Optimal Data Analysis
Vol. 8 (February 3, 2019), 48-52
Copyright 2019 by Optimal Data Analysis, LLC
2155-0182/10/$3.00

Multiple Regression *vs*. Novometric Analysis of a Contingency Table

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

This paper compares findings obtained using multiple regression analysis *vs*. novometric analysis to identify the relationship between the type of degree earned by 8th Grade math teachers and the training which they received in cultural and cognitive student diversity methods.

Data were obtained from the 1990 National Assessment for Educational Progress Mathematics 8th grade minifile.[1] Math teachers with a graduate degree were classified into one of three categories: undergraduate and graduate degree both in education (EE); undergraduate mathematics degree, graduate education degree (ME); or both degrees in mathematics (MM). Teachers were cross-classified with respect to if they received training to teach students of different cultures and/or possessing different cognitive learning styles: the results are presented in Table 1.

Table 1: Educational Background and Preparatory Courses for Student Diversity[1]

Educational Background	Diversity Preparatory Courses: Both	Cognitive	Cultural	None
EE	59	35	8	48
ME	23	44	2	25
MM	35	27	20	24

Multiple Regression Analysis

Analysis using a multiple regression-based approach[1] found: "teachers with undergraduate and graduate degrees in mathematics were significantly more likely to have received training in cultural diversity and less likely to have received training in cognitive diversity than those with graduate degrees in education" (pp. 86-87). This statistically significant effect accounted for 5.37% of shared variance.

A second statistically significant effect accounting for 2.77% of shared variance indicated that, among the teachers with graduate degrees in education, those with undergraduate degrees also in education were more likely to receive training in both cultural and cognitive disparity, compared to the teachers having undergraduate degrees in mathematics.

Novometric Comparison of Math Teachers' Educational Backgrounds

Exploratory novometric analysis was used to compare the educational background categories: all optimal effects reported herein had stable classification performance in training and LOO cross-generalizability analysis.[2-4]

The first analysis used ODA to directly contrast all three categories, treating background as an unordered (nominal) three-category class variable and the four diversity training catego-

Copyright 2019 by Optimal Data Analysis, LLC
2155-0182/10/$3.00

ries as nominal categorical attributes.[5] The ODA model was: if Training=Both or None then predict Background=EE; if Training=Cognitive then predict Background=ME; and if Training= Cultural then predict Background=MM. Thus, this model indicates math teachers having both undergraduate and graduate education degrees were the most likely to be trained in *both or neither* cultural and cognitive disparity; math teachers having undergraduate degree in math and graduate degree in education were most likely to be trained in cognitive disparity; and math teachers having both undergraduate and graduate degrees in math were most likely to be trained in cultural disparity. Classification performance achieved by the model (Table 2) is statistically significant (p<0.0001): as seen the level of accuracy achieved in classifying math teachers with an EE background compares well *vs*. chance—for which, if defined as a uniform random number, 50% accuracy is expected.[3] In contrast, the model performs slightly worse than expected by chance in classifying math teachers with ME backgrounds, and substantially worse for teachers with MM backgrounds. It is thus not surprising that the model ESS=18.50 shows this is a relatively weak[2] effect: D=13.22 is the additional number of attributes with equivalent mean ESS needed to obtain a perfect model.[6,7]

Table 2: Classification Performance of the Three-Category Comparison of Three Education Background Categories (*Sens*=Sensitivity)

		Predicted Category			
		EE	ME	MM	*Sens*
Actual	EE	107	35	8	71.33
Category	ME	48	44	2	46.81
	MM	59	27	20	18.87

For an omnibus model involving a class variable with more than two categories, and an attribute offering as many (or a few additional) possible response levels or categories (if ordered or qualitative, respectively), it is not unusual to obtain models in which excellent performance for one class category is obtained at the expense of poor performance for another, as a result of "insufficient room to vary" attributable to the combination of an imprecise attribute metric and a skewed response distribution.[3,8] Presently, the ODA model has 3 class categories, 4 response categories and a skewed response distribution.

