Optimal Data Analysis
2010, Vol. 1, Release 1 (September 17, 2010), 43-45

Copyright 2010 by Optimal Data Analysis, LLC
2155-0182/10/$3.00

Precision and Convergence of Monte Carlo Estimation of Two-Category UniODA Two-Tailed *p*

Paul R. Yarnold, Ph.D. and Robert C. Soltysik, M.S.
Optimal Data Analysis, LLC

Monte Carlo (MC) research was used to study precision and convergence properties of MC methodology used to assess Type I error in exploratory (*post hoc*, or two-tailed) UniODA involving two balanced (equal N) classes. Study 1 ran 10^6 experiments for each N, and estimated cumulative *p*'s were compared with corresponding exact *p* for all known *p* values. Study 2 ran 10^5 experiments for each N, and observed the convergence of the estimated *p*'s. UniODA cumulative probabilities estimated using 10^5 experiments are only modestly less accurate than probabilities estimated using 10^6 experiments, and the maximum observed error ($\pm$0.002) is small. Study 3 ran 10^5 experiments for Ns ranging as high as 8,000 observations in order to examine asymptotic properties of optimal values for balanced designs.

A recursive, closed-form solution for the theoretical distribution of optimal values for one-tailed "confirmatory" UniODA of random data is discovered, and associated computation time is linear in N.[1] In applications where an *a priori* alternative hypothesis has been specified, Type I error rate or alpha (*p*) can be computed for any combination of optimal value and N. For two-tailed "exploratory" applications, a closed-form solution for the distribution of optimal values has not yet been discovered. For *post hoc* UniODA the enumerable open-form solution for the theoretical distribution of optimal values is computationally intractable for N>30, but the one-tailed solution can be used to determine the two-tailed distribution if overall classification accuracy is at least 75%.[1] Other means are needed to estimate the two-tailed distribution if overall classification is less than 75%. This study uses Monte Carlo (MC) research to assess precision and convergence properties of MC methods used to estimate *p* for UniODA.

Precision

One million MC experiments were run for each balanced design of N$\leq$30. A design is balanced if the number of class 1 and 0 observations is identical for even N, or differs by one for odd N. In every experiment, the attribute was a uniform random number between zero and one.[2] For even N experiments the first N/2 observations were assigned to class 1, and the rest to class 0. For odd N experiments the first (N-1)/2 observations were assigned to class 1, and the rest to class 0. For each experiment the optimal value was determined and stored. For each N the estimated UniODA distribution was

Copyright 2010 by Optimal Data Analysis, LLC
2155-0182/10/$3.00

cumulated after 10^6 experiments were run. To compare estimated and known distributions, cumulative $p>0.001$ were rounded up to the nearest thousandth, and cumulative $p<0.001$ were rounded up on the second significant digit.

The results suggest that MC experiments accurately estimated known *post hoc* UniODA distributions. Over all N and possible optimal values, 170 of 238 (71.4%) estimated cumulative probabilities were identical to the exact value; 237 of 238 (99.6%) of the estimates were $\pm$0.001 of the exact value; and all estimates were $\pm$0.002 of the exact probability.

Estimated cumulative probabilities were most accurate when the exact probability was small. For example, for optimal values with associated exact cumulative probabilities of $0.05<p<0.001$, 45 of 50 (90%) of the estimated probabilities were identical to the corresponding exact probability; 49 of 50 (98%) of estimated probabilities were $\pm$0.001 of exact probability; and all estimated probabilities were $\pm$0.002 of the exact probability.

MC experiments also provided accurate estimates of exact cumulative probabilities for statistically marginal ($0.05<p<0.10$) effects: 13 of 15 (86.7%) of the estimated cumulative probabilities were identical to their corresponding exact values, and all estimated probabilities were $\pm$0.001 of the exact probability.

Cumulating 10^6 MC experiments for a given N provides an accurate approximation of the UniODA distribution, but the computational cost is high. Accordingly, Study 2 investigated convergence properties of MC methodology and was designed to determine the number of MC experiments that is sufficient to achieve stable, accurate estimates of UniODA distributions.

