Beyond ROC Curves: Finding Meaning and Balancing Trade-Offs

Karen Copeland, Statistician, Boulder Statistics
Tarek Zikry, Technical Intern, SAS
Mia Stephens, JMP Academic Ambassador, SAS

Agenda

- Binary Classification Models
- Thresholds/Cut-offs
- ROC Curves
- Confusion Matrices
- Performance Metrics
- Demo of *Classification Threshold Explorer*
- Questions/Discussion

Binary Classification Model

- Response has two categories
  - Yes/No, Disease/No Disease, Good/Bad, True/False
- One or more factors (predictors)
- Model types:
  - Logistic Regression
  - Partition Models
  - Discriminant Models
  - Neural Networks, Etc.
- Model output:
  - Probability
  - Score
  - Propensity Score

Binary Classification Model

The output is the “score” for each observation.

- How do you implement a score when you can only have one of two outcomes?
- How should you select the threshold?

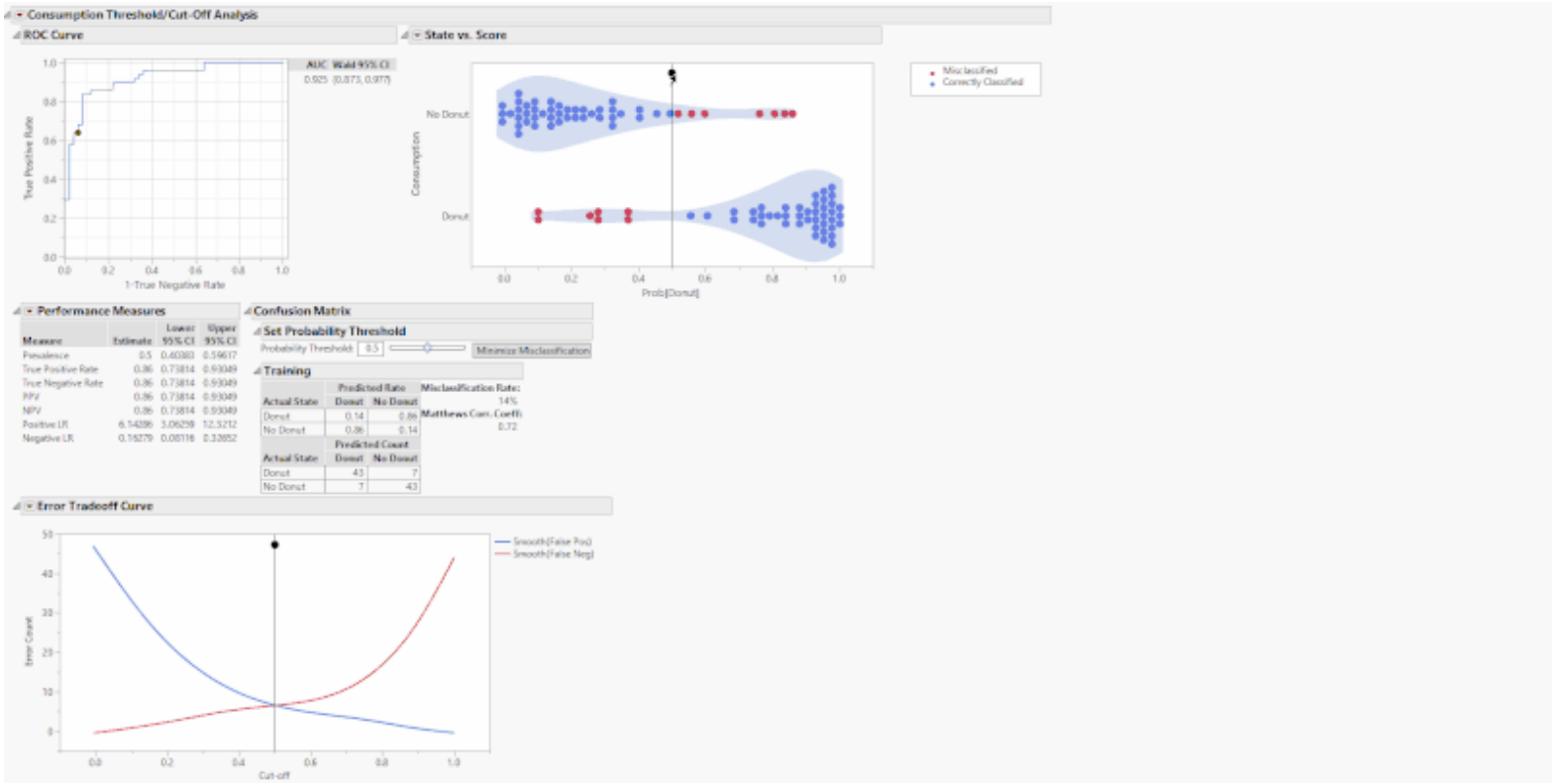
Examples:

