

分類木 (パーティション)

パーティションに基づいたモデル(決定木)を構築して、質的変数の応答を予測(分類)するのに最も重要な因子を特定します。作成した木は新しいオブザベーションの予測に用いられます。

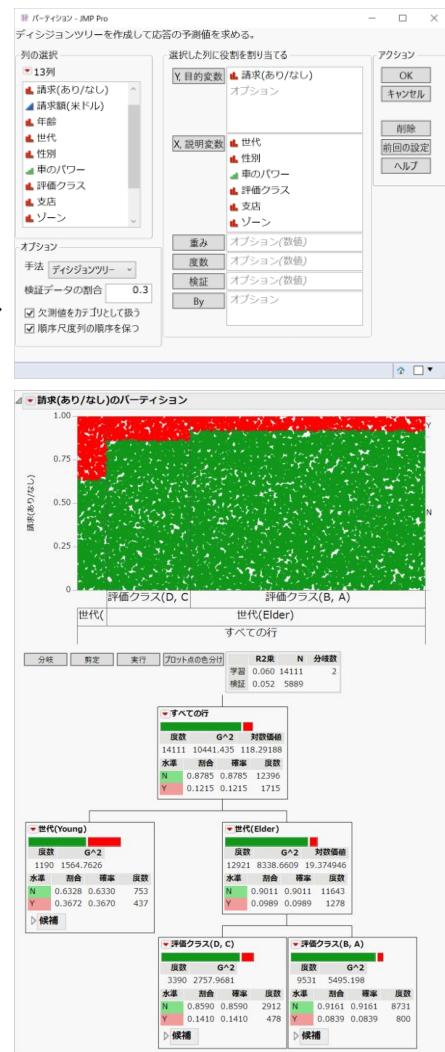
分類木

- データテーブルで分析 > 予測モデル > パーティションを選択します。
- 列の選択から名義尺度または順序尺度の応答変数を選択し、Y, 目的変数を選択します。
- 説明変数を選択し、X, 説明変数をクリックします。
- 必要な場合、検証データの割合を入力するか、検証用の列を選択し、検証(JMP® Proのみ)をクリックします。この例ではランダムに選択された30%の検証セットを用います。つまり70%が学習用です。
- JMP Proでは手法を選択可能です: ディシジョンツリー(デフォルト)、ブートストラップ森、ブースティングツリー、K近傍法、単純Bayes。
- OKをクリックします。応答変数の各水準の観測値の割合を示すグラフが表示されます。
- 分岐ボタンをクリックします。元の観測値が2つのノード(葉)に分岐されます。グラフは分岐を反映するように更新され、分岐の詳細を記述した樹形図が作成されます。

注意:一番上の赤い三角ボタンをクリックし、表示オプション > 割合を表示を選択すると、それぞれの葉の応答の水準に対する割合(観測値の割合)と確率(予測確率)を表示できます。

- さらに分岐をするには分岐をクリックします。分岐を除去するには剪定をクリックします。検証データを使用した場合、実行をクリックすると自動分岐を実施できます。

例: Auto Raw Data.jmp (ヘルプ > サンプルデータ)



最初の2つの分岐の解釈(この例では応答変数は、請求(あり/なし)):

- 左の葉の1,190例は、世代(Young)に対応しています。753例(63.3%)が「請求なし」で、437例(36.7%)が「請求あり」です。
- 右の葉の12,921例は、世代(Elder)に対応しています。11,643例(90.1%)が「請求なし」で、1,278例(9.9%)が「請求あり」です。
- 世代(Elder)の12,921例に対する2つめの分岐では、評価クラス変数に応じて、評価クラス(D, C)と評価クラス(B, A)に分岐します。

注意:

列の寄与、ROC曲線、リフトチャート等の追加のオプションを使うには一番上の赤い三角ボタンをクリックします。他の予測式の保存やSAS® DATAステップの作成等のオプションは、一番上の赤い三角ボタン > 列の保存から実行します。特定のノードでの分岐のオプションは、ノード上の赤い三角ボタンをクリックします。

検証、ブートストラップ森、ブースティングツリーを含む分類木のあてはめや評価の追加情報に関しては、JMPのヘルプか予測モデルおよび発展的なモデル(ヘルプ > JMPドキュメンテーションライブラリ以下)で「パーティション」と検索してご確認ください。