Optimal Data Analysis
Vol. 2, Release 2 (October 14, 2013), 63-68

2155-0182/13/$3.00

# Minimum Standards for Reporting UniODA Findings

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

As the number of researchers using any statistical method and the domain of disciplines they represent increases, the opportunity for and likelihood of the development of disparate traditions for the reporting of analytic findings also increases. The big advantage of establishing a minimum set of standards for reporting findings obtained using any method—is that researchers from *all fields* will be able to easily and clearly understand the fundamental statistical results of *any study* reporting findings using that method. This note discusses a tabular presentation of the minimum information which is required in order to understand a UniODA analysis.

Data for this exposition are taken from a survey-based study of patient satisfaction with care received in Emergency Department (ED).[1] In this example the class variable is the patient's satisfaction status (SATIS), on which the value 1 indicates satisfied, and the value 0 indicates dissatisfied.[2] The first attribute, PPHYS, is a binary indicator of whether the patient was treated by their primary care physician in the ED, with the value 1 indicating yes, and 0 indicating no.[2] The second attribute, WHO, is a categorical indicator of whether anyone accompanied the patient to the ED, with four mutually-exclusive levels: a value of 1 was used to indicate that the patient arrived with a family member; 2 to indicate a friend; 3 to indicate an employer; and 4 to indicate that the patient arrived alone. The third and final attribute, A13, used in this exposition is a Likert-type rating of the perceived helpfulness of the triage nurse, having five ordered levels: 1 through 5 indicate, respectively, ratings of very poor; poor; fair, good, and very good. In the analysis a total of 25,000 Monte Carlo experiments are used to estimate Type I error; missing values are indicated using the value -9; and leave-one-out (jackknife) validity analysis is requested.[3] Exploratory analysis predicting the class variable separately using each attribute is accomplished using the following UniODA software code[3] (control commands indicated in red).

```
OPEN EXAMPLE.DAT;
OUTPUT EXAMPLE.OUT;
VARS SATIS PPHYS WHO A13;
CLASS SATIS;
ATTR PPHYS WHO A13;
CAT PPHYS WHO;
MCARLO ITER 25000;
MISSING ALL (-9);
LOO;
GO;
```

2155-0182/13/$3.00

Table 1

Critical Information for UniODA Analyses

| Attribute | UniODA Model | N | % Satisfied | $p <$ | ESS |
|---|---|---|---|---|---|
| Was care delivered by personal physician | If NO, predict Dissatisfied | 1,670 | 88.0 | 0.25 | 2.3 |
| | If YES, predict Satisfied | 95 | 92.6 | | |
| Who accompanied the patient to ED | If family member or employer, predict Dissatisfied | 987 | 86.9 | 0.33 | 5.4 |
| | If friend or alone, predict Satisfied | 834 | 89.2 | | |
| Rated helpfulness of triage nurse | If fair, poor, or very poor, predict Dissatisfied | 215 | 46.1 | 0.0001 | 49.1 |
| | If good or very good, predict Satisfied | 1,549 | 94.0 | | |

Note: $p$ is Type I error, and ESS (effect strength for sensitivity) is a normed index of effect strength on which 0 represents the classification accuracy expected by chance, and 100 represents perfect, errorless classification.[3] All attributes were stable in leave-one-out validity analysis.[3]

Minimal sufficient information required to understand a UniODA finding for an attribute is the model; the number of observations and percent of class 1 membership in both predicted class categories; and model Type I error ($p$) and ESS in *training analysis* conducted for the total sample, and if performed then also for leave-one-out (LOO) *validity analysis*[3]. In Table 1 the classification results are indicated as stable in LOO analysis—the same as the results obtained in training analysis. Hypothetically, if the LOO performance was less than training performance, this would be indicated as illustrated in Table 2, in which the $p$ and ESS values obtained in LOO analysis are given beneath corresponding values obtained in training analysis.

Table 2

Modification to Table 1 Based on Hypothetical LOO-Instability of Ordered Attribute

| | | | | | |
|---|---|---|---|---|---|
| Rated helpfulness of triage nurse | If fair, poor, or very poor, predict Dissatisfied | 215 | 46.1 | 0.0001 | 49.1 |
| | If good or very good, predict Satisfied | 1,549 | 94.0 | 0.0003 | 32.8 |

Note: Model $p$ and ESS values for training analysis are provided in the first (top) row, and if LOO-unstable, then $p$ and ESS values for LOO validity analysis are given in the second (bottom) row.

