Course Notes for
EE 87021
Advanced Topics in Random Wireless Networks

Martin Haenggi

November 1, 2010

Contents

Part I

Point Process Theory

Chapter 1

Introduction

In the first part of the course, we give an introduction to stochastic geometry and, in particular, point process theory.

1.1 Motivation

Basic questions:

- How to describe a (random) collection of points in two or three dimensions?
- How about in one dimension? (In one dimension, renewal theory helps deal with processes with independent increments.
- How to generalize from one dimension to two dimensions? What is the main difference?
- Other random geometric objects: How to describe a "random line" or a "random triangle"?

Stochastic geometry (sometimes used synonymously with geometric probability) deals with random spatial patterns. Random point patterns are the most important objects, hence the *point process theory* is often considered to be the main sub-field of stochastic geometry.

We will use *point processes* to model the distributions of nodes (users, wireless terminals) in a wireless network where node locations are subject to uncertainty.

1.2 Asymptotic Notation

Let x tend to a. We write

$f(x) = O(g(x))$	if the ratio $f(x)/g(x)$ remains bounded.
$f(x) = o(g(x))$	if the ratio $f(x)/g(x)$ goes to 0.
$f(x) = \Theta(g(x))$	if $f(x) = O(g(x))$ and $g(x) = O(f(x))$.
$f(x) \sim g(x)$	if the ratio $f(x)/g(x)$ approaches 1.

More formally:

$$f(x) = O(g(x)) \text{ as } x \to a \text{ iff } \limsup_{x \to a} \left| \frac{f(x)}{g(x)} \right| < \infty$$

Always indicate the limit point.

Examples:

- $\ln x = O(x)$, $x^4 = O(e^x)$, $\sin x = O(1)$ as $x \to \infty$.
- $x^2 = O(x)$, $\sin x = O(x)$ as $x \to 0$.
- Taylor expansion: $e^x = 1 + x + O(x^2) = 1 + x + o(x) = 1 + x + o(x^{3/2}) = 1 + x + \Theta(x^2)$ (as $x \to 0$).
- As $x \to 0$, is $\cosh x = 1 + \Theta(x)$? Or $\cosh x = 1 + O(x)$? And/or $\cosh x = 1 + \Theta(x^2)$? For the sinh, Maple says: `taylor(sinh(x),x,3)=x+O(x^3)`.
- Let $x \to \infty$. What do we know about $f(x) = o(1/x)$? Is $1/\log x = o(1)$? Is $x^{O(2)} = O(e^x)$?

Note that $o(\cdot)$ implies and is stronger than $O(\cdot)$, *i.e.*, $f(x) = o(g(x)) \Rightarrow f(x) = O(g(x))$.

Chapter 2

Description of Point Processes

2.1 One-dimensional Point Processes

There is a number of ways to describe a point process $\{x_1, x_2, \ldots, \}$ in one dimension. Here are four of them:

1. Direct characterization of the points (or *arrival times* if the axis is a time axis) x_i. The problem is the dependency due to the ordering.

2. Using the *interarrival times* $S_i = x_{i+1} - x_i$. In the case of *renewal processes*, these increments are independent.

3. *Counting* the number of points falling in a set B:

$$N(B) = \sum_{i=1}^{\infty} \mathbf{1}\{x_i \in B\}$$

4. Using *vacancies*:

$$V(B) = \mathbf{1}\{N(B) = 0\}$$

2.2 General Point Processes

In the general case, a point process $\Phi = \{x_1, x_2, \ldots\}$ is a countable collection of randomly placed points in a general space. We will mostly consider the Euclidean spaces $\mathbb{R}^d$, such that $\Phi \subset \mathbb{R}^d$ is a *random closed set*.

Due to the lack of a natural ordering in higher-dimensional spaces, Methods 1 and 2 are cumbersome if not impossible to use in the general case. Methods 3 and 4, however, generalize in a straightforward manner since B can be taken to be a subset of any higher-dimensional space.