- Medical Diagnostic
  - False Positive could lead to more testing, worry, etc.
  - False Negative could lead to non-treatment
- Marketing
  - False Positive could lead to money lost (sending to uninterested customer)
  - False Negative could lead to money lost (not to an interested customer)

jmp

Model Classification Explorer

Add-In For Exploring Thresholds: Where we are headed….



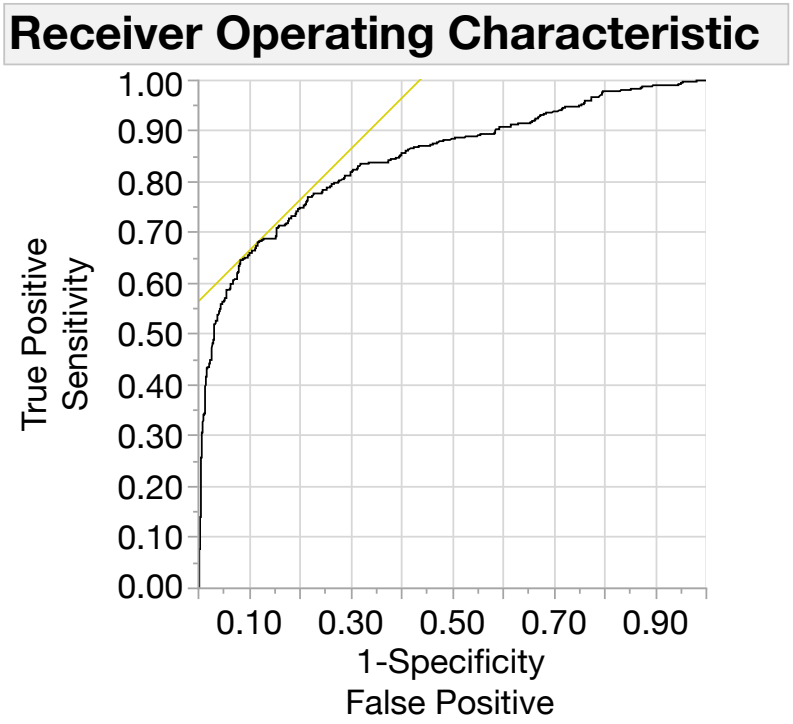
Classification Model Metrics

- ROC Curves and the AUC (Area Under the Curve)

The following metrics depend on a threshold/cut-off:

- Correct classification rates
- 2x2 Table of results
  - Contingency Table
  - Confusion Matrix
  - Cross Tabulation
- Performance Metrics
  - Se, Sp
  - TPR, FPR, TN, FN
  - Mathew's coefficient, etc.

ROC Curve



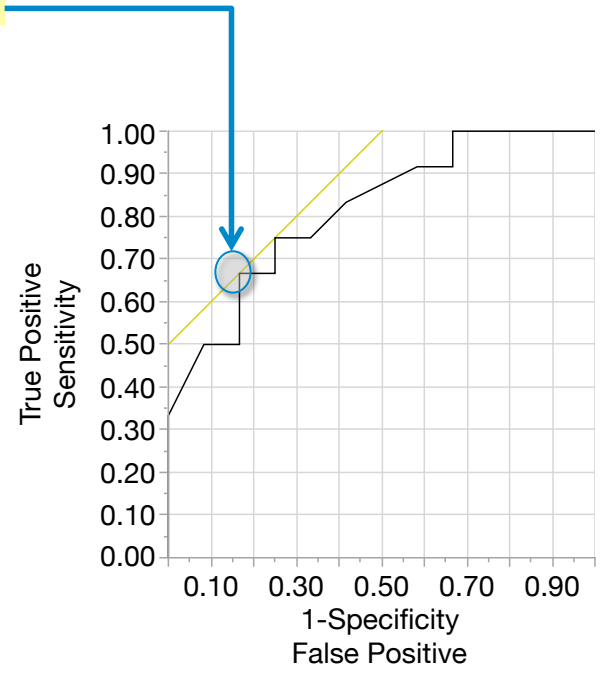
Using Survived='Yes' to be the positive level

AUC
0.84661

- Receiver operating characteristic (ROC) curve
- Describes the accuracy of a classification model across all possible thresholds.
- A plot of sensitivity versus 1 - specificity.
- JMP plots a yellow 45 degree tangent line to the curve.