2155-0182/13/$3.00

To illustrate where information in Tables 1 and 2 is found within UniODA output, Table 3 reproduces applicable output generated by the program code given earlier, for PPHYS. Note the UniODA model is indicated in the output as *ODA model*. By convention, for every attribute in the results table (Table 1), the model rule for predicting class category 0 should be listed first, and the model rule for predicting class category 1 should be listed second. Note also that in the results table the output codes were converted to text (0=Dissatisfied, No; 1=Satisfied, Yes).

The next column in the results table is N, which is located just beneath the classification performance table in the output. In Table 1 a total of 1,670 patients are predicted to be dissatisfied, and this is seen beneath the "Predicted 0" or left-hand-side of the output performance summary table. And, in Table 1, a total of 95 patients are predicted to be satisfied, as seen beneath the "Predicted 1" or right-hand-side of the output performance summary table.

Table 3

Selected UniODA Output for PPHYS Analysis (Binary Attribute)

```
ODA model:
----------
IF PPHYS = 0 THEN SATIS = 0
IF PPHYS = 1 THEN SATIS = 1


Fisher's exact test (nondirectional) training table  p = .248705


Classification performance summary:
-----------------------------------    Overall      Mean Sens
  Correct   Incorrect                 accuracy   across classes
    288        1477                    16.32%       51.13%

Class      Predicted
SATIS      0       1
       ----------------- NA      Sens
A      |       |       |
c     0|  200  |   7   | 207    96.62%
t      |       |       |
u     1| 1470  |  88   | 1558    5.65%
a      |       |       |
l      -----------------
   NP    1670     95
   PV    11.98%  92.63%  Mean PV  52.30%

Effect strength Sens   2.27%      Effect strength PV   4.61%


Results of leave-one-out analysis
---------------------------------
1765 observations

Fisher's exact test (directional) classification table  p = .112189

Effect strength Sens   2.27%      Effect strength PV   4.61%
```

2155-0182/13/$3.00

The next column in the results table is % Satisfied, which is located just beneath N in the output classification performance table. In Table 3, beneath the N for predicted class 0 (1,670), the percent of class 0 patients is given for the column, as 11.98%. For the left-hand column the percent of class 1 patients is 100% minus the value given: here, 100%-11.98%=88.02%. This value is indicated in Table 1, scientifically rounded to one decimal place. For the right-hand column the percent of class 1 patients is read directly from the output: rounded, 92.6%.

Table 4

Selected UniODA Output for WHO Analysis (Multi-Category Attribute)

```
ODA model:
----------
IF WHO = 1 THEN SATIS = 0
IF WHO = 2 THEN SATIS = 1
IF WHO = 3 THEN SATIS = 0
IF WHO = 4 THEN SATIS = 1


Monte Carlo summary (Fisher randomization):
-------------------------------------------
Iterations                Estimated p
----------                -----------
  25000                     .329800


Classification performance summary:
-----------------------------------    Overall      Mean Sens
  Correct   Incorrect                 accuracy   across classes
    873        948                     47.94%       52.67%

Class      Predicted
SATIS      0       1
       ----------------- NA      Sens
A      |       |       |
c     0|  129  |  90   | 219    58.90%
t      |       |       |
u     1|  858  |  744  | 1602   46.44%
a      |       |       |
l      -----------------
   NP     987     834
   PV    13.07%  89.21%  Mean PV  51.14%

Effect strength Sens   5.35%      Effect strength PV   2.28%


Results of leave-one-out analysis
---------------------------------
1821 observations

Fisher's exact test (directional) classification table  p = .077942

Effect strength Sens   5.35%      Effect strength PV   2.28%
```

2155-0182/13/$3.00

Table 5

Selected UniODA Output for A13 Analysis (Ordered Attribute)

```
ODA model:
----------
IF A13 <= 3.5 THEN SATIS = 0
IF 3.5 < A13 THEN SATIS = 1


Monte Carlo summary (Fisher randomization):
-------------------------------------------
Iterations                Estimated p
----------                -----------
  25000                     .000000


Classification performance summary:
-----------------------------------    Overall      Mean Sens
  Correct   Incorrect                 accuracy   across classes
    1572       192                     89.12%       74.57%

Class      Predicted
SATIS      0       1
       ----------------- NA      Sens
A      |       |       |
c     0|  116  |  93   | 209    55.50%
t      |       |       |
u     1|  99   | 1456  | 1555   93.63%
a      |       |       |
l      -----------------
   NP     215    1549
   PV    53.95%  94.00%  Mean PV  73.97%

Effect strength Sens  49.14%      Effect strength PV  47.95%


Results of leave-one-out analysis
---------------------------------
1764 observations

Fisher's exact test (directional) classification table  p = .217E-0062

Effect strength Sens  49.14%      Effect strength PV  47.95%
```

