Optimal Data Analysis
Vol. 6 (January 4, 2017), 13-17

2155-0182/10/$3.00

Novometric Comparison of Patient Satisfaction with Nurse Responsiveness Over Successive Quarters

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

A segmentation printout or "banner pass" constitutes a consolidated crosstabulation of user/customer responses to specific prompts, for desired reporting periods and/or organizational units. In this study a total of 6,005 hospital patients rated their satisfaction with time for a nurse to respond to the call button (1=very dissatisfied; 2=somewhat dissatisfied; 3=neutral; 4=somewhat satisfied; 5=very satisfied).[1] Responses were consolidated into successive fiscal quarters (1=3rd Quarter, 1989; 2=4th Quarter, 1989; 3=1st Quarter, 1990; 4=2nd Quarter, 1990). Novometric analysis[2-11] is used to evaluate the statistical and ecological significance of satisfaction gains observed over time.

Data (N for every rating level and fiscal quarter) analyzed herein are presented in Table 1.

Table 1: Patient Satisfaction with Nurse Response Time to Call Button[1]

	Fiscal Quarter			
Satisfaction Rating	3rd (1989)	4th (1989)	1st (1990)	2nd (1990)
5	596	586	626	716
4	357	480	412	557
3	432	345	313	287
2	60	75	57	32
1	45	15	14	0

For novometric analysis patient satisfaction rating was treated as the ordered attribute, and fiscal quarter as the class variable.[2]

Consistent Quarterly Increase

The first analysis used optimal data analysis (ODA)[2,12-18] to test the confirmatory hypothesis that satisfaction increased consistently across all four quarters in the study time span: that is, satisfaction was lowest in the 3rd quarter of 1989, next-to-lowest in the 4th quarter of 1989, next-to-highest in the 1st quarter of 1990, and highest in the 2nd quarter of 1990 (using the coding scheme the command specifying this test in UniODA[12] or MegaODA[19-21] software was: DIR < 1 2 3 4;). The resulting ODA model was not statistically reliable (p< 0.99), and it yielded training predictive accuracy lower than expected by chance (ESS= -1.33).[2,12] The confirmatory hypothesis that satisfaction rating increased consistently across all four quarters was therefore clearly unsupported.

2155-0182/10/$3.00

Consistent Quarter-to-Quarter Increases

This analysis evaluated the confirmatory hypothesis that satisfaction increased consistently quarter-to-quarter over the study.[2,12,15] Three quarter-to-quarter comparisons were evaluated: (a) 3rd *vs*. 4th quarter of 1989; (b) 4th quarter of 1989 *vs*. 1st quarter of 1990; and (c) 1st *vs*. 2nd quarter of 1990. Each binary class variable was modeled using exploratory-optimal classification tree analysis (EO-CTA).[2-7,22-27] The first (a) quarter-to-quarter comparison was not statistically reliable ($p<0.49$) and yielded very weak normed predictive accuracy (ESS= 0.96); the second (b) comparison was not statistically reliable ($p<0.88$) and yielded accuracy comparable to chance (ESS= -0.01); and the third (c) comparison was not statistically reliable ($p<0.99$) and yielded accuracy marginally worse than chance (ESS= -0.95). Therefore, the confirmatory hypothesis that satisfaction rating increased consistently in quarter-to-quarter comparisons was clearly not supported.

Increased Satisfaction over Quarters

Testing the confirmatory hypothesis that patient satisfaction increased at some point (to be identified by the novometric algorithm) over the study time span requires modeling quarter as an ordered class variable. This is accomplished using the partitioning algorithm[5-7] (PA) to create all possible binary parses of the class variable, and then conducting recursive EO-CTA using the novometric algorithm for each parse.[28] In this application the PA constructed three binary parses: (a) 3rd quarter of 1989=class 0, other quarters=class 1; (b) 3rd and 4th quarters of 1989 =class 0, other quarters= class 1; and (c) 3rd and 4th quarters of 1989, and 1st quarter of 1990= class 0, 2nd quarter of 1990=class 1.

Novometric theory asserts that—for a given application (the intersection of hypothesis, measures, and sample), statistical analysis may (a) fail to identify a reliable model (which has been the case thus far, presently); (b) identify one reliable model; or (c) identify two or more reliable models. The descendant family (DF) consists of all (one or more) optimal models identified for the application.[2,29] Models in the DF differ with respect to predictive accuracy normed *vs*. chance by the ESS statistic (0=the predictive accuracy expected by chance; 100= errorless prediction), and complexity defined as the number of sample strata (endpoints) in the model. Novometric analysis identifies the globally optimal (GO) model in the DF representing the "best" combination of predictive accuracy and parsimony (the antithesis of complexity, defined as ESS divided by the number of model strata). The distance of an empirical model from a theoretically ideal model (that achieves perfect accuracy with maximum possible parsimony for the application) is computed using the D statistic: normed over accuracy and parsimony, D facilitates direct comparison of model quality by evaluating every model relative to its corresponding theoretical ideal.[2,8]

