Spatial Statistics 11 (2015) 19–34

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Deviation test construction and power comparison for marked spatial point patterns

Mari Myllymäki [a,*], Pavel Grabarnik [b], Henri Seijo [a], Dietrich Stoyan [c]

[a] *Department of Biomedical Engineering and Computational Science, Aalto University School of Science, P.O. Box 12200, FI-00076, Aalto, Finland*

[b] *Institute of Physico-Chemical and Biological Problems in Soil Science, The Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia*

[c] *Institut für Stochastik, TU Bergakademie Freiberg, D-09596 Freiberg, Germany*

ARTICLE INFO

Article history:
Received 15 January 2014
Accepted 24 November 2014
Available online 3 December 2014

Keywords:
Deviation test
Marked point process
Mark-weighted K-function
Monte Carlo test
Random labeling
Simulation study

ABSTRACT

Deviation tests play an important role in testing distributional hypotheses in point process statistics. Such tests are based on differences between empirical summary functions and their theoretical counterparts, which depend on a distance variable r in a user-specified interval I. These differences are summarized to a single number, which serves then as the test statistic u. Statistical experience indicates that different distances r have different influence on u. We propose scalings of the differences to equalize the influence of the distances and show that the power of deviation tests can be improved by them. We further study how the power is affected by the other elements of deviation tests, which are the choice of the summary function, the deviation measure and the interval I. We consider in detail the construction of deviation tests for the particular case of testing the random labeling hypothesis, i.e. independence of the marks of a marked point process. By a large simulation study we come to clear statements about the role of the test elements. Furthermore, we demonstrate the potential of scaling by a data example from the literature.

* Corresponding author. Tel.: +358 50 512 4490; fax: +358 9 470 23182.
E-mail addresses: mari.myllymaki@aalto.fi (M. Myllymäki), gpya@rambler.ru (P. Grabarnik), henri.seijo@aalto.fi (H. Seijo), stoyan@math.tu-freiberg.de (D. Stoyan).

http://dx.doi.org/10.1016/j.spasta.2014.11.004
2211-6753/

1. Introduction

Testing statistical hypotheses is an important step in building statistical models. In point process statistics, typical hypotheses are complete spatial randomness (CSR), independent marking or some fitted model. Unlike in classical statistics, where null models are typically represented by a single hypothesis, hypotheses in spatial statistics have a spatial dimension and therefore a multiple character. Usually a summary function such as Ripley's K-function is employed in the test.

In this paper, we consider the tests in a generalized form: their basis is some test function $T(r)$, where r is a distance variable, and this $T(r)$ is not necessarily an unbiased (or ratio-unbiased) estimator of some summary function. The name "deviation" emphasizes that the tests are based on the differences between the empirical values of $T(r)$ and their expectation under the null hypothesis, which are called "residuals" in the following.

A problem is how to handle the residuals for different values of r. If there were *a priori* a single distance r^* which is of main interest, then one could proceed as in classical tests by comparing the empirical value $T(r^*)$ with the theoretical value $T_0(r^*)$ for this r^*, i.e. consider the residual $T(r^*) - T_0(r^*)$. However, since usually such a single special distance r^* is not given, one would like to consider the residuals simultaneously for all distances r in some interval $I = [r_{\min}, r_{\max}]$. Therefore, one is confronted with a situation typical for multiple hypothesis testing (or multiple comparisons), see Bretz et al. (2010).

The standard approach to resolve the multiple hypothesis testing problem in point process statistics is the deviation test suggested by Diggle (1979). In this test, the residuals for all r in I are summarized into a single number by some deviation measure, e.g. the maximum absolute residual in I. This approach has analogues in classical statistics, namely the Kolmogorov–Smirnov and Cramér–von Mises tests.

An alternative approach constructs envelopes around the theoretical function $T_0(r)$ and checks if the empirical function is completely between the envelopes. Difficulties of this popular method, which goes back to Ripley (1977), were discussed in detail by Loosmore and Ford (2006) and Grabarnik et al. (2011). The present paper concentrates solely on the deviation test.

Though Diggle's procedure is accepted as a standard in point process statistics, to our knowledge there are no studies which explore its properties systematically. Several power comparisons for different forms of deviation tests have been reported (e.g. Ripley, 1979; Gignoux et al., 1999; Thönnes and van Lieshout, 1999; Baddeley et al., 2000; Grabarnik and Chiu, 2002; Ho and Chiu, 2006, 2009), but these investigations concern only specific issues. In the present paper, we consider the construction of deviation tests in detail and systematically and come to general recommendations for their use.

Our particular premise is that if there is not an *a priori* interesting distance r^*, then it makes sense to give similar importance to all residuals on the chosen interval of distances I. In the classical deviation test this is not guaranteed because the distributions of the residuals for different distances r can differ greatly. Thus it makes sense to transform given test functions and to scale the residuals in order to obtain similar importance for all r in I. We demonstrate the effect of such modifications of deviation tests.

A summary function frequently used in point process statistics for stationary processes and in deviation tests is Ripley's K-function (Ripley, 1976, 1977). For this function, commonly the transformation $L(r) = \sqrt{K(r)/\pi}$ (for processes in $\mathbb{R}^2$) called L-function is used instead of K. This dates back to Besag (1977) who found that under CSR a standard estimator of the L-function has approximately constant variance over the distances r. This variance-stabilizing transformation leads to tests that are considered "better" than tests based on the K-function. Variance-stabilizing transformations are available also in some other cases, see e.g. Schladitz and Baddeley (2000) and Grabarnik and Chiu (2002), where the fourth root of an originally used summary function (a third order analogue of Ripley's K function) was applied. Moreover, the Aitkin–Clayton transformation $\arcsin(\sqrt{1-\cdot})$ stabilizes variances for the nearest neighbor distance distribution function (G-function) and the empty space function/spherical contact distribution (F-function) (see Aitkin and Clayton, 1980). In this paper, we employ the L-transformation to mark-weighted K-functions. Of course, other transformations are possible to make the variance more stable, for example employing the log transformation of the