Optimal Data Analysis
Vol. 10 (March 21, 2021), 24-29

2155-0182/10/$3.00

Implementing ODA from Within Stata: Assessing Parallel-Forms Reliability Using a Binary and an Ordered Attribute

Paul R. Yarnold, Ph.D. and Ariel Linden, Dr.P.H.
Optimal Data Analysis, LLC and Linden Consulting Group, LLC

This paper illustrates testing a directional (i.e., confirmatory) hypotheses for a parallel-forms reliability study using a binary and an ordered measure, via the Stata package for implementing ODA.

Recent papers[1-24] introduce the new Stata package called **oda**[25] for implementing ODA from within the Stata environment. This package is a wrapper for the MegaODA software system[26-28], so the MegaODA.exe file must be loaded on the computer for the **oda** package to work.[29] To download the **oda** package, at the Stata command line type: "ssc install oda" (without the quotation marks).

According to classical test theory, parallel forms are alternative equivalent forms of a measuring instrument which measure the same attribute of the subject of measurement. Every person completing a pair of parallel forms is hypothesized to receive identical scores on both forms.[30]

This paper demonstrates use of the **oda** package to evaluate a directional hypothesis for a design in which two different procedures for assessing Type A Behavior (TAB) are employed to assess subjects within a sample.[31]

Methods

Data

Matthews et al. presented data for assessing agreement between two of the most frequently reported measures of TAB.[32] Considered the "gold standard" in assessment of TAB, the Structured Interview is a standardized clinical interview scored using a binary variable (Type A = 1; Type B = 0). The other procedure is a self-report questionnaire measure called the Jenkins Activity Survey or JAS.[31] While the JAS is designed to be scored using an ordered scale ranging between 0 and 16, Matthews et al. arbitrarily coded JAS scores: scores of 0-3 were coded as 1; scores of 4 and 5 were coded as 5; scores of 6 and 7 were coded as 7; scores of 8 and 9 were coded as 9; scores of 10 and 11 were coded as 11; and scores of 12-16 were coded as 16. This practice has been denounced because it can reduce classification accuracy.[33]

2155-0182/10/$3.00

Using an unpublished 4-point ordinal scoring protocol for the Structured Interview, Matthews et al. reported a Pearson correlation of r=0.31 (p<0.001) with the arbitrarily coded JAS scores. This suggests that scores on the modified Structured Interview and on the modified JAS share 9.61% of their variance without accounting for the effect of chance. A correlation of this weak magnitude returns extremely inaccurate classifications.[34]

Analytic Process

We test the directional ("confirmatory") alternative hypothesis that subjects scored as Type A (coded as 1) on the binary Structured Interview assessment will have greater scores on the modified JAS scale. The null hypothesis is this is not true.[30,35] Analysis was accomplished using the following **oda** syntax (see the help file for **oda** for a complete description of syntax options):

```
oda interv jas , pathoda("C:\ODA\")
store("C:\ODA\") iter(10000)
direction(< 0 1)   loo
```

This syntax is explained as follows. Here "interv" (the "gold-standard" measure of TAB) is the binary *class* variable and the modified "jas" score is the ordered *attribute*. "C:\ODA\" is the directory path where the MegaODA.exe file exists on the computer, and where other files generated in analysis are stored; 10,000 iterations (repetitions) are used to obtain a permutation p-value; the directional hypothesis is that the Type B's have lower JAS scores than Type A's; and a leave-one-out (LOO) cross-generalizability analysis is to be conducted.[25] The **oda** package produces an extract of the total output produced by ODA software (the complete output is stored in the specified directory with the extension ".out").

As seen in the output for this analysis, the resulting ODA model is: if JAS<8 then predict interv=0 (Type B); otherwise, if JAS>8 then predict interv=1 (Type A).

```
ODA model:
----------
IF JAS <= 8.0 THEN INTERV = 0
IF 8.0 < JAS THEN INTERV = 1


Summary for Class INTERV  Attribute JAS
---------------------------------------

Performance Index          Train    LOO
-----------------          ------  ------
Overall Accuracy           61.83%  61.83%
PAC INTERV=0               65.08%  65.08%
PAC INTERV=1               60.16%  60.16%
Effect Strength PAC        25.24%  25.24%
PV INTERV=0                45.56%  45.56%
PV INTERV=1                77.08%  77.08%
Effect Strength PV         22.64%  22.64%
Effect Strength Total      23.94%  23.94%


Monte Carlo summary (Fisher randomization):
-------------------------------------------
Iterations:  10000
Estimated p: 0.000500


Results of leave-one-out analysis
---------------------------------
186 observations

Fisher's exact test (directional) classification table  p = .000909
```

Effect strength for sensitivity (ESS) is labelled in the output as "Effect Strength PAC" (Percentage Accurate Classification). For the exploratory hypothesis ESS is 25.2%, which barely exceeds the minimum criterion (ESS≥25) to be classified as a moderate effect.[30] However, an exact discrete confidence interval for this result would fall into the region considered to reflect a weak effect.[36,37] This effect was stable in LOO analysis, suggesting it may cross-generalize to an independent random sample.[30]

