Statistica Sinica **27** (2017), 207-227
doi:http://dx.doi.org/10.5705/ss.202015.0232

ON INDEPENDENCE AND SEPARABILITY BETWEEN POINTS AND MARKS OF MARKED POINT PROCESSES

Tonglin Zhang

Purdue University

Abstract: An important problem in statistical methods for marked point processes (MPPs) is to evaluate the relationship between points and marks, which can be developed under either the concept of independence or the concept of separability. Although both have been used, the connection between these two concepts is still unclear in the literature. The present article provides a way to evaluate such a connection, concluding that the concept of independence and the concept of separability are equivalent if the Kolmogorov consistency condition is satisfied, but not otherwise. We also provide a testing method to assess first-order independence between points and marks, where first-order independence is concluded if the test statistic is insignificant and first-order dependence is concluded if the test statistic is significant. The performance of the testing method is evaluated under simulation and case studies.

Key words and phrases: Counting measure, independence, Janossy measure, Kolmogorov consistency condition, marked point processes, separability.

1. Introduction

An important problem in statistical methods for marked point processes (MPPs) is to evaluate the relationship between points and marks, which can be considered under either the concept of independence or the concept of separability. Although both have been used (Schlather, Ribeiro and Diggle (2004); Schoenberg (2004)), the connection between the two concepts is still unclear in literature. The goal here is to provide a rigorous theory to evaluate such a connection. The theory will provide a basis for future development of statistical methods of MPPs in both model specifications and applications.

MPPs are commonly used in a wide variety of applications when observations are described by spatial or spatiotemporal locations (i.e. points) and their corresponding measurements (i.e. marks). Methods of MPPs are often used to model a number of natural hazard events or other phenomena located in space and time, and many successful applications can be found. Examples include earthquakes (Holden, Sannan and Bungum (2003); Ogata (1998); Zhuang, Ogata and Vere-Jones (2002)), forest wildfires (Peng, Schoenberg and Woods (2005);

Schoenberg (2004); Zhang and Zhuang (2014)), tree locations and sizes (Guan (2008)), and extreme events (Brown, Caesar and Ferro (1950); Hall and Tajvidi (2000)). Among these, a three-dimensional space is involved if spatiotemporal locations are considered, a two-dimensional space is involved if spatial locations are considered, and a one-dimensional space is involved if temporal locations are considered. The MPPs discussed here cover all the three cases.

Statistical approaches to MPPs rely on definitions. Classical geostatistical methods including variogram analysis, various kinds of kriging, and geostatistical techniques (Cressie (1993)) is used to analyze marked point patterns. These methods rely on the assumption often violated in applications that the dependence between points and marks can be ignored (Diggle, Ribeiro and Christensen (2003)). For instance, the relative positions of trees in a forest have repercussions on their size owing to their competition for light or nutrient (Schlather, Ribeiro and Diggle (2004)). Forest wildfire activities exhibit power-law relationships between frequency and burned area (Malamud, Millington and Perry (2005)).

For effective statistical approaches in applications, we need to describe and understand the relationship between points and marks. Two approaches have been proposed. The first is developed under the concept of independence (Guan and Afshartous (2007); Schlather, Ribeiro and Diggle (2004)) formulated under the framework of the distribution theory using the Janossy measure (Janossy (1950)). The second is developed under the concept of separability (Schoenberg (2004)) formulated under the framework of intensity theory using counting measure (Daley and Vere-Jones (2003)). The evaluation of the connection between the two is important in both theory and applications.

It is convenient in modeling, estimation, and prediction of marked point patterns if marks and points are independent or separable. Many commonly used Hawkes models, such as the epidemic-type aftershock sequences (ETAS) model (Ogata (1998)), assume marks and points are separable. Several **R** packages, such as the **spatstat** (Baddeley and Turner (2005)) and **PtProcess** (Harte (2010)), can be used if the assumption of independence or separability holds. If the assumption is violated, then intensity-dependent models can be considered (Ho and Stoyan (2008); Malinowski, Schlather and Zhang (2014); Myllymäki and Penttinen (2009)). Before using these methods, it is necessary to account for dependence. A few testing methods have been proposed. These include a test for stationarity and isotropy of an MPP using variograms (Assuncao and Maia (2007); Schlather, Ribeiro and Diggle (2004)), a nonparametric kernel-based test to assess the separability of the first-order intensity function (Schoenberg (2004)), a χ^2-based test to assess the interaction effect between points and marks (Guan and Afshartous (2007)), and a Kolmogorov-Smirnov type test for independence (Zhang (2014)). However, theoretical connections between the two concepts are still unclear.

We investigate the connection between the methods of the Janossy measure and counting measure for point processes. We find that the concepts of independence and separability are equivalent under a few weak regularity conditions but not otherwise, where the most important condition is Kolmogorov consistency.

The rest of the article is organized as follows. Section 2 provides the theory for the connection between independence and separability, and includes a review of the theories of MPPs based on the Janossy measure and the counting measure. Section 3 provides a testing method for first-order independence or separability. Section 4 records simulation results for the performance of the testing method. Section 5 applies the testing method to an earthquake study. Section 6 provides a discussion. Proofs of the theorems are given in the online supplementary material.

2. Marked Point Processes

The definition of MPPs can be found in many textbooks (Daley and Vere-Jones (2003); Karr (1991)). Overall, an MPP can be treated as an unmarked point process on the product space of points and marks, but the concept has its own life in applications. Let $\mathcal{S} \in \mathscr{B}(\mathbb{R}^d)$ and $\mathcal{M} \in \mathscr{B}(\mathbb{R}^q)$ be the domains of points and marks, respectively. Let $\mathscr{S} = \mathscr{B}(\mathcal{S})$, $\mathscr{M} = \mathscr{B}(\mathcal{M})$, $\mathcal{X} = \mathcal{S} \times \mathcal{M}$, and $\mathscr{X} = \sigma(\mathscr{S} \times \mathscr{M})$. An MPP $\mathcal{N}$ with points in $\mathcal{S}$ and marks in $\mathcal{M}$ is an unmarked point process on $\mathcal{X}$ with $\mathcal{N}_s(A) = \mathcal{N}(A \times \mathcal{M}) < \infty$ for any bounded $A \in \mathscr{S}$, where $\mathcal{N}(A \times B)$ is the number of events in $A \times B$. If $\mathcal{S}$ is bounded, then $N = \mathcal{N}(\mathcal{X})$ is an almost finite discrete random variable. Let n be the observed value of N. If $n \geq 1$ then, based on an artificial order, the observations of $\mathcal{N}$ can be expressed as $\{\mathbf{x}_i = (\mathbf{s}_i, \mathbf{m}_i) \in \mathcal{X} : i = 1, \ldots, n\}$.

The distribution of $\mathcal{N}$ can be defined using methods of unmarked point processes. Two methods have been proposed, based on Janossy measure (Janossy (1950)) and counting measure (Daley and Vere-Jones (2003)), respectively. Although the second is more popular, the first is also important.

2.1. Janossy measure

The distribution of unmarked point process using Janossy measure has been well-discussed (e.g., Moyal (1962); Daley and Vere-Jones (2003)); it can be easily modified to generate the distribution of MPPs in three steps. It generates the total number of events N in the first, the points in the second, and the marks in the third. We write $p_n = P(N = n)$ and, conditioning on $N = n$, the joint distribution of events, points and marks, is given by π_n. Then, $\sum_{n=0}^{\infty} p_n = 1$ if $\mathcal{S}$ is bounded and, for each $n \geq 1$, the probability measure π_n is defined on $\mathcal{X}^n$, the n-fold product space of $\mathcal{X}$. Let $\mathscr{X}^n$ be the minimal σ-field of sets in $\mathcal{X}^n$, $\Omega = \sum_{n=0}^{\infty} \mathcal{X}^n$, and $\mathscr{F} = \sum_{n=0}^{\infty} \mathscr{X}^n$, where $\mathcal{S}^0 = \{\mathbf{0}_d \in \mathbb{R}^d\}$, $\mathcal{M}^0 = \{\mathbf{0}_q \in \mathbb{R}^q\}$, $\mathcal{X}^0 = \{\mathbf{0}_{d+q} \in \mathbb{R}^{d+q}\}$. Then, $\mathscr{S}^0 = \{\phi, \{\mathbf{0}_d\}\}$, $\mathscr{M}^0 = \{\phi, \{\mathbf{0}_q\}\}$, and $\mathscr{X}^0 =$

$\{\phi, \{\mathbf{0}_{d+q}\}\}$, where ϕ is the empty set and $\mathbf{0}_k$ is the k-dimensional vector with all components equal to 0.

Let $\mathbf{x}^{(n)} = (\mathbf{x}_{1,n}, \ldots, \mathbf{x}_{n,n})$ be the ordered state and $\mathbf{x}^n = \{\mathbf{x}_{1,n}, \ldots, \mathbf{x}_{n,n}\}$ be the unordered state of individual events, respectively. Let $C^{(n)} = (C_1, \ldots, C_n)$ be the ordered sets and $C^n = \{C_1, \ldots, C_n\}$ be the unordered sets of subsets $C_1, \ldots, C_n \in \mathscr{X}$, respectively. Then $\mathscr{F}$ is the minimal σ-field of sets in Ω containing all sets $C^{(n)} \in \mathscr{X}^n$ for $n \in \mathbb{N} = \{0, 1, \ldots, \}$. The pair $(\Omega, \mathscr{F})$ is called the *measurable space* of $\mathcal{N}$ under the method of the Janossy measure, used to define the probability measure $\mathbb{P}$ of $\mathcal{N}$. A probability measure π_n is uniquely determined by

$$\pi_n(C^{(n)}) = P(\mathbf{x}^{(n)} \in C^{(n)}) = P(\mathbf{x}_1 \in C_1, \ldots, \mathbf{x}_n \in C_n). \tag{2.1}$$

If π_n is continuous with respect to Lebesgue measure on $\mathcal{X}^n$, then there is a unique nonnegative function $f_n(\mathbf{x}^{(n)})$ on $\mathcal{X}^n$ such that for any $C^{(n)} \in \mathcal{X}^n$ there is $\pi_n(C^{(n)}) = \int_{C^{(n)}} f_n(\mathbf{x}^{(n)}) d\mathbf{x}^{(n)}$. To be consistent with treating $\mathcal{N}$ as a theory of unordered sets, we assume that π_n is permutation invariant, $\pi_n(C^{(n)}) = \pi_n(C^n)$ or $\pi_n(C_1, \ldots, C_n) = \pi_n(C_{i_1}, \ldots, C_{i_n})$ for any permutation $(i_1, \ldots, i_n)$ of $(1, \ldots, n)$. Then, f_n is also permutation invariant.

