Optimal Data Analysis
Vol. 7 (November 22, 2018), 85-89

2155-0182/10/$3.00

Visualizing Application and Summarizing Accuracy of ODA Models

Paul R. Yarnold, Ph.D.
Optimal Data Analysis, LLC

This note illustrates visualizing an ODA optimal cutpoint used to classify observations in training or validity samples, and summarizing resulting accuracy using the confusion matrix and PAC, ESS and D indexes.

Exposition first addresses problems in which the attribute—known as a "dependent measure" in legacy methods[1,2]—is assessed using an *ordered* (ordinal, interval, or ratio) measurement scale, and then considers problems with the attribute assessed using a *nominal* measurement scale having two (binary) or more (multicategorical) qualitative class categories.[3-9]

Ordered Attribute

This example features data from a simulation evaluating the ability of an algorithm designed by the author to identify initial solitaire[10] hands which will win (i.e., more than ten cards-on-top, class=1) *vs*. lose (class=0). Table 1 gives results for 88 randomized hands—in each of which all moves were explored to explicitly maximize the final score. As seen, the algorithm was always correct when it predicted a hand would win: all predicted class 1 hands returned a score greater than ten (red font highlights the winning class=0 hands). Presently, no matter if the objective is to maximize overall percentage of accurate classification (PAC; 0=0% correct, 100=100% correct classification), or effect strength for sensitivity (ESS; 0=level of accuracy which is expected by

Table 1: Solitaire Simulation Results

Predicted Class	Score	Number
0	0	5
0	1	6
0	2	4
0	3	5
0	4	11
0	5	11
0	6	4
0	7	9
0	8	1
0	9	1
0	10	4
0	12	3
1	13	1
1	15	1
1	18	1
0	20	1
0	21	1
0	52	11
1	52	8

Optimal Data Analysis
Vol. 7 (November 22, 2018), 85-89

2155-0182/10/$3.00

chance, 100=perfect classification)[3-6,11,12], the identical ODA model is obtained[13-16]:

if Score≤12.5 then predict Class=0;

or

if Score>12.5 then predict Class=1.

Table 2 is the confusion table summarizing the accuracy achieved when this ODA model is applied to classify all 88 observations (hands) in the so-called *training* sample:

Table 2: Confusion Table for Solitaire Model: Training (Total Sample) Analysis

		Predicted Score ≤10	>10	Sensitivity
Actual	≤10	64	13	83.12
Score	>10	0	11	100.00
Predictive Value		100.00	45.83	

Considering overall accuracy the model correctly classified 64+11=75 of the total of 88 hands, yielding PAC=85.23%, $p<0.0001$. Considering accuracy normed against chance[17] the model correctly classified 83.12% of the losing hands and 100% of the winning hands, a strong[5] effect (ESS=83.12, $p<0.0001$): D=0.41 indicates that less than one more attribute with equivalent ESS is needed to obtain a perfect model.[18,19]

Table 3 is the confusion table summarizing accuracy that is achieved when the ODA model is used to classify observations in leave-one-out (LOO) single-sample jackknife *cross-generalizability* analysis:

Table 3: Confusion Table for Solitaire Model: LOO (Generalizability) Analysis

		Predicted Score ≤10	>10	Sensitivity
Actual	≤10	64	13	83.12
Score	>10	1	10	90.91
Predictive Value		98.46	43.48	

Considering overall accuracy the model correctly classified 74 of the total of 88 hands yielding PAC=84.09% ($p<0.0001$), and considering accuracy normed against chance the model correctly classified 83.12% of losing hands and 90.91% of winning hands yielding a strong effect: ESS=70.97 ($p<0.0001$), D=0.82.