We believe ODA should be considered the preferred statistical approach *vs*. alternative methods since it avoids statistical assumptions required of conventional models, is insensitive to skewed data or outliers, and has the ability to handle any variable metric including categorical, Likert-type integer, and real number measurement scales.[30] In contrast to other methods, only ODA can identify the optimal (maximum-accuracy) assignments (categorical attributes) or cutpoints (ordered attributes) that exist for the attribute, which in turn facilitates the use of measures of predictive accuracy. ODA can evaluate model reproducibility using multiple methods, allowing assessment of potential

2155-0182/10/$3.00

cross-generalizability of the model when it is applied to classify independent random samples.[30] For these reasons we recommend that researchers employ ODA and CTA frameworks to evaluate the statistical hypotheses which are explored in their laboratory and field research endeavors.[37-55]

References

[1]Linden A (2020). Implementing ODA from within Stata: An application to data from a randomized controlled trial (*Invited*). *Optimal Data Analysis*, *9*, 9-13.

[2]Linden A (2020). Implementing ODA from within Stata: Implementing ODA from within Stata: An application to estimating treatment effects using observational data (*Invited*). *Optimal Data Analysis*, *9*, 14-20.

[3]Linden A (2020). Implementing ODA from within Stata: An application to dose-response relationships (*Invited*). *Optimal Data Analysis*, *9*, 26-32.

[4]Linden A (2020). Implementing ODA from within Stata: assessing covariate balance in observational studies (*Invited*). *Optimal Data Analysis*, *9*, 33-38.

[5]Linden A (2020). Implementing ODA from within Stata: Evaluating treatment effects for survival (time-to-event) outcomes (*Invited*). *Optimal Data Analysis*, *9*, 39-44.

[6]Linden A (2020). Implementing ODA from within Stata: Evaluating treatment effects in multiple-group interrupted time series analysis (Invited). *Optimal Data Analysis*, *9*, 45-50.

[7]Linden A (2020). Implementing ODA from within Stata: identifying structural breaks in single-group interrupted time series designs (Invited). *Optimal Data Analysis*, *9*, 51-56.

[8]Linden A (2020). Implementing ODA from within Stata: Finding the optimal cut-point of a diagnostic test or index (Invited). *Optimal Data Analysis*, *9*, 74-78.

[9]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Exploratory hypothesis, binary class variable, and binary attribute. *Optimal Data Analysis*, *9*, 94-98.

[10]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Confirmatory hypothesis, binary class variable, and binary attribute. *Optimal Data Analysis*, *9*, 99-103.

[11]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Exploratory hypothesis, binary class variable, and binary attribute. *Optimal Data Analysis*, *9*, 104-108.

[12]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Exploratory hypothesis, binary class variable, and ordinal (rank) attribute. *Optimal Data Analysis*, *9*, 109-113.

[13]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: confirmatory hypothesis, binary class variable, and ordinal attribute. *Optimal Data Analysis*, *9*, 128-132.

[14]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Exploratory hypothesis, binary class variable, categorical ordinal attribute. *Optimal Data Analysis*, *9*, 133-136.

[15]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Nondirectional hypothesis, binary class variable, categorical ordinal attribute. *Optimal Data Analysis*, *9*, 137-140.

[16]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Directional hypothesis, binary class variable, ordinal attribute. *Optimal Data Analysis*, *9*, 141-145.

[17]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Confirmatory hypothesis, binary class variable, continuous attribute. *Optimal Data Analysis*, *9*, 146-151.

[18]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Nondirectional, multicategorical class variable, multicategorical attribute. *Optimal Data Analysis*, *9*, 152-156.

[19]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Directional hypothesis, multicategorical class variable and attribute. *Optimal Data Analysis*, *9*, 157-161.

[20]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: Directional hypothesis, multicategorical class variable, ordinal attribute. *Optimal Data Analysis*, *9*, 162-166.

[21]Yarnold PR, Linden A (2020). Implementing ODA from within Stata: *A Priori* hypothesis, three-category class variable, four-level (integer) attribute. *Optimal Data Analysis*, *9*, 167-171.

[22]Linden A, Yarnold PR (2020). Implementing ODA from within Stata: A reanalysis of the National Supported Work Experiment. *Optimal Data Analysis*, *9*, 178-182.

[23]Yarnold PR, Linden A (2021). Implementing ODA from within Stata: Exploratory hypothesis, three-category class variable, continuous attribute. *Optimal Data Analysis*, *10*, 3-9.

[24]Yarnold PR, Linden A (2021). Implementing ODA from within Stata: Confirmatory and exploratory inter-rater reliability hypothesis with a three-category ordinal rating. *Optimal Data Analysis*, *10*, 12-17

[25]Linden A (2020). ODA: Stata module for conducting Optimal Discriminant Analysis. *Statistical Software Components S458728, Boston College Department of Economics*.