ROC Table

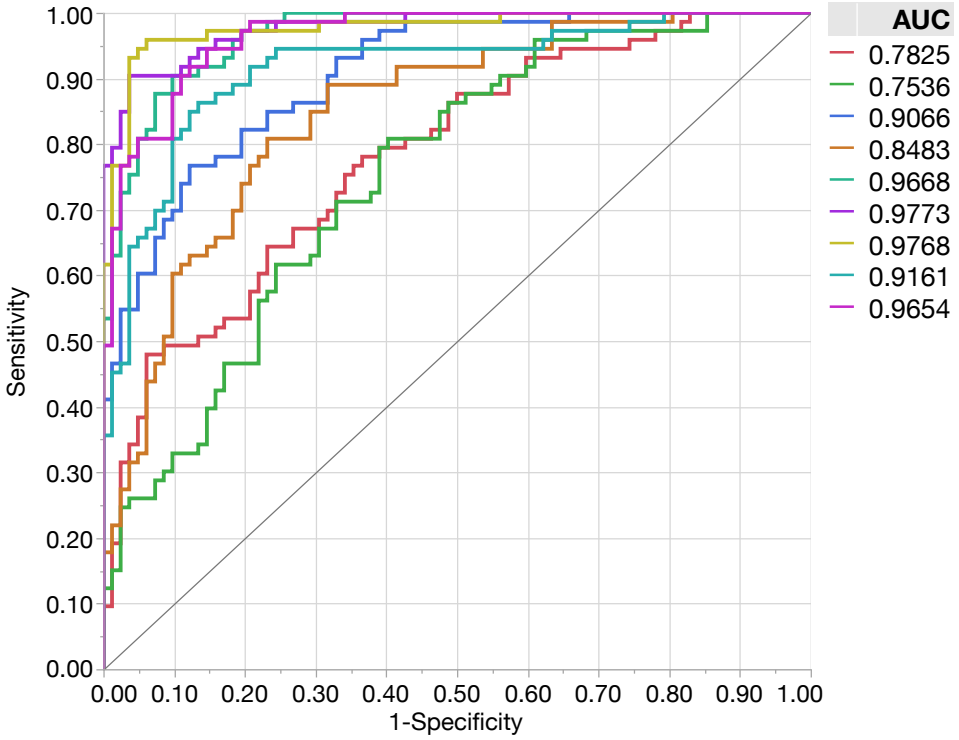
X	Prob	1-Specificity	Sensitivity	Sens-(1-Spec)		True Pos	True Neg	False Pos	False Neg
.	.	0.0000	0.0000	0.0000		0	12	0	12
23.60000	0.9508	0.0000	0.0833	0.0833		1	12	0	11
22.40000	0.8995	0.0000	0.1667	0.1667		2	12	0	10
22.00000	0.8738	0.0000	0.2500	0.2500		3	12	0	9
21.60000	0.8427	0.0000	0.3333	0.3333		4	12	0	8
20.80000	0.7622	0.0833	0.5000	0.4167		6	11	1	6
20.40000	0.7126	0.1667	0.5000	0.3333		6	10	2	6
20.00000	0.6573	0.1667	0.6667	0.5000	*	8	10	2	4
19.60000	0.5974	0.2500	0.6667	0.4167		8	9	3	4
19.20000	0.5344	0.2500	0.7500	0.5000	*	9	9	3	3
18.80000	0.4703	0.3333	0.7500	0.4167		9	8	4	3
18.40000	0.4071	0.4167	0.8333	0.4167		10	7	5	2
17.60000	0.2912	0.5833	0.9167	0.3333		11	5	7	1
17.20000	0.2412	0.6667	0.9167	0.2500		11	4	8	1
16.80000	0.1973	0.6667	1.0000	0.3333		12	4	8	0
16.40000	0.1598	0.7500	1.0000	0.2500		12	3	9	0
16.00000	0.1283	0.8333	1.0000	0.1667		12	2	10	0
14.80000	0.0638	0.9167	1.0000	0.0833		12	1	11	0
14.00000	0.0392	1.0000	1.0000	0.0000		12	0	12	0
14.00000	0.0392	1.0000	1.0000	0.0000		12	0	12	0



RidingMowers.jmp
Analyze > Fit Y by X
Y = Ownership, X = Lot_Size

ROC Curves and AUCs

What is the AUC? The metric that is requested from management...



