Area-Under-the-Curve Analysis

Douglas H. Spence, Ph.D.

Presented by
Justice Research and Statistics Association

4/20/2017

Area-Under-the-Curve Analysis

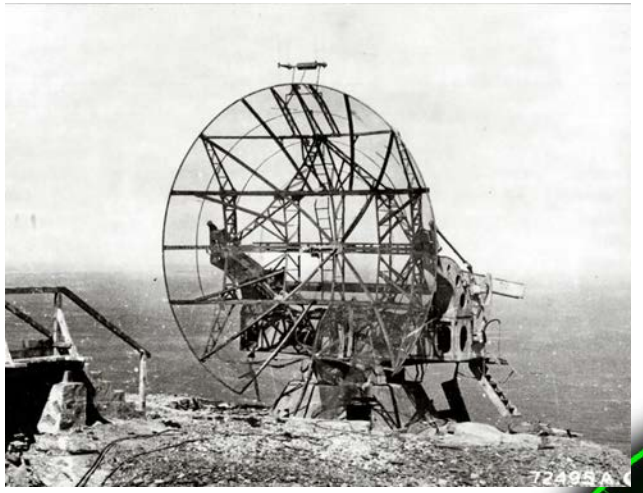
Douglas H. Spence, Ph.D.
Director, Office of Research and Strategic Planning (ORSP)
West Virginia Division of Justice and Community Services (DJCS)

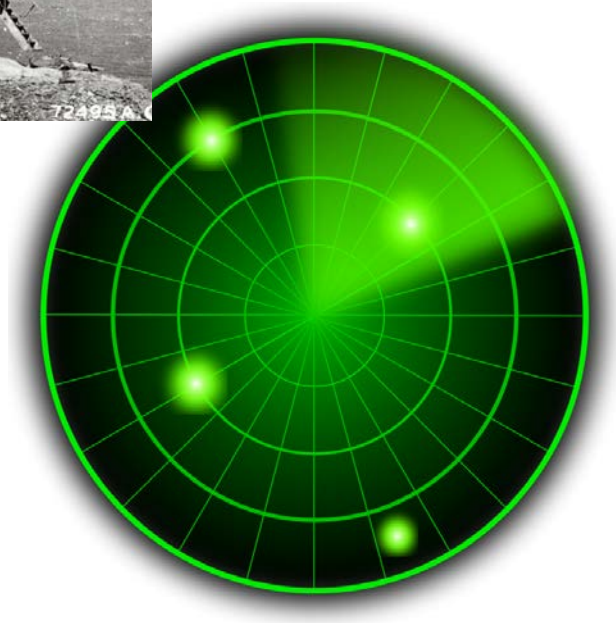
Webinar Overview

- Background and historical context
- Evaluating binary classifiers
- The receiver operating characteristic (ROC) curve
- The AUC statistic
- Conducting AUC analysis in SPSS
- Using AUC analysis to supplement and enhance analysis results

Background and Historical Context

- AUC analysis examines the area under a receiver operating characteristic (ROC) curve
- Originally developed to enhance signal detection during World War II, [Radar] ROC curves plot the performance of classifiers for binary outcomes
- They help identify which classifiers maximize the number of correct predictions while minimizing false positives
- ROC curves proved to be a highly effective means of measuring the performance of predictions, and their use spread to other fields (medicine, psychology, criminal justice, etc.)





Binary Classifiers

- Binary classifiers (aka predictors) are rules or procedures for classifying cases in terms of their predicted outcome (positive or negative).
- In statistics, classifiers are usually the predicted probabilities produced by logistic regression or some other analysis.
- When classifiers are probabilistic, their effectiveness will vary depending on the discrimination threshold.
- The discrimination threshold refers to the minimum predicted probability at which cases are classified as positive.

PRE_1	PredictedOutcome	ActualOutcome
.55302	1.00	1.00
.53893	1.00	.00
.34522	.00	1.00
.55302	1.00	1.00
.60834	1.00	1.00
.41198	.00	.00
.35818	.00	1.00
.41198	.00	.00
.64816	1.00	1.00
.56703	1.00	1.00
.53893	1.00	1.00
.53893	1.00	1.00
.41198	.00	.00
.43978	.00	1.00
.58094	1.00	.00
.59471	1.00	1.00
.60834	1.00	1.00
.28405	.00	.00

Evaluating Binary Classifiers

- For any given case, binary classifiers can have one of four possible results (TP, FP, TN, and FN).