The final two columns in Table 1 give statistics for the overall model, the first of which is *p*, or Type I error.[3] For this attribute estimated *p* is computed, as directed by use of the MC program command. However, for totally binary problems having both binary class variable and attribute, the result of Fisher's exact test and the randomization procedure used in ODA software are isomorphic.[3] Therefore, for totally binary problems Fisher's exact test is used to obtain *p* and Monte Carlo simulation is not needed: this can greatly speed solution time.

The final column in Table 1 gives model ESS, which is provided just beneath the output

2155-0182/13/$3.00

performance summary table. Presently this is 2.27% in Table 3, and rounded scientifically to one decimal place as 2.3% in Table 1. Note that the ESS achieved in LOO analysis was the same as obtained in training analysis: this is called a *LOO-stable model*. If ESS obtained in LOO analysis is lower than ESS achieved in training analysis, the model is called *LOO-unstable*.[3]

Table 4 reproduces applicable output generated by the program code given earlier, for a multi-level categorical attribute, WHO. In Table 1, output codes for WHO were converted into text to describe the UniODA model: 1= family member, 2=friend, 3=employer, and 4= no one. N for each predicted class category is read directly from the output as was done earlier (987 for predicted Class 0, 834 for predicted Class 1). As earlier, the % Satisfied is computed as 100%-13.07%=86.93% for predicted Class 0, but read directly from the output for predicted Class 1 (89.2%). For this attribute estimated *p* is computed by Monte Carlo simulation: as is seen, $p<0.33$, and (LOO-stable) ESS=5.4.

Table 5 reproduces the applicable output generated by the program code given earlier, for an ordered attribute, A13. Note that in Table 1, output codes for WHO were converted into text in order to describe the UniODA model (1=very poor, 2=poor, 3=fair, 4=good, 5=very good). As before, N for each predicted class category is read directly from the output: 215 for predicted Class 0, and 1,549 for predicted Class 1. As earlier, % Satisfied is computed as 100%-53.95%= 46.05% for predicted Class 0, and is directly read from the output for predicted Class 1, at 94.0%. For this attribute the estimated *p* is computed by Monte Carlo simulation: as seen, $p<0.0001$, with LOO-stable ESS=49.1.

## Discussion

The ordered attribute used presently, a Likert-type scale, is an example of what is also called an *ordinal categorical scale*: categorical because levels are discrete with clear meaning, ordinal because their meaning locates the levels on an ordered continuum.[3,4] In the event that the attribute is measured on a more precise interval or ratio scale[3], in the results table (Table 1) one simply substitutes the numerical cut-point from the UniODA output. For example, imagine that age, measured as the closest integer, was used as the attribute, and the model was: If age≤35 then predict Satisfied; otherwise If age>35 then predict Dissatisfied. In the present context the latter half of the UniODA model would be entered on the first line for Age in Table 1, and the first half of the UniODA model would be entered on the second line for Age in Table 1.

Generalizing these methods for problems involving multi-categorical class variables[3] with more than two categories is straightforward, except for the "% Outcome" column. The initial impression in our laboratory is that the cell entry in this column for such designs should be the percentage of correctly classified observations: research on this issue is currently underway.

## References

[1]Yarnold PR, Michelson EA, Thompson DA, Adams SL (1998). Predicting patient satisfaction: A study of two emergency departments. *Journal of Behavioral Medicine*, *21*: 545-563.

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

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

[4]Yarnold, PR, Soltysik RC (2010). UniODA vs. chi-square: Ordinal data sometimes feign categorical. *Optimal Data Analysis*, *1*: 62-66.

## Author Notes

E-mail: Journal@OptimalDataAnalysis.com.

Mail: Optimal Data Analysis, LLC
1220 Rosecrans St., #330
San Diego, CA 92106