Convergence

MC experiments were designed and data generated as in Study 1. For each N between 3 and 30 inclusive, 10^5 experiments were run in successive blocks of 1,000 experiments, and the UniODA distribution was cumulated at each block. Thus, 100 UniODA distributions were estimated for each N: the first based on 1,000 experiments, the second based on 2,000 experiments, and the 100th based on 10^5 experiments.

Many (56.9 percent) of the estimated *p*'s converged to their final value (i.e., their value at the end of the study) within 20,000 experiments, and most (86.3 percent) of the estimated *p*'s converged to their final value within 70,000 experiments.

After 10^5 experiments were completed, every estimated *p* in the range $0.001<p<0.10$ was identical to the corresponding estimated *p* based on 10^6 experiments (precision study).

Consistent with the first study, known UniODA distributions were accurately modeled. For probabilities in the range $0.001<p<0.05$: 35 of 50 (70%) estimated cumulative probabilities were identical to corresponding exact values; 49 of 50 (98%) estimated probabilities were $\pm$0.001 of exact; and all estimated probabilities were $\pm$0.002 of the exact value.

Thus, UniODA cumulative probabilities estimated using 100,000 MC experiments are only modestly less accurate than probabilities estimated using one million experiments, and the maximum observed error ($\pm$0.002) is small.

Asymptotic Convergence

A final study investigated convergence properties of interesting levels of classification performance for balanced two-category *post hoc* UniODA. MC experiments were designed and data generated as in Study 1. For all N between 1,000 and 8,000 inclusive, in steps of 1,000, a total of 10^5 MC experiments were run. Results of the simulation are presented in Table 1.

Tabled for the indicated value of *p* and N are the optimal value and the corresponding percentage accuracy in classification or PAC (top and bottom row, respectively). The optimal value is the maximum number of misclassifications possible to still achieve the *p* value. For example, for N=1,000 observations and $p<0.001$ a maximum of 438 misclassifications can be

Copyright 2010 by Optimal Data Analysis, LLC
2155-0182/10/$3.00

made, corresponding to 562 correct classifications, and thus to PAC=(562/1,000)x100%, or 56.2% (see Table 1).

Table 1: Maximum Optimal Value for 2-Tail *p* in Balanced 2-Category UniODA

	Two-Tail *p*<			
N	0.001	0.01	0.05	0.10
1,000	438	448	457	461
	56.2	55.2	54.3	53.9
2,000	912	927	939	945
	54.4	53.6	53.1	52.8
3,000	1393	1411	1425	1433
	53.6	53.0	52.5	52.2
4,000	1876	1896	1913	1922
	53.1	52.6	52.2	52.0
5,000	2361	2384	2403	2413
	52.8	52.3	51.9	51.7
6,000	2849	2874	2894	2905
	52.5	52.1	51.8	51.6
7,000	3336	3364	3386	3397
	52.3	51.9	51.6	51.5
8,000	3825	3853	3878	3890
	52.2	51.8	51.5	51.4

For $p<0.05$ a maximum of 457 misclassifications are possible, corresponding to PAC=(543/1,000)x100%, or 53.9%. For N=5,000 and $p<0.01$, a maximum of 2,384 misclassifications are possible, corresponding to PAC=[(5,000-2,384)/5,000]x100%, or 52.3%.

In balanced designs involving as few as 1,000 observations, a UniODA model performing only a modicum better than an unbiased flipped coin (i.e., obtaining at least 55.2% "heads") yields classification accuracy which is sufficient to achieve $p<0.001$. Therefore, as N increases in magnitude the significance of *p* as an index of performance rapidly diminishes to trivial levels.

References

[1]Soltysik RC, Yarnold PR. Univariable optimal discriminant analysis: one-tailed hypotheses. *Educational and Psychological Measurement* 1994, 54:646-653.

[2]Yarnold PR, Soltysik RC. *Optimal data analysis: a guidebook with software for Windows*. APA Books, Washington, DC, 2005.

Author Notes

Appreciation for computer resources used in this research is extended to the University of Illinois at Chicago. Mail correspondence to the authors at: Optimal Data Analysis, 1220 Rosecrans St., Suite 300, San Diego, CA 92106. Send E-mail to: Journal@OptimalDataAnalysis.com.