Such a point process can be described using two formalisms:

Random measure formalism. For a stochastic point process, we may *define* a point process as a collection of (integer-valued) random variables $N(B)$ indexed by $B \subset \mathbb{R}^2$. $N(B)$ denotes the number of points in $B \subset \mathbb{R}^d$, *i.e.*, $N(B) = \#\{\Phi \cap B\}$. The *vacancy indicators* $V(B)$ are defined as $V(B) = \mathbf{1}\{N(B) = 0\}$.

Random set formalism. A simple point process can be regarded as a random set $\Phi = \{x_1, x_2, \ldots\}$ consisting of random variables x_i as its elements.

If $N_1(B)$ and $N_2(B)$ are the numbers of nodes in B for two point processes Φ_1 and Φ_2, then $N(B) = N_1(B) + N_2(B)$ is the number of nodes in B of the superimposed process (additive property).

Similarly, if $V_1(B)$ and $V_2(B)$ are the vacancy indicators, then $V(B) = V_1(B)V_2(B)$ (multiplicative property).

In this course we are focusing on *simple* and *locally finite* processes, *i.e.*,

$$N(\{x\}) \leqslant 1 \quad \forall x \in \mathbb{R}^d$$

and

$$|B| < \infty \implies N(B) < \infty \quad \text{w.p. } 1\,.$$

If we know the value of the vacancy function $V(B)$ for all sets B, the PP is determined by the complement of the union of all vacant regions:

$$\Phi = \mathbb{R}^2 \setminus \bigcup \{B \subset \mathbb{R}^2 \mid V(B) = 1\}$$

2.3 Basic Point Processes

2.3.1 One-dimensional Poisson processes

Definition 2.1 (One-dimensional Poisson process) *The one-dimensional Poisson point process (PPP), with* uniform intensity β*, is a point process in* $\mathbb{R}$ *such that*

- *for every bounded interval* $(a, b]$*,* $N(a, b]$ *has a* Poisson *distribution with mean* $\beta(b - a)$*.*
- *if* $(a_1, b_1], (a_2, b_2], \ldots, (a_m, b_m]$ *are disjoint bounded intervals, then* $N(a_1, b_1], N(a_2, b_2], \ldots, N(a_m, b_m]$ *are independent random variables.*

The first property alone is sufficient to characterize a PP. Alfréd Rényi proved that the second property is a consequence of the first one if it holds for all Borel sets and not just intervals.

Other useful properties of the one-dimensional homogeneous PP are:

1. The inter-arrival times S_i are iid, and follow an exponential distribution with parameter β.
2. Let $R_{k,i} = x_{k+i} - x_i$ for $k > 0$. Then $R_{k,i}$ is Erlang (gamma) distributed with parameters k and β for all i.
3. The probability of two or more arrivals in a given interval is asymptotically of smaller order than the length of the interval:
$$\mathbb{P}(N(a, a + h] \geq 2) = o(h), \quad h \to 0.$$

We may also define the *inhomogeneous* PP, where the intensity function $\beta(t)$ is not a constant, in a similar way. For this process, the average number of points in any bounded interval $(a, b]$ is

$$\mathbb{E}N(a, b] = \int_a^b \beta(t)\mathrm{d}t,$$

and the number of arrivals in disjoint intervals are independent random variables.

2.3.2 Spatial Poisson processes

Definition 2.2 (Spatial point process) *The spatial PPP, with intensity β, is a point process in $\mathbb{R}^2$ such that*

- *for every bounded closed set B, $N(B)$ has a Poisson distribution with mean $\beta\lambda_2(B)$.*
- *if $B_1, B_2, \ldots, B_m$ are disjoint bounded sets, then $N(B_1), N(B_2), \ldots, N(B_m)$ are independent random variables.*

Here, $\lambda_2(B)$ denotes the area of the set B. β is the expected number of points of the process per unit area.

