Center for Spatially Integrated Social Science

Colin Loftin and Sally K. Ward: Application of Spatial Autocorrelation in Sociology

By Sam Ying

Background

In 1972 three sociologists, Omer R. Galle, Walter R. Gove, and Miller McPherson, developed a model (the GGM model) that displayed the correlation between population density and pathology (Galle et al., 1972). The article was published in *Science* and has since been discussed thoroughly. Using statistical tests, including least-squares estimates, Galle et al. produced results that seemed to suggest that "density has effects on pathology independent of other causes of pathology, notably social class and ethnicity" [Loftin and Ward, 1983]. The study area included "community areas" in Chicago, Illinois, constructed by modifying 1970 census tracts. The GGM model concluded that high population density has a great effect on human behavior and that overcrowding should be taken into consideration when sociologists research pathological behaviors. Table 1 displays the results of the least-squares estimates of the GGM model). According to the table, population density has an influence on the fertility rate within each study area. Although, the GGM model uses geographically defined units, the effects of spatial processes were not included in their analysis. The article, "A spatial autocorrelation model of the effects of population density on fertility," authored by Colin Loftin and Sally K. Ward, disputes the GGM model by introducing spatial autocorrelation as a way to take spatial interaction into consideration, introducing an alternative to the GGM model to study the relationship between population density and fertility.

Innovation

Spatial autocorrelation

In the simplest terms, spatial autocorrelation is based on the first law of geography: near things are more similar than things that are more distant. Positive spatial autocorrelation is exhibited when neighboring areas are similar or the same. A checker board is a good example of the absence of spatial autocorrelation; the squares adjacent to each white square is black, showing the nearest neighbors are not similar. (A very important consideration should be taken when introducing spatial autocorrelation into a project's analysis: It is scale dependent. Take Figure 1, for example. It is a checkerboard made of 4 squares where the squares adjacent to white squares are always black, displaying no spatial autocorrelation. Now look at Figure 2. It displays the same characteristics as Figure 1, except we have made the units smaller, where 4 squares in Figure 2 represent the same area as 1 square in Figure 1. Suddenly, there is now positive spatial autocorrelation where the squares in the corners are surrounded on two sides with squares of the same color. This is a basic concept that should be understood when studying spatial autocorrelation, but was not introduced in the Loftin and Ward article.

The Spatial Autocorrelation Model and Weighting Methods

The residuals from the GGM model were mapped, which showed clusters of the five intervals of residuals. Loftin and Ward used this map as evidence that a factor, in this case, spatial interaction, was causing the pattern in the distribution of residual values. The spatial autocorrelation model is best described as an alternative to the GGM, which "assumes that the disturbances in each community area are systematically related to those in adjacent areas" [Loftin and Ward, 1983]. In time series, data points are linear in that there are only two neighboring data points, one before and one after. Spatial data points complicate this issue because any area or point can have an unlimited number of neighbors. Loftin and Ward introduce the weight matrix, W, to represent the pattern of interactions between disturbances at locations i and j [Loftin and Ward, 1983]. Because of the asymmetry of time order, one would choose the first-order autoregressive time-series model. Unlike time-series, the weighting scheme to use is not so obvious when handling spatial data, therefore three different weighting schemes were used to evaluate the effects of neighboring areas.

1. **Common Boundary Weights**. Using this weighting method, areas with common boundaries are given weight, whereas areas that do not have common boundaries will not have weight. Relative to area i, j and k will be given the same weight.

2. **Standardized Common Boundary Weights**. In figure 3, the boundary between area A and area B is significantly longer than that between area B and area C. Using standardized common boundary weights, the weighting values are not binary-like common boundary weights; instead, the greater the length of the common border, the greater the weight. Relative to area i, area j will be given a greater weight than area k because of the longer common border.

3. **Standardized Distance Weights**. This weighting method does not use borders; rather it finds the distance between the geographic centers of the study areas, where the greater the distance between two centroids, the larger the area of the polygon, therefore the greater the weight given.

Conclusion

Loftin and Ward applied the three weighting methods to their analysis producing a table of estimates that showed significantly different results from those of the GGM model. They stated that the difference in results show that the GGM model's ordinary least-squared (OLS) estimates are misleading and that density has little effect on fertility. In fact, for three of the four density factors taken into consideration (*ln* rooms per unit, *ln* units per structure, *ln* structures per acre), three of the density coefficients were found to have true values of zero. Loftin and Ward conclude "when the model includes spatially autocorrelated disturbances, there is only one statistically significant effect of density on fertility" [1983]. Because of these results they believe the inclusion of spatial processes would significantly change all of the results produced by the GGM model. Although not mentioned within the article, the insignificance of the estimates within the spatial autocorrelation model may not only be due to the consideration of spatial autocorrelation, but could also be caused by the multi-colineality of the independent variables (ethnicity, class, and population density overlap in multiple ways; e.g., a poor neighborhood of blacks often has very high population density). Nevertheless, Loftin and Ward's discussion contributes greatly to the advocating of integrating spatial processes into all studies that use geographically defined areas.

Table 1. Least-Squares Estimates of the GGM Model

Variable	B-hat	S(B-hat)	B-hat/S(B-hat)
Intercept	46.04	37.96	1.21
ln Persons per Room	96.89	29.66	3.27
ln Rooms per Unit	65.88	26.47	2.49
ln units per Structure	18.94	8.72	2.17
Structures per Acre	7.12	4.01	1.78
Class Index	-.065	.020	-3.22
Ethnicity index	.003	.0038	.70
S.E.	15.08	---	---
R^2	.75	---	---

Fig. 1

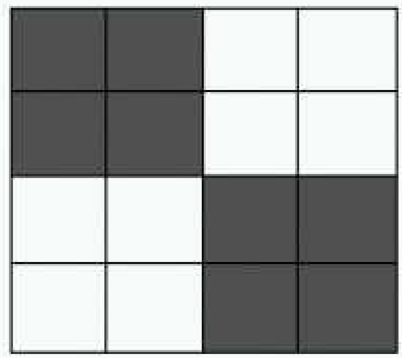

Fig. 2

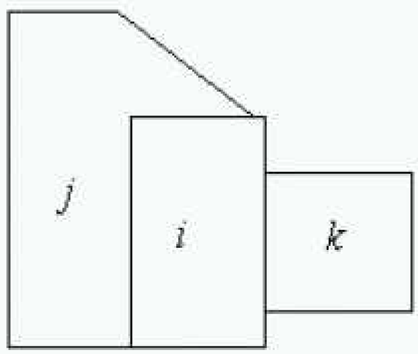

Fig. 3

Publications

Loftin, Colin. 1972. Galton's problem as spatial autocorrelation: Comments on Ember's empirical test. *Ethnology* 11:425–35.

Loftin, Colin and Sally K. Ward. 1981. Spatial autocorrelation models for Galton's problem. *Behavior Science Research* 16:105–28.

Ward, Sally K. 1975. Methodological considerations in the study of population density and social pathology. *Human Ecology* 3:275–86.

Related Works

Choldin, Harvey M. 1978. Urban density and pathology. *Annual Review of Sociology* 4:91–113.

Calhoun, John B. 1962. Population density and social pathology. *Scientific American* 206:139–48.

Ord, Keith. 1975. Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* 70:120–26.

McPherson, J. Miller. 1975. Population density and social pathology, a reexamination. *Sociological Symposium* 14:77–90.