Second-Order Analysis of Spatial Point Process

Tonglin Zhang
Outline

- Spatial Point Processes
- Intensity Functions
- Mean and Variance
- Pair Correlation Functions
- Stationarity
- K-functions
- Some Asymptotics
- Nonstationary Model
- Estimation
Definition

The definition of point processes is well-established and can be found in many textbooks. Overall, a point process is defined on a complete separable metric space $S \subseteq \mathbb{R}^d$. Let $N = N(A)$ be the number of points observed in $A \in \mathcal{I}$, where $\mathcal{I}$ is the collection of Borel subsets of $S$. Then, $N(A)$ is finite if $A$ is bounded.
The \( k \)-th order intensity function of \( N \) (if it exists) as

\[
\lambda_k(s_1, \cdots, s_k) = \lim_{|ds_i| \to 0, i=1,\cdots,k} \left\{ \frac{E[N(ds_1) \cdots N(ds_k)]}{|ds_1| \cdots |ds_k|} \right\},
\]

where \( s_i \) are distinct points in \( S \), \( ds_i \) is an infinitesimal region containing \( s_i \in S \), and \( |ds_i| \) is the Lebesgue measure of \( ds_i \).
The first and second-order intensity functions

It is of interest in the first and second-order intensity functions. The first-order intensity function is

\[ \lambda(s) = \lim_{|ds| \to 0} \frac{E[N(ds)]}{|ds|} \]

and the second-order intensity function is

\[ \lambda_2(s_1, s_2) = \lim_{|ds_1|, |ds_2| \to 0} \frac{E[N(ds_1)N(ds_2)]}{|ds_1| |ds_2|} \]

where \( s_1 \neq s_2 \).
Mean and Variance Functions

The mean measure of $N$ is

$$\mu(A) = \int_A \lambda(s) ds,$$

where $\lambda(s) = \lambda(s)$ is the first order intensity function. The second-order moment measure is

$$\mu^{(2)}(A_1, A_2) = \int_{A_1} \int_{A_2} \lambda_2(s_1, s_2) ds_2 ds_1.$$
Then, the mean structure of $N$ is

$$E[N(A)] = \mu(A)$$

and the covariance structure of $N$ is

$$\text{Cov}[N(A_1), N(A_2)] = \mu^{(2)}(A_1, A_2) + \mu(A_1 \cap A_2) - \mu(A_1)\mu(A_2)$$

$$= \int_{A_1} \int_{A_2} [\lambda_2(s_1, s_2) - \lambda(s_1)\lambda(s_2)] ds_2 ds_1 + \int_{A_1 \cap A_2} \lambda(s) ds.$$
Pair Correlation Function

Let

\[ g(s_1, s_2) = \frac{\lambda_2(s_1, s_2)}{\lambda(s_1)\lambda(s_2)}. \]

Then, \( g \) is called the pair correlation function and the covariance structure is

\[
\text{Cov}[N(A_1), N(A_2)] = \int_{A_1} \int_{A_2} [g(s_1, s_2) - 1] \lambda(s_1)\lambda(s_2) ds_2 ds_1 + \int_{A_1 \cap A_2} \lambda(s) ds.
\]
Strong Stationarity

A spatial point process $N$ is said \textit{strong stationary} if for any measurable $A_1, \cdots, A_k \in \mathcal{B}(\mathbb{R}^d)$ the joint distribution of

$$N(A_1 + s), \cdots, N(A_k + s)$$

does not depend on $s$, where

$$A_i + s = \{s' + s : s' \in A_i\}.$$
Second-Order Stationarity

If $\lambda(s)$ is a constant and

$$\lambda_2(s_1 - s_2) = \lambda_2(s_1 - s_2),$$

then $N$ is called second-order stationary. In addition, if

$$\lambda_2(s_1 - s_2) = \lambda_2(\|s_1 - s_2\|),$$

then $N$ is called isotropic.

If $N$ is isotropic, then

$$g(s_1, s) = g(\|s_1 - s_2\|).$$
K-functions

Suppose $N$ is stationary. Let $\lambda$ be the first-order intensity function. Then, the $K$-function is defined by

$$K(t) = \frac{1}{\lambda} E[\text{number of extra events within distance of } t \text{ of a randomly chosen event}].$$

The $L$-function is

$$L(t) = \sqrt{\frac{K(t)}{\pi}} - t.$$

In real application, $K(t)$ is more often used.
Estimation

Estimation of K-function, L-function and pair correlation can be found in $\mathbb{R}$. The method is

$$\hat{K}(t) = \frac{1}{\hat{\lambda}} \sum_i \sum_{j \neq i} w_{ij} \frac{I(d_{ij} < t)}{n},$$

where $d_{ij}$ is the distance, $I$ is the indicator function, $w_{ij}$ is the edge correlation. In general, we choose $w_{ij} = 1$ if the disc centered at $s_i$ with radius $t$ is completely inside. Otherwise, $w_{ij}$ is just the proportion of the disc inside the domain.
Strong mixing condition