- There are many ways to evaluate a classifier's performance:
  - True Positive Rate (aka Sensitivity) = TP/(TP + FN)
  - False Positive Rate = FP/(FP + TN)
  - True Negative Rate (aka Specificity) = TN/(FP + TN)
  - Accuracy = (TP + TN)/P + N

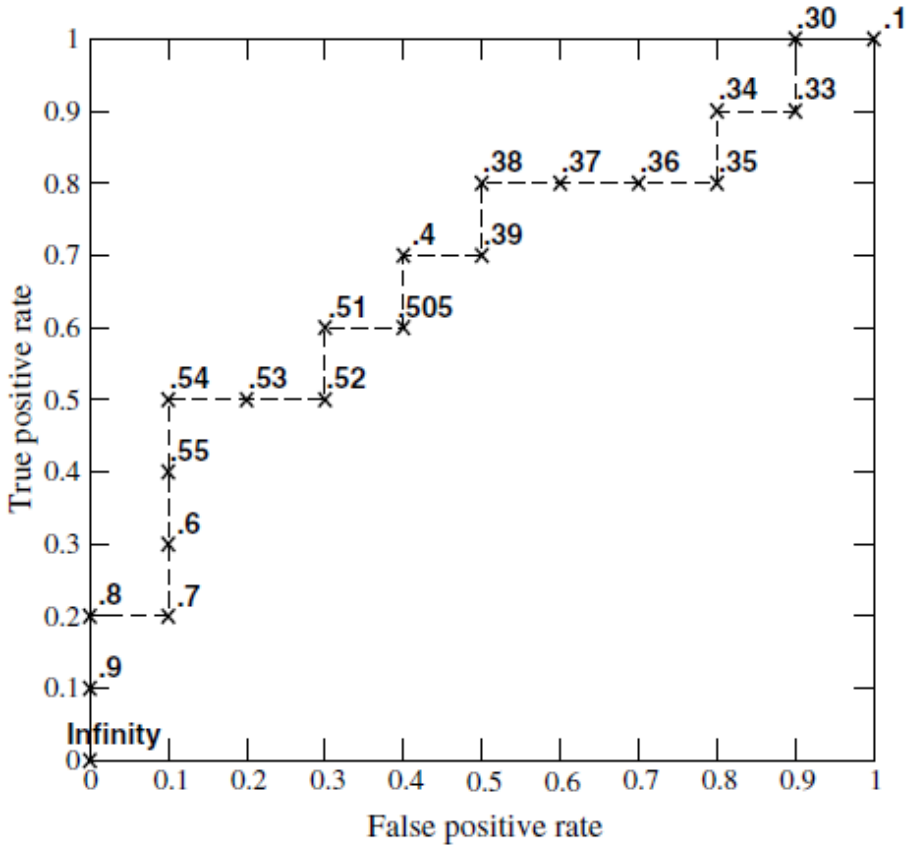
- We want a method that considers both true and false positives, and takes the discrimination threshold into account.

		Predicted class	
		P	N
Actual Class	P	True Positives (TP)	False Negatives (FN)
	N	False Positives (FP)	True Negatives (TN)

The ROC Curve

- ROC curves plot the true positive rate versus the false positive rate for a classifier as its discrimination threshold is varied.