If there exist $A_1, \ldots, A_n \in \mathscr{S}$ and $B_1, \ldots, B_n \in \mathscr{M}$ such that $C_i = A_i \times B_i$ for $i = 1, \ldots, n$, then there is $\pi_n(C^{(n)}) = \pi_{n,m|s}(\mathbf{m}_1 \in B_1, \ldots, \mathbf{m}_n \in B_n | \mathbf{s}_1, \ldots, \mathbf{s}_n)$ $\pi_{n,s}(\mathbf{s}_1 \in A_1, \ldots, \mathbf{s}_n \in A_n)$. According to *Dynkin*'s π-λ theorem (e.g., Theorem 3.3 of Billingsley (1995)), π_n is uniquely determined. Therefore, the distribution of $\mathcal{N}$ should contain a discrete probability distribution $\{p_n : n \in \mathbb{N}\}$ (with $\sum_{n=0}^{\infty} p_n = 1$), the conditional distribution (given n) of points $\pi_{n,s}$, and the conditional distribution (given n and $\mathbf{s}_1, \ldots, \mathbf{s}_n$) of marks $\pi_{n,m|s}$.

Theorem 1. *Let $P^{(n)}$ be a finite measure on $\mathscr{X}^n$ for a given $n \in \mathbb{N}$. If $\sum_{n=0}^{\infty} P^{(n)}(\mathcal{X}^n) = 1$, then the function $\mathbb{P}$ given by $\mathbb{P}(C) = \sum_{n=0}^{\infty} P^{(n)}(A^{(n)} \times B^{(n)})$ for any $C = \sum_{n=0}^{\infty} A^{(n)} \times B^{(n)}$ with $A^{(n)} \in \mathscr{S}^n$ and $B^{(n)} \in \mathscr{M}^n$ is the unique probability measure on $\mathscr{F}$ whose restriction on $\mathscr{X}^n$ agrees with $P^{(n)}$ for all $n \in \mathbb{N}$.*

For any probability measure $\mathbb{P}$ on $\mathscr{F}$, we can define its restriction $P^{(n)}$ to $\mathcal{X}^n$ as $P^{(n)}(C) = \mathbb{P}(C)$ if $C \in \mathscr{X}^n$. Then, $P^{(n)}$ is a measure on $\mathscr{X}^n$. Therefore, the expression of $\mathbb{P}$ provides a way to interpret a probability distribution of $\mathcal{N}$.

Let μ_s and μ_m be some σ-finite measures on $\mathscr{S}$ and $\mathscr{M}$, and μ_s^n and μ_m^n be their n-fold product measures on $\mathscr{S}^n$ and $\mathscr{M}^n$, respectively. If $\pi_{n,s}$ and $\pi_{n,m|s}$ are continuous in μ_s^n and μ_m^n, then using the Radon-Nykodym Theorem (Billingsley (1995)) there is

$$\mathbb{P}(C) = \sum_{n=0}^{\infty} p_n \int_{A^{(n)}} \left[\int_{B^{(n)}} f_{n,m|s}(\mathbf{m}^{(n)} | \mathbf{s}^{(n)}) d\mu_m^n \right] f_{n,s}(\mathbf{s}^{(n)}) d\mu_s^n, \tag{2.2}$$

where $C = \sum_{n=0}^{\infty} A^{(n)} \times B^{(n)}$, $\{\mathbf{s}\} = \{\mathbf{s}^{(n)} \in \mathcal{S}^n : n \in \mathbb{N}\}$, and $\{\mathbf{m}\} = \{\mathbf{m}^{(n)} \in \mathcal{M}^n : n \in \mathbb{N}\}$ such that $\omega = (\{\mathbf{s}\}, \{\mathbf{m}\}) \in \Omega$. The expression contains a discrete probability distribution for the total number of points $\{p_n : n \in \mathbb{N}\}$, a class of conditional densities of points on the total number of points $\{f_{n,s} : n \in \mathbb{N}\}$, a class of conditional densities of marks on point locations and total number of points $\{f_{n,m|s} : n \in \mathbb{N}\}$. A probability distribution of $\mathcal{N}$ can be almost surely expressed by these three classes as displayed in (2.2). If μ_s^n and μ_m^n are the Lebesgue measures on $\mathcal{S}^n$ and $\mathcal{M}^n$ for any $n \geq 1$, respectively, then $\mathbb{P}$ can be expressed by its derivatives with respect to the Lebesgue measure μ on $\mathscr{F}$. Combining μ with the ordinary counting measure on $\mathbb{N}$, a mixed counting-Lebesgue measure μ on $\mathscr{F}$ is derived. Based on μ, the mixed PMF-PDF function for $\mathbb{P}$ can be expressed as $f(\omega) = \sum_{n=0}^{\infty} p_n f_{n,m|s}(\mathbf{m}^{(n)}|\mathbf{s}^{(n)}) f_{n,s}(\mathbf{s}^{(n)})$. Then $\mathbb{P}(C) = \int_C f(\omega)\mu(d\omega)$ for any $C \in \mathscr{F}$. The method based on the mixed counting-Lebesgue measure μ is useful for continuous marks when $\mathcal{N}_s$ is simple.

As for independence between points and marks, note that the function $f_{n,m|s}$ depends on both n and point locations. If marks are independent of points, then $f_{n,m|s}$ should depend on n only.

Definition 1. (Independence) If the distribution of $\mathcal{N}$ is expressed by (2.2), points and marks of $\mathcal{N}$ are independent if $f_{n,m|s}(\mathbf{m}^{(n)}|\mathbf{s}^{(n)})$ is independent of $\mathbf{s}^{(n)}$ for any $n \in \mathbb{N}$, and $\mathcal{N}$ is an independent MPP if its points and marks are independent.

For $\mathcal{N}$ an independent MPP, for any $n \in \mathbb{N}$ there is $\pi_{n,m|s} = \pi_{n,m}$, where $\pi_{n,m}$ is the marginal distribution of marks by integrating out points, $\pi_{n,m}(B^{(n)}) = \pi_n(\mathcal{S}^n \times B^{(n)})$ for any $B^{(n)} \in \mathscr{M}^n$. Then, Definition 1 can also be expressed as

$$\pi_n(A^{(n)} \times B^{(n)}) = \pi_{n,s}(A^{(n)})\pi_{n,m}(B^{(n)}). \tag{2.3}$$

Thus for any $C \in \mathcal{F}$ there is $\mathbb{P}(C) = \sum_{n=0}^{\infty} p_n \pi_{n,m}(B^{(n)}) \pi_{n,s}(A^{(n)})$, and for any $\omega \in \Omega$ there is $f(\omega) = \sum_{n=0}^{\infty} p_n f_{n,m}(\mathbf{m}^{(n)}) f_{n,s}(\mathbf{s}^{(n)})$.

2.2. Counting measure

Instead of focusing on the distribution of $\mathcal{N}$, the method of counting measure focuses on the number of events using intensity functions. For any distinct $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathcal{X}$, $k \in \mathbb{N}^+ = \{1, 2, \cdots\}$, the kth order intensity function of N (if it exists) is defined as

$$\lambda_k(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \lim_{|d\mathbf{x}_i| \to 0, i=1,\ldots,k} \left\{ \frac{\mathbb{E}[\prod_{i=1}^k N(d\mathbf{x}_i)]}{\prod_{i=1}^k |d\mathbf{x}_i|} \right\}, \tag{2.4}$$

where $d\mathbf{x}_i$ is an infinitesimal region containing $\mathbf{x}_i \in \mathcal{S}$ and $|d\mathbf{x}_i|$ is its Lebesgue measure. For convenience, one often writes $\lambda(\mathbf{x}) = \lambda_1(\mathbf{x})$. The moments of

$\mathcal{N}$ can be derived using the method for unmarked point processes (Moller and Waagepetersen (2007)) that provides

$$\mu(C) = \mathbb{E}[N(C)] = \int_C \lambda(\mathbf{x})d\mathbf{x}$$

and

$$\mathbb{C}\mathrm{ov}[N(C_1), N(C_2)] = \int_{C_1}\int_{C_2} [g(\mathbf{x}_1, \mathbf{x}_2) - 1]\lambda(\mathbf{x}_1)\lambda(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2 + \mu(C_1 \cap C_2),$$

where $g(\mathbf{x}_1, \mathbf{x}_2) = [\lambda_2(\mathbf{x}_1, \mathbf{x}_2) - 1]/[\lambda(\mathbf{x}_1)\lambda(\mathbf{x}_2)]$ is the pair correlation function.

