

モデルの比較と選択

JMP PRO このプラットフォームでは、データにあてはめた複数のモデルの性能を要約して比較することができます。異なる統計モデルのあてはめ方の詳細については、該当する資料を確認してください。

モデルの比較 – 応答が連続変数の場合

例: サンプルデータの Body Fat.jmp を用いて、体脂肪率を予測します。データテーブルに保存されている様々なモデルの計算式が列パネルの **Prediction Formulas** 内にまとめられています。これらのモデルは検証データにおけるモデルの性能に基づいて選択されたものです。

1. データテーブルから、分析 > 予測モデル > モデルの比較を選択します。
2. 列の選択から予測式の列を選択し、Y, 予測子をクリックします。列が選択されていないときは、保存されている全ての予測式が用いられます。
3. 列の選択から検証列を選択して、By をクリックします。それから OK をクリックします。

異なるモデルの性能を比較する統計量が計算されます。(右図では検証列の結果を表示しています。):

- R2 乗(高いほど良い)
- RASE: 誤差の平均平方根(誤差の大きさの尺度)
- AAE: 平均絶対誤差(誤差の大きさの尺度)

モデルの比較 – 応答が質的変数の場合

例: サンプルデータの Equity.jmp を用いて、予測変数の関数として Bad Risk となる確率を予測します。この例ではデータが学習用、検証用、テスト用に分割されています。

検証列を指定して予測モデルを作成し、上記の 1 から 3 のステップに従ってモデル検証用の統計量を計算します。

- 誤分類率などの性能指標が表示されます。
- ROC 曲線、リフトチャート、混同行列及びその他のオプションを赤い三角ボタンから選択できます。
- プロファイラーを用いると、各モデルにおいて各因子が応答に与える影響を詳細に図示し、様々な因子の設定の下で予測値を比較することができます。

注意: 追加のオプションを赤い三角ボタンから選択できます。モデルの比較は計算式デポからも利用できます。モデルの比較と計算式デポの詳細については、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「モデルの比較」と検索してご確認ください。

作成方法	R2乗	RASE	AAE	度数
Fit Least Squares	0.7156	4.6405	3.8349	72
Fit Least Squares - 2次交差作用スティップライズ	0.5209	6.0233	4.7128	72
Fit Least Squares - リックの主成分スコアによる変換	0.7358	4.4725	3.7222	72
Fit Least Squares - スティップライズ	0.6435	5.1956	4.1678	72
Fit Least Squares - パーティション	0.6984	4.7770	4.0219	72
Fit Least Squares - PLS	0.6594	5.2694	4.0707	72
Fit Least Squares - パーティション	0.6034	5.4804	4.4486	72
Fit Least Squares - ブースティングツリー	0.6880	4.8609	3.7132	72
Fit Least Squares - ニューラル	0.6186	5.3742	4.1153	72
Fit Least Squares - ニューラル (0.0, 8.0, 0.0)	0.7275	4.5422	3.4976	72
Fit Least Squares - ブースティングツリー	0.7063	4.7157	3.7557	72

作成方法	エントロピーR2乗	一般化R2乗	平均 -Log p	RMSE	平均	絶対偏差	誤分類率	N
パーティション	0.4042	0.5357	0.3218	0.3082	0.1663	0.1166	1192	
ブースティングツリー	0.5325	0.6555	0.2564	0.2809	0.1707	0.1133	1192	
ブースティングツリー	0.3979	0.5290	0.3252	0.3126	0.2141	0.1351	1192	