

WHAT' S NEW IN SYSTAT 13

SYSTAT13の統計解析機能

新機能

時系列解析におけるARCH & GARCH モデル

従来の時系列分析処理は、任意の（誤差）期間の分散は時間経過にかかわらず一定であると仮定しています。しかし、実際には、特に金融の領域では、ある系列が期間によって異なるレベルの分散を持つ変動率を示すことがあります。この現象を取り込み、モデル化するために、ARCHモデル（自己回帰条件付き分散不均一モデル：Autoregressive Conditional Heteroskedasticity Model）が開発されました。このモデルでは、系列の過去の攪乱を使って、系列の各点での分散をモデル化します。

ARCHモデルは、誤差分散のダイナミクスを正しく取り込むために大量のパラメータを必要とします。GARCHモデル（一般化された自己回帰条件付き分散不均一モデル：Generalized Autoregressive Conditional Heteroskedasticity）は追加の自己解析期間の誤差分散を導入し、モデリング時にパラメータのparsimony（けちの原則）を適用することを助けます。

SYSTAT 13の時系列分析は次の機能が追加されています：

- ・ ARCH効果用の仮説検定： よく知られたマクラウドとラグランジュの乗数検定を提供します。
- ・ 収束基準のさまざまな設定をすることができる最尤法の実装ごと（BHHH、BFGS、Newton - Raphson）のARCHとGARCHモデルのパラメータの推定
- ・ パラメータ推定を用いた誤差分散の予測
- ・ 誤差の正規性のJarque - Bera検定

ベストサブセット回帰

重（線形）回帰モデルの作成に於いて、開発されたモデル内の予測変数の数を予測力を犠牲にすることなく小さくできるのであれば素晴らしいことです。ベストサブセット回帰がこれに対処します。

- ・ SYSTAT 13は、データセット内の1から最大数までの範囲の指定された数の予測変数に対する最適モデル（予測変数の選択）を検索します。



- ・ 最適モデルはR二乗、調整済みR二乗、Mallowの C_p 、MSE、AIC、AICC、BIC のようなさまざまな基準によって特定されます。
- ・ 次にSYSTAT 13は、ユーザが選択したデータセット（トレーニングセットと同様か、あるいは異なるもの）に対して、任意の基準で選択された最適モデルで完全な回帰分析を実行します。

確証的因子分析

SYSTAT 13の因子分析機能に確証的因子分析（CFA）が追加されました。

- ・ CFAは、観測変数（顕在変数）と潜在変数の関係についての先験的な知識（priori knowledge）に基づいた仮説因子構造を検証するために使うことができます。
- ・ CFAを使うと、SYSTATでユーザが観測変数、潜在変数の集合、それらの分散共分散構造を指定することができます。
- ・ パス図は仮説モデルを指定するために使うことができます。
- ・ SYSTAT 13は、次の推測オプションの1つを使ってCFAモデルのパラメータを推測します：最尤法、一般化最小二乗法、重み付き最小二乗法。
- ・ SYSTAT 13は、データに対する仮説因子モデルの適合の程度を測定するためのさまざまな適合度指標を提供しています。提供されている指標の中でよく使われるものとしては、適合度指標（GIF）、二乗平均平方根残差（RMR）、儉約性適合度指標（PGFI）、AIC、BIC、マクドナルドの確実性尺度、非標準適合度指標（NNFI）があります。

基本統計における環境変数

SYSTAT 13の基本統計モジュールに環境変数が新たに加わりました。

- ・ 環境変数は特定のセッション、特定のデータファイル、特定の変数で使うことができるさまざまな統計の計算済みの値を保存している変数です。
- ・ これらは統計計算をさらに続けて処理するときに、その返還命令で直接利用することができます。

多項式回帰

SYSTAT 13は独立した1つの変数における多項式回帰を直接計算することができます。主な機能は：

- ・ 8次までの多項式
- ・ 標準のフィッティング多項式に加え、直交多項式回帰を提供
- ・ 適合度統計（R二乗と調整済みR二乗）とすべてのモデルにおいてp値を持つANOVAをレポートします。ユーザが指定した次数から始め、リニア（次数=1）に下がっていきます
- ・ QuickGraphのような、推測とともに信頼区間と予測区間のプロット、残差と予測値のプロットを提供

機能拡張

分散分析

分散分析に新しく追加された機能：

- ・ 中央値に基づいて分散の均一性を検定する Levene の検定
- ・ 解析に必要な要因だけを分類する SUBCAT コマンド

基本統計

基本統計では次の機能が拡張されました：

- ・ 刈り込み平均のための標準誤差と信頼区間
- ・ ウィンザライズド平均とその標準誤差と信頼区間
- ・ 標本のモードとデータのカーネル密度推定に基づいたモード
- ・ 四分位範囲