In novometric theory, for a given application, in addition to point estimates of model ESS and D, exact discrete 95% confidence intervals (CIs) are developed for the model as well as for chance (vis-à-vis bootstrap and Monte Carlo simulation, respectively): if the 95% CIs for model and for chance overlap, then the model is judged to be statistically unreliable.[2,30]

The final Axiom of novometric theory stipulates that while the training analysis drives model development, estimated reproducibility of model performance is based on validity analysis (e.g., jackknife, K-fold, holdout) so as to inhibit bias attributable to overfitting.[2,28]

In the present application EO-CTA identified a single optimal model for each of the three constructed parses, and all of these models had identical accuracy in training and leave-one-out jackknife analysis.[2,15]

The first parse compared satisfaction in the 3rd Quarter of 1989 *vs*. the subsequent three Quarters. The optimal model was: if Satisfaction Score$\leq$2 (somewhat dissatisfied) then predict

2155-0182/10/$3.00

the 3rd Quarter of 1989, otherwise predict subsequent Quarters (p< 0.0001). This model indicates that satisfaction was lower in the initial Quarter than in (combined) subsequent Quarters. Table 2 is the confusion matrix for this model in training and jackknife analysis (Sens= model sensitivity). As seen, the model correctly predicted actual class status of 1 of 14 observations in the 3rd Quarter of 1989, and of 19 of 20 observations in the subsequent three Quarters. The model yielded very weak ESS=2.77 (for 10,000 bootstrap iterations, exact discrete 95% CI for *model* ESS=1.08-4.50; for 10,000 Monte Carlo experiments, exact discrete 95% CI for *chance* ESS=0.01-1.26): because the 95% CIs for model and chance ESS overlap, this model is judged to be statistically unreliable. For this model D=70.2 (exact discrete 95% CI=42.4-183.2). [8]

Table 2: Confusion Matrix for EO-CTA Model: First Parse of Four-Quarter Series

		Predicted Quarter		
		Q3-1989	Later	Sens
Actual	Q3-1989	105	1385	7.1
Quarter	Later	193	4322	95.7

The second parse compared satisfaction in the 3rd and 4th Quarters of 1989 *vs*. the subsequent two Quarters. The optimal model was: if Satisfaction Score$\leq$4 (somewhat satisfied) then predict the 3rd or 4th Quarter of 1989, otherwise predict subsequent Quarters (p<0.0001). This model indicates that satisfaction was lower in the initial two Quarters than in (combined) subsequent Quarters. Table 3 gives the confusion matrix for this model in training and jackknife analysis. The model correctly predicted actual class status of 3 of 5 observations in the 3rd and 4th Quarters of 1989, and of 4 of 9 observations in the subsequent two Quarters (for a binary parse, 50% sensitivity is expected by chance for both class variable categories). The model had very weak ESS=5.01 (for 10,000 bootstrap iterations, exact discrete 95% CI for *model* ESS= 2.02-7.98; for 10,000 Monte Carlo experiments, exact discrete 95% CI for *chance* ESS=0.01-2.52): because 95% CIs for model and chance ESS overlap, this model is not statistically reliable. For this model D=37.9 (exact discrete 95% CI=23.1-97.0). Computed using point estimates, compared to the GO model for the first parse (D =70.2), the GO model for the second parse (D= 37.9) is [(70.2/37.9)-1]x100% =85.2% closer to its corresponding theoretically ideal counterpart.

Table 3: Confusion Matrix for EO-CTA Model: Second Parse of Four-Quarter Series

		Predicted Quarter		
		Q3-4/89	Later	Sens
Actual	Q3-4/89	1809	1182	60.5
Quarter	Later	1672	1342	44.5

The third (final) parse compared satisfaction in the 3rd and 4th Quarters of 1989, and 1st Quarter of 1990, *vs*. the 2nd Quarter in 1990. The optimal model was: if Satisfaction Score$\leq$3 (neutral) predict initial three Quarters, otherwise predict 2nd Quarter of 1990 (p<0.0001). This model indicates satisfaction was lower in the initial three Quarters than in the final Quarter of the study. Table 4 is the confusion matrix for this model in training and jackknife analysis. The model correctly predicted actual class status of 2 of 5 observations in the first three Quarters studied, and 4 of 5 observations in the final Quarter of the study. The model had relatively weak ESS=20.0 (for 10,000 bootstrap iterations, exact discrete 95% CI for *model* ESS=17.1-22.9; for 10,000 Monte Carlo experiments, exact discrete 95% CI for *chance* ESS=0.06-2.71): because the 95% CIs for model and chance ESS do not overlap, this model is judged as statistically reliable. For this model D=7.99 (exact discrete 95% CI= 6.75-9.69). Computed using point estimates, the GO model for the

2155-0182/10/$3.00

third parse is 778.6% closer to its corresponding theoretically ideal counterpart *vs*. the first parse, 374.3% closer *vs*. the second parse.

