METMA V International Workshop on Spatio-Temporal Modelling
Santiago de Compostela
30 Jun - 2 Jul 2010

Testing for spatial stationarity in point patterns

J. Mateu[1], C. Comas[1,*] and M.A. Calduch[2]

[1] *Departamento de Matemáticas, Universitat Jaume I, Campus Riu Sec, E-12071 Castellón, Spain; mateu@mat.uji.es, comas@mat.uji.es*
[2] *Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, Universidad Politécnica de Valencia, mcalduch@eio.upv.es*
[*] *Corresponding author*

Abstract. *A common assumption while analyzing spatial point processes is translation and rotation invariance. However, in many practical situations nonstationarity is a more plausible assumption. Considering that second-order methods provide a natural starting point for the analysis of spatial point process data, we propose a formal nonparametric approach to test for stationarity based on a wide range of possible estimators for the first-order intensity function and the reduced second moment measure. We demonstrate the efficacy of the approach through simulation studies and real applications.*

Keywords. *First-order intensity function; Isotropy; K-function; Nonstationarity; Point patterns.*

1 Introduction

Methods for the statistical analysis of stationary spatial point process data are now well established, methods for nonstationary processes less so. One of many sources of nonstationary point process data is a case-control study in environmental epidemiology. It is clear that many practical situations are concerned with the analysis of spatial point process data when the underlying point process is nonstationary.

Cox (1972) discussed likelihood-based inference for log-linear inhomogeneous Poisson processes. Berman and Diggle (1989) extended to two-dimensional processes a method for nonparametric estimation of the intensity function of an inhomogeneous Poisson process that had been proposed in the one-dimensional case by Diggle (1985). Baddeley *et al.* (2000) extended Ripley's (1976) definition of a reduced second moment measure, or K-function, to include a class of nonstationary processes.

The first- and second-order intensities of an orderly spatial point process are denoted by $\lambda(x)$ and $\lambda_2(x,y)$, respectively. Informally, $\lambda(x)dx$ gives the probability that an event occurs within an infinitesimal region with center x and area dx, and $\lambda_2(x,y)dxdy$ similarly gives the probability that two events occur, one within each of two infinitesimal regions with centers x and y and areas dx and dy. The pair correlation function is defined as $g(x,y) = \lambda_2(x,y)/\lambda(x)\lambda(y)$. A point process is second-order stationary when its first-order and second-order properties are invariant under translation. Baddeley *et al.* (2000) defined a weaker form of stationarity for spatial point processes, called second-order intensity-reweighted stationarity. This allows a non-constant intensity, which is required to be bounded away from zero, and assumes that the spatial pair correlation function $g(x,y) = g(||x-y||)$. We also say that a spatial point process is homogeneous if $\lambda(x) = \lambda$ for some constant $\lambda > 0$, and inhomogeneous otherwise.

The reduced second moment measure, or K-function, of a second-order intensity-reweighted stationary and isotropic spatial point process is $K_I(s) = 2\pi\int_0^s ug(u)du$. For an inhomogeneous Poisson process, $K_I(s) = \pi s^2$. Processes for which $K_I(s) > \pi s^2$ or $K_I(s) < \pi s^2$ generate spatial distributions of events that are more strongly aggregated or more regular, respectively, than would be the case for an inhomogeneous Poisson process with first-order intensity $\lambda(x)$.

Baddeley *et al.* (2000) showed that if $\lambda(x)$ is known, an unbiased estimator of $K_I(s)$ is $\hat{K}_I(s;\lambda) = 1/|A|\sum_i\sum_{j\neq i}\frac{w_{ij}I(d_{ij}\leq s)}{\lambda(x_i)\lambda(x_j)}$. Here $|A|$ denotes the area of A on which the point process is defined, $I(\cdot)$ is the indicator function, $d_{ij} = ||x_i - x_j||$, and w_{ij} is an edge-correction. In practice, the intensity $\lambda(x)$ is unknown and must be estimated, either parametrically or nonparametrically. Estimation of the intensity function of spatial point processes is a fundamental problem. Cox (1972) discussed parametric estimation of $\lambda(x)$ for an inhomogeneous Poisson process. For nonparametric estimation, one widely used method is kernel smoothing. Diggle *et al.* (2007) used the kernel smoothing method with an edge-correction as in Berman and Diggle (1989), while Baddeley *et al.* (2000) used a closely related edge correction. Van Lieshout (2010) interprets the Delaunay tessellation field estimator as an adaptive kernel estimator. Guan (2008a) proposed a kernel estimator with a kernel weight a function of the difference between the covariate values.

So, there are now several available possibilities to estimate the first-order intensity function and, consequently, the corresponding K-function or pair-correlation function. These statistical tools are clearly necessary is a nonstationary context which is widely encountered in practical analysis of spatial point patterns. There are some papers on testing the isotropy condition (Guan *et al.*, 2006). However, very few papers have focussed on testing stationarity (Guan, 2008b; Comas *et al.*, 2009). Here we focus on this latter aspect by proposing several possible tests for inhomogeneity or nonstationarity.

