Optimal Data Analysis
Vol. 8 (January 17, 2019), 21-25

2155-0182/10/$3.00

The Structure of *Perfect* Optimal Models with a Two-Category Class Variable and Four or Fewer Endpoints

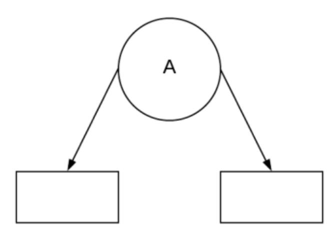
Paul R. Yarnold, Ph.D.
Optimal Data Analysis, LLC

An optimal model has a specific geometric configuration defined by the number of attributes ("independent variables"—schematically illustrated using circles) and endpoints (defined by response on attribute—indicated by rectangles). Branches direct attributes to endpoints via an if/then/else-based decision rule identified by the (ODA/CTA/novometric) algorithm and operationalized vis-à-vis numerical thresholds or categorical rosters which explicitly maximize (weighted) classification accuracy. In hopes of aiding in the visualization, pursuit and discovery of perfectly accurate statistical classification models, this paper presents schematic diagrams which correspond to combinations of number of attributes and endpoints that are possible for a range of optimal models commonly reported.[1]

Two-Strata Model

The least-granular model possible is a solution involving two endpoints, known as a two-strata model (Figure 1). Dummy-labeled "A" for clarity, the attribute may be ordered or (multi)categorical. Endpoints are defined by category (one for category X, one for category Y), or as being above *vs*. below a numerical threshold. It is not unusual to obtain perfect ODA models, that are stable in leave-one-out (LOO) jackknife analysis, in (interrupted) time-series applications.[1-4] Perfect two-strata models may occur in univariable analyses involving a single attribute[5], or in multivariable designs with multiple attributes.[6] When using CTA to create propensity scores, a special-purpose procedure is needed for endpoints with perfectly classified strata.[7] Also, it is possible that the exact discrete 95% confidence interval for accuracy yielded by a binary CTA model for one or both endpoints overlaps perfect classification.[8] However, binary attributes may be insufficiently sensitive (inadequately granular) to identify true maxima.[9]

Figure 1: One-Attribute Configuration Possible for Two-Strata Optimal Classification Model

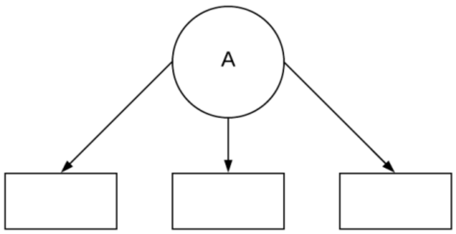


2155-0182/10/$3.00

Three-Strata Models

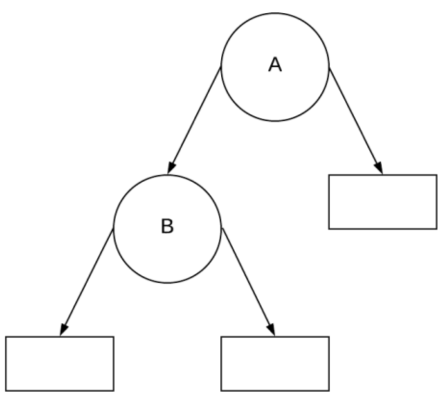
Linear and nonlinear findings of optimal models in applications with a binary class variable, a multicategorical or an ordered attribute, and a three-strata solution, are not uncommon.[10-14]

Figure 2: One-Attribute Configuration Possible for Three-Strata Optimal Classification Model



Among published optimal models for applications with a binary class variable and multiple categorical and/or ordered attributes, two-attribute three-strata CTA models (Figure 3) often reflect the best combination of *predictive accuracy* (assessed as ESS or PAC)[1] and *parsimony* (inhibiting model overfitting).[15,16] Attribute B may enter the model from either the left or right branch emanating from A.

Figure 3: Two-Attribute Configuration Possible for Three-Strata Optimal Classification Model



Four-Strata Models

Increasingly complex models involving more sample strata are sometimes optimal for the sample, and often are fertile hunting grounds for statistically-motivated (branch) sub-models that are productive in precision forecasting as well as model performance-boosting.[17] Four-strata models may involve one (Figure 4), two (Figures 5 and 6) or three attributes (Figure 7).

Figure 4: One-Attribute Configuration Possible for Four-Strata Optimal Classification Model

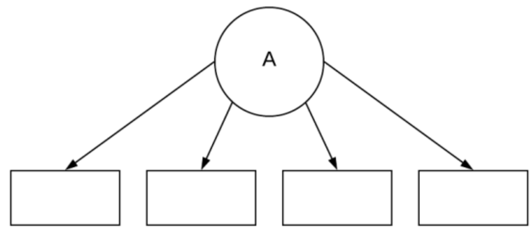


Figure 5: Two-Attribute Configuration Possible for Four-Strata Optimal Classification Model