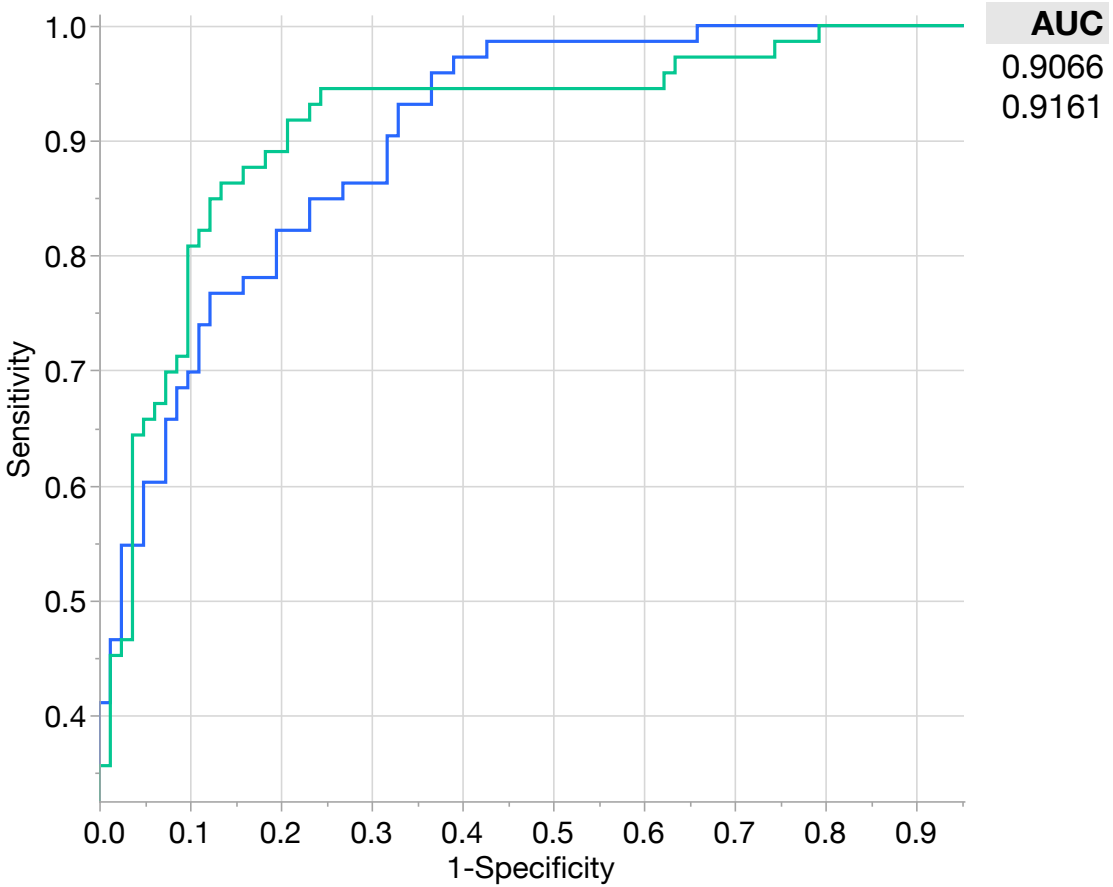
The area under the curve (AUC) is a measure of the overall classification ability of the model.

If AUC = 0.5 then you should just guess and not use your model.

If AUC = 1.0 your model is really good.

