Heriot-Watt University

Heriot-Watt University
Research Gateway

Faà di Bruno's formula and spatial cluster modelling

Clark, Daniel E; Houssineau, Jeremie

Published in:
Spatial Statistics

DOI:
10.1016/j.spasta.2013.09.002

Publication date:
2013

Spatial Statistics 6 (2013) 109–117

Contents lists available at ScienceDirect

Spatial Statistics

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

Faà di Bruno's formula and spatial cluster modelling☆

Daniel E. Clark *,[1], Jeremie Houssineau [2]

School of Engineering and Physical Sciences, Heriot–Watt University, United Kingdom

ARTICLE INFO

Article history:
Received 14 May 2013
Accepted 30 September 2013
Available online 16 October 2013

Keywords:
Faà di Bruno's formula
Spatial cluster modelling
Point processes
Variational calculus
Chain differentials

ABSTRACT

The probability generating functional (p.g.fl.) provides a useful means of compactly representing point process models. Cluster processes can be described through the composition of p.g.fl.s, and factorial moment measures and Janossy measures can be recovered from the p.g.fl. using variational derivatives. This article describes the application of a recent result in variational calculus, a generalisation of Faà di Bruno's formula, to determine such results for cluster processes.

1. Introduction

Neyman and Scott proposed the stochastic process of clustering as a mathematical model of galaxies (Neyman and Scott, 1958). Since then spatial clustering models have been investigated in various applications. A collection of articles on spatial cluster modelling reviewed current approaches for statistical inference of spatial and spatial cluster processes and their applications (Lawson and Denison, 2002).

The probability generating functional (Moyal, 1962; Bogolyubov, 1946), p.g.fl., provides a means of uniquely characterising a point process. In a similar way to the probability generating function, the probability measures and factorial moment measures of point processes can be found from the p.g.fl. by differentiation, with Gâteaux differentials (Gâteaux, 1919). The probability generating

* Corresponding author. Tel.: +44 1314513323.
E-mail addresses: declark@gmail.com, d.e.clark@hw.ac.uk (D.E. Clark), jh207@hw.ac.uk (J. Houssineau).

[1] Daniel Clark is a Royal Academy of Engineering/EPSRC Research Fellow.

[2] Jeremie Houssineau has a Ph.D. scholarship sponsored by DCNS and a tuition fees scholarship by Heriot–Watt University.

2211-6753/$ – see front matter
http://dx.doi.org/10.1016/j.spasta.2013.09.002

functional is well known within the point process literature (Daley and Vere-Jones, 2003, p. 15; Cressie, 1991, p. 627; Moller and Waagepetersen, 2003, p. 9; Cox and Isham, 1980, p. 38), yet the approach of functional differentiation is rarely used since it results in a combinatorial number of terms. To highlight this point, we note that cluster processes, such as the model by Neyman and Scott (1958), can be specified simply through the composition of p.g.fl.s, e.g. Cox and Isham (1980, p. 76), Daley and Vere-Jones (2003, p. 178) and Moyal (1962), yet the Janossy measure for arbitrary point processes does not appear to have been determined using this method. Specific parameterisations with Neyman–Scott processes have been studied, though researchers often prefer to work directly with the measures, e.g. Ripley (1988, p. 5), van Lieshout and Baddeley (2002), van Lieshout (2000, p. 140) and Illian et al. (2008, p. 368), rather than with the p.g.fl. representation.

In the aerospace and signal processing literature, the approach has become popular for deriving algorithms for tracking multiple targets from radar, following Mahler's method for Bayesian estimation with point processes (Mahler, 2003, 2007). Practical applications of these techniques were made possible through the development of sequential Monte Carlo (Vo et al., 2005) and Gaussian mixture implementations (Vo and Ma, 2006; Vo et al., 2007). To determine Bayes' theorem for point processes, Mahler (2003) proposed the use of functional derivatives of the probability generating functional. This approach often involves finding the parameterised form of the updated process, and proving its correctness by induction. The process can be quite involved and needs to be applied for each model (Mahler, 2003, 2007, 2009a,b; Swain and Clark, 2010). The construction of the models often involves composition of basic models, whose derivatives are easy to find, yet when composed, their form for higher derivatives becomes more unclear. In this paper we circumvent this problem by using a recently derived tool from variational calculus, Faà di Bruno's formula for variational calculus (Clark and Houssineau, 2013). We use this approach to determine the Janossy measures and factorial moment measures of cluster processes and illustrate the approach through an example with Matérn cluster processes.

