実験版 (Experimental Version) プログラムを公開したということであり、その形式は 2004 年 2 月の時点でも変化していないと思われる。1996 年にインターネットで無料配布を開始したのは正確には X-12-ARIMA( $\beta$ -Version) と呼ばれている版であり、 $\beta$ -Version と云う名前の通りその後たびたびプログラムの誤り (バグ) 等の修正が行なわれている。(その具体的な修正箇所・改善箇所についての情報はセンサス局の X-12-ARIMA についてのインターネット URL 上にファイル Corrections(訂正箇所) として公開されている。)

その後、X-12-ARIMA プログラムは 1998 年になり X-12-ARIMA として利用可能になったが、ここではこのプログラムを X-12-ARIMA(1998) と呼んでおくことにする。このときに  $\beta$ -Version と云う呼び名がなくなったことからプログラム X-12-ARIMA の基本的な部分については計算プログラム上の問題はほぼなくなったと理解してよいと判断されよう。2000 年 12 月時点で利用可能な版は 2000 年 5 月に修正した Version 0.2.7 であるのでここではこれを X-12-ARIMA(2000) と呼ぶことにするが、この解説稿のもとになっているマニュアルはこのプログラムに対応して 2000 年 5 月に作られた版のマニュアルである。なお、2002 年に X-12-ARIMA の最終版が確定したようであるので、ここでは X-12-ARIMA(2002) と呼ぶことにする。

本解説の第 1 章にもあるようにプログラム X-12-ARIMA を動かすにはまず二通りのやり方のどちらかで最新版の計算プログラムと関係ファイルをセンサス局の URL からダウンロードする必要がある。我々は最初に 2000 年 11 月頃にインターネットを利用して卓上計算機に転送したファイルをセンサス局のホームページに書かれている指示にしたがいプログラムを解凍し利用したが、その後、必要が生じたため 2002 年 11 月頃に最新版を再びダウンロードして利用した経験がある。本解説のもととなる英文で書かれた原稿はこの X-12-ARIMA(2000) のマニュアル原稿を直接に書いたセンサス局の X-12-ARIMA 開発の責任者 (デービッド・フィンドレー博士 Dr. David Findley)

から 2001 年 11 月頃にインターネットにより送られたものである。

なお、センサス局 X-12-ARIMA( $\beta$ -Version) に関する注意点や疑問点、あるいは季節調整法を巡る論争点などについては国友直人(1997)「季節調整法 X-12-ARIMA の特長と問題点」(東京大学経済学部 Discussion Paper No. J-97-1, 誤植訂正版 (<http://www.e.u-tokyo.ac.jp> 中の国友の URL からダウンロード可能)), 統計数理(1997)「季節調整法特集」(統計数理研究所)などに掲載された論文や討論などを参考にされたい。X-12-ARIMA プログラムの開発当事者による技術的部分の説明としては引用文献中の Findley et.al. (1998) が比較的詳しいので参考になろう。ただし、この論文の最終版は同名の論文(1996)の改訂版であるが、X-12-ARIMA プログラムに関する元々の技術的説明も少し変化していることに注意しておく。日本の官庁において実際に X-12-ARIMA プログラムを利用する場合の注意点については、国友直人「季節調整法 X-12-ARIMA(2000)の利用：法人企業統計の事例」(経済学論集 第 6 7 巻第 3 号 2 項～2 9 項 2 0 0 1 年 1 0 月、東京大学経済学部) が一つの方法を提案している。

## B. 季節調整法小史

経済時系列における季節的変動の分析は 19 世紀まで遡ることができる。スタンレー・ジェボンズを始めとする何人かの著名な経済学者が経済変数の変動において季節性がかなり大きな役割を演じていることに気がつき、主として記述統計的な方法で議論していた。

近代的な統計学の成立とともに時系列データの分析方法も発展を遂げていく中で統計的時系列分析 (statistical time series analysis) と呼ばれる分野も成立し発展してきたが、その中では季節変動を巡る問題は常に重要な意味を持ち続けて来たようである。特に戦間期から 1950 年代頃にかけて統計的時系列分析では記述統計的手法やスペクトル分析 (spectral analysis) におけるピリオドグラム (periodogram) 解析などの手法が開発されたのでその応用などが試みられている。

こうした中で特に官庁における季節調整法に関して注目すべきこととしては NBER (National Bureau of Economic Research) で経済時系列を研究していたジュリアス・シスキン (Julius Shiskin) が米国政府統計局 (センサス局) に研究者として招かれ、当時に利用可能であった統計的分析手法を使って季節調整プログラムの開発に着手した。当時に利用可能であった季節性調整法としては連環比率法や B L S (Bureau of Labor Statistics) 調整法などがあったが、シスキンは当時によりやく利用可能となりつつあった電子計算機を利用して汎用となる季節調整プログラムの開発に乗り出したわけである。こうした中で 1954 年当時にセンサス局の季節調整プログラムとして初めて実用化されたプログラムはセンサス局 I 法と呼ばれている。この方法では季節調整の基本的な方法としてよく知られるようになった移動平均法 (Moving average method) を利用して季節性を取り除き、原系列から季節調整値を計算しようとするものであった。1950 年代から 1960 年代にかけて国際的にも官庁統計では季節調整法を開発しようとする機運が高まり、ヨーロッパの当局やわが国の当局によっても幾つかの季節調整プログラムが開発されている。わが国では当時の通商産業省が MITI 法と呼ばれる季節調整プログラムを開発したのにつき、当時の経済企画庁も E P A 法という季節調整法を開発している。これら二つの季節調整プログラムはセンサス局 I 法と同様に移動平均法を中心のアイデアとしつつも、独自に様々な工夫を施した方法であった。なお、この E P A 法は 1978 年に運用を中止し、M I T I 法は改良された M I T I 法-III として 1998 年ま