2 Methods

The use of nonstationary point process techniques has to be motivated by the presence of nonhomogeneous point patterns. We compute the degree of nonstationarity of point patterns by obtaining the sum of the intersection areas of the estimated homogeneous Ripley's K-function and the inhomogeneous counterpart version of this function, applied to the same point pattern via

$$\int_0^r ||\hat{K}(v) - \hat{K}_{inh}(v)||dv,$$

where v is the range of possible distances. Thus for stationary point patterns this integral should be close to zero, whilst values far from zero denote nonstationarity in the point pattern. If we compare these two functions, we notice that for a given point pattern contained in a bounded region A, the only difference between them is the nature of the intensity function. Under stationarity, this intensity function is constant and equals λ (the intensity of the point process), whilst under the inhomogeneous version the point intensity depends on the spatial location, i.e. $\lambda(\mathbf{x})$. This suggests that a tentative measure of inhomogeneity for point patterns can be simply computed by comparing the intensity function $\lambda(\mathbf{x})$ and the point intensity λ (i.e. assuming stationarity in the point pattern)

$$S = \int_A \|\lambda - \lambda(\mathbf{x})\| d\mathbf{x}. \tag{1}$$

Thus if $S = 0$, then the point pattern is stationary, whilst if $S > 0$, the point pattern can be considered nonstationary. Note that in order to compute S for a given point process Φ contained in a bounded region A with intensity function $\lambda(\mathbf{x})$, we first need to obtain the point intensity of this point process under the hypothesis of stationarity. In practice, given a point pattern, an edge-corrected estimator for the measure S can be used instead of (**??**)

$$\hat{S} = \int_A \|\hat{\lambda} - \hat{\lambda}(\mathbf{x})\| d\mathbf{x}, \tag{2}$$

where $\hat{\lambda} = N/|A|$ and $\hat{\lambda}(\mathbf{x})$ is a tentative intensity estimator. Note that when dealing with $\hat{S}$, i.e with $\hat{\lambda} = N/|A|$ and $\hat{\lambda}(\mathbf{x})$, we are incorporating into this measure the intrinsic stochastic variability of these estimators. Hence, two considerations are important to note here: a) the behaviour of the measure S will depend on the correct choice and use of the intensity estimators (for example, the choice of the kernel function or the bandwidth parameter), and (b) if the theoretical measure S should be zero under stationarity, the empirical $\hat{S}$ is not expected to be exactly zero, but should take values close to zero. We thus need to analyse the behaviour of this measure.

We develop and show the properties of several stationarity tests based on the K-function, the first-order intensity function or the pair-correlation function. We use a collection of possible first-order intensity estimates and compare the corresponding performances under simulations and real data analysis. An interesting extension would be considering spatio-temporal point patterns. A natural starting point for analysis of spatio-temporal point process data is to investigate the nature of any stochastic interactions among the points of the process after adjusting for spatial and/or temporal inhomogeneity. Gabriel and Diggle (2009) propose a spatio-temporal inhomogeneous K-function and this could be used as the building block for new tests of space-time inhomogeneity.

References

[1] Baddeley, A., Møller, J., and Waagepetersen, R. (2000). Non and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, **54**, 329-350.

[2] Berman, M. and Diggle, P.J. (1989). Estimating weighted integrals of the second-order intensity of a spatial point process. *Journal of the Royal Statistical Society, Series B*, **51**, 81-92.

[3] Comas, C., Palahi, M., Pukkala, T. and Mateu, J. (2009). Characterising forest spatial structure through inhomogeneous second order characteristics. *Stochastic Environmental Research and Risk Assessment*, **23**, 387-397.

[4] Cox, D.R. (1972). The statistical analysis of dependencies in point processes. In *Stochastic Point Processes*, P.A.W. Lewis (ed), 55-66. New York: Wiley.

[5] Diggle, P.J. (1985). A kernel method for smoothing point process data. *Applied Statistics*, **34**, 138-147.

[6] Diggle, P.J. and Chetwynd, A.G. (1991). Second-order analysis of spatial clustering for inhomogeneous populations. *Biometrics*, **47**, 1155-1163.

[7] Diggle, P.J., Gomez-Rubio, J., Brown, P.E., Chetwynd, A.G. and Gooding, S. (2007). Second-order analysis of inhomogeneous spatial point processes using case-control data. *Biometrics*, **63**, 550-557.

[8] Gabriel, E. and Diggle, P.J. (2009). Second-order analysis of inhomogeneous spatio-temporal point process data. *Statistica Neerlandica*, **63**, 43-51.

[9] Guan, Y. (2008). On consistent nonparametric intensity estimation for inhomogeneous spatial point processes. *Journal of American Statistical Association*, **103**, 1238-1247.

[10] Guan, Y. (2008). A KPSS test for stationarity for spatial point process. *Biometrics*, **64**, 800-806.

[11] Guan, Y., Sherman, M. and Calvin J.A. (2006). Assessing isotropy for spatial point processes. *Biometrics*, **62**, 119-125.

[12] Ripley, B.D. (1976). The second order analysis of stationary point processes. *Journal of Applied Probability*, **13**, 255-266.

[13] van Lieshout, M.N.M. (2010). On estimation of the intensity function of a point process. Working paper.