Table 4: Confusion Matrix for EO-CTA Model: Third Parse of Four-Quarter Series

		Predicted Quarter		
		Earlier	Q2-1990	Sens
Actual	Earlier	1768	2645	40.1
Quarter	Q2-1990	319	1273	80.0

Novometric analysis unambiguously identified the effect: over the time span of the study, patient satisfaction with nurse responsiveness showed a statistically significant, relatively weak increase only in the final Quarter. That is, the 95% CIs for ESS and D overlapped for GO models identified for the first and second parses, and these CIs lay outside the 95% CI for the GO model identified for the third parse. For the first and second parses the 95% CIs for chance and model overlapped—and thus the models were judged to be statistically unreliable, however for the third parse the 95% CIs for chance and model did not overlap—and the model was therefore judged to be statistically reliable. In contrast, the 95% CIs for chance overlapped for all three parses.

References

[1]Steiber SR, Krowinski WJ (1990). *Measuring and managing patient satisfaction*. Chicago, IL: American Hospital Publishing, Inc. (pp. 156-157).

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

[3]Yarnold PR, Soltysik RC (2014). Globally optimal statistical classification models, I: Binary class variable, one ordered attribute. *Optimal Data Analysis, 3*, 55-77.

[4]Yarnold PR, Soltysik RC (2014). Globally optimal statistical classification models, II: Unrestricted class variable, two or more attributes. *Optimal Data Analysis, 3*, 78-84.

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

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

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

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

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

[10]Yarnold PR (2016). Using novometric analysis of transition matrices to ascertain Markovian order. *Optimal Data Analysis, 6*, 4-7.

[11]Yarnold, P.R. (2016). Novometric comparison of Markov transition matrices for heterogeneous populations. *Optimal Data Analysis, 6*, 8-11.

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

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

2155-0182/10/$3.00

[14]Linden A, Yarnold PR (2016). Using machine learning to assess covariate balance in matching studies. *Journal of Evaluation in Clinical Practice, 6*, 848-854.

[15]Linden A, Yarnold PR (2016). Using machine learning to identify structural breaks in single-group interrupted time series designs. *Journal of Evaluation in Clinical Practice, 6,* 855-859.

[16]Linden A, Yarnold PR, Nallamothu BK (2016). Using machine learning to model dose-response relationships. *Journal of Evaluation in Clinical Practice, 6,* 860-867.

[17]Linden A, Yarnold PR (2016). Combining machine learning and matching techniques to improve causal inference in program evaluation. *Journal of Evaluation in Clinical Practice, 6,* 868-874.

[18]Linden A, Yarnold PR (2016). Combining machine learning and propensity score weighting to estimate causal effects in multivalued treatments. *Journal of Evaluation in Clinical Practice*, *6*, 875-885.

[19]Soltysik RC, Yarnold PR (2013). MegaODA large sample and BIG DATA time trials: Separating the chaff. *Optimal Data Analysis*, *2*, 194-197.

[20]Soltysik RC, Yarnold PR (2013). MegaODA large sample and BIG DATA time trials: Harvesting the Wheat. *Optimal Data Analysis*, *2*, 202-205.

[21]Yarnold PR, Soltysik RC (2013). MegaODA large sample and BIG DATA time trials: Maximum velocity analysis. *Optimal Data Analysis*, *2*, 220-221.

[22]Linden A, Yarnold PR (2016). Using data mining techniques to characterize participation in observational studies. *Journal of Evaluation in Clinical Practice*, *6*, 839-847.

[23]Yarnold PR (1996). Discriminating geriatric and non-geriatric patients using functional status information: An example of classification tree analysis via UniODA. *Educational and Psychological Measurement*, *56*, 656-667.

[24]Yarnold PR, Soltysik RC, Bennett CL (1997). Predicting in-hospital mortality of patients with AIDS-related *Pneumocystis carinii* pneumonia: An example of hierarchically optimal classification tree analysis. *Statistics in Medicine, 16*, 1451-1463.

[25]Soltysik RC, Yarnold PR (2010). Automated CTA software: Fundamental concepts and control commands. *Optimal Data Analysis, 1*, 144-160.

[26]Yarnold PR, Bryant FB (2015). Obtaining an enumerated CTA model via automated CTA software. *Optimal Data Analysis*, *4,* 54-60.

[27]Yarnold PR (2015). Optimal statistical analysis involving a confounding variable. *Optimal Data Analysis*, *4*, 87-103.

[28]Yarnold PR (2016). Novometric theorem generalized to unrestricted class variables. *Optimal Data Analysis*, *5*, 102-103.

[29]Yarnold PR (2016). Refined definition of the descendant family in novometric theory. *Optimal Data Analysis*, *5*, 153.

[30]Yarnold PR, Soltysik RC (2014). Discrete 95% confidence intervals for ODA model- and chance-based classifications. *Optimal Data Analysis*, *3*, 110-112.

Author's Notes

Analyzed data are publically available, and no conflict of interest was reported.