

JMP Add-In: Alternate Cut-off Confusion Matrix

Statistical Design by Sue Walsh

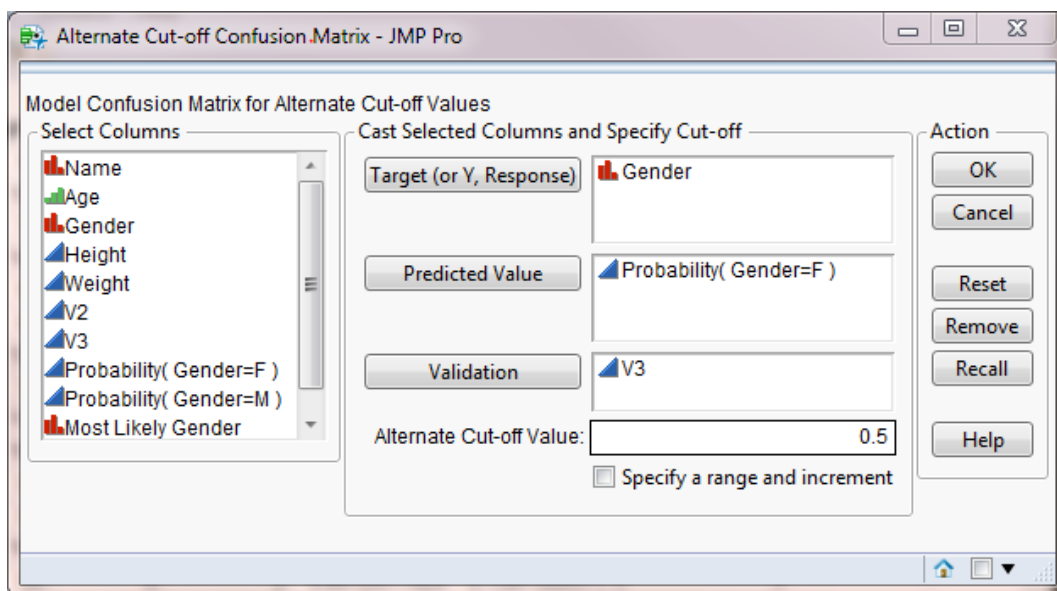
Developed by Wendy Murphrey

JMP can produce a confusion matrix in the Partition, Neural, and Fit Model platforms. Each of these confusion matrices is based upon a cut-off value of 0.5. The Alternate Cut-off Confusion Matrix JMP Add-In is designed to allow the user to specify a cut-off value or a range of values to generate a new confusion matrix for a binary response variable.

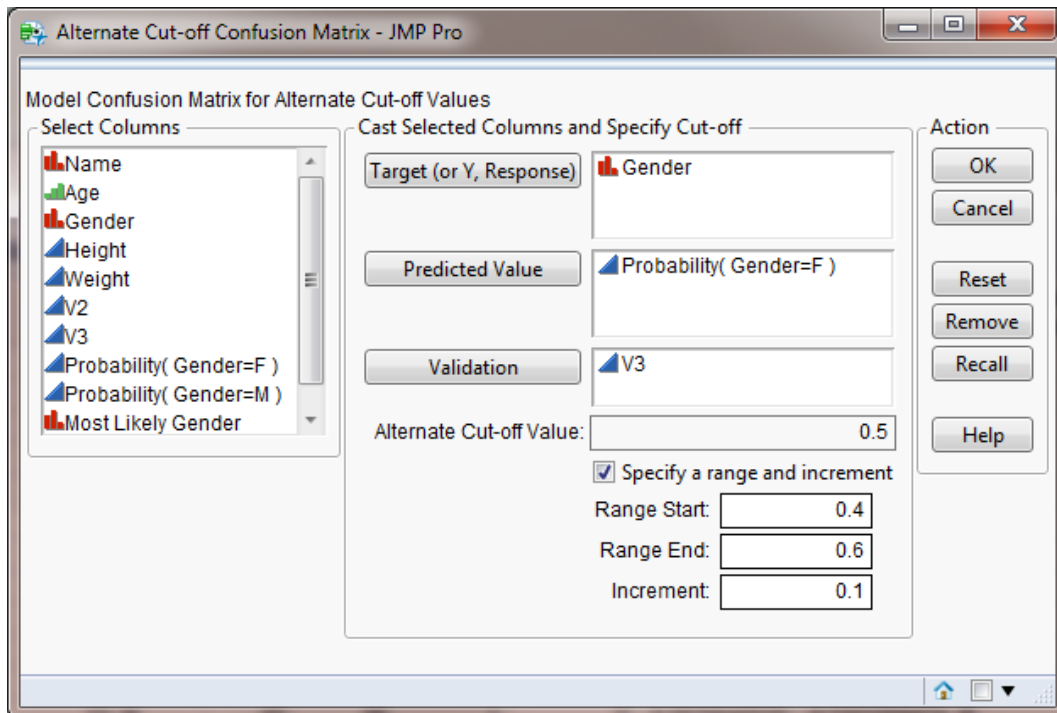
Before using this JMP Add-In, the prediction formulas must be saved to the data table from the Partition, Neural (JMP Pro only), or Fit Model report.

