J. Statist. Comput. Simul., 1986, Vol. 26, pp. 127–143
0094-9655/86/2602-0127 $18.50/0

Comments, Conjectures and Conclusions

Section Editor: I. J. GOOD

Please be succinct but lucid and interesting.
It is helpful if you submit relevant tidy rough work.

C263. SMALL-SAMPLE PROPERTIES OF SPATIAL AUTOCORRELATION

1. INTRODUCTION

Spatial patterns arise in various contexts such as agricultural field experiments, geological explorations, epidemic studies and satellite image processing. In this paper, we focus on data collected in regular grids, or contiguous quadrats, and examine the small sample properties of two tests for autocorrelation by some Monte Carlo studies. We consider first-order models for simultaneous autoregression and simultaneous moving averages. Models not directly considered include conditional autoregression and error-in-variables; see Cliff and Ord (1981) and Ripley (1981) for details and further references.

The purpose of this paper is to determine the small-sample power of Moran's I statistic (Moran, 1950) in comparison with the likelihood ratio (LR) statistic when the underlying model is misspecified. Haining (1978) found empirically that the LR test was more powerful than Moran's I, when the underlying model is correct. It is also known that the LR test is better than the I test when sample size is large and the underlying model is right.

2. NOTATION AND MODELS

Consider a rectangular $m \times n$ grid, with $\mathbf{Y} = \{Y_i\}, i = 1, \ldots, mn$ being the random vector of realised values for a spatial process. For

convenience, let $E(\mathbf{Y}) = \mathbf{0}$. Define (a) the simultaneous autoregressive model (SAR) as $(\mathbf{I} - \rho\mathbf{W})\mathbf{Y} = \mathbf{E}$, and (b) the simultaneous moving-average model (SMA) as $\mathbf{Y} = (\mathbf{I} + \theta\mathbf{W})\mathbf{E}$. Here $\mathbf{E} \sim \mathrm{MVN}(0, \sigma^2\mathbf{I})$ and $\mathbf{W} = \{W_{ij}\}$, is the $mn \times mn$ matrix of indicators for nearest neighbours. The likelihood-ratio test under both models are developed as in Cliff and Ord (1981). Moran's I statistic is defined as $I = [mn\mathbf{Y}'\mathbf{W}\mathbf{Y}]/[(\mathbf{1}'\mathbf{W}\mathbf{1})(\mathbf{Y}'\mathbf{Y})]$, where $\mathbf{1}$ is a vector of 1's and $\mathbf{Y}$ is adjusted by the sample mean. Both tests are used to test the presence of spatial autocorrelation. The LR tests whether the underlying parameter is zero or not where an extreme value of Moran I statistic indicates the presence of autocorrelation.

3. SIMULATION PROCEDURE AND RESULTS

Square grids of edge sizes 5 to 10 were simulated with $\sigma = 1$ and correlation parameter θ for the SMA data or ρ for the SAR data, having values between -0.2 and 0.2 in steps of 0.05. For each grid size and parameter combination, 1000 Monte Carlo simulations were performed. Normal variates were generated using IMSL and the decomposition for the $\mathbf{W}$ matrix was done by LINPACK. Each simulation yielded an I test statistic and an LR test statistic for the "wrong" model, i.e. the LR statistic for the SMA model was computed from the SAR data, and vice versa. The critical values were quantiles generated from 2000 simulations under the null case of normal white noise (see Table 1). Empirical power was calculated for $\alpha = 0.05$ and 0.01, with similar results. The simulations for $\alpha = 0.05$ are summarized in Tables 2 and 3. The main conclusions follow: (1) When sample size is small, the LR approach appears to yield a slightly more powerful test even if the model is wrong. Nevertheless, the I test performs just as well when the grid size is large, say larger than 7×7. (2) The power of the SAR data with SMA fit appears to be larger than the other "wrong" model for both tests. (3) In both models, when the true autocorrelation parameter is positive, the power of the I test is greater than the LR test. The opposite is observed when the true parameter is negative. This seemed to be caused by estimating the correlation parameter by MLE, which was noted by Brandsma and Ketellapper (1979).

This work was supported in part by grants from the U.S.

Department of Agriculture and NSF. The computing was done on the UW-Madison Statistics VAX 11/750 Research Computer.

References

Brandsma, A. S. and Ketellapper, R. H. (1979). Further evidence on alternative procedures for testing of spatial autocorrelation among regression disturbances. In *Exploratory and Explanatory Statistical Analysis of Spatial Data* (eds. C. P. A. Bartels and R. H. Ketellapper; Boston: Martinus Nijhoff).

Cliff, A. D. and Ord, J. K. (1981). *Spatial Processes Models and Applications*. London: Pion.

Haining, R. P. (1978). The moving average model for spatial interaction. *Transactions, Institute of British Geographers (New Series)*, 205–225.

Moran, P. A. P. (1950). Notes on continuous stochastic phenomenia. *Biometrika* 37, 17–23.

Ripley, B. D. (1981). *Spatial Statistics*. New York: John Wiley & Sons.

Winson Taam and Brian S. Yandell
University of Wisconsin-Madison, USA

C264. A GEOMETRICAL APPROACH TO THE AMALGAMATION PARADOX

This note was previously an appendix to Good & Mittal (1986) but was detached because we were requested to shorten that paper.

Let $\mathbf{a}=[a,b;c,d]$ denote a 2 by 2 "population contingency table" and let α denote an association measure. Most association measures α, and all those in Good & Mittal (1986) (called G & M later), are homogeneous functions of (a, b, c, d) of degree zero, that is $\alpha(\mathbf{a})=\alpha(\lambda\mathbf{a})$ for all positive values of λ. It is therefore convenient to represent $\mathbf{a}$ or $\lambda\mathbf{a}$ geometrically by the same point. In other words we may use three-dimensional *homogeneous coordinates* (a,b,c,d) (see, for example, McCrea, 1947, pp. 41–42). The reader might prefer to replace a, b, c, and d by letters at the end of the alphabet to make them look more like coordinates.

The equation $\alpha(\mathbf{a})=$constant thus represents a surface embedded in three dimensions. If the equation is quadratic this surface is part of a quadric: only a part because the coordinates are all positive. In these circumstances we say that the measure α is *quadric*. All the