Optimal Data Analysis
Vol. 9 (March 13, 2020), 74-78

2155-0182/10/$3.00

Implementing ODA from Within Stata: Finding the Optimal Cut-Point of a Diagnostic Test or Index (*Invited*)

Ariel Linden, Dr.P.H.
Linden Consulting Group, LLC

Maximizing the discriminatory accuracy of a diagnostic or screening test is paramount to ensuring that individuals with or without the disease (or disease marker) are correctly identified as such. In this paper, I describe how the new Stata package for implementing ODA can be used to determine the optimal cut-point along the continuum of test values that maximally discriminates between those with and without the disease.

Prior papers[1-7] introduced the new Stata package called **oda**[8] for implementing ODA from within the Stata environment. This package is a wrapper for the MegaODA software[9], and therefore the MegaODA.exe file must be loaded on the computer for the **oda** package to work (MegaODA software is available at https://odajournal.com/resources/). To download the **oda** package, at the Stata command line type: "ssc install oda" (without the quotation marks).

This paper demonstrates how the **oda** package can be used for identifying the optimal cut-point on a diagnostic test to maximally discriminate between those with and without the disease (or marker) under study.

Maximizing the discriminatory accuracy of a diagnostic or screening test is paramount to ensuring that individuals with or without the disease are correctly identified as such. An individual is classified as disease positive if their test value surpasses a pre-established cut-point, and is classified as disease negative if their test value falls below that cut-point.[10] Model sensitivity is the proportion of true disease positives (individuals correctly predicted by the test as having the disease), and model specificity is the proportion of true disease negatives (individuals correctly predicted as not having the disease).[11]

Receiver Operating Characteristic or ROC analysis is a popular approach for assessing and displaying the overall discriminatory accuracy of diagnostic tests, which involves plotting sensitivity *vs*. 1-specificity across the full range of values for the sample. Area under the curve (AUC) is then computed to assess the test's overall discriminatory ability.

2155-0182/10/$3.00

A test with perfect discriminatory ability has AUC=1.0, and a test unable to distinguish between individuals with *vs*. without disease has AUC=0.50.[11]

Using values computed in ROC analysis, several approaches have been developed to find the "optimal" cut-point on a diagnostic test at which individuals are "best" classified. The "nearest to (0,1) method" finds the cut-point on the ROC curve closest to the upper left-hand corner (i.e., the point with perfect sensitivity and specificity).[12] The "concordance probability method" by Liu[13] defines the optimal cut-point as the point maximizing the product of sensitivity and specificity. And the "Youden index (*J*) method" defines the optimal cut-point as the point maximizing the difference between the true positive rate and false positive rate over all possible cut-point values.[14]

In **oda** we identify the optimal cutpoint for a diagnostic test by simply specifying the reference variable (indicating the true state of the observation, such as diseased and non-diseased or normal and abnormal) as the *class* variable and the rating (or score) of the diagnostic test (or index) is specified as the *attribute*.

Methods

Data

For this example, we will re-examine previously evaluated data[15] from 13,667 newly enrolled HMO Medicare members who completed a health risk assessment called the PRA test[16] where the results classified them as being either at low or high risk of hospitalization (using a cut-point of .50 on a 0-1.0 scale) over the course of the following 25-month period. At the end of 25 months, 2,920 members had been hospitalized.

Analytic process

To identify the optimal cut-point in this study, **oda** is implemented with the following syntax (see the help file for **oda** for a complete description of the syntax options):

```
oda event prascore , pathoda("C:\ODA\")
store("C:\ ODA\output") iter(25000) loo
```

The above syntax is explained as follows: The outcome variable "event" is the *class* variable; the PRA score predicting hospitalization is specified as the *attribute*; the directory path where the MegaODA.exe file is located on my computer is "C:\ODA\"; the directory path where the output and other files generated during the analysis should be stored is "C:\ODA\output"; the number of iterations (repetitions) for computing a permutation *P*-value is 25,000; and leave-one-out analysis is used to assess cross-generalizability.

The **oda** package produces an extract of the total output produced by oda software (the complete output is stored in the specified directory with the extension ".out")