これらの統計はすべて、行だけでなく、列に対しても行うことができます。

ブートストラップ分析

SYSTATは、ほとんどの統計解析機能において、各標本とそれに対して生成された完全な出力のそれぞれにおいて、同じ分析を実行することで、ブートストラップを含む再サンプリング機能を提供してきました。これまでに、ブートストラップ出力の分析、ヒストグラムによる出力の主要な部分の抽出、ブートストラップ推定のさまざまな形式の計算とそのバイアス、標準誤差、信頼区間、p-値を提供してきました。SYSTAT 13ではこの機能が仮説検定と最小二乗回帰に追加されました。

- ・仮説検定機能にブートストラップベースの p-値とブートストラップ標本から得られた検定統計のヒストグラムが新たに加われました。通常の検定の p-値とは別に、平均（1 標本 z-検定、1 標本 t-検定、2 標本 z-検定、2 標本 t-検定、対応のある t-検定、ポワソン）、および分散（1 標本、2 標本および複数標本）についての検定すべてで利用することができます。
- ・最小二乗回帰でブートストラップを実行する場合、残差を選択できるようになりました。回帰係数、バイアス、標準誤差および信頼区間のブートストラップ推定は、これらをもとに計算されます。

クロス集計

クロス集計 (XTAB) 機能に次のものが追加されました：

- ・ 相対危険度：相対危険度は 2x2 分割表において、行または列に定義された 2 群の結果が正（ポジティブ）であるケースの比率です。相対危険度は 2 値変数の関連性を評価するのに広く利用されます。
- ・ モード：一元表（度数分布）のはじめの N カテゴリのみをリストとして表示させるオプションが提供されます。XTAB 内の PLENGTH コマンドに MODE=N オプションを追加することでこれを実行できます。
- ・ 結果の保存：
 - Multiway: Standardize で指定されたすべての列を保存
 - 保存された結果ファイルの列に対応する入力変数の値ラベル
- ・ 分割表タイプにもとづいて適切に分類され、統計量が再編成されたアウトプット。

フィッティング分布

SYSTAT 13 では、最尤法を使用して以下の分布パラメータを評価できます：

- ・ ベータ分布
- ・ カイ二乗分布
- ・ アーラン分布
- ・ ガンマ分布
- ・ ゴンペルツ分布
- ・ ガンベル分布
- ・ ロジスティック分布
- ・ 対数ロジスティック分布
- ・ 負の二項分布

- ・ ワイブル分布
- ・ Zipf分布

仮説検定

- ・ 仮説検定では多変量データの平均ベクトルについての検定を追加し、機能が強化されました。
 - － ホテルングの1標本 T2 検定。多変量データの平均ベクトルが既知のベクトルと等しいかどうかを検定。
 - － ホテルングの2標本 T2 検定。2つの多変量データの平均ベクトルが等しいかどうかを検定。
- ・ 2標本 z-検定、2標本 t-検定、2分散の検定において、2つの標本データが異なる列に配置されている入力データをレイアウトに直接入力することができます。現行のインデックス型配置か、このオプションのいずれかを選択可能。

最小二乗回帰

最小二乗回帰から得られる標準誤差と信頼区間を保存する機能が追加されました。

ロジスティック回帰

2値モデル、多項式モデル、条件付きモデル、離散選択モデルを従来よりも直感的な方法で解析できるようになりました。特に：

- ・ 2値、多項式、条件付き、離散選択モデルそれぞれについてユーザーインターフェースとコマンドライン構造が簡素化され、操作性が向上しました。
- ・ 2値および多項式応答モデルにおいて基準レベルを設定できるオプションを追加。
- ・ セット毎に1つのケースと任意数のコントロールをもつ標本数が同じケース-コントロール調査を分析するための容易なデータ入力。
- ・ 離散選択モデルにおける2種類のデータ配置の入力：「Choice set」と「By choice」。選択の特性に応じて個人の選択をモデル化。
 - － 「Choice set」では、変数をグループ化して名前を付けた選択集合 (choice set) を定義できます。変数は作成、編集、削除することができます。
 - － 「By choice」フレームワークにおいては、あらかじめ定義された選択集合 (choice set) を分析のデータに利用できます。

ノンパラメトリック検定



ノンパラメトリック検定に以下の機能が追加されました：

- ・ Mann - Whitney 検定の代替としてのJonckheere - Terpstra検定：処理群が応答に関して順序付けされた時に使われます。この検定は $k(k - 1)/2$ Mann - Whitney カウント (k処理群) の合計に基づいています。
- ・ Mann - Whitney 検定の代替としてのFligner - Wolfe検定：片側検定を使って他のすべての処理群と比較して応答が等価であることを検定するために、処理群の1つが対象として振る舞うときに使われます。これは対象が1群で、他のすべての処理群が他の群を形成しているような2つの群のMann - Whitney 検定です。
- ・ Dwass - Steel - Critchlow - FlignerとConover - InmanのPost hoc (多重比較) 検定は、Kruskal - Wallis 検定が意味深いことを示したときのフォローとして使うことができます。
 - Dwass - Steel - Critchlow - Fligner 検定は漸近のステューデント化された範囲の分布に基づいています。
 - Conover - Inman 検定は漸近のステューデント分布 (t分布) に基づいています。
- ・ Friedman検定が有意差を示したときのフォローとしてのConoverによって多重比較検定を使うことができます。