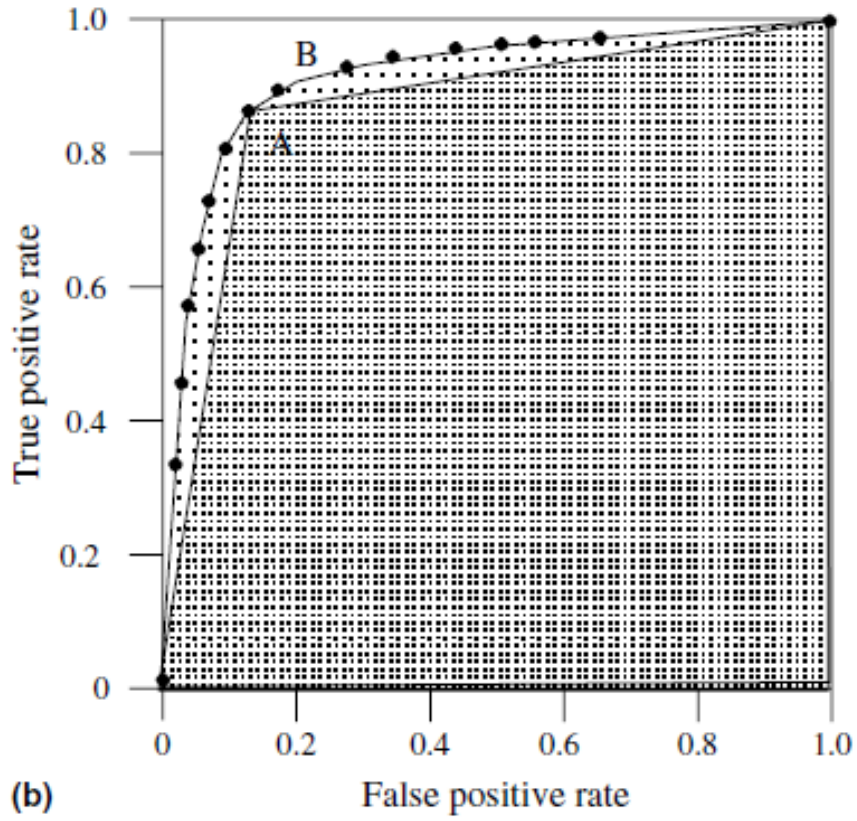
- The best possible classifier would yield a point in the upper left corner of the ROC plot. This would indicate that all true positives were predicted correctly, with no false positives.



Source: Fawcett, 2006

The AUC Statistic

- The AUC statistic represents the area of the plot that falls under a given ROC curve.
- Summarizes entire ROC curve across all discrimination thresholds
- Ranges from 0 to 1 and is equal to the probability that the classifier will rank a randomly chosen positive case (TP) higher than a randomly chosen negative one (TN)
- Predictors are commonly considered to be effective if AUCs are 0.7 or above. An AUC of 0.5 is considered no better than chance.