The next three analyses also contrast all three class categories—however each analysis combines a different pair of categories.

The first analysis used ODA to contrast EE *vs*. combined ME and MM categories, treating background as an unordered (nominal) two-category class variable and the diversity training categories as nominal categorical attributes. The ODA model was: if Training=Both or None predict Background=EE; if Training=Cognitive or Cultural predict Background=ME+MM. This model indicates math teachers with both degrees in education were most likely to be trained in *both or neither* cultural and cognitive disparity, and teachers with any degree in math were most likely to be trained in cultural *or* cognitive disparity. Model classification performance (Table 3) is statistically significant (p<0.0026), but only the accuracy achieved in classifying math teachers with EE background compares well *vs*. chance: relatively weak ESS=17.83, D=9.22.

Table 3: Classification Performance of Two-Category Comparison of EE *vs*. (ME & MM) Education Background Categories

		Predicted Category		
		EE	ME, MM	*Sens*
Actual	EE	107	43	71.33
Category	ME, MM	107	93	46.50

The second analysis used ODA to contrast ME *vs*. combined EE and MM categories. The ODA model was: if Training=Cognitive predict Background=ME; if Training=Both, None or Cultural predict Background=EE+MM. This model indicates math teachers with an

Copyright 2019 by Optimal Data Analysis, LLC
2155-0182/10/$3.00

undergraduate degree in math and a graduate degree in education were the most likely to be trained in cognitive disparity, while teachers having an undergraduate education degree, or graduate math degree, were more likely to be trained in cultural disparity, and both cultural and cognitive disparity—or neither. Achieved classification performance (Table 4) is statistically significant ($p<0.0005$), but only the accuracy achieved in classifying math teachers with an undergraduate education degree or a graduate math degree compared well against chance: relatively weak ESS=22.59, D=6.85.

Table 4: Classification Performance of Two-Category Comparison of ME *vs*. (EE & MM) Education Background Categories

		Predicted Category		
		ME	EE, MM	*Sens*
Actual	ME	44	50	46.81
Category	EE, MM	62	194	75.78

The third analysis compared MM *vs*. EE and ME categories. The ODA model was: if Training=Cultural predict Background=MM; if Training=Both, None or Cognitive predict Background=EE+ME. The model indicates teachers with a math graduate degree were more likely to be trained in cultural disparity, and those with a graduate education degree were more likely to be trained in cognitive disparity, cultural and cognitive disparity, or neither. Classification performance of the model (Table 5) is statistically significant ($p<0.031$): accuracy in classifying teachers having a graduate degree in education was excellent, but otherwise accuracy was much worse than expected by chance: relatively weak ESS=14.77, D=11.54.

The final set of three analyses evaluated all of the pairwise comparisons among the three educational background categories.

The first pairwise analysis compared EE *vs*. ME class categories. The ODA model was: if Training=Cultural predict Background=ME,

Table 5: Classification Performance of Two-Category Comparison of MM *vs*. (EE & ME) Education Background Categories

		Predicted Category		
		EE, ME	MM	*Sens*
Actual	EE, ME	234	10	95.90
Category	MM	86	20	18.87

otherwise predict Background=EE. This model indicates teachers with an undergraduate degree in math and a graduate degree in education were more likely to be trained in cultural disparity, and teachers having both degrees in education were more likely to be trained in cognitive disparity, or in both—or in neither—cultural and cognitive disparity. Classification performance (Table 6) is statistically significant ($p<0.0008$), however only the accuracy classifying teachers with only education degrees exceeded chance: relatively weak ESS=23.48, D=6.52.