SYSTAT 13の新しいグラフィックス機能

- ・RGB (赤・緑・青) の各色成分の値によって色指定できるようになりました。グラフ要素、軸、枠の色指定に利用可能。
- ・ダイアログボックスによるサーフェス・グラデーションの指定。
- ・ドット (サマリー) チャートにおけるラベルドット
- ・組み込み色が前バージョンの12色から45色に
- ・隣接する要素間のコントラストが高くなるように重ね合わせグラフ、円グラフ、積み上げ棒グラフの色指定が可能
- ・グループ化変数を考慮する積み上げ棒グラフ。積み上げられたグラフは、グループ毎に描画され、すべてのグラフは同一の枠内に配置されます
- ・プロットに表示されるボーダーの個別指定。二次元プロットにおける任意のボーダーについて、それを非表示にしたり、2つのボーダーにそれぞれ異なる表示法を指定することができます
- ・円グラフの複数スライス表示
- ・プロット、多変量表示およびマップ内のラベル設定で数値変数を指定できるようになりました。

他の新機能

- ・ ワークスペースとコマンドスペースを自動的に非表示にします。
- ・ トークン用ダイアログボックスを使用してコマンドスクリプトにリンクした選択トークンを指定できるようになりました。
- ・ Data Editor に Data Edit Bar というツールバーが新たに組み込まれ、データ値を表示／編集できるようになりました。
- ・ マウスでクリックするだけでデータファイル情報にアクセスできるFile Comments ボタン
- ・ コマンドファイルの保存形式を選択できるようになりました。
- ・ コンテンツのドラッグをサポートしている外部エディタから、SYSTAT の Data Editor にテキストをドラッグ&ドロップできるようになりました。
- ・ Format Bar、Data Edit Bar、Graph Editing toolbar がViewspaceの各タブに埋め込まれるようになりました。
- ・ ファイルの先頭にVERSION コマンドを挿入すればレガシーなコマンドファイルを開いて実行することができます。
- ・ Viewツールバーを View メニューから呼び出すことができるようになりました。
- ・ SYSTAT の Data/Variable Editor (Data->Transformでダイアログボックスが開き、DataでCasesダイアログボックスを選択します) がWindows XPスタイルグリッドに対応しました。
- ・ 文字列データにおける前後の空白を削除できるようになりました。
- ・ Data Editor のコンテキストメニューまたは CLOSE コマンドを使用して、データファイルを閉じることができるようになりました。
- ・ 数値の出力形式に関わらず、Data Editor で新規に追加する数値変数のデフォルト形式を設定できるようになりました。
- ・ データファイルを View モードで保存できるようになりました。
- ・ データソースに「Business Objects Universe」を利用し、ビジネス・オブジェクトをインポートできるようになりました。
- ・ Output エディタで数字を表示させる場合、使用するロケール（地域別定義）を選択できるようになりました。
- ・ NODE コマンドを使用して Output Organizer のノードと折り畳み型リンクの見出しを設定できるようになりました。
- ・ ACTIVE コマンドでビューモードのファイルをアクティブにできるようになりました。



- ・ 新しい組み込み関数が追加されました。
- ・ Data Editor、Graph Editor、Output Editor の表示を切り替える FOCUS コマンドが新たに加わりました。
- ・ コマンドスクリプトでマクロを定義し、呼び出すことができるようになりました。
- ・ コマンドの引数、オプション、オプション値が自動的に補完されます。
- ・ コマンドスペースにおいてコマンドキーワードの色分けはデフォルトを指定するオプションで設定できるようになりました。
- ・ ダイアログボックスのタブが縦方向の配置になり、マウスをワンクリックするだけでアクセスできるようになりました。
- ・ Send Reportダイアログボックスを使ってセッションの復旧を試みる機能が追加されました。
- ・ ショートカットキーが追加されました。
- ・ ステータスバーから Insert、Caps Lock、Num Lock、Scroll Lock などキーボードの状態を切り替えることができるようになりました。
- ・ インストールしたいテーマを選択するDownload Themes ダイアログボックスが追加されました。
- ・ テーマファイルにバージョンが付き、SYSTAT のサーバーに新しいバージョンのテーマがアップされたらいつでもそれを更新することができます。
- ・ テーマを適用する際は、現在使用中のテーマを保存するかどうかのダイアログが表示されます。
- ・ Data/Variable Editor内のセルをコピーし、列にコピーすることができます。
- ・ 複数のデータファイルを処理を行った順や入力データファイルに基づいて開くことができるようになりました。
- ・ ELSEオプションは、与えられた値集合以外の値を、指定された値に再コード化することのできるオプションです。
- ・ アクティブなデータファイル、値ラベルの表示形式、変数ラベルの表示形式に対する設定の保存と回復ができるようになりました。
- ・ PAGE NONE コマンドを使用することでページ幅を無制限に設定できるようになりました。