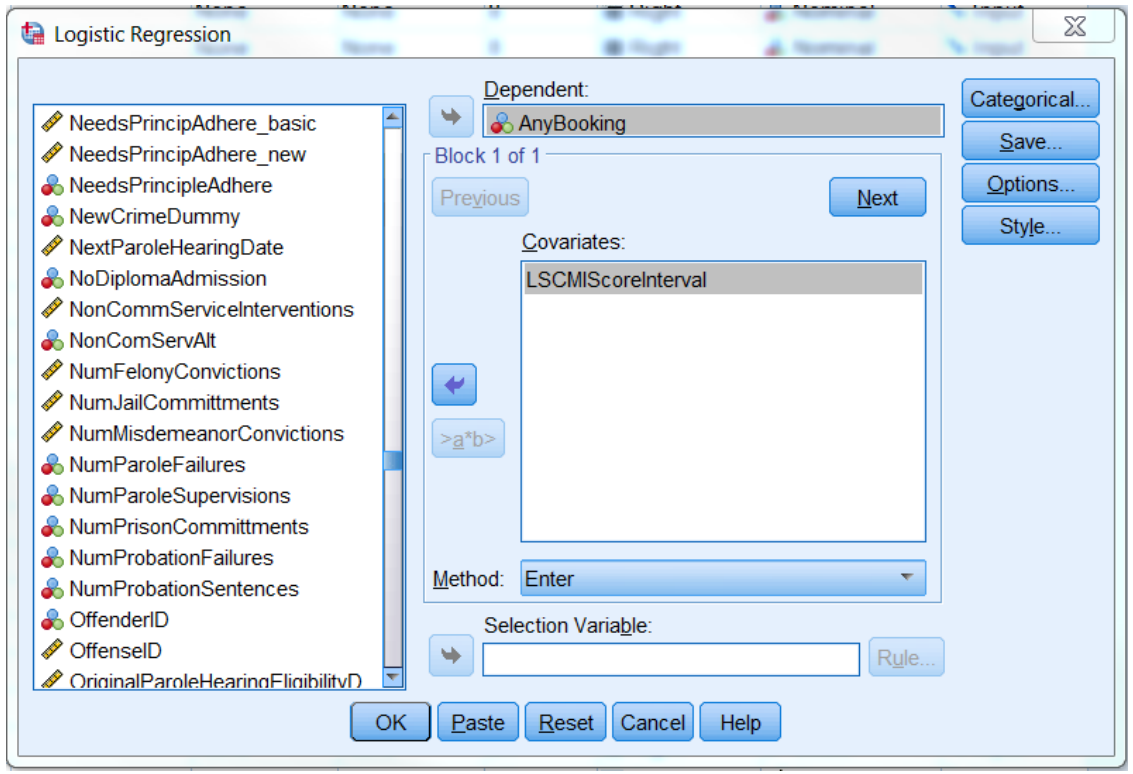


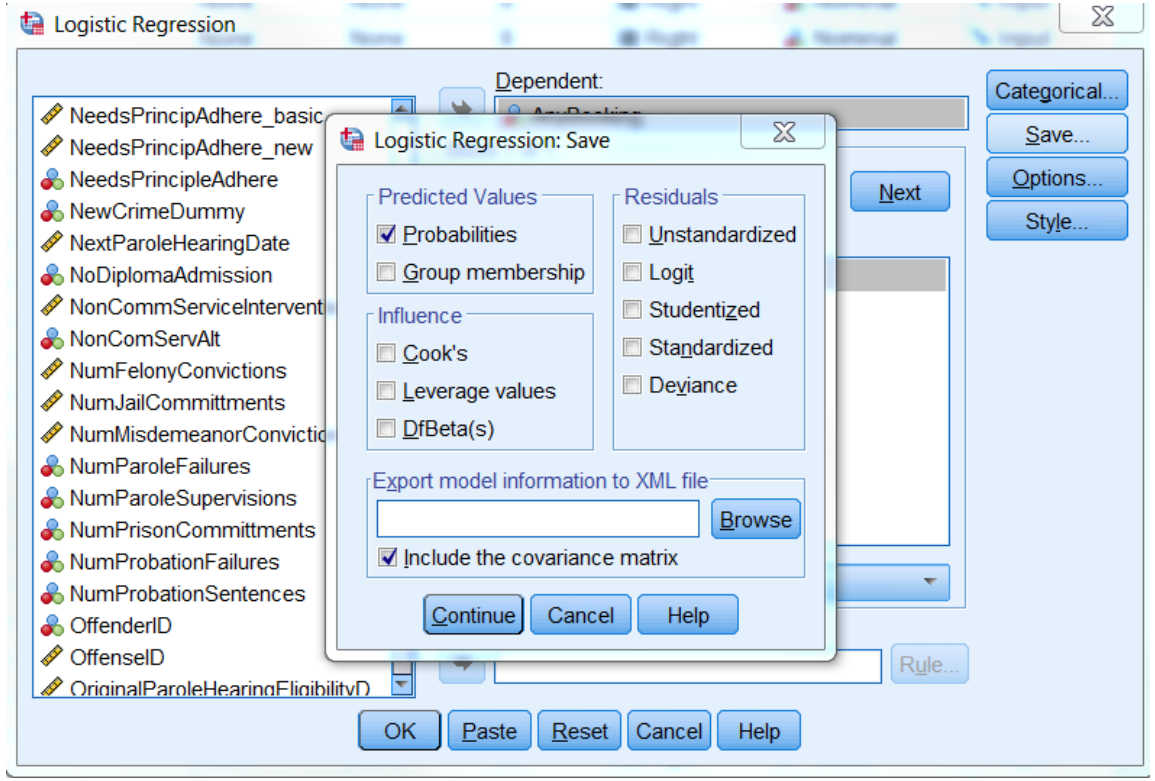
Source: Fawcett, 2006

Conducting AUC Analysis After Logistic Regression in SPSS

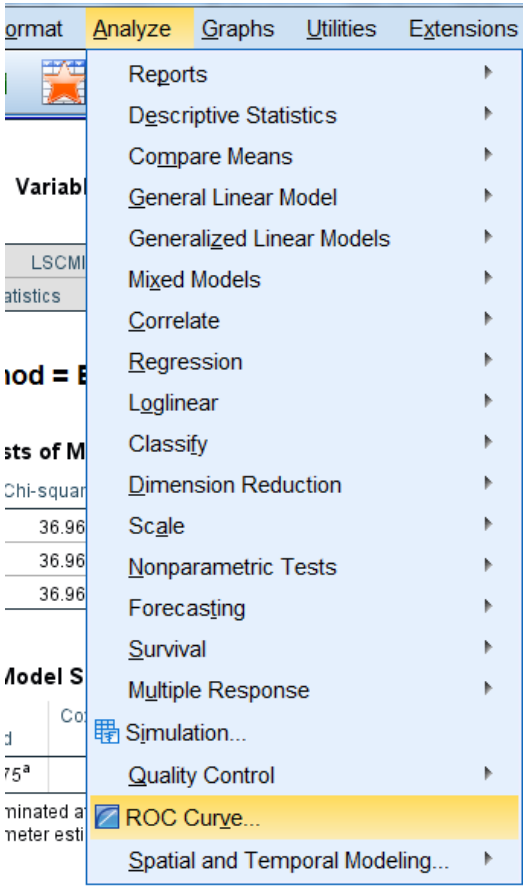
1. Set up logistic regression model by selecting **Analyze-Regression-Binary Logistic** from the pull down menu.
2. In the window select the save button to open to Logistic Regression: Save window and check the box for probabilities.