To compare the method of counting measure with the method of the Janossy measure, we restrict our attention to the case that N is almost surely finite. If the total number of events is n with points and marks being denoted by $\mathbf{x}_{1,n}, \ldots, \mathbf{x}_{n,n}$ then, based on an artificial order of events, there is

$$\mathcal{N}(C|\mathbf{x}^{(n)}) = \sum_{i=1}^{n} I_C(\mathbf{x}_{i,n}), \tag{2.5}$$

where $I_C(\mathbf{u})$ is the indicator function and the right side of (2.5) is permutation invariant. As the counting measure uses $\mathcal{N}(C)$ without conditioning on the total number of events n, (2.5) cannot be directly used. To be consistent, we express $\mathcal{N}(C)$ as

$$\mathcal{N}(C) = \mathcal{N}(C|\omega) = \sum_{n=0}^{\infty} \mathcal{N}(C|\mathbf{x}^{(n)}) = \sum_{n=0}^{\infty}\sum_{i=1}^{n} I_C(\mathbf{x}_{i,n}) \tag{2.6}$$

for any $\omega = \{\mathbf{x}^{(n)} : n \in \mathbb{N}\} \in \Omega$. Using the distribution of ω, which is determined by the distribution of $\mathbf{x}^{(n)}$ for each n, we have $\mathbb{E}[\mathcal{N}(C)] = \mathbb{E}\{\mathbb{E}[\mathcal{N}(C|\omega)]\} = \sum_{n=1}^{\infty}\int_{\mathcal{X}^n}\sum_{i=1}^{n} I_C(\mathbf{x}_{i,n})dP^{(n)} = \sum_{n=1}^{\infty} nP^{(n)}(C \times \mathcal{X}^{n-1})$ for any $C \in \mathscr{X}$. This provides the expression of $\lambda(\mathbf{x})$ if we choose $C = d\mathbf{x} \times \mathcal{X}^{n-1}$ above. Then,

$$\lambda(\mathbf{x}) = \sum_{n=1}^{\infty} np_n f_{1,n}(\mathbf{x}), \tag{2.7}$$

where $f_{k,n}(\mathbf{x})$ with $1 \leq k < n$ and $n \geq 1$ is the kth-order marginal density function of $f_n(\mathbf{x}^{(n)})$ given by $f_{k,n}(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \int_{\mathcal{X}^{n-k}} f_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)d\mathbf{x}_{k+1}\cdots d\mathbf{x}_n$.

We also derive the expression for the kth order intensity function of $\mathcal{N}$ for any other $k \in \mathbb{N}^+$. Let $C_1, \ldots, C_k$ be disjoint subsets of $\mathcal{X}$ containing distinct $\mathbf{x}_1, \ldots, \mathbf{x}_k$, respectively. According to Moyal (1962), there is $\prod_{i=1}^{k}\mathcal{N}(C_i) = \sum_{n=k}^{\infty}\sum_{i_1\neq\cdots\neq i_k}\prod_{j=1}^{k} I(\mathbf{x}_{i_j,n} \in C_j)$ and $\mathbb{E}[\prod_{i=1}^{k}\mathcal{N}(C_i)] = \sum_{n=k}^{\infty} P^{(n)}(C_1 \times \cdots \times C_k \times \mathcal{X}^{n-k})[n!/(n-k)!]$. Using the same method for the derivation of (2.7) with the definition given by (2.4), there is

$$\lambda_k(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} p_n f_{k,n}(\mathbf{x}_1, \ldots, \mathbf{x}_k). \tag{2.8}$$

Therefore, $\lambda_k(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ is well defined if the right side of (2.8) always converges.

Theorem 2. *Assume π_n is continuous with respect to the Lebesgue measure on $\mathcal{X}^n$ for every $n \in \mathbb{N}^+$. Let f_n be the density of π_n. If $\mathbb{E}(N^k)$ exists and for every $r \leq k$ there exists a finite function H_r on $\mathcal{X}^r$ such that $f_{r,n}(\mathbf{x}_1, \ldots, \mathbf{x}_r) \leq H_r(\mathbf{x}_1, \ldots, \mathbf{x}_r)$, then $\lambda_r(\mathbf{x}_1, \ldots, \mathbf{x}_r)$ is well-defined for all $r \leq k$. If $\mathbb{E}(N^k)$ exists for every $k \in \mathbb{N}^+$, then $\lambda_k(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ is always well-defined.*

Definition 2 (Separability)**.** If $\lambda_k(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ exists for every $k \in \mathbb{N}^+$, points and marks of $\mathcal{N}$ are separable if for any $k \in \mathbb{N}^+$ there exist positive $\lambda_{k,s}(\mathbf{s}_1, \ldots, \mathbf{s}_k)$ and $\lambda_{k,m}(\mathbf{m}_1, \ldots, \mathbf{m}_k)$ with $\mathbf{s}_1, \ldots, \mathbf{s}_k \in \mathcal{S}$ and $\mathbf{m}_1, \ldots, \mathbf{m}_k \in \mathcal{M}$ such that

$$\frac{\lambda_k(\mathbf{x}_1, \ldots, \mathbf{x}_k)}{\lambda_{k,s}(\mathbf{s}_1, \ldots, \mathbf{s}_k)\lambda_{k,m}(\mathbf{m}_1, \ldots, \mathbf{m}_k)} \tag{2.9}$$

does not depend on $\mathbf{x}_1, \ldots, \mathbf{x}_k$; $\mathcal{N}$ is a separable MPP if its points and marks are separable.

One can compare the concept of separability given by Definition 2 with the concept of separability of space-time correlation or covariance functions in geostatistics (Cressie and Huang (1999)). Let $c(\mathbf{h}, u)$ be a stationary space-time correlation or covariance function in geostatistics, where $\mathbf{h} \in \mathbb{R}^d$ represents space and $u \in \mathbb{R}$ represents time. If $c(\mathbf{h}, u)$ is chosen as a correlation function, then separability means $c(\mathbf{h}, u) = c_s(\mathbf{h})c_t(u)$ for any $\mathbf{h}$ and u, where $c_s(\mathbf{h}) = c(\mathbf{h}, 0)$ and $c_t(u) = c(\mathbf{0}, u)$. If $c(\mathbf{h}, u)$ is chosen as a covariance function, then separability means $c(\mathbf{h}, u) = c_s(\mathbf{h})c_t(u)/c(\mathbf{0}, 0)$ for any $\mathbf{h}$ and u, where $c(\mathbf{0}, 0)$ is the variance of the variable of interest. In Definition 2, we usually do not have $\lambda_k(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \lambda_{k,s}(\mathbf{s}_1, \ldots, \mathbf{s}_k)\lambda_{k,m}(\mathbf{m}_1, \ldots, \mathbf{m}_k)$. A quick example is the Poisson case where $\lambda_1(\mathbf{x}_1) = [\lambda_{1,s}(\mathbf{s}_1)\lambda_{1.m}(\mathbf{m}_1)]/\mathbb{E}(N)$ under the assumption of separability.

2.3. The relationship

As the formal concepts of independence and separability may be too strong for applications, we provide their weaker versions.

Definition 3. If the distribution of $\mathcal{N}$ is expressed as (2.2), points and marks of $\mathcal{N}$ are rth-order independent if $f_{n,m|s}(\mathbf{m}^{(n)}|\mathbf{s}^{(n)})$ is independent of $\mathbf{s}^{(n)}$ for every $n \leq r$.

Definition 4. If $\lambda_k(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ positively exists for every $k \leq r$, points and marks of $\mathcal{N}$ are rth-order separable if (2.9) holds for every $k \leq r$.

The concepts of independence and separability are equivalent under some conditions.

Regularity Conditions:

(C1) π_n satisfies the Kolmogorov consistency condition $\pi_{n+k}(C^{(n)} \times \mathcal{X}^k) = \pi_n(C^{(n)})$ for every $n, k \geq 1$ and $C^{(n)} \in \mathscr{X}^n$;

(C2) π_n is continuous with respect to Lebesgue measure on $\mathcal{X}^n$ for any $n \geq 1$;

(C3) $\mathbb{E}(N^k)$ exists for a certain $k \in \mathbb{N}^+$.

The Kolmogorov consistency condition (C1) is the most important to justify in practice. In modeling MPPs one does not need to consider change in distribution if events with small mark values are ignored.

Theorem 3. *If* (C1)$-$(C3) *hold, a sufficient and necessary condition for $\mathcal{N}$ to be kth-order separable is that $\mathcal{N}$ be kth-order independent.*

Corollary 1. *If* (C1)$-$(C3) *hold with* (C3) *holding for every $k \in \mathbb{N}^+$, a sufficient and necessary condition for $\mathcal{N}$ to be separable is that $\mathcal{N}$ be independent.*

2.4. Examples

We provide six examples to evaluate the importance of (C1). The condition is satisfied in the first five examples but not in the last. We have $\mathcal{S}\times\mathcal{M} \subseteq \mathbb{R}^d \times \mathbb{R}^q$ with almost surely finite N in all of these examples. We use $\mathcal{N}_s$ to represent the unmarked underlying point process of $\mathcal{N}$ that ignores the marks.

Example 1 (*Poisson Model*)**.** If $\mathcal{N}$ is Poisson, then $\lambda_k(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \prod_{i=1}^k \lambda(\mathbf{x}_i)$, $f_n(\mathbf{x}^{(n)}) = \prod_{i=1}^n f(\mathbf{x}_i)$, and $p_n = \kappa^n e^{-\kappa}/n!$, where $\kappa = \int_{\mathcal{X}} \lambda(\mathbf{x})d\mathbf{x}$ and $f(\mathbf{x}) = \lambda(\mathbf{x})/\kappa$. Then

$$f_{k,n}(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \prod_{i=1}^{k} f(\mathbf{x}_i) \int_{\mathbb{R}^{n-k}} \prod_{i=k+1}^{n} f(\mathbf{x}_i) d\mathbf{x}_{k+1} \cdots d\mathbf{x}_k = f_k(\mathbf{x}_1, \ldots, \mathbf{x}_k) \tag{2.10}$$

for any $k \leq n$, and f_n satisfies the Kolmogorov consistency condition. Therefore, for Poisson MPPs, the concept of independence and the concept of separability are equivalent.