2. Variational calculus and the higher-order chain rule

This section describes differentials and the general form of Faà di Bruno's formula required to determine the results in the following section. We adopt a restricted form of Gâteaux differential, known as the chain differential (Bernhard, 2005), in order that a general chain rule can be determined (Clark and Houssineau, 2013). Following this, we describe the general higher-order chain rule.

Definition 2.1 (*Chain Differential, from Bernhard, 2005*)**.** The function $f : X \to Y$, where X and Y are normed spaces, has a *chain differential* $\delta f(x; \eta)$ at $x \in X$ in the direction $\eta \in X$ if, for any sequence $\eta_n \to \eta$ in X, and any sequence of real numbers $\theta_n \to 0$, it holds that

$$\delta f(x; \eta) = \lim_{n\to\infty} \frac{1}{\theta_n} \left(f(x + \theta_n \eta_n) - f(x)\right). \tag{1}$$

The nth-order chain differential can be defined recursively as

$$\delta^n f\,(x; \eta_1, \ldots, \eta_n) = \delta\left(\delta^{n-1} f\,(x; \eta_1, \ldots, \eta_{n-1})\,; \eta_n\right). \tag{2}$$

Applying nth-order chain differentials on composite functions can be an extremely laborious process since it involves determining the result for each choice of function and proving the result by induction. For ordinary derivatives, the general higher-order chain rule is normally attributed to Faà di Bruno (1855). The following result generalises Faà di Bruno's formula to chain differentials.

Theorem 2.1 (*General Higher-Order Chain Rule, from Clark and Houssineau, 2013*)**.** *Let T, U and V be normed spaces. Assume that $g : T \to U$ has higher order chain differentials in any number of directions in the set $\{\eta_1, \ldots, \eta_n\}$, with $\eta_1, \ldots, \eta_n \in T$ and that $f : U \to V$ has higher order chain differentials in any number of directions in the set $\{\delta^m g(x; S_m)\}_{m=1:n}$, $S_m \subseteq \{\eta_1, \ldots, \eta_n\}$. Assuming additionally that for all $1 \le m \le n$, $\delta^m f(y; \xi_1, \ldots, \xi_m)$ is continuous on an open set $\Omega \subseteq Y^{m+1}$ and linear with respect to the* *directions $\xi_1, \ldots, \xi_m$, the nth-order variation of composition $f \circ g$ in directions $\eta_1, \ldots, \eta_n$ at point $x \in T$ is given by*

$$\delta^n (f \circ g)(x; \eta_1, \ldots, \eta_n) = \sum_{\pi \in \Pi(\eta_1, \ldots, \eta_n)} \delta^{|\pi|} f \left(g(x); \delta^{|\omega|} g\,(x; \xi : \xi \in \omega) : \omega \in \pi \right),$$

where $\Pi(A)$ represents the set of partitions of the discrete set A, $|\pi|$ denotes the cardinality of the set π and where $h(y : y \in \pi) = h(y_1, \ldots, y_m)$ when $\pi = \{y_1, \ldots, y_m\}$.

3. Point processes

In this section we describe point processes via probability generating functionals, and apply the general form of Faà di Bruno's formula for cluster process probability generating functionals to determine Janossy measures and factorial moment measures.

3.1. Univariate point processes

Let $\mathbf{X}$ be the complete separable metric space in which points of the process are located, e.g. $\mathbb{R}^d$. Let Φ be a point process defined as the measurable mapping

$$\Phi : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbf{X}^{\cup}, \mathcal{B}(\mathbf{X}^{\cup})), \tag{3}$$

from the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to the measurable space $(\mathbf{X}^{\cup}, \mathcal{B}(\mathbf{X}^{\cup}))$ where

$$\mathbf{X}^{\cup} = \mathbf{X}^0 \cup \mathbf{X}^1 \cup \mathbf{X}^2 \ldots, \tag{4}$$

with $\mathbf{X}^0$ corresponding to the empty configuration, and where $\mathcal{B}(\mathbf{X}^{\cup})$ is the Borel σ-algebra over $\mathbf{X}^{\cup}$.

For any $n \geq 1$, the symmetric probability measure $P_{\Phi}^{(n)}$, defined on $\mathcal{B}(\mathbf{X}^n)$, describes the probability for the point process Φ to be composed of n points, and the distribution of these points. By extension, $P_{\Phi}^{(0)}$ is the probability for the point process Φ to be empty. For any $n \geq 0$, $J_{\Phi}^{(n)}$ denotes the nth-order Janossy measure (Daley and Vere-Jones, 2003, p. 124), and is defined as

$$J_{\Phi}^{(n)}(B_1 \times \cdots \times B_n) = \sum_{\sigma \in S_n} P_{\Phi}^{(n)}(B_{\sigma_1} \times \cdots \times B_{\sigma_n}) \tag{5}$$

$$= n! P_{\Phi}^{(n)}(B_1 \times \cdots \times B_n), \tag{6}$$

where S_n is the permutation group on n letters. Let $\mathcal{U}(\mathbf{X})$ be the space of real-valued bounded measurable functions on the complete separable metric space $\mathbf{X}$, equipped with the supremum norm

$$\|u\| = \sup_{x \in \mathbf{X}} |u(x)|.$$

The normed space $\mathcal{U}(\mathbf{X})$ is a Banach space.