Theorem 2.1 (Conditional property) *Consider a homogeneous PPP in $\mathbb{R}^2$ with intensity $\beta > 0$. Let $W \subset \mathbb{R}^2$ be any region with $0 < |W| < \infty$. Given that $N(W) = n$, the conditional distribution of $N(B)$ for $B \subseteq W$ is binomial:*

$$\mathbb{P}\left(N(B) = k \mid N(W) = n\right) = \binom{n}{k} p^k (1-p)^{n-k},$$

where $p = |B|/|W|$.

Proof: See [1, Lemma 1.1]. □

The only distinction between a binomial process and a PPP in W is that different realizations of the PPP will consist of different number of points.

2.3.3 General Poisson point processes

A uniform (homogeneous) PPP in $\mathbb{R}^d$, $d \geq 1$, or an inhomogeneous PPP in $\mathbb{R}^d$, or a PPP on some other space S, is defined as:

Definition 2.3 (Poisson point process (PPP)) *The PPP on a space S with intensity measure Λ is a point process such that*

- *for every compact set $B \subset S$, $N(B)$ has a Poisson distribution with mean $\Lambda(B)$:*

$$\mathbb{P}(N(B) = k) = \exp\Big(-\int_B \lambda(x)dx\Big) \cdot \frac{\Big(\int_B \lambda(x)dx\Big)^k}{k!}$$

- *if $B_1, B_2, \ldots, B_m$ are disjoint compact sets, then $N(B_1), N(B_2), \ldots, N(B_m)$ are independent.*

2.4 Distributional Characterization

Like any random process, a point process can be described in statistical terms by defining the space of possible outcomes and then specifying the probabilities of different events.

2.4.1 The distribution of a point process

The space of realizations of a point process in $\mathbb{R}^d$ is N, the set of all counting measures on $\mathbb{R}^d$, where a counting measure is non-negative, integer-valued and finite on compact sets.

Basic events: A basic event $E_{B,k}$ about the point process is the event that there are exactly k points in the region B, for compact $B \subset \mathbb{R}^d$ and non-negative integer $k = 0, 1, 2, \ldots$.

$$E_{B,k} = \{N(B) = k\} = \{N \in \mathsf{N} : N(B) = k\}.$$

Canonical space: Let N be the set of all counting measures on $\mathbb{R}^d$ and $\mathcal{N}$ be the σ-field of subsets of N generated by all events of the form $E_{B,k}$. The space N equipped with its σ-field $\mathcal{N}$ is called the **canonical space** or **outcome space** for a point process in $\mathbb{R}^d$.

A point process Φ may now be defined formally, using its counting measure $N = N_\Phi$, as a measurable map $N : \Omega \to \mathsf{N}$ from a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ to the outcome space $(\mathsf{N}, \mathcal{N})$. Thus, each elementary outcome $\omega \in \Omega$ determines an outcome $N^\omega \in \mathsf{N}$ for the entire point process.

The *distribution* of N_Φ is

$$\mathsf{P}(E) = \mathbb{P} \circ N^{-1}(E) = \mathbb{P}(N \in E) \qquad \forall E \in \mathcal{N}\,.$$

Measurability requires that $N^{-1}(E) \in \mathcal{A}$. Note that the counting random variables are indexed by the Borel sets B.

A general RV (not necessarily numeric) is also referred to as a *random element*. So, a point process is a random element $(\Omega, \mathcal{A}, \mathbb{P}) \to (\mathsf{N}, \mathcal{N})$.

Note:

- A possible source of confusion is the fact that in the case of numerical RVs, the Borel sets B are events. In the case of PPs, the Borel sets are indices of the counting random variables, or arguments if the counting RVs are viewed as functions on Borel sets.
- In principle we should distinguish the random point set Φ and the counting measure N_Φ. However, in the literature often the point process random element Φ is used itself for its associated counting measure for notational convenience.

