Novometric Theorem Generalized to Unrestricted Class Variables

Paul R. Yarnold, Ph.D.

Optimal Data Analysis, LLC

Novometric theory originally consisted of four axioms and applied to designs with a binary class variable and an unrestricted attribute.[1] This note adds a fifth axiom generalizing novometrics to designs involving unrestricted class variables.

Novometric statistical analysis is used to identify the descendant family of optimal models within which resides the globally optimal (GO) model for the sample. As originally postulated the novometric theorem applies to applications involving a binary class variable.[1] However, recent developments in optimal analysis of designs with ordered class variables[2-7] require the theorem to be generalized to designs that involve unrestricted class variables.

Binary Class Variable

As originally postulated the novometric theorem consists of four axioms:

1. Analysis is conducted for a sample of sufficient size to provide adequate statistical power to test the statistical hypothesis.

2. The minimum set of attributes that produce the GO model (which has the smallest D statistic in the descendant family) in the sample is identified by structural decomposition analysis.

3. The descendant family is identified by initiating analysis with an unrestricted CTA analysis using attributes identified in axiom 2, then applying the minimum denominator selection heuristic.

4. Model predictive accuracy is estimated based on cross-generalizability analysis, not on total sample (training) analysis.

Unrestricted Class Variable

The generalized novometric theorem consists of five axioms.

1. Initial (binary class variable) Axiom 1

2. Class variables measured on ordered or multicategorical scales are deconstructed into a set of binary dummy codes using the partitioning algorithm employed with time- and value-ordered series. Each of the constructed dummy class variables is subjected to novometric analysis.

3. Initial (binary class variable) Axiom 2

4. Initial (binary class variable) Axiom 3

5. Initial (binary class variable) Axiom 4

References

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

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

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

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

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

[6]Yarnold PR (2016). Novometrics *vs*. regression analysis: Modeling patient satisfaction in the Emergency Room. *Optimal Data Analysis*, *5*, 86-92.

[7]Yarnold PR (2016). Matrix display of pairwise novometric associations for ordered variables. *Optimal Data Analysis*, *5*, 93-101.

Author Notes

No conflict of interest was reported. By-by, Mom.

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