The probability generating functional (p.g.fl.) is a useful tool in point process theory (Moyal, 1962) and is defined as follows.

Definition 3.1 (*Probability Generating Functional*)**.** Let $\mathcal{V}(\mathbf{X})$ be the space of functions $v \in \mathcal{U}(\mathbf{X})$ such that $1 - v$ is vanishing outside some bounded set and is satisfying $0 \leq v(x) \leq 1$, for any $x \in \mathbf{X}$. The probability generating functional G_{Φ} of a point process Φ can be written for $v \in \mathcal{V}(\mathbf{X})$ as

$$G_{\Phi}(v) = J_{\Phi}^{(0)} + \sum_{n \geq 1} \frac{1}{n!} \int v(x_1) \ldots v(x_n) J_{\Phi}^{(n)}\big(\mathrm{d}(x_1, \ldots, x_n)\big),$$

where $J_{\Phi}^{(0)}$ is the probability that there is no point within the support of $1 - v$.

The space $\mathcal{V}(\mathbf{X})$ is a Banach space since it is closed and it is a subset of the Banach space $\mathcal{U}(\mathbf{X})$. Banach spaces are particular normed spaces so that the chain differential can be applied. Taking the

kth-order variation of G_Φ in the directions $\xi_1, \ldots, \xi_k \in \mathcal{V}(\mathbf{X})$ and at point $v \in \mathcal{V}(\mathbf{X})$, we have (see, for example Srinivasan, 1973, p. 21),

$$\delta^k G_\Phi(v; \xi_1, \ldots, \xi_k) = \sum_{n \geq k} \frac{1}{(n-k)!} \int \xi_1(x_1) \ldots \xi_k(x_k) v(x_{k+1}) \ldots v(x_n) J_\Phi^{(n)}\big(\mathrm{d}(x_1, \ldots, x_n)\big). \quad (7)$$

It is useful to consider the cases when we set $v = 1$ or $v = 0$, where "$v = 0$" means "$v(x) = 0$ for any point x within the support of $1 - v$". The differential (7) at points $v = 1$ or $v = 0$ can be expressed as

$$\delta^k G_\Phi(0; \xi_1, \ldots, \xi_k) = \int \xi_1(x_1) \ldots \xi_k(x_k) J_\Phi^{(k)}\big(\mathrm{d}(x_1, \ldots, x_k)\big), \quad (8)$$

$$\delta^k G_\Phi(1; \xi_1, \ldots, \xi_k) = \int \xi_1(x_1) \ldots \xi_k(x_k) M_\Phi^{(k)}\big(\mathrm{d}(x_1, \ldots, x_k)\big), \quad (9)$$

where $M_\Phi^{(k)}$ is the kth-order factorial moment measure, defined as (Stoyan et al., 1995, p. 111)

$$\int f(x_1, \ldots, x_k) M_\Phi^{(k)}\big(\mathrm{d}(x_1, \ldots, x_k)\big)$$
$$= \sum_{n \geq k} \frac{1}{n!} \int \sum_{\{y_1, \ldots, y_k\} \subseteq \{x_1, \ldots, x_n\}}^{\neq} f(y_1, \ldots, y_k) J_\Phi^{(n)}\big(\mathrm{d}(x_1, \ldots, x_n)\big), \quad (10)$$

where f is a non-negative measurable function on $\mathbf{X}^k$ and where $\sum^{\neq}$ is the sum over k-tuples of distinct points. For example, the Janossy and factorial moment measures are recovered by setting the directions to be indicator functions, so that $\xi_i = \mathbf{1}_{B_i}$, $B_i \in \mathcal{B}(\mathbf{X})$, $1 \leq i \leq k$, and hence

$$\delta^k G_\Phi(0; \mathbf{1}_{B_1}, \ldots, \mathbf{1}_{B_k}) = J_\Phi^{(k)}(B_1 \times \cdots \times B_k) \quad (11)$$

$$\delta^k G_\Phi(1; \mathbf{1}_{B_1}, \ldots, \mathbf{1}_{B_k}) = M_\Phi^{(k)}(B_1 \times \cdots \times B_k). \quad (12)$$