Example 2 (*Compound Model*)**.** If $\mathcal{N}$ is a compound MPP, then $f_n(\mathbf{x}^{(n)}) = \prod_{i=1}^n f(\mathbf{x}_i)$ and $\sum_{n=0}^\infty p_n = 1$ for some PMF $\{p_n : n \in \mathbb{N}\}$. Then, $\lambda(\mathbf{x}) = f(\mathbf{x}) \sum_{n=1}^\infty n p_n$ and $\lambda_k(\mathbf{x}_1, \ldots, \mathbf{x}_k) = [\prod_{i=1}^k f(\mathbf{x}_i)][\sum_{n=k}^\infty n! p_n/(n-k)!]$. Therefore, λ_k is well-defined if $\mathbb{E}(N^k)$ is finite. Using (2.10), there is $f_k = f_{k,n}$ for any $k \leq n$. Therefore, the Kolmogorov consistency condition is satisfied. The concept of independence and the concept of separability are equivalent.

Example 3 (*Point Cluster Model*)**.** If $\mathcal{N}_s$ is a cluster point process, then $\mathcal{N}$ is a cluster MPP. A particular form of a cluster point process is the Neyman-Scott process on a bounded $\mathcal{S}$ (Neyman and Scott (1958)), where the parent points form a stationary Poisson process with Poisson numbers of offspring points independently from a multivariate Gaussian distribution around their parent points. In the general case, the parent points may not be stationary and the distribution of offspring points may not be Gaussian. Let the intensity function of the parent points be $\lambda_0(\mathbf{c})$ and the distribution of offspring points be $\psi(\mathbf{s}-\mathbf{c})$ with the total number Poisson(γ), where $\mathbf{c} \in \mathcal{S}$ is the parent point. Then, $f_{n,s}(\mathbf{s}^{(n)}) = \sum_{r=1}^{\infty}\{\prod_{i=1}^{n}[\frac{1}{r}\sum_{j=1}^{r}\psi(\mathbf{s_i}-\mathbf{c}_j)]\}\{\prod_{j=1}^{r} f_0(\mathbf{c}_j)\}\{\kappa_0^r e^{-\kappa_0}/r!\}$ for $n \geq 1$ and $p_n = \sum_{r=1}^{\infty}[(r\gamma)^n e^{-r\gamma}/n!][\kappa_0^r e^{-\kappa_0}/r!]$, where $f_0(\mathbf{c}) = \lambda_0(\mathbf{c})/\kappa_0$ and $\kappa_0 = \int_{\mathcal{S}}\lambda_0(\mathbf{c})d\mathbf{c}$. Thus, $f_{n,s}(\mathbf{s}^{(n)})$ satisfies the Kolmogorov consistency condition. If marks are independently distributed with conditional density $f(\mathbf{m}_0|\mathbf{s}_0)$ ($\mathbf{m}_0 \in \mathcal{M}$ and $\mathbf{s}_0 \in \mathcal{S}$) given points, then $f_n(\mathbf{x}^{(n)}) = [\prod_{i=1}^{n} f(\mathbf{m}_i|\mathbf{s}_i)]f_{n,s}(\mathbf{s}^{(n)})$, which implies that $f_{1,n}(\mathbf{x}) = f(\mathbf{m}|\mathbf{s})f_{1,s}(\mathbf{s})$. Therefore, $\mathcal{N}$ satisfies the Kolmogorov consistency condition and the concept of independence and the concept of separability are equivalent. If $f(\mathbf{m}_0|\mathbf{s}_0)$ does not depend on $\mathbf{s}_0$, then $\mathcal{N}$ is independent and also separable; otherwise $\mathcal{N}$ is neither independent nor separable.

Example 4 (*Mark Geostatistical Model)***.** The model has been previously considered by Schlather, Ribeiro and Diggle (2004). When $q = 1$, individual marks are univariate. Let $m^{(n)}$ be the ordered state of marks. Suppose the Kolmogorov consistency condition is satisfied in $\mathcal{N}_s$ and given $\mathcal{N}_s$ marks follow a mean zero Gaussian process with a stationary variance σ^2 and a correlation function $c(\mathbf{s},\mathbf{s}')$, where $\mathbf{s}$ and $\mathbf{s}'$ represent any two points observed in $\mathcal{N}_s$. Then $m^{(n)}|\mathbf{s}^{(n)}$ follows a multivariate normal distribution with mean zero and covariance $\mathbf{R}_{\mathbf{s}^{(n)}}$, where the (i,j)th entry of $\mathbf{R}_{\mathbf{s}^{(n)}}$ is $r_{ij} = \sigma^2 c(\mathbf{s}_i,\mathbf{s}_j)$. We find

$$f_{k,n}(\mathbf{x}^{(k)}) = f_k(\mathbf{x}^{(k)}) = \frac{e^{-(1/2\sigma^2)m^{(k)}\mathbf{R}^{-1}_{\mathbf{s}^{(k)}}m^{(k)}}}{(2\pi)^k\sigma^k|\det(\mathbf{R}_{\mathbf{s}^{(k)}})|}f_{k,s}(\mathbf{s}^{(k)}).$$

Then $f_1(\mathbf{x}) = [e^{-m^2/(2\sigma^2)}/(\sqrt{2\pi}\sigma)]f_{1,s}(\mathbf{s})$ and

$$f_2(\mathbf{x}^{(2)}) = \frac{e^{-(m_1^2-2c(\mathbf{s}_1,\mathbf{s}_2)m_1m_2+m_2^2)/(2\sigma^2(1-c^2(\mathbf{s}_2,\mathbf{s}_2)))}}{2\pi\sigma^2\sqrt{1-c^2(\mathbf{s}_1,\mathbf{s}_2)}}f_{2,s}(\mathbf{s}^{(2)}).$$

Therefore, $\mathcal{N}$ is first-order but not second-order independent.

Example 5 (*Intensity-Dependent Model)***.** Intensity dependent models have been previously used to account for the influence of points on marks (Ho and Stoyan (2008); Malinowski, Schlather and Zhang (2014); Myllymäki and Penttinen (2009)). Suppose $q = 1$. Use $m^{(n)}$ as in the previous example. Assume $\mathcal{N}_s$ satisfies the Kolmogorov consistency condition and the distribution of marks depends on points via the intensity function of points. If marks occur independently with density $f(m|\lambda_s(\mathbf{s}))$, then $f_{k,n}(\mathbf{x}^{(n)}) = f_k(\mathbf{x}^{(k)}) = [\prod_{i=1}^{k} f(m_i|\lambda_s(\mathbf{s}_i)]f_{k,s}(\mathbf{s}^{(k)})$. For $\mathcal{N}_s$ stationary, marks and points are independent. Since the Kolmogorov consistency condition is satisfied, the concept of independence and the concept of separability are equivalent.

Example 6 (*Violation of the Kolmogorov Consistency Condition)***.** Let $\mathcal{N}$ be an MPP defined on $\mathcal{X} = \mathcal{S} \times \mathcal{M} = [0,1]^d \times \mathbb{R}^q$ and $\{p_n : p_n > 0 \ \forall n \in \mathbb{N}\}$ be the PMF of N. Assume conditioning on $N = n$, the events $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are iid. Then, $f_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \prod_{i=1}^n f_{1,n}(\mathbf{x}_i)$ for each $n \geq 1$. Assume there is a positive integer n_0 such that $f_{1,n}(\mathbf{s}, \mathbf{m}) = g_1(\mathbf{s})h_1(\mathbf{m})$ if $n \leq n_0$ and $f_{1,n}(\mathbf{s}, \mathbf{m}) = g_2(\mathbf{s})h_2(\mathbf{m})$ if $n > n_0$. Then $\lambda(\mathbf{s}, \mathbf{m}) = g_1(\mathbf{s})h_1(\mathbf{m}) \sum_{n=1}^{n_0} np_n + g_2(\mathbf{s})h_2(\mathbf{m}) \sum_{n=n_0+1}^{\infty} np_n$. Here $\mathcal{N}$ is first-order independent but not first-order separable. Since $\mathcal{N}$ does not satisfies the Kolmogorov consistency condition, the concepts of independence and separability are not equivalent.

It is important to know whether the Kolmogorov consistency condition holds in the analysis of MPP data. If marks occur independently, then the justification of the Kolmorogov consistency condition for $\mathcal{N}_s$ is enough. If the condition holds, then we can conclude that points and marks are independent provided that distributions of individual marks do not depend on their locations. However, if marks are spatially correlated, then points and marks are generally second-order dependent. We expect that the Kolmogorov consistency condition is violated if the distribution of an MPP has a significant change after a few events are removed from the data set. Since this usually cannot happen, we believe that the Kolmogorov consistency condition generally holds in practice.

3. Testing Methods

If the Kolmogorov consistency condition holds, then another interest is to test the independence between points and marks. If independence is concluded, then points and marks can be modeled separately; otherwise they must be modeled jointly. To test independence, we provide a method which can be treated as a modification from the classical Kolmogorov-Smirnov test for multivariate independence, where a similar idea has been previously used (Zhang and Zhuang (2014)).