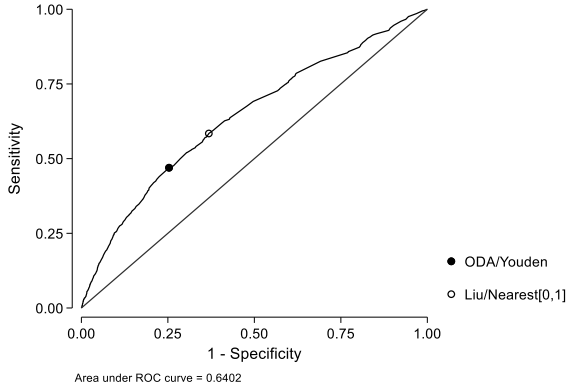
```
ODA model:
----------
IF PRASCORE <= 0.2895 THEN EVENT = 0
IF 0.2895 < PRASCORE THEN EVENT = 1


Summary for Class EVENT  Attribute PRASCORE
-----------------------------------------------

Performance Index           Train    LOO
-----------------          ------  ------
Overall Accuracy            68.74%  68.73%
PAC EVENT=0                 74.66%  74.66%
PAC EVENT=1                 46.95%  46.92%
Effect Strength PAC         21.61%  21.58%
PV EVENT=0                  83.82%  83.81%
PV EVENT=1                  33.49%  33.47%
Effect Strength PV          17.31%  17.28%
Effect Strength Total       19.46%  19.43%


Monte Carlo summary (Fisher randomization):
-----------------------------------------------
Iterations:  25000
Estimated p: 0.000000


Results of leave-one-out analysis
---------------------------------
13667 observations

Fisher's exact test (directional) classification table  p = .454E-0106
```

As shown in the **oda** output, the ODA model can be interpreted as follows: "if PRA score <= 0.2895, then predict that event is 0 (no hospitalization). If PRA score is > 0.2895, then predict that the event is 1 (hospitalization)." The effect strength for sensitivity (ESS) is labelled in the output as "Effect Strength PAC". The ESS is 21.61% (a relatively weak effect)[17] in the

2155-0182/10/$3.00

training analysis and in LOO. The sensitivity is 46.95% and the specificity is 74.66% (with identical results for LOO). The permutation *P*-values for both training sample and LOO are statistically significant $P < 0.0001$. In summary, ODA was able to discriminate between the (non)hospitalized patient groups on the basis of their score on the PRA test.

The ROC graph below shows a comparison of the ODA results to that of the three other methods.[12-14] ODA and Youden derived identical estimates, and Liu and the nearest [0,1] approach derived identical estimates. The latter methods derived a cut-point of 0.2435, sensitivity at the cut-point of 0.5842, and specificity at the cut-point of 0.6316.



Discussion

This paper demonstrates how ODA can be used to identify the optimal cut-point on a diagnostic test to maximally discriminate between those with and without the disease (or any other outcome).

ODA should be considered the preferred approach over other methods because ODA avoids many of the statistical assumptions required of conventional models, is insensitive to skewed data or outliers, and has the ability to handle any variable metric including categorical, Likert-type integer, and real number measurement scales. Moreover, in contrast to other methods, ODA also has the distinct ability to ascertain where the optimal (i.e., maximum-accuracy) cutpoints are on the outcome variable, which in turn, facilitates the use of measures of predictive accuracy. Moreover, ODA can perform cross-validation using LOO which allows for assessing the cross-generalizability of the model to potentially new study participants or non-participants.[18]

Finally, the findings continue to support our recommendation to employ the ODA and CTA frameworks to evaluate the efficacy of health-improvement interventions and policy initiatives.[19-39]

References

[1]Linden A (2020). Implementing ODA from Within Stata: An Application to Data From a Randomized Controlled Trial (*Invited*). *Optimal Data Analysis,* 9, 9-13.

[2]Linden A (2020). Implementing ODA from Within Stata: Implementing ODA from Within Stata: An Application to Estimating Treatment Effects using Observational Data (*Invited*). *Optimal Data Analysis,* 9, 14-20.

[3]Linden A (2020). Implementing ODA from Within Stata: An Application to Dose-Response Relationships (*Invited*). *Optimal Data Analysis*, 9, 26-32.

[4]Linden A (2020). Implementing ODA from Within Stata: Assessing Covariate Balance in Observational Studies (*Invited*). *Optimal Data Analysis*, 9, 33-38.

[5]Linden A (2020). Implementing ODA from Within Stata: Evaluating Treatment Effects for Survival (Time-to-Event) Outcomes (*Invited*). *Optimal Data Analysis,* 9, 39-44.

[6]Linden A (2020). Implementing ODA from Within Stata: Evaluating Treatment Effects in Multiple-Group Interrupted Time Series Analysis (Invited). *Optimal Data Analysis*, 9, 45-50.

2155-0182/10/$3.00

[7]Linden A (2020). Implementing ODA from Within Stata: Identifying Structural Breaks in Single-Group Interrupted Time Series Designs (Invited). *Optimal Data Analysis*, 9, 51-56.

[8]Linden A (2020). ODA: Stata module for conducting Optimal Discriminant Analysis. *Statistical Software Components S458728, Boston College Department of Economics*.

[9]Yarnold PR, Soltysik RC (2016). *Maximizing Predictive Accuracy*. Chicago, IL: ODA Books. DOI: 10.13140/RG.2.1.1368.3286

[10]Linden A, Yarnold PR (2018). Identifying maximum-accuracy cut-points for diagnostic indexes via ODA. *Optimal Data Analysis*, 7, 59-65.