3.2. Bivariate point processes

Bivariate point processes are useful for modelling the dependencies between processes, and processes that can be modelled through conditioning. In cluster modelling, we have a parent process describing the distribution of cluster centres, and a daughter process that describes the distribution of points conditioned on a parent point. Let $\mathbf{X}$ and $\mathbf{Y}$ be two complete separable metric spaces and let Φ_1 and Φ_2 be two point processes such that

$$\Phi_1 : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbf{X}^\cup, \mathcal{B}(\mathbf{X}^\cup)) \quad (13)$$

$$\Phi_2 : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbf{Y}^\cup, \mathcal{B}(\mathbf{Y}^\cup)). \quad (14)$$

A bivariate point process $\Phi_c = (\Phi_1, \Phi_2)$ can be defined as

$$\Phi_c : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbf{X}^\cup \times \mathbf{Y}^\cup, \mathcal{B}(\mathbf{X}^\cup) \otimes \mathcal{B}(\mathbf{Y}^\cup)), \quad (15)$$

where "$\times$" and "$\otimes$" are respectively the Cartesian product and the σ-algebra product.

Definition 3.2 (*Joint Probability Generating Functional*)**.** The joint probability generating functional G_{Φ_c} of the bivariate process Φ_c with Janossy measures $J_{\Phi_c}^{(n,m)}$ on $\mathcal{B}(\mathbf{X}^n) \otimes \mathcal{B}(\mathbf{Y}^m)$ with $n, m \in \mathbb{N}$, can be written for $v \in \mathcal{V}(\mathbf{X})$ and $w \in \mathcal{V}(\mathbf{Y})$ as

$$G_{\Phi_c}(v, w) = \sum_{n,m \geq 0} \frac{1}{n!m!} \int \left(\prod_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} v(x_i) w(y_j) \right) J_{\Phi_c}^{(n,m)}\big(\mathrm{d}(x_1, \ldots, x_n) \times \mathrm{d}(y_1, \ldots, y_m)\big), \quad (16)$$

where $J_{\Phi_c}^{(0,0)}$ is the probability that there is no point within the support of the functions $1 - v$ and $1 - w$.

We can consider similar results to the univariate point processes, e.g.

$$\delta^{k+l} G_{\Phi_c}\big(0, 0; (\xi_1, \dots, \xi_k), (\eta_1, \dots, \eta_l)\big)$$
$$= \int \xi_1(x_1) \dots \xi_k(x_k) \eta_1(y_1) \dots \eta_l(y_l) J^{(k,l)}_{\Phi_c}\big(\mathrm{d}(x_1, \dots, x_k), \mathrm{d}(y_1, \dots, y_l)\big). \tag{17}$$

In particular, we can use the bivariate p.g.fl. to determine a conditional p.g.fl. from which we can recover Janossy measures and factorial moment measures which is useful in Bayesian estimation (Mahler, 2003). Assuming that the conditional Janossy measure $J_{\Phi_2}(\cdot|X)$, for any $X \in \mathbf{X}$, is absolutely continuous with respect to the reference measure in $(\mathbf{Y}^{\cup}, \mathcal{B}(\mathbf{Y}^{\cup}))$, and denoting $\hat{J}_{\Phi_2}(\cdot|X)$ the associated conditional probability density, we can write

$$G_{\Phi_1|\Phi_2}(v|z_1, \dots, z_m) = \frac{\delta^m G_{\Phi_c}(v, 0; \emptyset, (\delta_{z_1}, \dots, \delta_{z_m}))}{\delta^m G_{\Phi_c}(1, 0; \emptyset, (\delta_{z_1}, \dots, \delta_{z_m}))},$$

where δ_z is the Dirac delta function at point z such that $\delta_z(x) \neq 0$ iif $z = x$ and $\int \delta_z(x)\mathrm{d}x = 1$, and where the differentiation in the direction $(\emptyset, (\mathbf{1}_{B_1}, \dots, \mathbf{1}_{B_m}))$ means that G_{Φ_c} is differentiated 0 times with respect to its first variable and m times with respect to its second. Similarly, the conditional Janossy and factorial moment measures are then computed with

$$J^{(n)}_{\Phi_1|\Phi_2}(B_1 \times \dots \times B_n|z_1, \dots, z_m)$$
$$= \frac{\delta^{n+m} G_{\Phi_c}\big(0, 0; (\mathbf{1}_{Bf_1}, \dots, \mathbf{1}_{B_n}), (\delta_{z_1}, \dots, \delta_{z_m})\big)}{\delta^m G_{\Phi_c}\big(1, 0; \emptyset, (\delta_{z_1}, \dots, \delta_{z_m})\big)}$$
$$= \frac{\int_{B_1 \times \dots \times B_n} \hat{J}^{(m)}_{\Phi_2}(z_1, \dots, z_m|x_1, \dots, x_n) J^{(n)}_{\Phi_1}(\mathrm{d}(x_1, \dots, x_n))}{\sum_{k \geq 0} \frac{1}{k!} \int \hat{J}^{(m)}_{\Phi_2}(z_1, \dots, z_m|x'_1, \dots, x'_k) J^{(k)}_{\Phi_1}(\mathrm{d}(x'_1, \dots, x'_k))},$$