3.1. Kolmogorov-Smirnov test for multivariate independence

The Kolmogov-Smirnov test for multivariate independence appeared in Blum, Kiefer and Rosenblatt (1961). Let $\mathbf{y}$ be a continuous random vector on $\mathbb{R}^p$ with a joint CDF F and marginal CDFs F_j for $j = 1, \ldots, p$. Suppose one wants to test $H_0 : F(y_1, \ldots, y_p) = \prod_{j=1}^p F_j(y_j)$ for all $(y_1, \ldots, y_p) \in \mathbb{R}^p$. The Kolmogorov-Smirnov statistic for multivariate independence is

$$T_{n,i} = \sup_{y_1, \ldots, y_p \in \mathbb{R}} \sqrt{n} \left| \hat{F}(y_1, \ldots, y_p) - \prod_{j=1}^{p} \hat{F}_j(y_j) \right|, \tag{3.1}$$

where n is the sample size, $\hat{F}$ and $\hat{F}_j$ are the joint and marginal sample CDFs, respectively. If observations are identically and independently collected, then under H_0, $T_{n,i} \xrightarrow{D} \|\mathbb{W}_p\|_\infty = \sup_{\mathbf{t}\in[0,1]^p} |\mathbb{W}_p(\mathbf{t})|$, where $\mathbb{W}_p$ is the p-dimensional (standard) *Brownian pillow* (explained in the online supplementary material). The α-level test rejects H_0 if $T_{n,i} > \|\mathbb{W}_p\|_\alpha$, where $\|\mathbb{W}_p\|_\infty$ is the upper α quantile of $\|\mathbb{W}_p\|_\alpha$. As neither the exact nor the approximate distribution of $\mathbb{W}_p$ is known (Keifer (1958); Koning and Protasov (2003); Kruglova (2008, 2010)), a simulation method is used. We find $\|\mathbb{W}_2\|_{0.1} = 0.7298$, $\|\mathbb{W}_2\|_{0.05} = 0.7948$, $\|\mathbb{W}_2\|_{0.01} = 0.9234$, and $\|\mathbb{W}_2\|_{10^{-6}} < 1.5$.

3.2. Modification to marked point processes

Since F and F_j are not well defined in an MPP, $T_{n,i}$ given by (3.1) cannot be directly used. Therefore, one needs to to modify $T_{n,i}$ for MPPs. This idea was previously used (Zhang (2014)).

To modify the test, one can consider a testing problem based on (2.3) with $n = 1$, with the null hypothesis $H_0 : \pi_1(A \times B) = \pi_{1,s}(A)\pi_{1,m}(B)$ for any $A \in \mathscr{S}$ and $B \in \mathscr{M}$, and the alternative hypothesis $H_1 : \pi_1(A \times B) \neq \pi_{1,s}(A)\pi_{1,m}(B)$ for some $A \in \mathscr{S}$ and $B \in \mathscr{M}$. One can test whether $\pi_1(A\times B) = \pi_{1,s}(A)\pi_{1,m}(B)$ for any $A \in \mathcal{A}$ and $B \in \mathcal{B}$, where $\mathcal{A} \subseteq \mathscr{S}$ and $\mathcal{B} \subseteq \mathscr{M}$ are π-systems in $\mathcal{S}$ and $\mathcal{M}$, respectively. If $n \geq 1$, then the test statistic can be

$$T_{n,\mathcal{A},\mathcal{B}} = \sup_{A\in\mathcal{A},B\in\mathcal{B}} \sqrt{n}|\hat{\pi}_1(A \times B) - \hat{\pi}_{1,s}(A)\hat{\pi}_{1,m}(B)|, \tag{3.2}$$

where $\hat{\pi}_1(A \times B) = \mathcal{N}(A \times B)/n$, $\hat{\pi}_{1,s}(A) = \mathcal{N}_s(A)/n$, and $\hat{\pi}_{1,m} = \mathcal{N}_m(B)/n$. If $T_{n,\mathcal{A},\mathcal{B}}$ is large, H_0 is rejected.

There are many choices for $\mathcal{A}$ and $\mathcal{B}$. To find a p-value, a convenient way is to choose $\mathcal{A}$ and $\mathcal{B}$ such that both $\mathcal{A}$ and $\mathcal{B}$ can be generated by a univariate function from $\mathcal{S}$ to $\mathbb{R}$ or from $\mathcal{M}$ to $\mathbb{R}$, respectively. Then one can reject the null hypothesis if $T_{n,\mathcal{A},\mathcal{B}} > \|\mathbb{W}_2\|_\alpha$ if the significance level of the test is α. For example, if $\mathcal{S} \subseteq \mathbb{R}^2$ (or $\mathcal{S} \subseteq \mathbb{S}^2$) and $\mathcal{M} = \mathbb{R}$, then one can choose $\mathcal{A} = \mathcal{A}_{\mathbf{s}_0} = \{\{\mathbf{s} \in \mathcal{S} : \|\mathbf{s} - \mathbf{s}_0\| \leq a\} : a \geq 0\}$ and $\mathcal{B} = \{(-\infty, b] : b \in \mathbb{R}\}$, inducing

$$T_{n,\mathbf{s}_0} = \sup_{a\geq 0,b\in\mathbb{R}} \sqrt{n}\Big|\frac{1}{n}\sum_{i=1}^n I_{\|\mathbf{s}_i-\mathbf{s}_0\|\leq a,m_i\leq b} - \Big(\frac{1}{n}\sum_{i=1}^n I_{\|\mathbf{s}_i-\mathbf{s}_0\|\leq a}\Big)\Big(\frac{1}{n}\sum_{i=1}^n I_{m_i\leq b}\Big)\Big|. \tag{3.3}$$

We recommend rejecting H_0 if $T_{n,\mathbf{s}_0} \geq 0.7948$ at the significance level 0.05 (Zhang (2014)). In practice, one can choose $\mathbf{s}_0$ as a corner point of the study area in the construction of $\mathcal{A}_{\mathbf{s}_0}$.

With rejection of H_0, one concludes that the points and marks are not independent. Since $\sigma(\mathcal{A}_{\mathbf{s}_0})$ is much smaller than the collection of all Borel subsets in $\mathcal{S}$, rejection of H_0 is not enough for independence.

Table 1. Upper α quantiles of $\|\mathbb{W}_{F_1,F_2}\|_\infty$ when F_1 is the CDF of p_1-dimensional mean zero normal distribution with all variances equal to one and all correlations equal to ρ for selected ρ, and F_2 is the CDF of uniform distribution on $[0,1]$, where results were derived based on 10,000 replications.

	Values of $\|\mathbb{W}_{F_1,F_2}\|_\alpha$ for selected α					
	$p_1 = 2$ for selected α			$p_1 = 3$ for selected α		
ρ	0.1	0.05	0.01	0.1	0.05	0.01
hline 0.0	0.8874	0.9507	1.0803	0.9975	1.0566	1.1842
0.5	0.8651	0.9267	1.0509	0.9556	1.0221	1.1515
0.9	0.8191	0.8819	1.0084	0.8711	0.9317	1.0570
0.95	0.8006	0.8650	0.9925	0.8392	0.9096	1.0481
0.99	0.7717	0.8342	0.9587	0.8028	0.8691	1.0012
0.999	0.7583	0.8247	0.9552	0.7743	0.8408	0.9652

3.3. A new method

We provide another method to modify the classical Kolmogorov-Smirnov test for multivariate independence. The method replies on a test for the independence between two continuous random vectors $\mathbf{y}_1 \in \mathbb{R}^{p_1}$ and $\mathbf{y}_2 \in \mathbb{R}^{p_2}$, and is different from the classical problem.

Let the joint CDF of $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$ be F, the marginal CDF of $\mathbf{y}_1$ be F_1, and the marginal CDF of $\mathbf{y}_2$ be F_2. Consider the test for $H_0 : F(\mathbf{y}_1, \mathbf{y}_2) = F_1(\mathbf{y}_1)F_2(\mathbf{y}_2)$ for any $\mathbf{y}_1 \in \mathbb{R}^{p_1}$ and $\mathbf{y}_2 \in \mathbb{R}^{p_2}$. Let $(\mathbf{y}_{11}, \mathbf{y}_{12}), \ldots, (\mathbf{y}_{n1}, \mathbf{y}_{n2})$ be n iid copies of $(\mathbf{y}_1, \mathbf{y}_2)$. Then, the test statistic is

$$T_{n,p_1,p_2} = \sup_{\mathbf{y}_1 \in \mathbb{R}^{p_1}, \mathbf{y}_2 \in \mathbb{R}^{p_2}} \sqrt{n} \left| \hat{F}(\mathbf{y}_1, \mathbf{y}_2) - \hat{F}_1(\mathbf{y}_1)\hat{F}_2(\mathbf{y}_2) \right|,$$

where $\hat{F}$, $\hat{F}_1$, and $\hat{F}_2$ are empirical distributions of F, F_1, and F_2, respectively. Under H_0, there is $T_{n,p_1,p_2} \stackrel{D}{\rightarrow} \|\mathbb{W}_{F_1,F_2}\|_\infty$, where $\mathbb{W}_{F_1,F_2}$ is the (F_1, F_2)-*functional Brownian pillow* on $\mathbb{R}^{p_1+p_2}$. Since the asymptotic distribution of $\|\mathbb{W}_{F_1,F_2}\|_\infty$ depends on F_1 and F_2, the computation of the p-value of T_{n,p_1,p_2} is complicated. The interpretation of $\mathbb{W}_{F_1,F_2}$ is given in our online supplementary material.

We carried out a simulation study to investigate the performance of the limiting distribution of $\|\mathbb{W}_{F_1,F_2}\|_\infty$. We chose F_1 as the CDF of the p_1-dimensional mean zero normal distribution with all variances equal to one and all correlations equal to ρ, and F_2 as the uniform distribution on $[0,1]$. We computed the values of $\|\mathbb{W}_{F_1,F_2}\|_\alpha$ for selected α (Table 1). We found the values of $\|\mathbb{W}_{F_1,F_2}\|_\alpha$ almost constant for ρ not close to 1. At 0.05 significance, we can simply reject H_0 if $T_{n,p_1,p_2} \geq 0.9507$ for $p_1 = 2$ or $T_{n,p_1,p_2} \geq 1.0566$ for $p_1 = 3$ when ρ is not close to 1. This is enough in most applications.

