Optimal Data Analysis
Vol. 8 (April 10, 2019), 72-75

2155-0182/10/$3.00

ODA *vs*. χ^2, *r*, and τ: Trauma Exposure in Childhood and Duration of Participation in Eating-Disorder Treatment Program

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

This note illustrates the disorder and confusion attributable to analytic ethos whereby a smorgasbord of different statistical tests are used to test identical or parallel statistical hypotheses. Herein four classic methods are used for an application with a binary class (dependent) variable and an ordered attribute (independent variable) measured using a five-point scale. Legacy methods reach different conclusions—which is correct? In absolute contrast, for a given sample and hypothesis novometric analysis identifies every statistically viable model (models vary as functions of precision and complexity) which *reproducibly maximizes the predictive accuracy* for the sample.

Data in Table 1 are from a tutorial exposition involving estimating the likelihood of a patient remaining-in (Remain) *vs*. dropping-out (Drop) of an eating disorder treatment program as a function of the number of traumatic events experienced in childhood.[1]

Table 1: Data

	Number of Traumatic Events				
Class	0	1	2	3	4+
Drop	25	13	9	10	6
Remain	31	21	6	2	3

The study objective is to determine if participants who experienced more traumatic events during childhood are more or less likely to drop out of treatment. The association of class and the number of traumatic events was estimated by using four classic methods. Table 2 summarizes the findings.[1]

Table 2: Legacy Methods Tried: Training Analysis Results

Method	$p\leq$
Pearson Chi-Square	0.051
Agresti Chi-Square	0.016
Pearson Correlation	0.016
Kendall's Tau	0.040

Neither indices of effect strength nor findings of cross-generalizability analysis were reported, nor was violation *vs*. satisfaction of the method-specific distributional assumptions considered. Other legacy tests used to compare the two classes are *t*-test[2,3] and ANOVA[4-10], polychoric correlation[11,12], and discriminant function and logistic regression analysis.[13-25]

2155-0182/10/$3.00

In utter contrast to the maelstrom of the legacy tradition, the ODA algorithm—a marvel of efficiency—precisely finds the most accurate, reproducible solution(s) for any application and hypothesis.[26-29]

In this exposition the exploratory alternative hypothesis—Drop *vs*. Remain subjects may be discriminated on the basis of the number of traumatic events experienced in childhood—was evaluated in training (model development) analysis: the null hypothesis is Drop *vs*. Remain subjects can't be discriminated on this basis. The outcome (Remain *vs*. Drop) was treated as a two-category class variable, and the number of times eating/day (0-4) as an ordered attribute.

The ODA model was: if number of times eating/day≤1 predict group=Remain; otherwise predict group=Drop. Table 3 is the confusion table summarizing predictive accuracy achieved by the ODA training model.

Table 3: Training Classification Performance

Actual Category	Predicted Category: Drop	Remain	*Sensitivity*
Drop	25	38	39.68
Remain	11	52	82.54
Predictive Value	69.44	57.78	

Note that the model was accurate when classifying Remain observations (50% accuracy is expected by chance for a uniform random number), and that accuracy classifying the Drop observations was less than expected by chance. Overall accuracy achieved by the ODA model was statistically significant, $p<0.0292$. For this model the effect strength for sensitivity or ESS= 22.22%, indicating a relatively weak effect (by convention, ESS<25 is relatively weak; ESS<50 is moderate; ESS<75 is relatively strong, etc.).[26]

In novometric theory[28] the final (fourth) axiom is that model performance is evaluated in cross-generalizability (reproducibility) analysis, not in training (model development) analysis. Presently, cross-generalizability analysis using a one-sample ("leave one out" or LOO) jackknife analysis revealed that the classification accuracy achieved by the ODA model in training analysis was stable (unchanged), yielding $p<0.0050$.

In novometric research, model residuals (the misclassified observations) are of *utmost* importance—however this is NOT in deference to distributional assumptions required by classic methods, which sample data may violate.[30,31]

In the ODA paradigm the misclassified observations are considered to be the subsample for which more research is obviously needed. Successful future research should integrate an evolved theoretical orientation with accordingly more precise measurement and experimental design.[32] Iteration of this process constitutes a research methodology algorithm for efficient elimination of model residuals to an *a priori* specified criterion level of accuracy.

References

[1]Howell DC. Treatment of ordered variables in a contingency table. https://www.uvm.edu/~dhowell/methods/Supplements/OrdinalChiSq.html

[2]Yarnold PR (2014). UniODA *vs*. t-Test: Comparing two migraine treatments. *Optimal Data Analysis*, *3*, 6-8.