Analyze > Predictive Modeling > Model Comparison

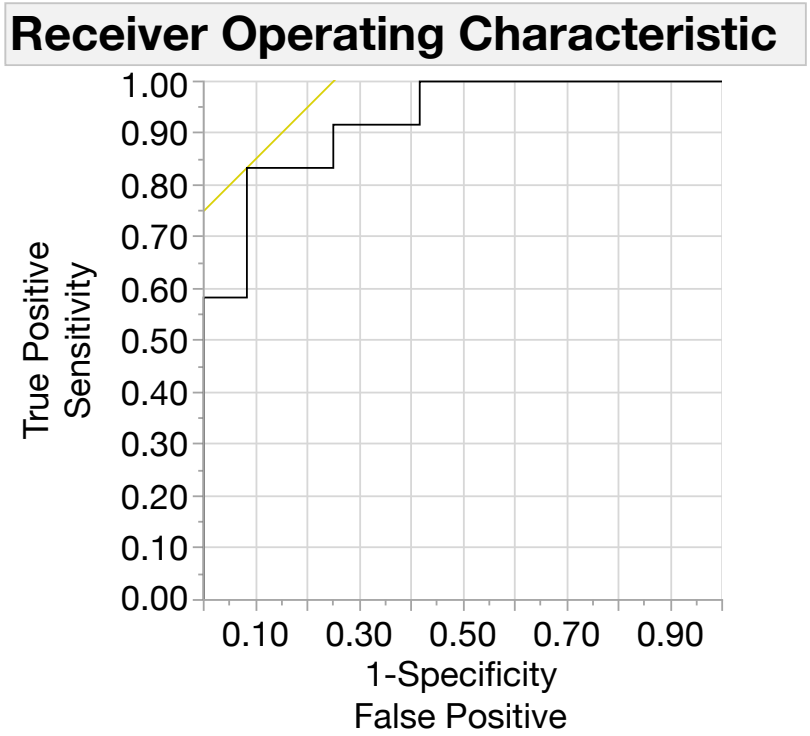
AUC and Shape of the Curve



AUCs are about the same.

Se/Sp trade-offs differ.