Based on our findings in Table 1, we recommend using critical values from the (p_1, p_2)-*standard Brownian pillow* on $[0,1]^{p_1} \times [0,1]^{p_2}$ in our method, where the definition is also in our online supplementary material. Let $\|\mathbb{W}_{p_1,p_2}\|_\alpha$ be the upper α-quantile of $\mathbb{W}_{p_1,p_2}$. Then, $\|\mathbb{W}_{1,1}\|_\alpha = \|\mathbb{W}_2\|_\alpha$. Based on Table 1, we have $\|\mathbb{W}_{2,1}\|_{0.1} = 0.8874$, $\|\mathbb{W}_{2,1}\|_{0.05} = 0.9507$, $\|\mathbb{W}_{2,1}\|_{0.01} = 1.0803$, $\|\mathbb{W}_{3,1}\|_{0.1} = 0.9975$, $\|\mathbb{W}_{3,1}\|_{0.05} = 1.0566$, and $\|\mathbb{W}_{3,1}\|_{0.01} = 1.1842$, roughly.

For our testing method, assume $\mathcal{S}$ is bounded such that N is almost surely finite. Let $A_\mathbf{s} = \{\mathbf{s}' \in \mathcal{S} : \mathbf{s}' \preceq \mathbf{s}\}$ and $B_\mathbf{m} = \{\mathbf{m}' \in \mathcal{M} : \mathbf{m}' \preceq \mathbf{m}\}$, $F(\mathbf{x}) = \mathbb{E}[\mathcal{N}(A_\mathbf{s} \times B_\mathbf{m})]/\kappa$, $F_s(\mathbf{s}) = \mathbb{E}[\mathcal{N}(A_\mathbf{s} \times \mathcal{M})]/\kappa$, and $F_m(\mathbf{m}) = \mathbb{E}[\mathcal{N}(\mathcal{S} \times B_\mathbf{m})]/\kappa$, where $\kappa = \mathbb{E}(N)$. Let $f(\mathbf{x}) = \lambda(\mathbf{x})/\kappa$, $f_s(\mathbf{s}) = \int_\mathcal{M} \lambda(\mathbf{x})d\mathbf{m}$, and $f_m(\mathbf{m}) = \int_\mathcal{S} \lambda(\mathbf{x})d\mathbf{s}$. Then, f, f_s, and f_m are probability density functions on $\mathbb{R}^{d+p}$, $\mathbb{R}^d$, and $\mathbb{R}^p$, respectively. If $\mathcal{N}$ satisfies the Kolmogorov consistency condition, then $F(\mathbf{x}) = \int_{A_\mathbf{s}\times B_\mathbf{m}} f(\mathbf{u})d\mathbf{u}$, $F_s(\mathbf{s}) = \int_{A_\mathbf{s}} f_s(\mathbf{u})d\mathbf{u}$, and $F_m(\mathbf{m}) = \int_{B_\mathbf{m}} f_m(\mathbf{u})d\mathbf{u}$ are CDFs on $\mathbb{R}^{d+q}$, $\mathbb{R}^d$, and $\mathbb{R}^q$, respectively. We consider $H_0 : F(\mathbf{x}) = F_s(\mathbf{s})F_m(\mathbf{m})$ for all $\mathbf{s} \in \mathcal{S}$ and $\mathbf{m} \in \mathcal{M}$. Our test statistic is

$$T_n = \sup_{\mathbf{s}\in\mathcal{S},\mathbf{m}\in\mathcal{M}} \sqrt{n}\left|\frac{\mathcal{N}(A_\mathbf{s}\times B_\mathbf{m})}{n} - \frac{\mathcal{N}(A_\mathbf{s}\times\mathcal{M})}{n}\frac{\mathcal{N}(\mathcal{S}\times B_\mathbf{m})}{n}\right| \tag{3.4}$$

if $n \geq 1$, and $T_n = 0$ if $n = 0$. We recommend rejecting H_0 if $T_n > \|\mathbb{W}_{d,q}\|_\alpha$ at the α significance level.

3.4. Asymptotics

We provide the asymptotic properties of T_n under the framework of increasing domain asymptotics (Guan (2006)). Let $\mathcal{M} = \mathbb{R}^q$. Assume $\mathcal{S} \subseteq \mathbb{R}^d$ is bounded. The properties of T_n are investigated under $|\mathcal{S}| \to \infty$ with $|\partial\mathcal{S}| = 0$, where $\partial\mathcal{S}$ is the boundary of $\mathcal{S}$. We assume that $\mathcal{S} = \mathcal{S}_\eta = \{\eta\mathbf{s} : \mathbf{s} \in \mathcal{S}_0\}$, where $\mathcal{S}_0$ is a fixed bounded subset of $\mathbb{R}^d$ with $|\mathcal{S}_0| > 0$ and $|\partial\mathcal{S}_0| = 0$. We take $\eta \to \infty$ and assume that $\mathcal{N}$ is observed from a larger MPP $\mathcal{N}_L$ on $\mathbb{R}^d \times \mathbb{R}^q$ when only points in $\mathcal{S}$ are observed.

A critical condition in the derivation of the asymptotic distribution of T_n is the mixing condition (Rosenblatt (1956)) for $\mathcal{N}_{L,s}$, the points for $\mathcal{N}_L$. Here, we use the mixing coefficient $\alpha_\eta(v)$ for $\mathcal{N}_L$ defined as

$$\begin{aligned}\alpha_\eta(v) = \sup\{&|P_s(C_1\cap C_2) - P_s(C_1)P_s(C_2)| : C_1 \in \sigma(\mathcal{N}_s(A_1)),\\ &C_2 \in \sigma(\mathcal{N}_s(A_2)), \rho(A_1, A_2) \geq v, A_1, A_2 \subseteq [-\eta,\eta]^d\},\end{aligned} \tag{3.5}$$

where P_s is the distribution of $\mathcal{N}_{L,s}$ and ρ is the distance function between sets. The coefficient of strong mixing first introduced by Rosenblatt (1956), and later modified by Ivanoff (1982), is $\alpha(v) = \sup_{\eta\geq 0}\alpha_\eta(v)$. The definition of the mixing condition can be similarly explained by the approach provided by Herrdnorf (1984).

We need regularity conditions.

(A1) There exist positive c_1 and c_2 such that $c_1 < \lambda_s(\mathbf{s}) < c_2$ for all $\mathbf{s} \in \mathbb{R}^d$.

(A2) $\lambda_s(\mathbf{s})$ and $\lambda_{2,s}(\mathbf{s}_1, \mathbf{s}_2)$ are positive, uniformly bounded, and invariant as $\eta \to \infty$.

(A3) Conditioning on points, the marks are independently observed with a conditional density function $f_{m|s}(\mathbf{m}|\mathbf{s})$ on $\mathbb{R}^q$.

(A4) There exists a $\beta > 2$ such that $\int_{\mathbb{R}^d} \alpha^{1-2/\beta}(\|\mathbf{u}\|)d\mathbf{u} < \infty$.

(A5) $\mathcal{N}$ satisfies the Kolmogorov consistency condition.

Theorem 4. *If* (A1)$-$(A5) *hold and* $f_{m|s}(\mathbf{m}|\mathbf{s}) = f_m(\mathbf{m})$, *then* $T_n \xrightarrow{D} \|\mathbb{W}_{F_1,F_2}\|$ *as* $\eta \to \infty$.

4. Simulation

We carried out simulation studies to compare the performance of T_n and $T_{n,\mathbf{s}_0}$. We simulated realizations from MPP on $[0,1]^2 \times \mathbb{R}$, where $\mathcal{S} = [0,1]^2$ and $\mathcal{M} = \mathbb{R}$. We considered Poisson MPPs and Poisson cluster MPPs. We chose these processes here due to their popularity in modeling geological and ecological data. For Poisson MPPs, we first generated Poisson point processes on $[0,1]^2$ with $\lambda(\mathbf{s}) = \kappa\beta(s_x;\gamma)\beta(s_y;\gamma)$, where $\beta(u;\gamma)$ was the density of the $Beta(\gamma,\gamma)$ distribution and κ was the expected number of points. After that, we independently generated marks at each points from $N(\mu(\mathbf{s}_i), 1)$, where the values of $\mu(\mathbf{s}_i)$ were equal to $\delta/2$, $-\delta/2$, δ, and $-\delta$ according to the cases of $\mathbf{s}_i$ displayed in Figure 1, respectively. For Poisson cluster MPPs, we first generated parent points from a Poisson point process with $\lambda(\mathbf{s}) = (\kappa/k)\beta(s_x;\gamma)\beta(s_y;\gamma)$ and then we independently generated offspring points around parent points, where each parent point had a Poisson(k) number of offsprings and the locations of offspring points to their parent points were determined by a bivariate Gaussian distribution with the standard deviation σ. We fixed $k = 4$ and $\sigma = 0.01$. After points were derived, we used the same method in the Poisson MPP for marks. Then, points and marks were independent if and only if $\delta = 0$.

We used T_n and $T_{n,\mathbf{s}_0}$ to test the first-order independence between points and marks, which could be equivalently expressed as $H_0 : \delta = 0$ against $H_1 : \delta \neq 0$. To study the impact of the choice of $\mathbf{s}_0$, we chose $\mathbf{s}_0 = (0,0)$ and $\mathbf{s}_0 = (0,1)$ in the definition of $T_{n,\mathbf{s}_0}$. We used $T_{n,1}$ to represent $T_{n,\mathbf{s}_0}$ when $\mathbf{s}_0 = (0,0)$ and $T_{n,2}$ to represent $T_{n,\mathbf{s}_0}$ when $\mathbf{s}_0 = (0,1)$. According to Table 1, if T_n was used, we rejected H_0 if $T_n \geq 0.9507$. According to Zhang (2014), if $T_{n,\mathbf{s}_0}$ was used, then we rejected H_0 if $T_{n,\mathbf{s}_0} \geq 0.7948$. We evaluated the type I error probabilities and the power functions of T_n, $T_{n,1}$, and $T_{n,2}$ with respect to $\gamma = 1$ and $\gamma = 2$ in our simulation at 0.05 significance.

