Optimal Data Analysis
Vol. 5 (November 19, 2016), 225-228

2155-0182/10/$3.00

Novometric *vs*. Logit Analysis: Abortion Attitude by Religion and Time

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

Prior research[1] using logit analysis to model abortion attitude (oppose=0; favor=2) as a function of religion (Protestant=1; Catholic=2; Jewish=3; Other=4) and time (1972=72; 1978=78) found: "The best fitting model…has separate effects for being Catholic or non-Catholic and for being Protestant or non-Protestant. The categories of Jewish and Other have no separate effects and are implicitly grouped or collapsed together. The result is a religious trichotomy. …The odds on a favorable response are identical in both years: .89 for Protestants, .64 for Catholics, and 3.44 for Jews and Others" (pp. 70-72). For these data exploratory novometric analysis[2-28] predicting abortion attitude (class variable) as a function of religion (multicategorical attribute) and time (ordered attribute) identified a parsimonious, relatively weak model with stable classification training and LOO performance.

SAS™ code used to construct the data analyzed herein[1] is given in the Appendix. Novometric analysis identified a single two-strata model that had stable classification performance in LOO analysis: if religion=Protestant or Catholic predict attitude=oppose; otherwise predict attitude=favor. Table 1 is the confusion matrix for this model: relatively weak ESS=11.39, D=15.56, $p<0.001$).

Table 1: Optimal Model Confusion Matrix

		Predicted Attitude		
		Oppose	Favor	
Actual	Oppose	1464	63	95.9%
Attitude	Favor	1182	217	15.1%

As seen, the model accurately classified 19 in 20 people who oppose abortion, but only 1 in 7 people who favor abortion (50% sensitivity is expected by chance for each class category in two-category designs that do not use analytic weights[2,30-35]). The deficient performance of the model for predicting favorable ratings is likely attributable to the category "other", notorious for riddling the literature with paradoxically-confounded results.[2,36-38]

References

[1]Knoke D, Burke PJ (1980). *Log-linear models*. Beverly Hills, CA: Sage (pp. 47-48).

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

2155-0182/10/$3.00

[3]Yarnold PR, Linden A (2016). Novometric analysis with ordered class variables: The optimal alternative to linear regression analysis, *Optimal Data Analysis*, *5*, 65-73.

[4]Yarnold PR, Bennett CL (2016). Novometrics *vs*. correlation: Age and clinical measures of PCP survivors, *Optimal Data Analysis*, *5*, 74-78.

[5]Yarnold PR, Bennett CL (2016). Novometrics *vs*. multiple regression analysis: Age and clinical measures of PCP survivors, *Optimal Data Analysis*, *5*, 79-82.

[6]Yarnold PR (2016). Novometrics *vs*. regression analysis: Literacy, and age and income, of ambulatory geriatric patients. *Optimal Data Analysis*, *5*, 83-85.

[7]Yarnold PR (2016). Novometrics *vs*. regression analysis: Modeling patient satisfaction in the Emergency Room. *Optimal Data Analysis*, *5*, 86-93.

[8]Yarnold PR (2016). Matrix display of pairwise novometric associations for ordered variables. *Optimal Data Analysis*, *5*, 94-101.

[9]Yarnold PR, Batra M (2016). Matrix display of pairwise novometric associations for mixed-metric variables. *Optimal Data Analysis*, *5*, 104-107.

[10]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.

[11]Yarnold PR (2016). Parental smoking behavior, ethnicity, gender, and the cigarette smoking behavior of high school students. *Optimal Data Analysis*, *5*, 136-140.

[12]Yarnold PR (2016). Using gender of an imaginary rated smoker, and subject's gender, ethnicity, and smoking behavior to identify perceived differences in peer-group smoking standards of American high school students. *Optimal Data Analysis*, *5*, 141-143.

[13]Yarnold PR (2016). Novometric models of smoking habits of male and female friends of American college undergraduates: Gender, smoking, and ethnicity. *Optimal Data Analysis*, *5*, 146-150.

[14]Yarnold PR (2016). Predicting daily television viewing of senior citizens using education, age and marital status. *Optimal Data Analysis*, *5*, 151-152.

[15]Yarnold PR (2016). Novometric statistical analysis and the Pearson-Yule debate. *Optimal Data Analysis*, *5*, 162-165.