and

$$M^{(n)}_{\Phi_1|\Phi_2}(B_1 \times \dots \times B_n|z_1, \dots, z_m)$$
$$= \frac{\delta^{n+m} G_{\Phi_c}\big(1, 0; (\mathbf{1}_{B_1}, \dots, \mathbf{1}_{B_n}), (\delta_{z_1}, \dots, \delta_{z_m})\big)}{\delta^m G_{\Phi_c}\big(1, 0; \emptyset, (\delta_{z_1}, \dots, \delta_{z_m})\big)}$$
$$= \frac{\int_{B_1 \times \dots \times B_n} \hat{J}^{(m)}_{\Phi_2}(z_1, \dots, z_m|x_1, \dots, x_n) M^{(n)}_{\Phi_1}(\mathrm{d}(x_1, \dots, x_n))}{\sum_{k \geq 0} \frac{1}{k!} \int \hat{J}^{(m)}_{\Phi_2}(z_1, \dots, z_m|x'_1, \dots, x'_k) J^{(k)}_{\Phi_1}(\mathrm{d}(x'_1, \dots, x'_k))}.$$

We shall use these results in the next section to determine results for Bayesian estimation for spatial cluster models using Faà di Bruno's formula on composite functionals.

4. Cluster processes

In the following, we adopt an approach similar to the approach of van Lieshout and Baddeley (2002), which considered the use of "germ-grain" models (Stoyan et al., 1995, Chapter 6, p. 186) for the modelling of spatial cluster processes.

Definition 4.1 (*Cluster Processes*)**.** Let $\mathbf{X}$ and $\mathbf{Y}$ be two complete separable metric spaces, and let Φ_c be a joint point process in $(\mathbf{X}^{\cup} \times \mathbf{Y}^{\cup}, \mathcal{B}(\mathbf{X}^{\cup}) \otimes \mathcal{B}(\mathbf{Y}^{\cup}))$. If Φ_c can be decomposed into a parent point process Φ_p in $(\mathbf{X}^{\cup}, \mathcal{B}(\mathbf{X}^{\cup}))$ and a conditional point process $\Phi_d(X)$, $X \in \mathbf{X}^{\cup}$, in $(\mathbf{Y}^{\cup}, \mathcal{B}(\mathbf{Y}^{\cup}))$, and if Φ_d is such that

$$\Phi_d(X) = \bigcup_{x \in X} \Phi'_d(x), \tag{18}$$

with $\Phi'_d(x)$, $x \in \mathbf{X}$, a conditional point process on $(\mathbf{Y}^{\cup}, \mathcal{B}(\mathbf{Y}^{\cup}))$ such that $\Phi'_d(x)$ is independent of $\Phi'_d(x')$ whenever $x \neq x'$, then Φ_c is called a cluster process.

The p.g.fl. of a cluster process can be found easily, as demonstrated in the following theorem. The proof of similar results can be found in Moyal (1962) and Daley and Vere-Jones (2003).

Theorem 4.1 (*P.g.fl. of a Cluster Process*)**.** *The joint p.g.fl.* G_{Φ_c} *of the cluster process* Φ_c *representing independent clusters is given by*

$$G_{\Phi_c}(v, w) = G_{\Phi_p}\left(vG_{\Phi'_d}(w|\cdot)\right). \tag{19}$$

For the sake of compactness, the following shorthand notation is used: let μ and ν be two measures on the same measurable space $(E, \mathcal{E})$, the equation "$\mu(\mathrm{d}x) = \nu(\mathrm{d}x)$" is equivalent to "for any measurable function f on $(E, \mathcal{E})$, $\int f(x)\mu(\mathrm{d}x) = \int f(x)\nu(\mathrm{d}x)$".

