Optimal Data Analysis
Vol. 4 (December 24, 2015), 190-191

2155-0182/10/$3.00

GO-CTA *vs*. Logit Models: Gender and Desirability of Divorce

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

In prior reseach[1] a sample of N = 601 male and N = 783 female parents in unhappy relationships used a five-point Likert-type scale ranging from 0 (much better to divorce) to 4 (much worse to divorce) to respond to a survey item: "When a marriage is troubled and unhappy, do you think it is generally better for the children if the couple stays together or gets divorced?" No logit model achieved acceptable fit in modeling the responses of men and women on the desirability of divorce scale.[1] GO-CTA revealed that a single model underlies the data, yielding a statistically significant but relatively weak effect.

Response to the survey item was treated as an ordered attribute, *gender* was used as the class variable, and the minimum-denominator selection algorithm (MDSA) was used to identify the descendant family of all possible enumerated-optimal CTA (EO-CTA) models that exist in this application (the sample provided sufficient statistical power).[2] The unrestricted initial (most granular) model in the family was identified via the following CTA software[3] syntax:

```
OPEN divorce.dat;
OUTPUT divorce.out;
VARS gender response;
CLASS gender;
ATTRIBUTE response;
MC ITER 5000 CUTOFF .05 STOP 99.9;
PRUNE .05;
ENUMERATE;
GO;
```

The descendant family for a single attribute may also be identified manually using UniODA software[4] to conduct MDSA vis-à-vis hierarchically-optimal CTA.[5] The unrestricted initial model in the descendant family would be identified via UniODA using the CTA software syntax provided above, but substituting "MC ITER 25000;" for the three syntax lines that precede the GO command. Regardless whether CTA or UniODA software is used, only one optimal (maximum-accuracy) model exists for this application—illustrated in Figure 1. Model accuracy is statistically significant ($p < 0.0001$), but the effect is relatively weak: D = 12.3; ESS = 14.0; exact discrete 95% CI for ESS = 8.9 – 19.4 for the model, 0.18 – 7.3 for chance.[2]

2155-0182/10/$3.00

Figure 1: Globally Optimal (GO) Model Discriminating Gender

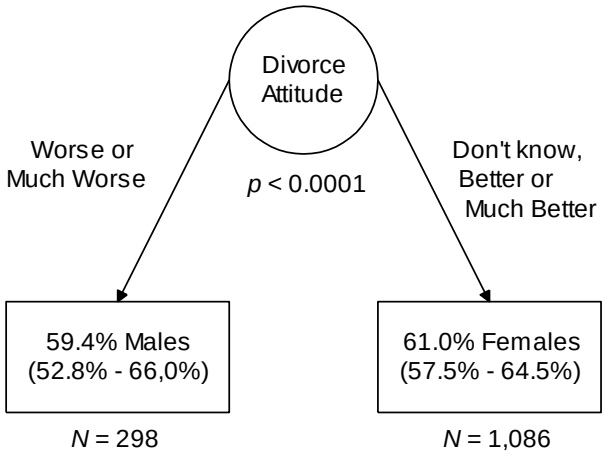

Table 1 presents the confusion table summarizing the predictive accuracy of the model.

Table 1: GO Model Predictive Accuracy

		Predicted Gender		
		F	M	Sensitivity
Actual	F	662	121	84.6%
Gender	M	424	177	29.4%
Predictive Value		61.0%	59.4%	

The model represents the responses of females in the sample well, and the responses of males in the sample very poorly. However, the point predictions made using the model yield mediocre predictive value for both genders.

References

[1]CC Clogg, ES Shihadeh (1994). Statistical models for ordinal variables. Thousand Oaks, CA: Sage (pp. 155-158).

[2]Yarnold PR, Soltysik RC (In Review). *Maximizing predictive accuracy*. Chicago, IL: ODA Books.

[3]Soltysik RC, Yarnold PR (2010). Automated CTA software: Fundamental concepts and control commands. *Optimal Data Analysis*, *1*, 144-160. URL: http://odajournal.com/2013/09/19/62/

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

[5]Yarnold PR, Bryant FB (2015). Obtaining a hierarchically optimal CTA model via UniODA software. *Optimal Data Analysis*, *4*, 36-53. URL: http://optimalprediction.com/files/pdf/V4A11.pdf

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
USA