Figure 1: Visualizing Training ODA Model for Ordered Attribute (Red=Misclassified)

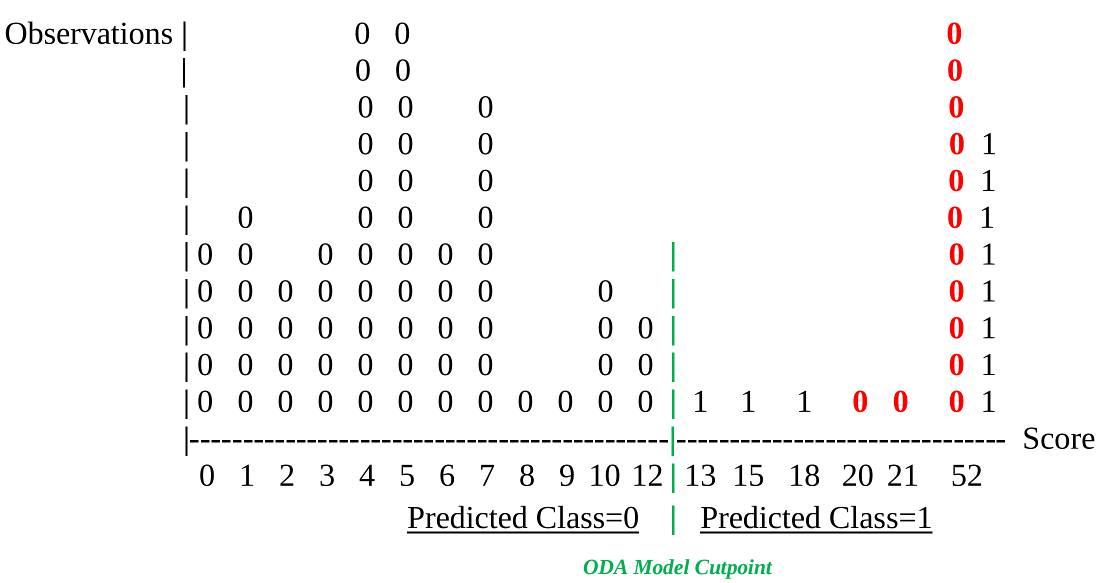

Figure 1 is a visual representation of the distribution of class=1 (predicted to win by the solitaire algorithm) and class=0 (not predicted to win) hands (observations): every "1" in the Figure represents one class=1 hand and each "0" is one class=0 hand. **Green** font highlights the cutpoint derived in training analysis: all hands which were misclassified in training analysis are indicated in **red** font. The predicted class 1 hand that was correctly classified in training analysis but misclassified in LOO analysis is the score of 13 located just to the right of the ODA cutpoint.

In its current configuration the solitaire algorithm returned strong performance: all the losing hands yielding ten or fewer points were not predicted to win, while 11 (45.8%) of the 24 winning hands were predicted to win. However, the current algorithm misclassified three hands *left* of the ODA model cutpoint yielding a win (12 points) as being members of class 0, and it also misclassified 13 winning hands *right* of the cutpoint. Training and LOO analyses found D<1 thus indicating that including one more attribute (e.g., number of aces, or of redundant color-plus-value active cards on the initial deal) may yield a perfectly accurate, minimum-complexity CTA model.[3,6,20-34]

Nominal (Multicategorical) Attribute

This example tests Foa's *a priori* hypothesis on the exchange of psychological and material resources of six qualitative types: love, status, information, money, goods and services.[35] By placing the six learning styles into a hexagonal pattern within a two-dimensional space formed by orthogonal intersection of two meta-dimensions (particularism and concreteness), Foa's similarity hypothesis posits that responses to a message reflecting a given type are more similar to that type than to other types, requiring two ODA analyses—one travelling in the clockwise direction around the hexagon (Love → Services → Goods → Money → Information → Status), and the other in the counter-clockwise direction (Love → Status → Information → Money → Goods → Services). While these data could not be analyzed successfully by parametric methods due to issues relating to structural zeros in the major diagonal and marginal imbalance arising from skewed data, ODA offered statistically significant support for the *a priori* hypothesis: ESS=37.45, D=10.02, PAC=66.67, *p*'s<0.0001.[5]

References

[1]Grimm LG, Yarnold PR (Eds.). *Reading and Understanding Multivariate Statistics*. Washington, DC: APA Books, 1995.

[2]Grimm LG, Yarnold PR (Eds.). *Reading and Understanding More Multivariate Statistics*. Washington, DC: APA Books, 2000.

[3]Yarnold PR (2017). What is optimal data analysis? *Optimal Data Analysis*, *6*, 26-42.

[4]Yarnold PR, Soltysik RC (1991). Theoretical distributions of optima for univariate discrimination of random data. *Decision Sciences*, *22*, 739-752.

