Optimal Data Analysis
Vol. 5 (October 16, 2016), 141-143

2155-0182/10/$3.00

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

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

Novometric analysis is used to discriminate perceived peer-group standards for girls seeing boys smoke, *vs*. for boys seeing girls smoke (class variable), of 3,220 Anglo-American, 936 Mexican-American, and 723 Indian-American (multicategorical attribute) high-school students.[1] Subjects rated their opinion about boys seeing girls smoke, and about girls seeing boys smoke, using a three-point categorical ordinal scale (ordered attribute): approve, do not care, disapprove (coded using 3-1, respectively). Additional categorical attributes were nominal measures of gender and whether or not the subject smoked. The globally optimal model in this application selected only approval rating as an attribute (relatively weak ESS=21.4; D=7.3; $p<0.001$): 76.3% of ratings of girl smokers indicated disapproval, compared with 54.9% of boy smokers.

Data[1] were originally analyzed using disintegrated chi-square analysis[2] (legacy analyses recommended for this design are the log-linear model[3-10] or logistic regression analysis[3,4,9-17]), and findings were reported using undocumented chi-square analysis.[1,18]

It was concluded that: "For the Anglo-American group, differences between males and females responding "do not care" were… statistically significant. …Data for Mexican-American smokers parallel those for the Anglo-Americans; however a significantly lower proportion of the non-smokers (of both sexes) believe that girls disapprove of boys smoking. …Among Indian smokers, these differences between males and females are not significant" (p. 167)

Table 1 summarizes the descendant family of optimal models that was obtained for

2155-0182/10/$3.00

this application: model three has the lowest D statistic and thus is the globally-optimal (GO) model in this application.[10]

Table 1: Descendant Family of Optimal Models Predicting Perceived Peer-Group Standards for Boy *vs*. Girl High-School Smokers

Step	ESS	Strata	Efficiency	D	Minimum Endpoint N
1	23.73	6	4.0	19.3	43
2	23.62	3	7.9	9.7	721
3	21.42	2	10.7	7.3	3,347

Effects identified were relatively weak and highly uninspiring. The large D statistic for the GO model indicates the solution obtained is substantially distant from representing either a theoretically cohesive synthesis or pragmatically useful model of phenomena as they are presently conceptualized and measured.[9,10]

References

[1]Zagona SV (1967). Psycho-social correlates of smoking behavior and attitudes for a sample of Anglo-American, Mexican-American, and Indian-American high school students. In: Zagona SV (Ed.), *Studies and issues in smoking behavior*. Tucson, AZ: University of Arizona Press (pp. 157-180).

[2]Yarnold PR (2016). CTA *vs*. disintegrated chi-square: Integrated *vs*. piecemeal analysis. *Optimal Data Analysis*, *5*, 118-120.

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

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

[5]Yarnold PR (2015). UniODA-based structural decomposition *vs*. log-linear model: Statics and dynamics of intergenerational class mobility. *Optimal Data Analysis*, *4*, 179-181.

[6]Yarnold PR (2015). Modeling religious mobility by UniODA-based structural decomposition. *Optimal Data Analysis*, *4*, 192-193.

[7]Yarnold PR (2015). UniODA-based structural decomposition *vs*. legacy linear models: Statics and dynamics of intergenerational occupational mobility. *Optimal Data Analysis*, *4*, 194-196.

[8]Yarnold PR (2015). UniODA-based structural decomposition *vs*. legacy linear models: Statics and dynamics of intergenerational occupational mobility. *Optimal Data Analysis*, *4*, 194-196.

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

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

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

[12]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.

[13]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.

[14]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.

2155-0182/10/$3.00

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

[16]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.

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

[18]Yarnold PR (2016). ODA *vs*. undocumented chi-square: Clarity *vs*. confusion. *Optimal Data Analysis*, *5*, 121-123.

Author Notes

The study analyzed de-individuated data and was exempt from Institutional Review Board review. No conflict of interest was reported.

Mail: Optimal Data Analysis, LLC
6348 N. Milwaukee Ave., #163
Chicago, IL 60646