Optimal Data Analysis
Vol. 2, Release 2 (October 14, 2013), 71

2155-0182/13/$3.00

UniODA and Small Samples

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

Statistical hypotheses investigated by researchers representing a vast domain of empirical disciplines often involve discrimination and prediction of discrete outcomes: life versus death in medicine, innocent versus guilty in law, profit versus loss in finance, victory versus defeat in military science, etcetera. Although some studies feature large samples of a thousand or more observations, most of the literature employs much smaller samples. Preliminary research and examination of theoretical distributions indicates that UniODA is capable of identifying statistically significant effects in applications offering very small samples.

The smallest multi-observation samples (n=8) for which a statistically reliable effect was obtained by UniODA both involved a confirmatory hypotheses with an ordered attribute and a binary class variable. The two studies included a prospective evaluation of cost savings in asthma management occurring after treatment[1], and an analysis of the relationship between oil leverage and gross domestic product.[2]

For an ordered attribute the smallest sample size for which it is theoretically possible to obtain a generalized ("per-comparison") *two-tailed* p<0.05 is n=8—half from each of the two compared groups[3] (p<0.029), and the smallest sample size for which it is theoretically possible to obtain generalized *one-tailed* p<0.05 is n=6—again half from each of two compared groups[4] (p<0.05). Of course, such minute samples offer minute statistical power.[5]

References

[1]Levenson T, Grammer LC, Yarnold PR, Patterson R (1997). Cost-effective management of malignant potentially fatal asthma. *Allergy and Asthma Proceedings*, *18*, 73-78.

[2]Yarnold PR (2013). Percent oil-based energy consumption and average percent GDP growth: a small sample UniODA analysis. *Optimal Data Analysis*, *2*, 60-61.

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

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

[5]Soltysik RC, Yarnold PR (2013). Statistical power of optimal discrimination with one attribute and two classes: One-tailed hypotheses. *Optimal Data Analysis*, *2*, 26-30

Author Notes

E-mail: Journal@OptimalDataAnalysis.com.