3. Run logistic regression.
4. Select **Analyze-ROC Curve** from the pull down menu.
5. In ROC Curve window, move the predicted probability from your logistic regression into the test variable box and your dependent variable into the state variable box.
6. Check the box to display ROC curve and for the diagonal reference line if desired.

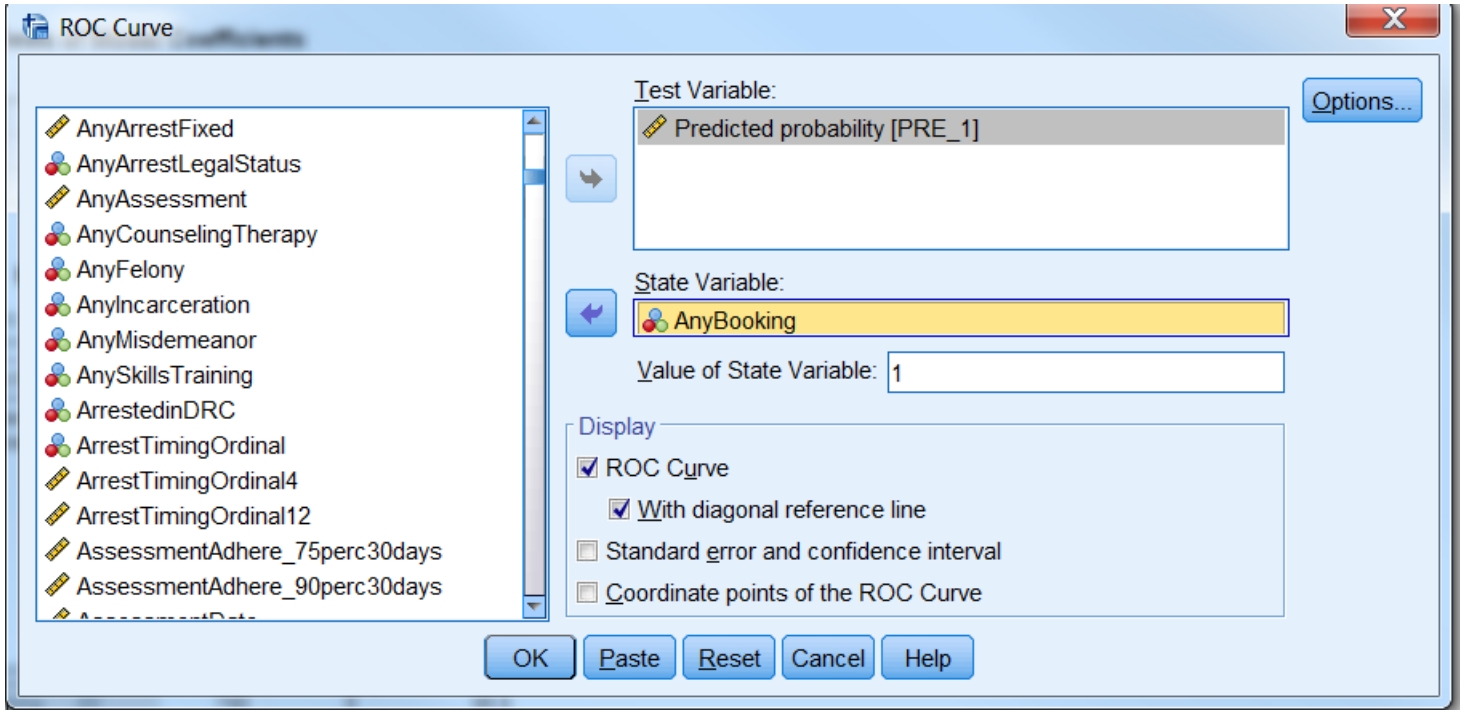
Conducting AUC Analysis After Logistic Regression in SPSS