The following theorem gives the higher-order Janossy and factorial moment measures for cluster processes.[3]

Theorem 4.2. *The higher-order Janossy measure* $J_{\Phi_c}^{(n,m)}$ *and the higher-order factorial moment measure* $M_{\Phi_c}^{(n,m)}$ *of the cluster process* Φ_c *are given by*

$$\begin{aligned} &J_{\Phi_c}^{(n,m)}\big(\mathrm{d}(x_1, \dots, x_n) \times \mathrm{d}(y_1, \dots, y_m)\big) \\ &\quad = \sum_{\pi \in \Pi_n(y_1,\dots,y_m)} \left(\prod_{i=1}^{n} J_{\Phi'_d}^{(|\pi_i|)}(\mathrm{d}\pi_i|x_i)\right) J_{\Phi_p}^{(n)}(\mathrm{d}(x_1, \dots, x_n)), \end{aligned} \tag{20}$$

and

$$\begin{aligned} &M_{\Phi_c}^{(n,m)}\big(\mathrm{d}(x_1, \dots, x_n) \times \mathrm{d}(y_1, \dots, y_m)\big) \\ &\quad = \sum_{\pi \in \Pi_n(y_1,\dots,y_m)} \left(\prod_{i=1}^{n} M_{\Phi'_d}^{(|\pi_i|)}(\mathrm{d}\pi_i|x_i)\right) M_{\Phi_p}^{(n)}(\mathrm{d}(x_1, \dots, x_n)), \end{aligned} \tag{21}$$

where $\Pi_n(z_1, \dots, z_m)$ *is the set of partitions of size n of the set* $\{z_1, \dots, z_m\}$*.*

Additionally, the following result is of interest:

$$\begin{aligned} &J_{\Phi_d}^{(m)}(\mathrm{d}(y_1, \dots, y_m)|x_1, \dots, x_n) M_{\Phi_p}^{(n)}(\mathrm{d}(x_1, \dots, x_n)) \\ &\quad = \sum_{\pi \in \Pi_n(y_1,\dots,y_m)} \left(\prod_{i=1}^{n} J_{\Phi'_d}^{(|\pi_i|)}(\mathrm{d}\pi_i|x_i)\right) M_{\Phi_p}^{(n)}(\mathrm{d}(x_1, \dots, x_n)). \end{aligned} \tag{22}$$

Proof. As in (11), the higher-order Janossy measure $J_{\Phi_c}^{(n,m)}$ can be recovered from the p.g.fl. G_{Φ_c} through the following functional differentiation

$$\begin{aligned} &J_{\Phi_c}^{(n,m)}\big((B_1 \times \cdots \times B_n) \times (A_1 \times \cdots \times A_m)\big) \\ &\quad = \delta^{n+m} G_{\Phi_c}\big(0, 0; (\mathbf{1}_{B_1}, \dots, \mathbf{1}_{B_n}), (\mathbf{1}_{A_1}, \dots, \mathbf{1}_{A_m})\big), \end{aligned} \tag{23}$$

where $A_i \in \mathcal{B}(\mathbf{Y})$, $1 \le i \le m$, and $B_i \in \mathcal{B}(\mathbf{X})$, $1 \le i \le n$.

Let $\hat{G}$ be the bivariate conditional functional defined as

$$\hat{G}(v, w|x) = v(x)G_{\Phi'_d}(w|x). \tag{24}$$

[3] A different approach for determining posterior measures of a Poisson cluster process was taken in van Lieshout and Baddeley (2002, p. 71). However, obtaining compact expressions for the factorial moments may be complicated without the general chain rule to determine the general structure.

Using the general higher-order chain rule, as described in Theorem 2.1, the r.h.s. of (23) is found to be

$$\delta^{n+m}G_{\Phi_c}\left(0,0;(\mathbf{1}_{B_1},\ldots,\mathbf{1}_{B_n}),(\mathbf{1}_{A_1},\ldots,\mathbf{1}_{A_m})\right)$$
$$=\sum_{\pi\in\Pi((B_1,\ldots,B_n)\cup(A_1,\ldots,A_m))}\delta^{|\pi|}G_{\Phi_p}\left(\hat{G}(0,0|\cdot);\delta^{|\omega|}\hat{G}\left(0,0|\cdot;\mathbf{1}_\xi:\xi\in\omega\right):\omega\in\pi\right). \tag{25}$$

Because of the linearity of $\hat{G}$ with respect to its first argument, and considering that $\hat{G}(0,w|x)=0$ for any function $w\in\mathcal{U}(\mathbf{Y})$ and point $x\in\mathbf{X}$, the differentiation carried in the r.h.s. of (25) gives 0 except if $|\pi|=n$, therefore, using (8),

$$\delta^{n+m}G_{\Phi_c}\left(0,0;(\mathbf{1}_{B_1},\ldots,\mathbf{1}_{B_n}),(\mathbf{1}_{A_1},\ldots,\mathbf{1}_{A_m})\right)$$
$$=\sum_{\pi\in\Pi_n(A_1,\ldots,A_m)}\int_{B_1\times\cdots\times B_n}\prod_{i=1}^{n}\delta^{|\pi_i|}G_{\Phi'_d}\left(0|x_i;\mathbf{1}_\xi:\xi\in\pi_i\right)J^{(n)}_{\Phi_p}(\mathrm{d}(x_1,\ldots,x_n)). \tag{26}$$