[3]Yarnold PR (2015). UniODA *vs*. *t*-Test: Low brain uptake of L-[11C]5-hydroxytryptophan in major depression. *Optimal Data Analysis*, *4*, 146-147.

[4]Yarnold PR, Brofft GC (2013). ODA range test *vs*. one-way analysis of variance: Comparing strength of alternative line connections. *Optimal Data Analysis*, *2*, 198-201.

[5]Yarnold PR (2013). ODA range test *vs*. one-way analysis of variance: Patient race and lab results. *Optimal Data Analysis*, *2*, 206-210.

[6]Yarnold, PR (2015). UniODA *vs*. doubly incomplete three-factor ANOVA: Production

2155-0182/10/$3.00

failure attributable to acid corrosion. *Optimal Data Analysis*, *4*, 165-167.

[7]Yarnold PR (2015). Reverse CTA vs. Five-Factor Factorial ANOVA: Purifying a crystalline product. *Optimal Data Analysis*, *4*, 184-185.

[8]Yarnold PR (2016). Novometrics *vs*. ODA *vs*. One-Way ANOVA: Evaluating comparative effectiveness of sales training programs, and the importance of conducting LOO with small samples. *Optimal Data Analysis*, *5*, 131-132.

[9]Yarnold PR (2016). Novometric analysis *vs*. MANOVA: MMPI codetype, gender, setting, and the MacAndrew Alcoholism scale. *Optimal Data Analysis*, *5*, 177-178.

[10]Yarnold PR (2018). ANOVA with one between-groups factor *vs*. novometric analysis. *Optimal Data Analysis*, *7*, 23-25.

[11]Yarnold PR (2016). Novometrics *vs*. polychoric correlation: Number of lambs born over two years. *Optimal Data Analysis*, *5*, 184-185.

[12]Yarnold PR (2016). Novometric *vs*. ODA reliability analysis *vs*. polychoric correlation with relaxed distributional assumptions: Inter-rater reliability of independent ratings of plant health. *Optimal Data Analysis*, *5*, 179-183.

[l3]Yarnold PR, Soltysik RC (1991). Refining two-group multivariable classification models using univariate optimal discriminant analysis. *Decision Sciences*, *22*, 1158-1164.

[14]Soltysik RC, Yarnold PR (1994). The Warmack-Gonzalez algorithm for linear two-category multivariable optimal discriminant analysis. *Computers and Operations Research*, *21*, 735-745.

[15]Yarnold PR, Hart LA, Soltysik RC (1994). Optimizing the classification performance of logistic regression and Fisher's discriminant analyses. *Educational and Psychological Measurement*, *54*, 73-85.

[16]Yarnold PR, Soltysik RC, McCormick WC, Burns R, Lin EHB, Bush T, Martin GJ (1995). Application of multivariable optimal discriminant analysis in general internal medicine. *Journal of General Internal Medicine*, *10*, 601-606.

[17]Yarnold PR, Soltysik RC, Lefevre F, Martin GJ (1998). Predicting in-hospital mortality of patients receiving cardiopulmonary resuscitation: Unit-weighted MultiODA for binary data. *Statistics in Medicine*, *17*, 2405-2414.

[18]Soltysik, R.C., & Yarnold, P.R. (2010). Two-group MultiODA: Mixed-integer linear programming solution with bounded *M*. *Optimal Data Analysis*, *1*, 31-37.

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

[20]Yarnold PR (2014). UniODA *vs*. logistic regression analysis: Serum cholesterol and coronary heart disease and mortality among middle aged diabetic men. *Optimal Data Analysis*, *3*, 17-18.

[21]Yarnold PR (2015). UniODA *vs*. logistic regression and Fisher's linear discriminant analysis: Modeling 10-year population change. *Optimal Data Analysis*, *4*, 139-145.

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

[23]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.

2155-0182/10/$3.00

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

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

[26]Yarnold PR, Soltysik RC (2005). *Optimal data analysis: Guidebook with software for Windows*. Washington, D.C.: APA Books.

[27]Yarnold PR, Soltysik RC (2010). Optimal data analysis: A general statistical analysis paradigm. *Optimal Data Analysis*, *1*, 10-22.

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

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

[30]Grimm LG, Yarnold PR, Eds. (1995). *Reading and understanding multivariate statistics*. Washington, D.C.: APA Books.

[31]Grimm LG, Yarnold PR, Eds. (2000). *Reading and understanding more multivariate statistics*. Washington, D.C.: APA Books.

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

Author Notes

No conflict of interest was reported.