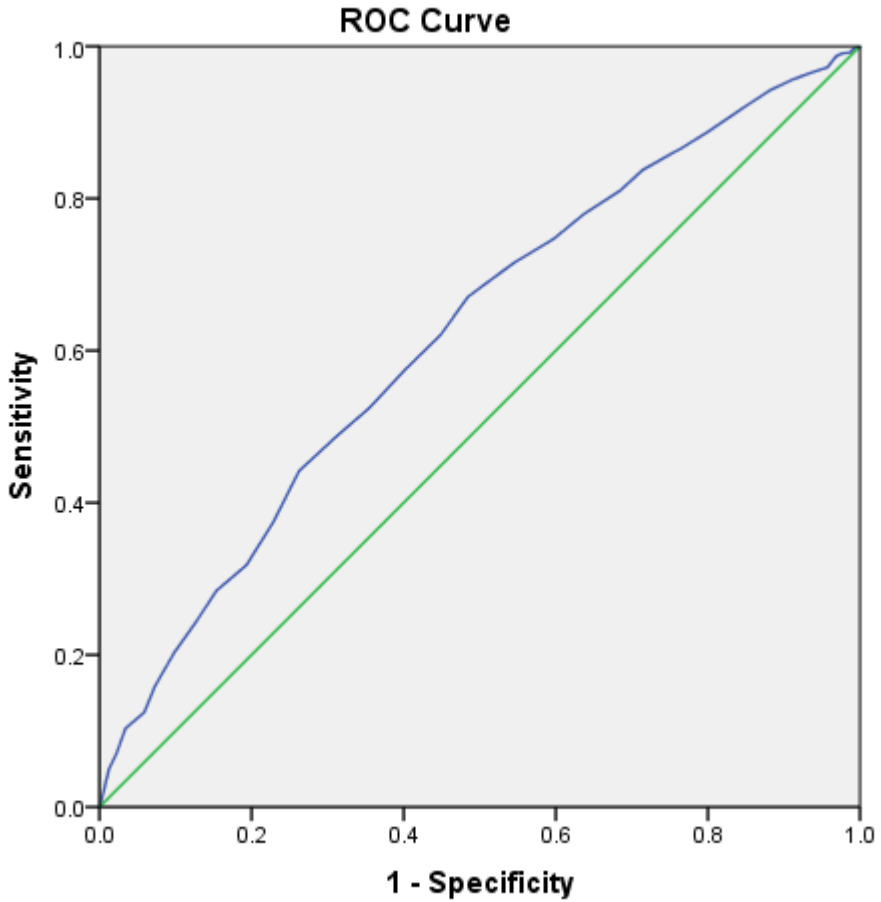


Conducting AUC Analysis After Logistic Regression in SPSS





Conducting AUC Analysis After Logistic Regression In SPSS



Diagonal segments are produced by ties.