2.4.2 Comparison with numerical random variables

Numerical random variables. Let X be a RV on $(\Omega, \mathcal{A}, \mathbb{P})$, *i.e.*, an $\mathcal{A}$-measurable function on Ω. The *distribution* of X is the measure

$$P \triangleq \mathbb{P} \circ X^{-1}$$

on $(\mathbb{R}, \mathcal{B})$, defined by

$$P(B) = \mathbb{P} \circ X^{-1}(B) = \mathbb{P}(X \in B) \quad \forall B \in \mathcal{B}\,,$$

where the *pre-image* of B is the set $X^{-1}(B) = \{\omega \in \Omega\colon X(\omega) \in B\}$.

Elementary events in $\mathcal{B}$ are the intervals $(a, b]$, $b > a$, and $\mathcal{B}$ includes all intersections and unions of countably many intervals.

Measurability is the requirement that $X^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}$.

For a numerical random variable $X : \Omega \to \mathbb{R}$, we almost always focus on the *distribution function* of X, where $B = (-\infty, x]$:

$$F(x) \triangleq P((-\infty, x]) = \mathbb{P}(X \leqslant x) \qquad \text{(right-continuous)}$$

The distribution function always exists and fully describes the distribution.

Comparison with point processes:

	Numerical random variable	Point process
Probability space	$(\Omega, \mathcal{A}, \mathbb{P})$	$(\Omega, \mathcal{A}, \mathbb{P})$
Measurable space	$(\mathbb{R}, \mathcal{B})$	$(\mathsf{N}, \mathcal{N})$
Random element	$X \in \mathbb{R}$	$N \in \mathsf{N}$
Events	$B \in \mathcal{B}$	$E \in \mathcal{N}$
Distribution	$P(B) = \mathbb{P} \circ X^{-1}(B) = \mathbb{P}(X \in B)$	$\mathsf{P}(E) = \mathbb{P} \circ N^{-1}(E) = \mathbb{P}(N \in E)$
Measurability condition	$X^{-1}(B) = \{\omega \in \Omega \colon X(\omega) \in B\} \in \mathcal{A}$	$N^{-1}(E) = \{\omega \in \Omega \colon N^{\omega} \in E\} \in \mathcal{A}$
Measure space	$(\mathbb{R}, \mathcal{B}, P)$	$(\mathsf{N}, \mathcal{N}, \mathsf{P})$
Distribution function	$F(x) = P((-\infty, x])$	
Range of counting measure		$N(B) \in \mathbb{N}_0$

We write N^{ω}, since $N(\omega)$ may be confused with $N(B)$.

Most often, the underlying probability space can be taken to be the generic one on the unit interval: $(\Omega, \mathcal{A}, \mathbb{P}) = ([0,1], \mathcal{B}, |\cdot|)$, where $|\cdot|$ denotes the Lebesgue measure.

The number of points in B can be expressed as

$$N(B) = \int_B N(\mathrm{d}x)\,.$$

2.4.3 Distribution of a point process viewed as a random set

As mentioned previously, a point process can be interpreted as a random measure or a random set or sequence. Let φ be a simple a locally finite countable subset of $\mathbb{R}^d$. Then the outcome space N is the set of all φ, and a point process Φ is a random choice of one of the φ in N. We also have

$$\int_{\mathsf{N}} \mathsf{P}(\mathrm{d}\varphi) = 1\,,$$

in the same way that for numerical random variables with distribution function (cdf) $F(x)$ and pdf $f(x)$ we have

$$\int_{\mathbb{R}} P(\mathrm{d}x) = \int_{\mathbb{R}} \mathrm{d}F(x) = \int_{\mathbb{R}} f(x)\mathrm{d}x = 1\,.$$

We will make extensive use of this notation later.

2.4.4 Finite-dimensional distributions and capacity functional

Definition 2.4 (Finite-dimensional distribution) *The finite-dimensional distributions (fidis) of a point process are the joint probability distributions of*

$$(N(B_1), ..., N(B_m))$$

for all finite integers $m > 0$ and all compact $B_1, B_2, \ldots$.