[16]Yarnold PR (2016). Comparing WAIS-R qualitative information for people 75 years and older, with *vs*. without brain damage. *Optimal Data Analysis*, *5*, 166-170.

[17]Yarnold PR (2016). Using novometrics to disentangle complete sets of sign-test-based multiple-comparison findings. *Optimal Data Analysis*, *5*, 175-176.

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

[19]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.

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

[21]Yarnold PR (2016). Comparing MMPI-2 *F-K* Index normative data among male and female

2155-0182/10/$3.00

psychiatric and head-injured patients, individuals seeking disability benefits, police and priest job applicants, and substance abusers. *Optimal Data Analysis*, *5*, 186-193.

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

[23]Yarnold PR (2016). Novometric analysis predicting voter turnout: Race, education, and organizational membership status. *Optimal Data Analysis*, *5*, 194-197.

[24]Yarnold PR (2016). Novometrics *vs*. Yule's Q: Voter turnout and organizational membership. *Optimal Data Analysis*, *5*, 198-199.

[25]Yarnold PR (2016). Novometric *vs*. recursive causal analysis: The effect of age, education, and region on support of civil liberties. *Optimal Data Analysis*, *5*, 200-203.

[26]Yarnold PR (2016). Novometric analysis *vs*. GenODA *vs*. log-linear model: Temporal stability of the association of presidential vote choice and party identification. *Optimal Data Analysis*, *5*, 204-207.

[27]Yarnold PR (2016). Novometric analysis *vs*. ODA *vs*. log-linear model in analysis of a two-wave panel design: Assessing temporal stability of Catholic party identification in the 1956-1960 SRC panels. *Optimal Data Analysis*, *5*, 208-212.

[28]Yarnold PR (2016). Novometric *vs*. logit *vs*. probit analysis: Using gender and race to predict if adolescents ever had sexual intercourse. *Optimal Data Analysis*, *5*, 218-222.

[29]Bryant FB, Harrison PR (2013). How to create an ASCII input data file for UniODA and CTA software. *Optimal Data Analysis*, *2*, 2-6.

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

[31]Linden A, Yarnold PR (In Press). Combining machine learning and propensity score weighting to estimate causal effects in multi-valued treatments. *Journal of Evaluation in Clinical Practice*.

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

[33]Linden A, Yarnold PR (In Press). Using machine learning to assess covariate balance in matching studies. *Journal of Evaluation in Clinical Practice*. DOI: 10.1111/jep.12538

[34]Linden A, Yarnold PR (In Press). Using data mining techniques to characterize participation in observational studies. *Journal of Evaluation in Clinical Practice*.

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

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

[37]Soltysik RC, Yarnold PR (2010). The use of unconfounded climatic data improves atmospheric prediction. *Optimal Data Analysis*, *1*, 67-100.

[38]Yarnold PR (2015). Estimating inter-rater reliability using pooled data induces paradoxical confounding: An example involving Emergency Severity Index triage ratings. *Optimal Data Analysis*, *4*, 21-23.

Author Notes

This study analyzed publically available data. No conflict of interest was reported.

Optimal Data Analysis
Vol. 5 (November 19, 2016), 225-228

2155-0182/10/$3.00

Appendix

SAS™ Code used to Construct (Reproduce[1]) the Data File for Analysis by ODA Software[2,29]

```
data real;
infile datalines;
input group row
column;
cards;
1 1 1
;
Data example;
Do n=1 to 460;
put '1 72 1';
end;
Do n=1 to 498;
put '1 72 0';
end;
Do n=1 to 424;
put '1 78 1';
end;
Do n=1 to 501;
put '1 78 0';
end;
Do n=1 to 147;
put '2 72 1';
end;
Do n=1 to 240;
put '2 72 0';
end;
Do n=1 to 151;
put '2 78 1';
end;
Do n=1 to 225;
put '2 78 0';
end;
Do n=1 to 41;
put '3 72 1';
end;
Do n=1 to 10;
put '3 72 0';
end;
Do n=1 to 23;
put '3 78 1';
end;
Do n=1 to 6;
put '3 78 0';
end;
Do n=1 to 65;
put '4 72 1';
end;
Do n=1 to 17;
put '4 72 0';
end;
Do n=1 to 88;
put '4 78 1';
end;
Do n=1 to 30;
put '4 78 0';
end;
Output;
Run;
```