The form (20) of the higher-order Janossy measure is then proved by using (11), and (22) is directly proved by considering $v=1$ and using (9). The same principle can be applied to prove the result for the higher-order factorial moment measure $M^{(n,m)}_{\Phi_c}$ by considering that $G_{\Phi_d}(1|\cdot)=1$ and using (9) and (12). □

Example 4.1 (*Poisson Cluster Processes*). Consider G_{Φ_p} to be the p.g.fl. of a Poisson point process with rate λ, so that

$$G_{\Phi_p}(v)=\exp\left(\int\lambda(v(x)-1)P_{\Phi_p}(\mathrm{d}x)\right). \tag{27}$$

Then G_{Φ_c} describes a Poisson cluster process, i.e.

$$G_{\Phi_c}(v,w)=\exp\left(\int\lambda(v(x)G_{\Phi'_d}(w|x)-1)P_{\Phi_p}(\mathrm{d}x)\right).$$

Suppose additionally that $G_{\Phi'_d}(w|x)$ is a Poisson process conditioned on parent point $x\in\mathbf{X}$, so that

$$G_{\Phi'_d}(w|x)=\exp\left(\int\mu(x)(w(y)-1)P_{\Phi'_d}(\mathrm{d}y|x)\right). \tag{28}$$

Then Φ_c is doubly Poisson. The Janossy and factorial moment measure of the cluster process are found by substituting the following into Theorem 4.2 (some of these results are known and can be found in Daley and Vere-Jones (2003, Section 6.3)),

$$J_{\Phi_p}(\mathrm{d}(x_1,\ldots,x_n))=\exp(-\lambda)\lambda^nP_{\Phi_p}(\mathrm{d}x_1)\ldots P_{\Phi_p}(\mathrm{d}x_n) \tag{29a}$$
$$M_{\Phi_p}(\mathrm{d}(x_1,\ldots,x_n))=\lambda^nP_{\Phi_p}(\mathrm{d}x_1)\ldots P_{\Phi_p}(\mathrm{d}x_n) \tag{29b}$$

for the parent process, and

$$J_{\Phi'_d}(\mathrm{d}(z_1,\ldots,z_m)|x)=\exp(-\mu(x))\mu(x)^mP_{\Phi'_d}(\mathrm{d}z_1|x)\ldots P_{\Phi'_d}(\mathrm{d}z_m|x) \tag{30a}$$
$$M_{\Phi'_d}(\mathrm{d}(z_1,\ldots,z_m)|x)=\mu(x)^mP_{\Phi'_d}(\mathrm{d}z_1|x)\ldots P_{\Phi'_d}(\mathrm{d}z_m|x), \tag{30b}$$

for the daughter process.

If we receive an observation $Z=(z_1,\ldots,z_m)\in\mathbf{Y}^\cup$ of the global daughter process Φ_d, we need to assume that the conditional measures $P_{\Phi'_d}(\cdot|x)$ and $J_{\Phi'_d}(\cdot|x)$, for any $x\in\mathbf{X}$, are absolutely continuous with respect to a reference measure in $(\mathbf{Y}^\cup,\mathcal{B}(\mathbf{Y}^\cup))$ and we denote respectively $\hat{P}_{\Phi'_d}(\cdot|x)$ and $\hat{J}_{\Phi'_d}(\cdot|x)$ the associated probability densities. Then substituting (29) and (30) into Theorem 4.1 and the Bayesian estimation formulae in Section 3, we can determine the posterior Janossy and factorial moment measures of the parent of a Poisson cluster process given the states of the daughter process with

$$J^{(n)}_{\Phi_p|\Phi_d}(\mathrm{d}(x_1,\dots,x_n)|Z) = \frac{\lambda^n}{C(Z)} \sum_{\pi\in\Pi_n(Z)} \prod_{i=1}^{n} e^{-\mu(x_i)}\mu(x_i)\left(\prod_{z\in\pi_i} \hat{P}_{\Phi'_d}(z|x_i)\right) P_{\Phi_p}(\mathrm{d}x_i),$$

and

$$M^{(n)}_{\Phi_p|\Phi_d}(\mathrm{d}(x_1,\dots,x_n)|Z) = e^{\lambda} J^{(n)}_{\Phi_p|\Phi_d}(\mathrm{d}(x_1,\dots,x_n)|Z),$$

with

$$C(Z) = \sum_{k\geq 0} \frac{\lambda^k}{k!} \sum_{\pi\in\Pi_k(Z)} \prod_{i=1}^{k} \int e^{-\mu(x_i)}\mu(x_i)\left(\prod_{z\in\pi_i} \hat{P}_{\Phi'_d}(z|x_i)\right) P_{\Phi_p}(\mathrm{d}x_i).$$