ROC Curve Visualization of a Classification Model

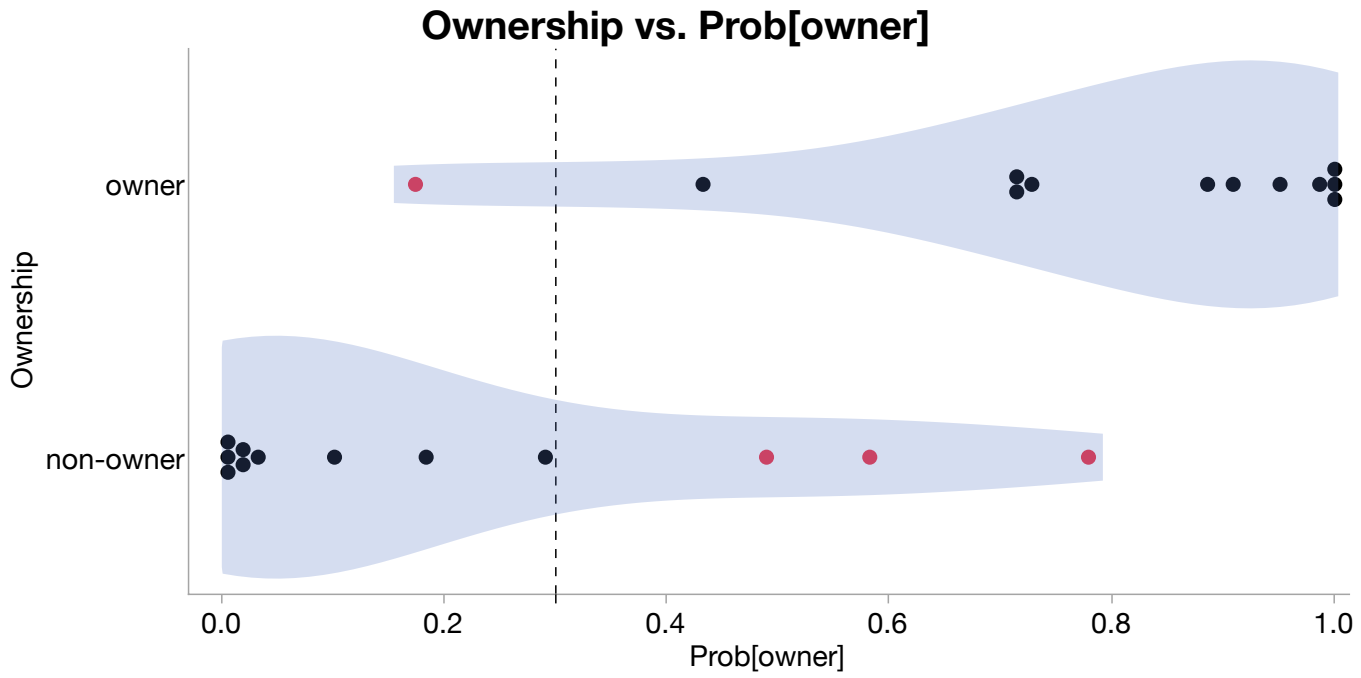


Using Ownership='owner' to be the positive level

AUC
0.92361

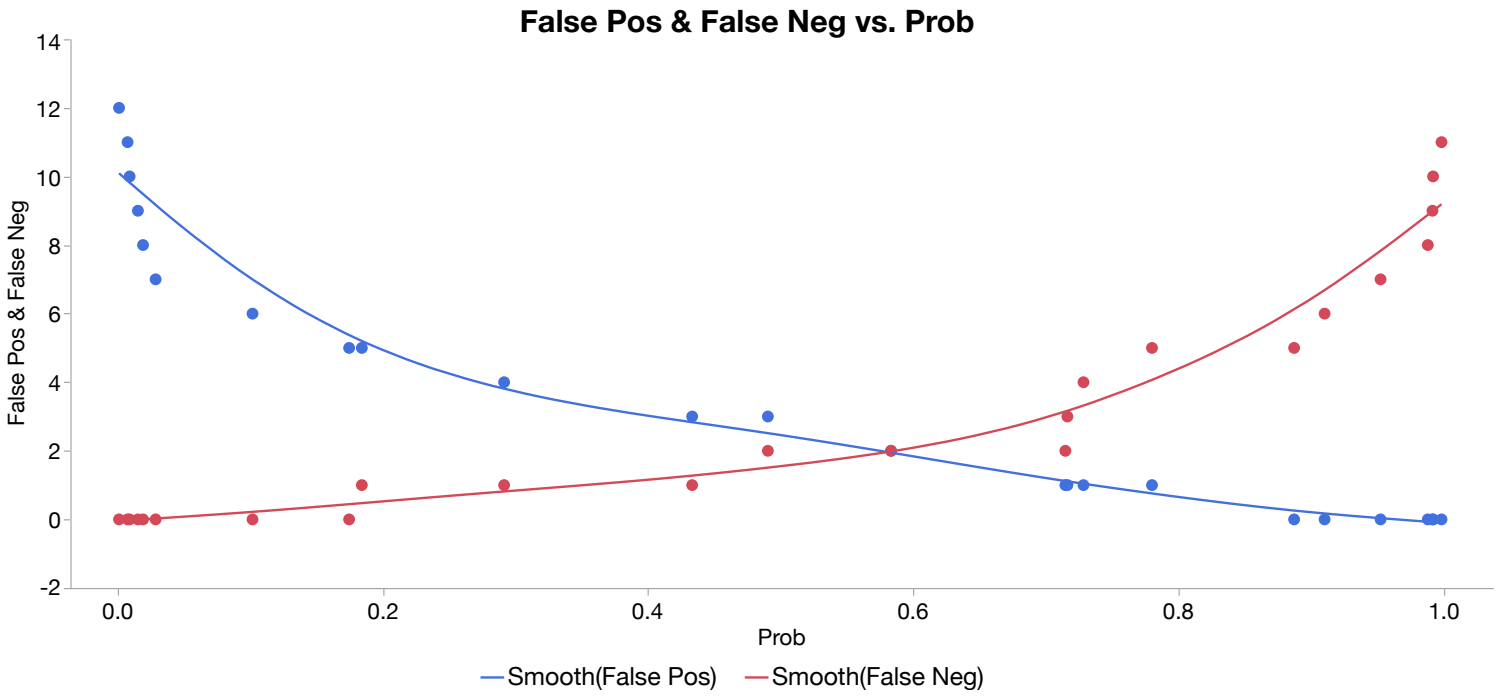
Alternative Visualization: State vs. Score Plot

The State vs. Score plot is used to visualize the predicted scores within each category.



Alternative Visualization: Error Tradeoff Curve

The tradeoff curve is a plot of the false positive counts and false negative counts across all thresholds.



Confusion Matrix or Contingency Table

The confusion matrix shows the 2x2 table of counts based on a *specific* threshold.

Confusion Matrix

Training

Actual Ownership	Predicted Count owner	Predicted Count non-owner
owner	10	2
non-owner	2	10

Beware of the arrangement of the table as to not to become confused!