で運用を続けていた。ヨーロッパでもドイツのブンデス・バンク法や英国中央銀行の季節調整法などが当時から開発されていたが、近年ではヨーロッパ統計局を中心に研究が行なわれている。

一方、米国センサス局ではセンサス I 法を開発した後もシスキンを中心として精力的に改良を重ね、次々に新しく改良された方法を研究し開発を続けていた。これらの一連の方法は実験用 (Experimental methods) という意味でセンサス局 II 法の X シリーズという名前になった。数々の試行錯誤を重ねプログラム修正を重ねた末に 1965 年 (解説マニュアルは内部資料 Shiskin et al. (1967) として発行) にセンサス局 X-11 法と呼ばれる季節調整法が開発された。このプログラムは当時、一般に利用可能となってきたメイン・フレームの電子計算機の利用を想定したものでフォートラン (FORTRAN) 言語でプログラムが書かれたが、そのプログラムを関係者が配布を受けて利用できる形をとった。こうした中で 1970 年ごろの時期からわが国では日本銀行をはじめ季節性を持つ経済データを公表している経済関係の中央官庁がこのセンサス局法 X-11 法を採用して季節調整系列を公表することが多くなった。ただし、センサス局 X-11 法はその内部にかなり複雑なオプションを含んでいるので X-11 法を採用した各官庁がどの様に季節調整法を利用しているかについては当時も完全に一般に公開されていたわけではない。

さて、センサス局では季節調整法としては X-11 法を開発してからしばらくの間は特に重要な研究・開発はされなかったようである。他方、統計的時系列分析の分野では 1970 年代にはボックス・ジェンキンスが提唱した ARMA (自己回帰移動平均) モデルにもとづく予測の方法が特にその実用性の観点からかなりの注目を浴びるようになっていた。こうした機運の中でカナダ・センサス局の E. Dagun を中心とするグループはセンサス局 X-11-ARIMA と呼ばれる季節調整法プログラムを 1975 年頃に開発した。このプログラムが X-11 法と異なる主要な機能は ARIMA モデルを用いた予測系列を利用する事で季節調整の際に生じる末端処理の問題を改善しようとしたことにある。この X-11-ARIMA 法はカナダ統計局ではかなり実用的にも用いられていたが、米国や日本を始めとする他の国々の官庁統計ではそれほど用いられなかったようである。

時代が 1980 年代半ばになるとよく知られているように計算機を取り巻く環境が劇的に変化し始めて来た。こうした中で米国センサス局においても 1980 年代末頃から再び季節調整法を検討しようとする機運が高まって来た。より具体的にはセンサス局の研究セクションの中に時系列解析研究グループがデビット・フィンドレー (David Findley) を中心に形成され、X-12 開発の計画が具体化しはじめていった。時系列研究グループが開発した経緯を文献から判断すると従来から存在している X-11 法をそれに対してかけられてきた幾つかの批判をその枠内で解決しようとしたと見ることができよう。基本的にはカナダ・センサス局が開発した X-11-ARIMA を基礎として、それに更に新たに幾つかの機能を付け加えることが研究開発の内容である。より具体的には本解説の第 1 章～第 5 章に述べたように従来の季節調整法 X-11 法に様々な機能を付け加えることにより利用者が様々なオプションを比較的簡単な操作で実行する事ができるようにしたことが大きな改訂点である。また、それとともに 1990 年代になって実現し始めたインターネットによる不特定多数のユーザーに対するプログラム配布という形で公開したが、この方法は官庁統計のあり方に対する一つの画期的な新しい出来事であろう。

ここで季節調整法についてはセンサス局法以外にもこれまでに様々な方法が開発されていることにも最後に言及しておこう。特に 1970 年代末にシカゴ大学の A. Zellner 教授の季節調整法比較プロジェクトがあり、その中で報告し注目を浴びたわが国の統計数理研究所の赤池弘次 (X-12-ARIMA

マニュアルに登場する H.Akaike と同一の研究者)・石黒真木夫が開発したプログラム BAYSEA が専門家の間ではよく知られている。この季節調整法はセンサス局法が基礎とする移動平均法とは異なり、時系列において季節性に関する滑らかさの事前情報を直接的に活用すると云う統計学的な新しいアイデアにより季節性を取り除くユニークな方法である。この季節調整プログラム BAYSEA はその後同研究所の北川源四郎により改良され、1987年に状態空間モデルに基づく季節調整プログラム DECOMP が研究・開発されている(北川源四郎(1993)「時系列プログラミング」(岩波書店)を参照)。さらに、同研究所の佐藤整尚により1998年に Web-decomp が開発され、インターネットの Web 上 (<http://ssnt.ism.ac.jp/inets2/title.html>) で具体的な統計計算を容易に実行できるという新しい方法により公開されている。