

単純 Bayes

JMP PRO この予測モデルの手法を使い、複数の予測変数の関数として、質的変数の応答を予測(分類)することができます。この手法はベイズの定理を条件付き確率に応用してオブザベーションを分類します。

単純 Bayes

- データテーブルから、分析 > 予測モデル > 単純 Bayes 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。
- 候補予測変数を選択し、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、右図のように検証列を選択して検証をクリックします。
- OK をクリックすると、JMP は以下の内容を表示します:

- 誤分類されたオブザベーションの度数と割合。
- 分類性能の詳細を示した混同行列。
- ROC 曲線と AUC(右図には検証データの結果のみを表示)。

注意: 一番上の赤い三角ボタンをクリックし、プロファイルを選択すると、予測変数の水準の関数として各クラスの予測確率を示した対話的なグラフを表示できます。他のオプション、例えば変数重要性の評価は予測プロファイルの横の赤い三角ボタンから選択します。右下の図は 3 つの予測変数 HDL、BMI 及び LTG のプロファイルです。

442 人の糖尿病患者から水準(High/Low)を予測した単純 Bayes 分類器の結果:

- 学習データは 309 例。66 例(21%)が誤分類された。 $41/(185+41) = 18\%$ の Low のオブザベーションが High と誤分類された。また $25/(25+58) = 30\%$ の High のオブザベーションが Low と誤分類された。
- 検証データは 133 例。32 例(24%)が誤分類された。 $25/(70+25) = 26\%$ の Low のオブザベーションが High と誤分類された。また $7/(7+31) = 18\%$ の High のオブザベーションが Low と誤分類された。

注意:

リフトチャート、予測値の保存、予測式の保存、確率の計算式の保存、確率の計算式の発行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。

単純 Bayes 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「単純 Bayes」と検索してご確認ください。

例: Diabetes.jmp (ヘルプ > サンプルデータライブラリ)



Y 2値			検証		
学習			検証		
度数	誤分類率	誤分類	度数	誤分類率	誤分類
309	0.21359	66	133	0.24060	32
混同行列					
学習			検証		
実測値	予測値	度数	実測値	予測値	度数
Y 2値	Low High		Y 2値	Low High	
Low	185 41		Low	70 25	
High	25 58		High	7 31	
実測値 予測値 割合			実測値 予測値 割合		
Y 2値	Low High		Y 2値	Low High	
Low	0.819 0.181		Low	0.737 0.263	
High	0.301 0.699		High	0.184 0.816	