JMP Confusion Matrices/Contingency Tables

Confusion Matrix

Logistic Regression

Training

Actual Ownership	Predicted Count owner	non-owner
owner	10	2
non-owner	2	10

Confusion Matrix

Set Probability Threshold

Probability Threshold 0.71

Training

Actual Ownership	Predicted Rate non-owner	owner
non-owner	0.917	0.083
owner	0.167	0.833

Misclassification Rate: 0.1250

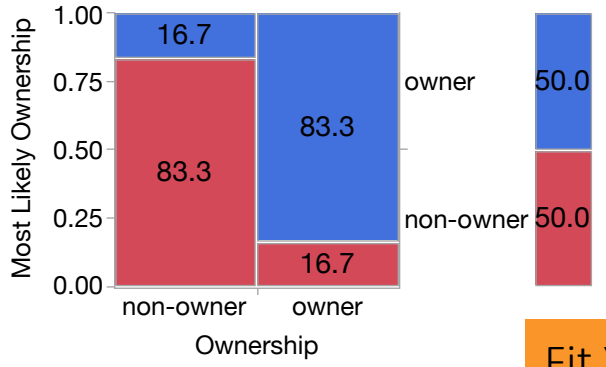
Actual Ownership	Predicted Count non-owner	owner
non-owner	11	1
owner	2	10

Gen Reg

Tabulate

	Most Likely Ownership					
	non-owner			owner		
Ownership	N	Column %	Row %	N	Column %	Row %
non-owner	10	83%	83%	2	17%	17%
owner	2	17%	17%	10	83%	83%

Mosaic Plot



Fit Y by X

Contingency Table

Most Likely Ownership

Ownership	Count / Total % / Col % / Row %	non-owner	owner	Total
non-owner		10	2	12
		41.67	8.33	50.00
		83.33	16.67	
		83.33	16.67	
owner		2	10	12
		8.33	41.67	50.00
		16.67	83.33	
		16.67	83.33	
Total		12	12	24
		50.00	50.00	

Performance Metrics

		Predicted		
		Yes	No	
Actual	Yes	TP	FN	TP+FN= #Yes
	No	FP	TN	FP+TN = #No
		TP+FP	FN+TN	N

Table 1.2 Performance Metrics

General	Marketing/Business	Medical/Diagnostics	Calculation
Prevalence	Prevalence	Prevalence	(TP+FN) /N
True Positive Rate	Recall	Sensitivity	TP/(TP+FN)
True Negative Rate	Specificity	Specificity	TN/(FP+TN)
PPV	Precision	PPV	TP/(TP+FP)
NPV	NPV	NPV	TN/(FN+TN)
Positive LR	Positive LR	Positive LR	Sensitivity/FPR
Negative LR	Negative LR	Negative LR	FNR/Specificity

Add-In

Exploring Model Classification Thresholds

How Did it Come to Be?

- Causal conversation between a consultant and an educator over classification models.
- Leverage the power of JMP to create interactive tool to explore thresholds.
- Interest from JMP Life Science Team.
- JMP Summer intern program!!!
- Classification models is an area that both Karen and Mia have worked with over the years.
  - Performance metrics Add-in:
    - https://community.jmp.com/t5/JMP-Add-Ins/Performance-Summary-for-Diagnostic-Tests/ta-p/22524
  - Books: *Building Better Models with JMP Pro* and *Data Mining for Business Analytics with JMP Pro*
  - Blog on model graphics
    - https://community.jmp.com/t5/JMP-Blog/Why-model-visualization-is-integral-to-model-building/ba-p/30704

DEMO TIME!!!

Want the App? Go to community.jmp.com!!

Add-in (File Exchange): Exploring Model Classification Thresholds
https://community.jmp.com/t5/JMP-Add-Ins/Exploring-Model-Classification-Thresholds/ta-p/66964

Blog: "*Beyond ROC curves: Exploring probability thresholds and error trade-offs in predictive models*"
https://community.jmp.com/t5/JMPer-Cable/Beyond-ROC-curves-Exploring-probability-thresholds-and-error/ba-p/66706

Slides: Go to the Discovery Summit Series community area

Contacts

karen@boulderstats.com

Tarek@live.unc.edu

Mia.Stephens@jmp.com

Closing Remarks

- Large file size
  - Be patient
  - Don’t use with really large files
  - Watch for update
- Potential Users
  - Anyone with a classification model
  - Educators

Thank you!

Questions?