[11]Linden A (2006). Measuring diagnostic and predictive accuracy in disease management: an introduction to receiver operating characteristic (ROC) analysis. *Journal of Evaluation in Clinical Practice*, 12, 132-139.

[12]Perkins NJ, Schisterman EF (2006). The inconsistency of "optimal" cut-points using two ROC based criteria. *American Journal of Epidemiology*, 163, 670–675.

[13]Liu X. Classification accuracy and cut point selection (2012). *Statistics in Medicine*, 31, 2676–2686.

[14]Youden WJ (1950). Index for rating diagnostic tests. *Cancer*, 3, 32–35.

[15]Linden A, Adams J, Roberts N (2004). Evaluating disease management program effectiveness: an introduction to survival analysis. *Disease Management*, 7, 180-190.

[16]Boult C, Dowd B, McCaffrey D, et al. (1993). Screening elders for risk of hospital admission. Journal of the American Geriatric Society, 41, 811–817.

[17]Yarnold PR, Soltysik RC (2016). *Maximizing Predictive Accuracy*. Chicago, IL: ODA Books. DOI: 10.13140/RG.2.1.1368.3286

[18]Linden A, Adams J, Roberts N (2004). The generalizability of disease management program results: getting from here to there. *Managed Care Interface*, 17, 38-45.

[19]Linden A, Adler-Milstein J (2008). Medicare disease management in policy context. *Health Care Finance Review*, 29, 1-11.

[20]Linden A, Roberts N (2005). A Users guide to the disease management literature: recommendations for reporting and assessing program outcomes. *American Journal of Managed Care*, 11, 81-90.

[21]Linden A, Yarnold PR (2017). Using classification tree analysis to generate propensity score weights. *Journal of Evaluation in Clinical Practice, 23*, 703-712.

[21]Linden A, Yarnold PR, Nallomothu BK (2016). Using machine learning to model dose-response relationships. *Journal of Evaluation in Clinical Practice, 22*, 860-867.

[21]Yarnold PR, Linden A (2016). Theoretical aspects of the D statistic. *Optimal Data Analysis*, 22, 171-174.

[26]Linden A, Yarnold PR (2017). Modeling time-to-event (survival) data using classification tree analysis. *Journal of Evaluation in Clinical Practice, 23*, 1299-1308.

[27]Linden A, Yarnold PR (2018). Identifying causal mechanisms in health care interventions using classification tree analysis. *Journal of Evaluation in Clinical Practice, 24*, 353-361.

[28]Linden A, Yarnold PR (2018). Estimating causal effects for survival (time-to-event) outcomes by combining classification tree analysis and propensity score weighting. *Journal of Evaluation in Clinical Practice, 24*, 380-387.

2155-0182/10/$3.00

[29]Linden A, Yarnold PR (2016). Using machine learning to assess covariate balance in matching studies. *Journal of Evaluation in Clinical Practice*, 22, 848-854.

[30]Linden A, Yarnold PR (2016). Combining machine learning and propensity score weighting to estimate causal effects in multivalued treatments. *Journal of Evaluation in Clinical Practice*, *22*, 875-885.

[31]Linden A, Yarnold PR (2016). Using machine learning to identify structural breaks in single-group interrupted time series designs. *Journal of Evaluation in Clinical Practice*, *22*, 855-859.

[32]Linden A, Yarnold PR (2016). Combining machine learning and matching techniques to improve causal inference in program evaluation. *Journal of Evaluation in Clinical Practice*, *22*, 868-874.

[33]Linden A, Yarnold PR (2017). Minimizing imbalances on patient characteristics between treatment groups in randomized trials using classification tree analysis. *Journal of Evaluation in Clinical Practice*, *23*, 1309-1315.

[34]Linden A, Yarnold PR (2018). Using machine learning to evaluate treatment effects in multiple-group interrupted time series analysis. *Journal of Evaluation in Clinical Practice*, *24*, 740-744.

[35]Yarnold PR, Linden A (2017). Computing propensity score weights for CTA models involving perfectly predicted endpoints. *Optimal Data Analysis*, *6*, 43-46.

[36]Yarnold PR, Linden A (2016). Using machine learning to model dose-response relationships via ODA: Eliminating response variable baseline variation by ipsative standardization. *Optimal Data Analysis*, *5*, 41-52.

[37]Linden A, Yarnold PR (2018). The Australian gun buy-back program and the rate of suicide by firearm. *Optimal Data Analysis*, *7*, 28-35.

[38]Linden A, Yarnold PR (2018). Using ODA in the evaluation of randomized controlled trials. *Optimal Data Analysis*, *7*, 46-49.

[39]Linden A, Yarnold PR (2018). Using ODA in the evaluation of randomized controlled trials: Application to survival outcomes. *Optimal Data Analysis*, *7*, 50-53.

Author Notes

No conflict of interest was reported.