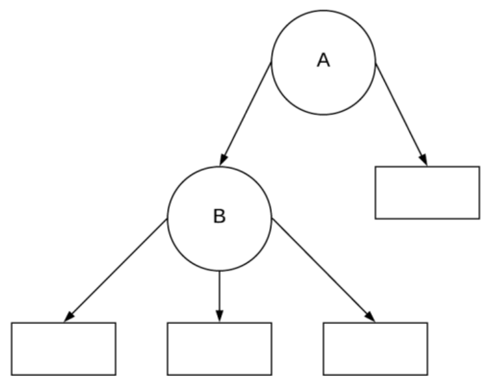
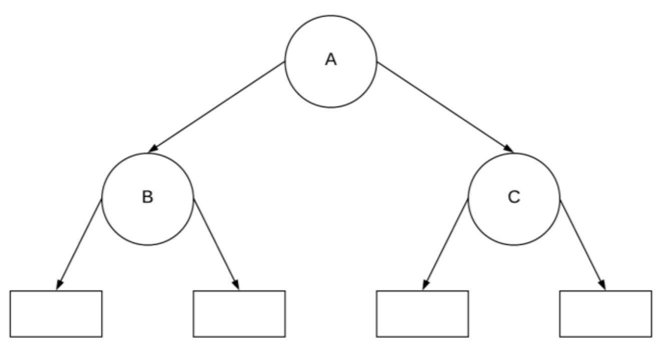
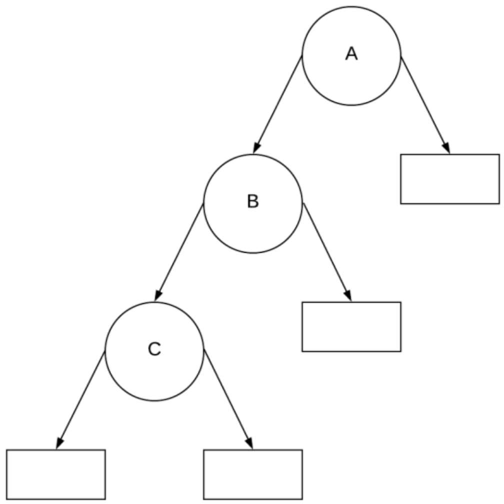


Figure 6: Two-Level Three-Attribute Configuration Possible for Four-Strata Optimal Classification Model



2155-0182/10/$3.00

Figure 7: Three-Level Three-Attribute Configuration Possible for Four-Strata Optimal Classification Model



In Figure 5 attribute B may enter the model from the left or right branch emanating from attribute A. In Figures 6 and 7 attribute C may enter the model from the left or the right branch emanating from attribute B.

Comments

Configurations illustrated here reflect structure typically identified by (occasionally perfect) optimal models obtained in primarily exploratory *research*.[17] Optimal exploratory *methods* have also been developed, such as the optimal analogue to principal components analysis, called structural decomposition analysis, which maximizes accuracy (ESS or PAC) rather than variance.[18] And, the minimum denominator selection algorithm[19] is employed to identify all of the unique, statistically viable optimal models with differing complexity and statistical power existing in the sample. Using ipsative standardization in serial designs limits or prevents paradoxical confounding.[20,21] The list goes on.[17]

By design, development of the ODA paradigm was predicated on providing unlimited flexibility to precisely, exactly model any *exploratory* or *confirmatory* hypothesis. Early examples abound, and include Foa's "spokes of a wheel" model of the exchange of psychological and material resources (p. 196) and "circumferential rotation of a wheel" model of temporal stability of learning style (p. 218); Loo's "rotating square" model of learning style transitions (p. 217); and Goodman's "rotating triangle" model of generalizable transitions in voting intentions.[10] Most recently, novometric models have been shown to outperform a host of legacy methods including general linear model, logistic regression, log-linear and Cox survival analysis, interrupted time series and Markov modelling, reliability and validity analyses, structural equation modeling, and a suite of causal inference methodologies.[22-32] The stage is set, the tools are available, for a host of new discoveries to be made using old as well as new data.

References

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

[2]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.

[3]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.

[4]Linden A, Yarnold PR (2018). The Australian Gun Buyback Program and rate of suicide by Firearm. *Optimal Data Analysis*, *7*, 28-35.

[5]Harvey RL, Roth EJ, Yarnold PR, Durham JR, Green D (1996). Deep vein thrombosis in stroke: The use of plasma D-dimer level as a screening test in the rehabilitation setting. *Stroke*, *27*, 1516-1520.

[6]Stalans LJ, Yarnold PR, Seng M, Olson DE, Repp M. (2004). Identifying three types of violent offenders and predicting violent recidivism while on probation: A classification tree analysis. *Law & Human Behavior*, *28*, 53-271.

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

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

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

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

[11]Yarnold PR (2010). UniODA *vs*. chi-square: Ordinal data sometimes feign categorical. *Optimal Data Analysis*, *1*, 62-65.

[12]Yarnold PR, Soltysik RC (2014). Globally optimal statistical classification models, I: Binary class variable, one ordered attribute. *Optimal Data Analysis*, *3*, 55-77.

[13]Bryant FB, Yarnold PR (2014). Finding joy in the past, present, and future: The relationship between Type A behavior and savoring beliefs among college undergraduates. *Optimal Data Analysis*, *3*, 36-41.

[14]Yarnold PR (2015). An example of nonlinear UniODA. *Optimal Data Analysis*, *4*, 124-128.

[15]Yarnold PR (2015). Distance from a theoretically ideal statistical classification model defined as the number of additional equivalent effects needed to obtain perfect classification for the sample. *Optimal Data Analysis*, *4*, 81-86.

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

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

[18]Yarnold PR (2015). Modeling intended and awarded college degree vis-à-vis UniODA-based structural decomposition. *Optimal Data Analysis*, *4*, 177-178.

[19]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.

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

[21]Yarnold PR, Soltysik RC (2013). Ipsative transformations are *essential* in the analysis of serial data. *Optimal Data Analysis*, *2*, 94-97.

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

[23]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.

2155-0182/10/$3.00

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

[25]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.

[26]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.

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

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

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

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

[31]Yarnold PR (2019). Maximum-precision Markov transition table: Successive daily change in closing price of a utility stock. *Optimal Data Analysis*, *8*, 3-10.

[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.

Author Notes

No conflict of interest was reported.