Upon launching the Alternate Cut-off Confusion Matrix JMP Add-In, you will be prompted for the following options:

- *Target*: This column must have a nominal modeling type and contain only two levels.
- *Predicted Value*: This column must have a continuous modeling type and should be one of the predicted (or probability) columns saved from the report.
- *Validation*: This column must be numeric and should only be used if a validation column was used in the analysis. If a validation column was used in the analysis, the same validation column should also be specified here. Validation is only available in JMP Pro.
- *Alternate Cut-off Value*: You can specify a cut-off value by typing the value in the dialog:



Or you can specify a range of values with an increment:



Upon clicking OK, a new window will appear containing the desired confusion matrix and confusion rates. The launch dialog will remain open so that you can enter additional cut-off values. If an Alternate Confusion Matrix report is already open, the new results will be added to the existing window.

NOTES:

- If an entire level has no observations, then the level will not be represented in the Alternate Cut-off Confusion Matrix report.
- In JMP Pro, specifying a cut-off of 0.5 should generate a confusion matrix with identical results to the confusion matrix in the JMP report. This is also the case for the Partition and Logistic Regression platforms in JMP. However if the data table contains missing values, the confusion matrix for a decision tree model may not match the one seen in the platform. This is due to the random assignment of observations with missing values in the split variable.
- The confusion matrix made with this JMP Add-In will not include observations that have a missing predicted probability.

Below is an example from the Neural platform in JMP Pro, which demonstrates identical results when the cut-off of 0.5 is specified:

Big Class - Neural of Gender by Height - JMP Pro

Validation Column: V3

Model Launch

Model NTanH(3)

Training

Gender

Measures	Value
Generalized RSquare	0.5042997
Entropy RSquare	0.3427675
RMSE	0.3868242
Mean Abs Dev	0.3179471
Misclassification Rate	0.25
-LogLikelihood	10.933412
Sum Freq	24

Confusion Matrix

Actual \ Predicted	F	M
Gender F	8	4
Gender M	2	10

Confusion Rates

Actual \ Predicted	F	M
Gender F	0.66667	0.33333
Gender M	0.16667	0.83333

Validation

Gender

Measures	Value
Generalized RSquare	0.3367161
Entropy RSquare	0.2144572
RMSE	0.3977611
Mean Abs Dev	0.2995482
Misclassification Rate	0.125
-LogLikelihood	4.1574901
Sum Freq	8

Confusion Matrix

Actual \ Predicted	F	M
Gender F	5	0
Gender M	1	2

Confusion Rates

Actual \ Predicted	F	M
Gender F	1.00000	0.00000
Gender M	0.33333	0.66667

Test

Gender

Measures	Value
Generalized RSquare	-13.01358
Entropy RSquare	-2.740873
RMSE	0.7096801
Mean Abs Dev	0.6326385
Misclassification Rate	0.625
-LogLikelihood	11.275595
Sum Freq	8

Confusion Matrix

Actual \ Predicted	F	M
Gender F	1	0
Gender M	5	2

Confusion Rates

Actual \ Predicted	F	M
Gender F	1.00000	0.00000
Gender M	0.71429	0.28571

Confusion Matrix - JMP Pro

Confusion Matrix for Cut-off = 0.5

Target: Gender
Predictor: Neural

	Modeling Data					
	Training		Validation		Test	
	Predicted F	Predicted M	Predicted F	Predicted M	Predicted F	Predicted M
Gender F	8	4	5	0	1	0
Gender M	2	10	1	2	5	2

Confusion Rates: Training (0.5)

Gender	Predicted	
	F	M
Gender F	66.67%	33.33%
Gender M	16.67%	83.33%

Confusion Rates: Validation (0.5)

Gender	Predicted	
	F	M
Gender F	100.00%	0.00%
Gender M	33.33%	66.67%

Confusion Rates: Test (0.5)

Gender	Predicted	
	F	M
Gender F	100.00%	0.00%
Gender M	71.43%	28.57%

Names Default To Here(bool) Determines where unresolved names are stored, either as a global/local