Area Under the Curve

Test Result Variable(s): Predicted probability

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.621	.014	.000	.592	.649

Advantages of AUC Statistic

- Provides a good summary of the predictive accuracy of limited dependent variable models (e.g., logit, probit) that is unaffected by base rates
- Can be interpreted directly and can be compared across models, even when outcomes are measured in different units
- Very useful for:
  - Assessing model specification for logistic regression (similar to r-squared for OLS)
  - Comparing the performance of different predictors (such as risk assessment tools) for a given outcome
  - Measuring the efficacy of a predictor for different outcomes or populations

Using AUC Analysis to Enhance Logistic Regression Results

Table 5

Logistic Regression Estimates for Predictive Factors Associated with Recidivism (N = 1,495)

	Arrest		*Regional Jail Booking*		*Incarceration*	
Variable	*B (SE)*	*Odds Ratio*	*B (SE)*	*Odds Ratio*	*B (SE)*	*Odds Ratio*
Age	-0.031***	0.969	-0.020*	0.980	-0.009	
	(0.009)		(0.009)		(0.020)	
Female	-0.403*	0.668	-0.613**	0.542	-0.363	
	(0.194)		(0.197)		(0.405)	
LS/CMI Risk Score	0.047***	1.070	0.042***	1.043	0.084***	1.087
	(0.012)		(0.012)		(0.025)	
Length of Stay	0.001		0.000		-0.002	
	(0.001)		(0.001)		(0.001)	
Successful Program Completion	-0.706***	0.494	-0.861***	0.423	-1.087*	0.337
	(0.195)		(0.189)		(0.384)	
N	651		766		766	
Nagelkerke R-Squared	0.147		0.123		0.212	
AUC	0.688		0.680		0.820	

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Using AUC Analysis to Enhance Logistic Regression Results

Table 8

Area-Under-the-Curve Statistics for Level of Service/Case Management Inventory (LS/CMI) Total and Section Scores for Arrests, Jail Bookings, and Incarcerations (N = 1,495)

	Arrest	*Jail Booking*	*Incarceration*
Total Risk and Needs Score	0.623	0.613	0.679
LS/CMI Section			
Procriminal Attitude	0.504	0.505	0.524
Companions	0.592	0.563	0.669
Antisocial Pattern	0.561	0.562	0.598
Criminal History	0.606	0.603	0.678
Family/Marital	0.540	0.537	0.631
Education/Employment	0.581	0.581	0.542
Leisure/Recreation	0.551	0.548	0.561
Alcohol/Drug Problem	0.591	0.567	0.593

Note: All AUC statistics are derived from separate logistic regression models that contain only the relevant LS/CMI subcomponent score.

Using AUC Analysis to Enhance Logistic Regression Results

TABLE 6: Area Under the Curve Values for COMPAS Risk Models Predicting Any Offense, Offenses Against Persons, and Felony Offenses

	Total Sample (N = 2,328)			*Women (n = 449)*			*Men (n = 1,879)*		
Model	*Any*	*Person*	*Felony*	*Any*	*Person*	*Felony*	*Any*	*Person*	*Felony*
COMPAS I	.66	.72	.70	.69	.78	.68	.67	.71	.71
COMPAS II	.68	.73	.72	.72	.80	.69	.68	.72	.73
Recidivism Risk III	.68	.71	.70	.65	.76	.66	.68	.70	.71

Note. COMPAS Model I includes COMPAS base scales; COMPAS II adds age at first arrest to Model I.

Source: Brennan, Dietrich, & Ehret, 2009

Conclusion

- ROC curves and AUC statistics provide an effective means for measuring the accuracy of predictions with many criminal justice applications.
- They provide a useful complement to logistic regression analysis that can be generated easily by most software packages.
- However, what constitutes a “good” AUC value in a particular context is not always clear. An AUC value of 0.7 is the most common, but standards vary significantly across fields.

References

Brennan, T., Dietrich, W., & Ehret, B. (2009). Evaluating the predictive validity of the COMPAS risk and needs assessment system. *Criminal Justice and Behavior, 36*(1), 21-40.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(4), 861-874.

Spence, D., & Haas, S. M. (2016). *Recidivism by Direct Sentence Clients Released From Day Report Centers in 2011: Predictors and Patterns over Time*. Report published by the WV Statistical Analysis Center. http://www.djcs.wv.gov/ORSP/SAC/Documents/WV_DRC%20Recidivism%20Jan%202016%20Final.pdf