Optimal Data Analysis
Vol. 7 (November 14, 2018), 74-75

2155-0182/10/$3.00

Using ODA to Determine if a Markov Transition Process is Second Order

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

This note demonstrates the use of ODA to test the hypothesis of an underlying second-order Markovian process.

Taken from a study of social class identification, data for this exposition (Table 1) were used to illustrate Goodman's chi-square procedure for testing the hypothesis that a Markov transition process is second order, and not first order.[1]

Table 1: Turnover Data[1] Used to Test Second Order Process Hypothesis

		1960	
1956	1958	Middle	Working
Middle	Middle	216	70
Working	Middle	56	75
Middle	Working	42	92
Working	Working	47	549

Chi-square may be used to evaluate the hypothesis that the transition process is second order. This is accomplished by computing chi-square for the table formed by the first pair of lines in Table 1 (42.6); for the table formed by the second pair of lines (56.2); summing the chi-square values (98.2); and evaluating the overall chi-square statistic (with 2 degrees of freedom) for statistical significance ($p<0.0001$). A statistically significant omnibus chi-square statistic results in rejection of the null hypothesis that a first order Markovian process underlies the data.

ODA offers an exact nonparametric analogue to Goodman's test in this application.[2-4] As in Goodman's method, two ODA analyses are used, however individual ODA results are considered separately (the ESS values aren't summed). For both ODA analyses (first and second line sets in Table 1), year is treated as a two-category class variable, and Middle *vs*. Working response as a categorical attribute.

Table 2: ODA Second Order Process Tests

	Model Sensitivity			
Line Pair	Middle	Working	ESS	$p<$
First Set	75.5	57.2	32.8	0.0001
Second Set	31.3	92.1	23.5	0.0001

Both ODA models achieved essentially moderate ($25<ESS\leq50$) ESS values—indicating moderate discrimination between the transition matrices, and therefore moderate support for the hypothesis that the modeled process is a second order Markovian.

2155-0182/10/$3.00

References

1. Markus GB (1979). *Analyzing Panel Data*. Beverly Hills, CA: Sage.

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

3. Yarnold PR, Soltysik RC (2016). *Maximizing Predictive Accuracy*. Chicago, IL: ODA Books. DOI: 10.13140/RG.2.1.1368.3286

4. Yarnold PR (2017). What is optimal data analysis? *Optimal Data Analysis*, *6*, 26-42.

Author Notes

No conflict of interest was reported by the author.