We say \( N \) satisfies the \textit{strong mixing condition} if

\[
\lim_{a \to \infty} \alpha(ar, ad) = 0,
\]

where

\[
\alpha(r, d) = \sup_{\rho(E_1, E_2) \geq r} \sup_{\substack{E_1 \in \mathcal{F}(E_1) \leq d, \ \ E_2 \leq d}} \left| P(U_1 \cap U_2) - P(U_1)P(U_2) \right|,
\]

\[
d(E) = \sup_{s, s' \in E} \rho(s, s') \text{ is the diameter of } E,
\]

\[
d(E_1, E_2) = \sup_{s \in E, s' \in E'} \rho(s, s') \text{ is the maximum distance between disjoint sets of points.}
\]
Some Results

(Theorems 4.1 and 4.3 in Ivanoff (1982)). Let $N$ be a strong stationary spatial point process and also satisfies the strong mixing condition. If

$$0 < C^2 = \int_{\mathbb{R}^d} \Gamma(0, s) ds < \infty,$$

where $\Gamma(0, s)$ is the covariance function of $N$, then

$$M_\eta(A) = \frac{N(\eta A) - E[N(\eta A)]}{\eta^{d/2}} \xrightarrow{D} N(0, |A|)$$

for any $A \subseteq \mathbb{R}^d$ and

$$(M_\eta(A_1), \ldots, M_\eta(A_k)) \xrightarrow{D} (M(A_1), \ldots, M(A_k))$$

for any disjoint Borel $A_1, \ldots, A_k \subseteq \mathbb{R}^d$, where $M(A_1), \ldots, M(A_k)$ are independent normal random variables.
(Corollary 7.2 in Ivanoff (1982)). If the $k$th-order factorial cumulant density function $Q_k$ of $N$ satisfies

$$\int_{s_1, \ldots, s_k \in \mathbb{R}^d} |Q_k(s_1, \ldots, s_k)| \, ds_1 \cdots ds_k < C_1, \quad k = 2, 3, 4,$$

and

$$\int_{s_1, \ldots, s_k \in \mathbb{R}^d} |Q_k(s_1, \ldots, s_k)| \, ds_1 \cdots ds_{k-1} < C_2, \quad k = 2, 3, 4$$

for some constants $C_1, C_2 \in \mathbb{R}$, then $M_\eta(\prod_{i=1}^d (0, t_i])$ with $t_i > 0$ for $i = 1, \ldots, d$ weakly converges to a Gaussian process which can be determined by the previous conclusion.
The Second-Order Intensity-Reweighted Stationary Model

Baddeley, Moller and Waagepetersen (2000) propose the model. It requires that the pair correlation function has

\[ g(s_1, s_2) = \frac{\lambda_2(s_1, s_2)}{\lambda(s_1)\lambda(s_2)} = g(s_1 - s_2) \]

some times, it even requires that

\[ g(s_1, s_2) = g(\|s_1 - s_2\|). \]
If the process has an isotropic pair correlation function, then

\[ K(s) = 2\pi \int_0^s u g(u) \, du, \quad s > 0. \]

For an inhomogeneous Poisson process, \( K(s) = \pi s^2 \). For a cluster process, \( K(s) > \pi s^2 \). For an inhibition process, \( K(s) < \pi s^2 \).
Stochastic Integral

Let $S$ be a set. Then,

$$N(S) = \int_A N(ds)$$

and

$$E[N(S)] = \int_S E[N(ds)] = \int_A \lambda(s) ds.$$
We consider

\[ T = \sum \sum \sum \frac{1}{|S_{s_i - s_j}|} \]

\[ = \int_S \int \int \frac{1}{|S_{s_i - s_j}|} N(ds')N(ds). \]
Then,

\[ E(T) = \int \int_{S} \int_{\|s-s'\| \leq \epsilon, s \neq s'} \frac{1}{|S_s-s'|} E[N(ds')N(ds)] \]

\[ = \int \int_{S} \int_{\|s-s'\| \leq \epsilon} \lambda_2(s-s') \frac{1}{|S_s-s'|} ds' ds \]

\[ = \int \int_{t \in S_s, \|s-t\| \leq \epsilon} \frac{\lambda_2(t)}{|S_s-s'|} ds' dt \]

\[ = \pi \epsilon^2. \]

Therefore, if we assume \( N \) is strong stationary, then

\[ \hat{C}^2 = \frac{1}{\hat{\lambda}} \sum \sum \frac{1}{|S_{s_i-s_j}|} - \pi \epsilon^2 \hat{\lambda} + 1 \]

where \( \hat{\lambda} \) is \( N(S)/|S| \).
Using the same idea, we can show that the estimate of $K$ function under the inhomogeneous model is

$$
\hat{K}(t) = \frac{1}{|S|} \sum_{s-s' \in S, s \neq s'} \sum \frac{I(\|s - s'\| \leq t)}{\hat{\lambda}(s)\hat{\lambda}(s')}.
$$

This is a moment estimator of the $K$-function.
Likelihood function for the first-order intensity function

If $N$ is Poisson process, then the loglikelihood function is

$$\ell(\beta) = \sum_{i=1}^{n} \log[\lambda(s_i, \beta)] - \int_S \lambda(s, \beta) ds.$$
Second-Order Parametric and Nonparametric Inference

- There is no second-order parametric model.
- Even there is a parametric model, it is still unknown how to write down the likelihood function.
- Although there are a few method to estimate the second-order parameter, the behavior is still unknown.