[5]Yarnold PR, Soltysik RC (2005). *Optimal data analysis: A guidebook with software for Windows*, Washington, DC, APA Books.

[6]Yarnold PR, Soltysik RC (2016). *Maximizing predictive accuracy*. Chicago: ODA Books.

[7]Yarnold PR (2013). Univariate and multivariate analysis of categorical attributes with many response categories. *Optimal Data Analysis*, *2*, 177-190.

[8]Yarnold PR (2018). Minimize usage of binary measurement scales in rigorous classical research. *Optimal Data Analysis*, *7*, 3-9.

[9]Telles CR (2018). Geometrical information flow regulated by time lengths: an initial approach. *Symmetry*, *10*, 645.
DOI: 10.3390/sym10110645

2155-0182/10/$3.00

[10]Dawson J (2018). *Solitaire: The ultimate guide to mastering the solitaire card game in 30 minutes or less!* Kindle Edition.

[11]Yarnold PR (2016). Identifying the descendant family of HO-CTA models by using the minimum denominator selection algorithm: Maximizing ESS versus PAC. *Optimal Data Analysis*, *5*, 53-57.

[12]Yarnold PR (2016). Pruning CTA models to maximize PAC. *Optimal Data Analysis*, *5*, 58-61.

[13]The ODA[5] and MegaODA[14-16] code used to do this analysis was:

```
OPEN solitaire.dat;
VARS predict score;
CLASS predict;
ATTR score;
DIR < 0 1;
MCARLO ITER 25000;
LOO;
GO;
TITLE ess analysis;
PRIORS OFF;
GO;
TITLE pac analysis;
```

[14]Soltysik RC, Yarnold PR (2013). MegaODA large sample and BIG DATA time trials: Separating the chaff. *Optimal Data Analysis*, 2, 194-197.

[15]Soltysik RC, Yarnold PR (2013). MegaODA large sample and BIG DATA time trials: Harvesting the Wheat. *Optimal Data Analysis*, 2, 202-205.

[16]Yarnold PR, Soltysik RC (2013). MegaODA large sample and BIG DATA time trials: Maximum velocity analysis. *Optimal Data Analysis*, 2, 220-221.

[17]The formula for computing effect strength for sensitivity is: ESS=100x(Mean Sensitivity Over Classes–c*)/(100–c*), where c*=100/Number of Class Categories.[5,6] In problems involving an unweighted two-category class variable and a single ordered attribute, the area under the ROC curve (from signal detection literature) is equal to ESS/100.[18,19] The formula for computing the D statistic is: D=100/(ESS/Strata)–Strata, where Strata is the number of endpoints in the model (here Strata=2 for PAC and ESS models).[6,20]

[18]Yarnold PR (2014). UniODA vs. ROC analysis: Computing the “optimal” cut-point. *Optimal Data Analysis*, *3*, 117-120.

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

[20]Yarnold PR, Linden A (2016). Theoretical aspects of the D statistic. *Optimal Data Analysis*, *5*, 171-174.

[21]Yarnold PR (1996). Discriminating geriatric and non-geriatric patients using functional status information: An example of classification tree analysis via UniODA. *Educational and Psychological Measurement*, *56*, 656-667.

[22]Yarnold PR (1996). Characterizing and circumventing Simpson’s paradox for ordered bivariate data. *Educational and Psychological Measurement*, *56*, 430-442.

[23]Linden A, Yarnold PR (2016). Using data mining techniques to characterize participation in observational studies. *Journal of Evaluation in Clinical Practice*, *22*, 839-847.

[24]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.

[25]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.

[26]Yarnold PR, Linden A (2016). Novometric analysis with ordered class variables: The optimal alternative to linear regression analysis. *Optimal Data Analysis*; *6*, 65-73.

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

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

[29]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.

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

[31]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.

[32]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.

[33]Yarnold PR, Soltysik RC (2014). Discrete 95% confidence intervals for ODA model- and chance-based classifications. *Optimal Data Analysis*, *3*, 110-112.

[34]Yarnold PR (2018). Assessing statistical significance by comparing exact discrete 95% CIs for ESS and D of model *vs*. chance. *Optimal Data Analysis*, *7*, 82-84.

[35]Foa U (1971). Interpersonal and economic resources. *Science*, *171*, 345-351.

Author Notes

No conflict of interest was reported.