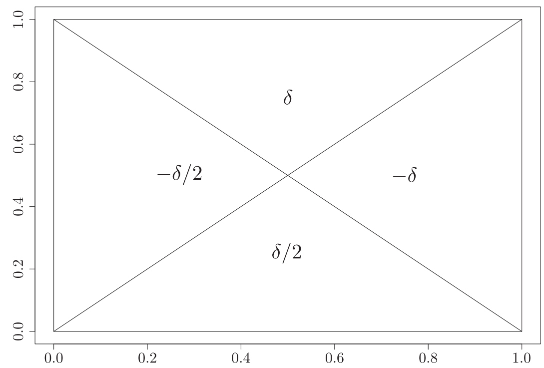

Figure 1. $\mu(\mathbf{s}_i)$ as functions of δ for $\mathbf{s}_i$ within the four specific regions, respectively.

We simulated 1,000 realizations from the Poisson MPP and the Poisson cluster MPP for each selected κ, γ, and δ in our simulation, and the rejection rates of T_n, $T_{n,1}$, $T_{n,2}$ were computed (Table 2). Results showed that the type I error probabilities (when $\delta = 0$) were all close to 0.05 but the power functions (when $\delta > 0$) were significantly different from each other. The power functions of T_n were higher than the power functions of $T_{n,1}$ and $T_{n,2}$ in all cases that we studied. In comparisons between $T_{n,1}$ and $T_{n,2}$, we found that that the rejection rates of $T_{n,1}$ were all close to the significance level but the rejection rates of $T_{n,2}$ increased as δ increased, which meant that the performance of $T_{n,\mathbf{s}_0}$ was significantly affected by the choice of $\mathbf{s}_0$. For a particular choice of $\mathbf{s}_0$, we were not able to conclude that the points and marks were independent if $T_{n,\mathbf{s}_0}$ was not significant. Therefore, different choices of $\mathbf{s}_0$ should be considered. This was expected because the null hypothesis of $T_{n,\mathbf{s}_0}$ was not equivalent to the first-order independence between points and marks. As the null hypothesis and the alternative for T_n were equivalent to the first-order independence and the first-order dependence between points and marks, respectively, it seems more convenient to use T_n in applications. In addition to the cases in Table 2, we studied similar issues for smaller κ values (not included). We concluded that our modified Kolmogorov-Smirnov test could still be used if the count of points was low (e.g. around 100).

5. Application

We expected the concept of independence or separability and the corresponding tests to have wide applications in natural hazard studies. As earthquakes are considered most important natural hazard phenomena, we applied our test for independence to earthquake data. Many sources of earthquake data are available via the internet and can be downloaded for free. Examples include the websites of

Table 2. Simulations (with 1,000 replications) for Type I error probabilities ($\delta = 0$) and power functions ($\delta > 0$) of T_n and $T_{n,\mathbf{s}_0}$ at 0.05 significance level in Poisson MPPs and Poisson cluster MPPs, where $T_{n,1} = T_{n,\mathbf{s}_0}$ for $\mathbf{s}_0 = (0,0)$ and $T_{n,2} = T_{n,\mathbf{s}_0}$ for $\mathbf{s}_0 = (0,1)$.

			δ for $\gamma = 1$				δ for $\gamma = 2$			
Process	κ		0.0	0.1	0.2	0.3	0.0	0.1	0.2	0.3
Poisson	1,000	T_n	0.032	0.082	0.175	0.452	0.038	0.074	0.168	0.422
		$T_{n,1}$	0.036	0.047	0.052	0.050	0.034	0.049	0.051	0.050
		$T_{n,2}$	0.038	0.080	0.178	0.353	0.038	0.069	0.167	0.348
	2,000	T_n	0.045	0.119	0.339	0.862	0.040	0.115	0.379	0.822
		$T_{n,1}$	0.049	0.055	0.053	0.055	0.048	0.051	0.049	0.064
		$T_{n,2}$	0.051	0.117	0.309	0.585	0.055	0.112	0.348	0.636
Cluster	1,000	T_n	0.041	0.201	0.776	0.991	0.047	0.189	0.746	0.981
		$T_{n,1}$	0.041	0.062	0.103	0.198	0.053	0.054	0.091	0.206
		$T_{n,2}$	0.032	0.185	0.541	0.801	0.052	0.196	0.574	0.847
	2,000	T_n	0.045	0.415	0.988	1.000	0.037	0.414	0.983	1.000
		$T_{n,1}$	0.044	0.055	0.102	0.293	0.054	0.055	0.114	0.276
		$T_{n,2}$	0.046	0.318	0.827	0.975	0.048	0.346	0.851	0.982

the United States National Geophysical (USGS) data center, the Northern California Earthquake Data Center (NCEDC), and the GeoCommunity data center.

We used the data from the NCEDC website on global earthquake activities since 1900. We focused on Japanese earthquakes because Japan is considered the highest risk country for earthquakes in the World. We found that most earthquakes occurred in an area between latitude 30 and latitude 45 North, and longitude 130 and 150 East. With this area as the study region, we collected earthquakes with magnitude greater than or equal to 4.0 from January 1, 2000 to December 31, 2014. There were 11,493 earthquakes in the dataset, 150 of them with magnitude between 6 and 7, 12 of them with magnitude between 7 and 8, and 2 of them with magnitude $\geq$ 8.0. A serious earthquake was the *Great Tohoku Earthquake* occurred in March 11, 2011 at 38.30 latitude North and 142.37 longitude East with magnitude 9.1 (Stein, Geller and Liu (2012)).

An important issue was to analyze earthquake clusters caused by aftershocks. A number of statistical methods have been proposed for earthquake clusters. As important framework for this is the epidemic-type aftershock sequences (ETAS) model (Ogata (1998); Zhuang, Ogata and Vere-Jones (2002); Ogata and Zhuang (2006)), which has widely been applied recently (Bansia, Dimri and Babu (2013); Console, Murru and Falcone (2010); Vere-Jones and Zhuang (2010)). Based on a few assumptions, the conditional intensity function of the ETAS model can be expressed as

$$\lambda(\mathbf{s}, t, M|\mathcal{H}_t) = j(M)[\mu(\mathbf{s}) + \sum_{k:t_k<t} \kappa(M_k^*)u(t-t_k)v(\mathbf{s}-\mathbf{s}_k^*|M_k^*)], \tag{5.1}$$

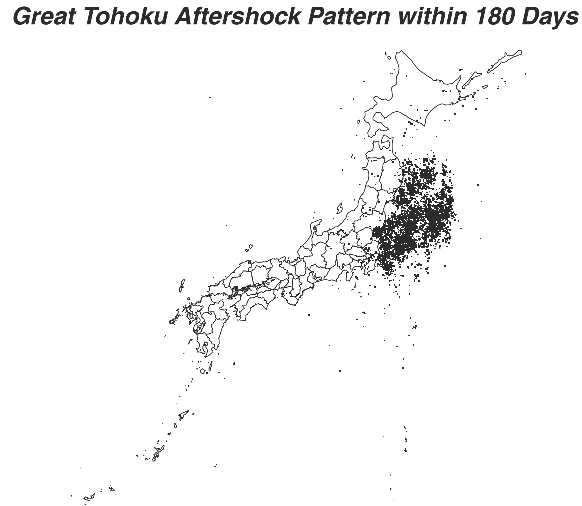

Figure 2. Aftershock occurrences with magnitude $\geq$ 4.0 within 180 days after the Great Tohoku Earthquake (March 11, 2011, magnitude 9.1).

where $j(M)$ is a standardized term, $\mu(\mathbf{s})$ is the background intensity function, $\mathcal{H}_t$ denotes the space-time magnitude occurrence history of earthquakes up to time t, and $\kappa(M_k^*)$ is the expected number of aftershocks from a mainshock ancestor. If an extremely large mainshock earthquake occurs then, within a short time period, the performance of the ETAS model is primarily dominated by its aftershock earthquakes, implying that the impact of the background and other mainshock earthquakes can be ignored. The magnitude and the spatiotemporal location of an extremely large mainshock earthquake is denoted by M^* and $(\mathbf{s}^*, t^*)$, respectively. Then, the conditional intensity function can be approximately expressed as

$$\lambda^*(\mathbf{s}, t, M) \approx j(M)\kappa(M^*)u(t - t^*)v(\mathbf{s} - \mathbf{s}^*|M^*). \tag{5.2}$$

Since the ETAS model assumes that each mainshock earthquake produces aftershock earthquakes independently, aftershock occurrences caused by an extremely large mainshock earthquake can be roughly treated as a Poisson MPP with the first-order intensity function given by $\lambda^*(\mathbf{s}, t, M)$.

We treated locations of occurrence as points and magnitudes as marks. We used T_n and $T_{n,\mathbf{s}_0}$ to test the first-order independence between points and marks, where we chose $\mathbf{s}_0$ at 30 latitude North and 130 longitude East. We tested the first-order independence between points and marks in the whole period as well as the period for the aftershocks of the *Great Tohoku Earthquake*. In the test for the whole fifteen year study period, we had $T_n = 2.961$ and $T_{n,\mathbf{s}_0} = 3.054$ with both p-values almost 0. Therefore, we concluded that the points and marks

Table 3. Test for the first-order independence between locations and magnitudes of aftershock earthquakes within the first 180 days produced by the *Great Tohoku Earthquake*, where each period contains 30 days.