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

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

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

[29]Rhodes NJ, Yarnold PR. 2020. ODA: a package and R-interface for the MegaODA software suite. R package version 1.0.1.3. Available: https://github.com/njrhodes/ODA

[30]Yarnold PR, Soltysik RC (2005). *Optimal data analysis: Guidebook with software for Windows*. Washington, D.C.: APA Books.

[31]Yarnold PR, Bryant F.B. (1988). A note on measurement issues in Type A research: Let's not throw out the baby with the bath water. *Journal of Personality Assessment*, *52*, 410-419.

[32]Matthews KA, Krantz DS, Dembroski TM, MacDougall JM (1982). Unique and common variance in Structured Interview and Jenkins Activity Survey measures of the Type A behavior pattern. *Journal of Personality and Social Psychology*, *42*, 303-313.

[33]Yarnold PR (2014). "Breaking-up" an ordinal variable can reduce model classification accuracy. *Optimal Data Analysis*, *3*, 19

[34]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.

2155-0182/10/$3.00

[35]Bryant FB, Harrison PR (2013). How to create an ASCII input data file for UniODA and CTA software (*Invited*). *Optimal Data Analysis, 2*, 2-6.

[36]Yarnold PR (2018). Comparing exact discrete 95% CIs for model *vs*. chance ESS to evaluate statistical significance. *Optimal Data Analysis*, *7*, 82-84.

[37]Rhodes JN, Yarnold PR (2020). Generating novometric confidence intervals in R: Bootstrap analyses to compare model and chance ESS. *Optimal Data Analysis*, *9*, 172-177.

[37]Linden A, Yarnold PR, Nallomothu BK (2016). Using machine learning to model dose-response relationships. *Journal of Evaluation in Clinical Practice*, *22,* 860-867.

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

[39]Yarnold PR, Linden A (2016). Theoretical aspects of the D statistic. *Optimal Data Analysis*, *22*, 171-174.

[40]Linden A, Yarnold PR (2017). Using classification tree analysis to generate propensity score weights. *Journal of Evaluation in Clinical Practice*, *23*, 703-712.

[41]Linden A, Yarnold PR (2017). Modeling time-to-event (survival) data using classification tree analysis. *Journal of Evaluation in Clinical Practice, 23*, 1299-1308.

[42]Linden A, Yarnold PR (2018). Identifying causal mechanisms in health care interventions using classification tree analysis. *Journal of Evaluation in Clinical Practice*, *24*, 353-361.

[43]Linden A, Yarnold PR (2017). Minimizing imbalances on patient characteristics between treatment groups in randomized trials using classification tree analysis. *Journal of Evaluation in Clinical Practice, 23*, 1309-1315.

[44]Linden A, Yarnold PR (2018). Estimating causal effects for survival (time-to-event) outcomes by combining classification tree analysis and propensity score weighting. *Journal of Evaluation in Clinical Practice*, *24*, 380-387.

[45]Linden A, Yarnold PR (2016). Using machine learning to assess covariate balance in matching studies. *Journal of Evaluation in Clinical Practice*, 22, 848-854.

[46]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*, *22*, 855-859.

[47]Linden A, Yarnold PR (2016). Combining machine learning and matching techniques to improve causal inference in program evaluation. *Journal of Evaluation in Clinical Practice*, *22*, 868-874.

[48]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*, *22,* 875-885.

[49]Linden A, Yarnold PR (2018). Using machine learning to evaluate treatment effects in multiple-group interrupted time series analysis. *Journal of Evaluation in Clinical Practice*, *24*, 740-744.

[50]Rhodes NJ (2020). Statistical power analysis in ODA, CTA and Novometrics (Invited). *Optimal Data Analysis*, *9*, 21-25.

2155-0182/10/$3.00

[51]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.

[52]Yarnold PR (2013). ODA range test *vs*. one-way analysis of variance: Patient race and lab results. *Optimal Data Analysis*, *2*, 206-210.

[53]Yarnold PR (2014). How to assess the inter-method (parallel-forms) reliability of ratings made on ordinal scales: Evaluating and comparing the Emergency Severity Index (Version 3) and Canadian Triage Acuity Scale. *Optimal Data Analysis*, *3*, 50-54

[54]Yarnold PR (2016) Causality of adverse drug reactions: The upper-bound of arbitrated expert agreement for ratings obtained by WHO and Naranjo algorithms. *Optimal Data Analysis*, *5*, 37-40.

[55]Yarnold PR (2016). Novometric *vs*. ODA reliability analysis *vs*. polychoric correlation with relaxed distributional assumptions: Inter-rater reliability of independent ratings of plant health. *Optimal Data Analysis*, *5*, 179-183.

Author Notes

No conflicts of interest were reported.