Table 6: Classification Performance of Two-Category Comparison of the EE *vs*. ME Education Background Categories

		Predicted Category		
		EE	ME	*Sens*
Actual	EE	115	35	76.67
Category	ME	50	44	46.81

The second pairwise analysis compared EE *vs*. MM teachers. The ODA model was: if Training=Both or Neither Cultural and Cognitive predict Background=EE; if Training= Cultural or Cognitive predict Background= MM. This model indicates teachers with both degrees in math were more likely to be trained in either cultural disparity or cognitive diversity, and teachers having both degrees in education were more likely to be trained in both, or in neither, cultural and cognitive disparity. Model classification performance (Table 7) is statistically significant ($p<0.040$): only teachers with education

Copyright 2019 by Optimal Data Analysis, LLC
2155-0182/10/$3.00

degrees were classified with accuracy exceeding chance: relatively weak ESS= 15.67, D=10.76.

Table 7: Classification Performance of Two-Category Comparison of the EE *vs*. ME Education Background Categories

		Predicted Category		
		EE	ME	*Sens*
Actual	EE	107	43	71.33
Category	ME	59	47	44.34

The final pairwise analysis compared the ME *vs*. MM class categories. The ODA model was: if Training=Cognitive or None predict Background=ME; if Training=Cultural or Both predict Background=MM. This model indicates teachers having both degrees in math were more likely to be trained in cultural disparity or both cognitive and cultural disparity, while teachers having an undergraduate degree in math and a graduate degree in education were more likely to be trained in cognitive disparity or in neither cognitive nor cultural disparity. Model classification performance (Table 8) is statistically significant ($p<0.0014$), however only the accuracy achieved in classifying teachers with a graduate degree in education exceeded accuracy expected by chance: *moderate* ESS=25.29, D=5.91.

Table 8: Classification Performance of Two-Category Comparison of the ME *vs*. MM Education Background Categories

		Predicted Category		
		ME	MM	*Sens*
Actual	ME	69	25	73.40
Category	MM	51	55	51.89

Comments

An advantage of novometric analysis, clearly illustrated herein, is the ability to identify every unique statistically significant model which exists in the sample: the family of such models is known as the descendant family.[3,9]

The model in the descendant family that achieved the greatest ESS (predictive accuracy adjusted for chance) in this study emerged for the pairwise comparison of ME *vs*. MM math teachers. This was the only model in this study which achieved *moderate* effect strength: all other reported effects were *relatively weak*. It is interesting that the strongest effect identified by multiple regression analysis included a portion of this maximum-accuracy ODA model (MM teachers have more training in cultural disparity *vs*. EE teachers): that a regression model missed the optimal solution is not surprising, if for no other reason than by formulation such models intrinsically regress toward the mean and thus inherently return suboptimal ESS.[10,11]

The model in the descendant family that achieved lowest D (ESS adjusted for parsimony) in this study similarly emerged for the pairwise comparison of the ME *vs*. MM teachers. Taken as a whole these findings suggest that a reliable moderate difference exists in the training of ME *vs*. MM math teachers, meriting more research. However it appears little is to be gained vis-à-vis replicating the present methodology: more specific and more granular measures of possible points of training differences are needed.

References

[1]Beasley TM, Schumacker RE (1995). Multiple regression approach to analyzing contingency tables: Post hoc and planned comparison procedures. *The Journal of Experimental Education*, *64*, 79-93.

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

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

Copyright 2019 by Optimal Data Analysis, LLC
2155-0182/10/$3.00

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

[5]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.

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

[7]Yarnold PR (2019). The structure of *perfect* optimal models with a two-category class variable and four or fewer endpoints. *Optimal Data Analysis*, *8*, 21-25.

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

[9]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.

[10]Linden A, Yarnold PR (2018). Comparative accuracy of a diagnostic index modeled using (optimized) regression *vs*. novometrics. *Optimal Data Analysis*, *7*, 66-71.

[11]Yarnold PR, Bryant FB, Soltysik RC (2013). Maximizing the accuracy of multiple regression models via UniODA: Regression *away from* the mean. *Optimal Data Analysis*, *2*, 19-25.

Author Notes

No conflict of interest was reported.