	Number of Aftershocks				
Period	Total	Strong	Major	T_n(p-value)	$T_{n,\mathbf{s}_0}$(p-value)
1	3047	46	3	1.394(< 0.001)	1.203(< 0.001)
2	554	5	0	0.778(0.279)	0.594(0.321)
3	317	2	0	0.717(0.420)	0.596(0.317)
4	203	1	0	0.651(0.603)	0.472(0.646)
5	213	4	0	0.570(0.807)	0.412(0.802)
6	169	2	0	0.681(0.515)	0.599(0.310)
Total	4503	60	3	1.457(< 0.001)	1.183(< 0.001)

were not independent in the whole study period. This was expected because the ETAS model given by (5.1) was not separable in the whole study period. We also tested the first-order independence between points and marks for the aftershock pattern within the first 180 days after the occurrence of the *Great Tohoku Earthquake* (Figure 2 and Table 3). Our results showed that first-order independence was rejected in the whole period. To understand the pattern, we partitioned the 180 days into six periods. Each contained 30 days. The results showed that first-order independence was rejected only in the first period; the ETAS model (5.2) was roughly separable in this period. We conclude that strong dependence between points and marks present just after the occurrence of mainshock earthquake. Because the dependence between points and marks disappeared quickly, we conclude that the ETAS model is still appropriate for modeling the patterns of earthquake clusters and aftershock activities.

In the end, we evaluated the Kolmogorov consistency condition in the *Great Tohoku Earthquake* data and concluded it was not violated because if it were violated the earthquake pattern would be significantly affected by a few events. To evaluate whether the Kolmogorov consistency condition was a concern, we looked at whether larger earthquake patterns, magnitude ≥ 5.0 or ≥ 6.0, could be affected by smaller earthquake events if they were removed from the data set. It was clear that the patterns were not affected if smaller earthquake events were removed.

6. Discussion

The justification of Kolmogorov consistency is important in practice. If it is satisfied, then the distribution of a portion of the events is not affected by the occurrences of other events. Since the distribution of large events is generally more important than the distribution of small events, it is then not necessary to consider the information of small events if one attempts to study the distribution of large events. However, if Kolmogorov consistency condition is violated, then one must consider the information of small events, which implies that the analysis becomes more complicated. According to the nature of such applications as earthquakes or forest wildfires, we do not think occurrences of a few events can significantly change the entire pattern.

We expect that the connection between the concepts of independence and separability to clarify research on natural hazards, infectious diseases, forestry, and social sciences where MPP data are common. Many natural hazard phenomena can be described by SMPPs or STMPPs. It is important to know whether points and marks are independent in such data. If Kolmogorov consistency condition is satisfied, methods based on distribution functions and methods based on intensity functions are equivalent. Otherwise, one should carefully address the difference between the methods.

Methods based on the count measure are more popular than methods based on the Janossy measure, the interaction effect between points and marks is mostly focused on the intensity dependent model (Ho and Stoyan (2008); Malinowski, Schlather and Zhang (2014); Myllymäki and Penttinen (2009)). If Kolmogorov consistency condition is satisfied, then it does not matter how one attempts to account for the interaction and one should look to unify the two approaches to modeling.

Supplementary Materials

The online supplememntary material includes the definitions of the Brownian sheet, pinned Brownian sheet, and Brownian pillow, the defintions of (F_1, F_2)-functional Brownian pillow and (p_1, p_2)-standard Brownian pillow. It contains the proofs of Theorem 1, Theorem 2, Theorem 3, Corollary 1, and Theorem 4, as well as the associated lemmas.

Acknowledgement

The author appreciates the suggestive comments from an associate editor and two anonymous referees, that significantly improves the quality of the article.

References

Assuncao, R. and Maia, A. (2007). A note on testing separability in spatial-temporal marked point processes. *Biometrics* **63**, 290-294.

Baddeley, A. and Turner, R. (2005). **spatstat**: an R package for analyzing spatial point patterns. *Journal of Statistical Software* **12**, 1-42.

Bansal, A. R., Dimri, V. P. and Babu, K. K. (2013). Epidemic type aftershock sequence (ETAS) modeling of northeastern Himalayan seismicity. *Journal of Seismology* **17**, 255-264.

Billingsley, P. (1995). *Probability and Measure*. Wiley, New York.

Blum, J. R., Kiefer, J. and Rosenblatt, M. (1961). Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist.* **32**, 485-496.

Brown, S. J., Caesar, J. and Ferro, C. A. T. (1950). Global changes in extreme daily temperature since 1950. *J. Geophys. Res.* **113**, D05115.

Console, R., Murru, M. and Falcone, G. (2010). Probability gains of an epidemic-type aftershock sequence model in retrospective foresting of $M \geq 5$ earthquakes in Italy. *J. Seism.* **14**, 9-26.

Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York.

Cressie, N. and Huang, H. C. (1999). Classes of nonseparable, spatio-temporal stationary covariance functions. *J. Amer. Statist. Assoc.* **94**, 1330-1340.

Daley, D.J. and Vere-Jones, D. (2003). *An Introduction to Theory of Point Processes: Volume I: Elementary Theory and Methods, 2nd Edition*. Springer, New York.

Diggle, P. J., Ribeiro, J. P. J. and Christensen, O. (2003). An introduction to model-based geostatistics. In *Spatial Statistics and Computational Methods* (Edited by J. Moller), 43-86. Springer, New York.

Guan, Y. (2006). A composite likelihood approach in fitting spatial point process models. *J. Amer. Statist. Assoc.* **101**, 1502-1512.

Guan, Y. and Afshartous, D. R. (2007). Test for independence between marks and points of marked point processes: a subsampling approach. *Environ. Ecolog. Statist.* **14**, 101-111.

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

Hall, P. and Tajvidi, N. (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statist. Sci.* **15**, 153-167.

Harte, D. (2010). **PtProcess**: an R Package for modelling marked point processes indexed by time. *J. Statist. Software* **35**, 1-32.

Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequqnces of random variables. *Ann. Probab.* **12**, 141-153.

Ho, L.P. and Stoyan D. (2008). Modeling marked point patterns by intensity-marked Cox processes. *Statist. Probab. Lett.* **78**, 2831-2842.

Holden, L., Sannan, S. and Bungum, H. (2003). A stochastic marked point process model for earthquakes. *Natural Hazards and Earth System Sci.* **3**, 95-101.

Ivanoff, G. (1982). Central limit theorems for point processes. *Stochastic Process. Appl.* **12**, 171-186.

Janossy, L. (1950). On the absorption of a nucleon cascade. *Proc. Roy. Irish Academy A* **53**, 181-888.

Karr, A. F. (1991). *Point Processes and Their Statistical Inference*. 2nd edition. Marcel Dekker, Inc. New York.

Keifer, J. (1958). On the derivations of the empirical distribution function of vector chance variables. *Trans. Amer. Math. Soc.* **87**, 173-186.

Koning, A. J. and Protasov, V. (2003). Tail behavior of Gaussian processes with applications to the Brownian pillow. *J. Multivariate Anal.* **87**, 370-397.

Kruglova, N. V. (2008). Distribution of the maximum of the Chentsov random field. *Theory of Stochastic Processes* **14**, 76-81.

Kruglova, N.V (2010). Asymptotic behavior of the distribution of the maximum of a Chentsov field on polygonal lines. *Theory Probab. Math. Statist.* **81**, 101-115.

Malamud, B. D., Millington, J. D. A. and Perry, G. L. W. (2005). Characterizing wildfire regimes in the United States. *Proc. Natl. Acad. Sci. USA* **102**, 4694-4699.

Malinowski, A., Schlather, M. and Zhang, Z. (2014). Intrinsically weighted means of marked point processes. *Ann. Inst. Statist. Math.*, DOI 10.1007/s10463-014-0485-6.

Moller, J. and Waagepetersen, R. (2007). Modern statistics for spatial point processes. *Scand. J. Statist.* **34**, 643-684.

Moyal, J. E. (1962). The general theory of stochastic population processes. *Acta Math.* **108**, 1-31.

Myllymäki, M. and Penttinen, A. (2009). Conditionally heteroscedastic intensity-dependent marking of log Gaussian Cox processes. *Statist. Neerlandica* **63**, 450-473.

Neyman, J. and Scott, E. L. (1958). Statistical approach to problems of cosmology. *J. Roy. Statist. Soc. Ser. B* **20**, 1-43.

Ogata, Y. (1998). Space-time point process models for earthquake occurrences. *Ann. Inst. Statist. Math.* **50**, 379-402.

Ogata, Y. and Zhuang, J. (2006). Space-time ETAS models and an improved extension. *Tectonophysics* **413**, 13-23.

Peng, R. D., Schoenberg, F. P. and Woods, J. A. (2005). A space-time conditional intensity model for evaluating a wildfire hazard index. *J. Amer. Statist. Assoc.* **100**, 26-35.

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* **42**, 43-47.

Schlather, M., Ribeiro, P. J. and Diggle, P. J. (2004). Detecting dependence between marks and locations of marked point processes. *J. Roy. Statist. Soc. Ser. B* **66**, 79-93.

Schoenberg, F. P. (2004). Testing separability in spatial-temporal marked point processes. *Biometrics* **60**, 471-481.

Stein, S., Geller, R. J. and Liu, M. (2012). Why earthquake hazard maps often fail and what to do about it. *Tectonophysics* **562-563**, 1-25.

Vere-Jones, D. and Zhuang, J. (2010). Distribution of the largest event in the critical epidemic-type aftershock-sequence model. *Phys. Rev. E* **78**, 047102.

Zhang, T. and Zhuang, Q. (2014). On the local odds ratio between points and marks in marked point processes. *Spatial Statist.* **9**, 20-37.

Zhang, T. (2014). A Kolmogorov-Smirnov type test for independence between marks and points of marked point processes. *Electronic J. Statist.* **8**, 2557-2584.

Zhuang, J., Ogata, Y. and Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. *J. Amer. Statist. Assoc.* **97**, 369-380.

Department of Statistics, Purdue University, 250 North University Street, 172 West Lafayette, IN 47907-2066, USA.

E-mail: tlzhang@purdue.edu

(Received July 2015; accepted December 2015)