5. Conclusion

A certain class of spatial cluster models can be compactly represented through the composition of probability generating functionals. In order to make this representation useful in practice, it is useful to be able to compute variational derivatives to determine Janossy measures and factorial moment measures. In this article, we introduce a recently derived tool in variational calculus, a general form of Faà di Bruno's formula, in order that these measures can be determined. We demonstrate the application of this result on the composition of point processes and illustrate through a Poisson cluster process parameterisation. The approach can be applied to other point process parameterisations if the Janossy measures of the parent and daughter processes defining the cluster process are known.

References

Bernhard, P., 2005. Chain differentials with an application to the mathematical fear operator. Nonlinear Analysis 62, 1225–1233.
Bogolyubov, N.N., 1946. Problems of a dynamical theory in statistical physics. In: N.N. Bogolyubov Selected Works, Part I: Dynamical Theory (1990). Gordon and Breach Science Publications.
Clark, D.E., Houssineau, J., 2013. Faà di Bruno's formula for chain differentials. arXiv:1310.2833.
Cox, D.R., Isham, V., 1980. Point Processes. Chapman and Hall.
Cressie, N., 1991. Statistics for Spatial Data. Wiley.
Daley, D., Vere-Jones, D., 2003. An Introduction to the Theory of Point Processes, Vol. I. Springer-Verlag.
Faà di Bruno, F., 1855. Sullo Sviluppo delle Funzioni. Annali di Scienze Matematiche e Fisiche 6, 479–480.
Gâteaux, R., 1919. Fonctions d'une infinité de variables indépendantes. Bulletin de la Société Mathématique de France 47, 70–96.
Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns. Wiley & Sons.
Lawson, A.B., Denison, D.G.T., 2002. Spatial Cluster Modelling. CRC Press.
Mahler, R., 2003. Multitarget Bayes filtering via first-order multitarget moments. IEEE Transactions on Aerospace and Electronic Systems 39 (4), 1152–1178.
Mahler, R., 2007. PHD filters of higher order in target number. IEEE Transactions on Aerospace and Electronic Systems 43 (4), 1523–1543.
Mahler, R., 2009a. PHD filters for nonstandard targets, I: extended targets, In: IEEE International Conference on Information Fusion, pp. 915–921.
Mahler, R., 2009b. The multisensor PHD filter part I: general solution via multitarget calculus. Proceedings of SPIE 7336.
Moller, J., Waagepetersen, R.P., 2003. Statistical Inference and Simulation for Spatial Point Processes. Chapman and Hall/CRC Press.
Moyal, J.E., 1962. The general theory of stochastic population processes. Acta Mathematica 108, 1–31.
Neyman, J., Scott, E.L., 1958. Statistical approach to problems of cosmology. Journal of the Royal Statistical Society 20 (1), 1–29.
Ripley, B.D., 1988. Statistical Inference for Spatial Processes. Cambridge University Press.
Srinivasan, S.K., 1973. Stochastic Point Processes and their Applications. In: Griffin's Statistical Monographs and Courses.
Stoyan, D., Kendall, W., Mecke, J., 1995. Stochastic Geometry and its Applications, second ed. Wiley, New York.
Swain, A., Clark, D.E., 2010. Extended object filtering using spatial independent cluster processes. In: IEEE International Conference on Information Fusion, pp. 1–8.
van Lieshout, M.N.M., 2000. Markov Point Processes and their Applications. Imperial College Press.
van Lieshout, M.N.M., Baddeley, A.J., 2002. Extrapolating and interpolating spatial point patterns. In: Lawson, A.B., Denison, D.G.T. (Eds.), Spatial Cluster Modelling. CRC Press, pp. 61–86.
Vo, B., Ma, W.K., 2006. The Gaussian mixture probability hypothesis density filter. IEEE Transactions on Signal Processing 54 (11), 4091–4104.

Vo, B.-N., Singh, S., Doucet, A., 2005. Sequential Monte Carlo methods for multitarget filtering with random finite sets. IEEE Transactions on Aerospace and Electronic Systems 41 (4), 1224–1245.
Vo, B.T., Vo, B.-N., Cantoni, A., 2007. Analytic implementations of probability hypothesis density filters. IEEE Transactions on Signal Processing 55 (7), 3553–3567. Part 2.