Equivalently the fidis specify the probabilities of all events of the form

$$\{N(B_1) = k_1, \ldots, N(B_m) = k_m\}$$

involving finitely many regions.

Definition 2.5 (Capacity functional) *The capacity functional of a simple point process Φ is the functional*

$$T(K) = \mathbb{P}(N(K) > 0)\,, \quad K \text{ compact}\,.$$

Facts:

- If the fidis of two point processes X, Y are identical then X and Y have the same distribution.
- If the capacity functionals of two simple point processes X, Y are identical then they have the same distribution[1].
- A simple point process is a homogeneous PPP of intensity λ on $\mathbb{R}^d$ if and only if its capacity functional
$$T(K) = 1 - \exp(-\lambda \|K\|)\,, \qquad \text{for all compact } K \subset \mathbb{R}^d\,.$$
- $X \subset \mathbb{R}^d$ is stationary if for any fixed $v \in \mathbb{R}^d$ the distribution of $X + v$ is the same as that of X.
- A simple point process is stationary if $T(K) = T(K + v)$ for $\forall v \in \mathbb{R}^d$ and all compact K.

Theorem 2.2 (Rényi's Theorem) *Let Φ be a PP and $\lambda : \mathbb{R}^d \to \mathbb{R}$ be a non-negative function such that $\Lambda(B) = \int_B \lambda(x)\mathrm{d}x < \infty$ for all bounded B. If*

$$\mathbb{P}(\Phi \cap B = \emptyset) = \exp(-\Lambda(B))$$

for any Borel B then Φ is Poisson with intensity function λ.

Proof: See Grimmett/Stirzaker p. 288f.
Idea: Partition B into a finite number n of subsets of equal and decreasing size such that in the limit $n \to \infty$, the union of the disjoint subsets is exactly B. From the condition (2.2) it follows that the indicator RVs that the subsets are not empty are independent. Use that independence to calculate the probability generating function, bound it, and use monotone convergence to show that the pgf is that of a Poisson process. □

2.4.5 Measurable decomposition

Consider point processes with an infinite number of points a.s. Any N can be decomposed measurably as

$$N = \sum_{i=1}^{\infty} \delta_{x_i}$$

$\delta_X(B) = \mathbf{1}_B(x)$ is the Dirac measure for $B \in \mathcal{B}$.

The Dirac measure and thus N is an *atomic measure*, *i.e.*, it is concentrated on a countably infinite collection of points. A counting measure is a special atomic measure that gives each point a mass of zero or one. In contrast, the Lebesgue measure is a *diffuse measure* since it gives zero mass to every point. Both are *Radon measures* since they are finite on finite sets. In fact, every translation-invariant Radon measure on $\mathbb{R}^d, \mathcal{B}^d$ is a constant multiple of the Lebesgue measure.

[1] This holds more generally for random closed sets.

2.4.6 Intensity measure

Definition 2.6 (Intensity (mean) measure) *The intensity or mean measure is defined by as*

$$\Lambda(B) \triangleq \mathbb{E}(\Phi(B))$$

for a PPP Φ.

If there is an intensity function $\lambda(x)$, we have

$$\Lambda(B) = \int_B \lambda(x)\mathrm{d}x\,.$$

Using the distribution of the point process, we can also write

$$\Lambda(B) = \int_{\mathsf{N}} \varphi(B)P(\mathrm{d}\varphi)\,.$$

In general, we have

$$\Lambda(\{x\}) = 0,\ \forall x \in \mathbb{R}^d\,.$$

This means that we do not expect to have a point on a specific location (unless the PP has at least one deterministically placed point). In other words, Λ is a *diffuse measure*. It cannot be used for conditioning on points at specific locations without some precaution.

The intensity (function) $\lambda(x)$ may not exist. If it does, we can retrieve it from the intensity measure by calculating the limit

$$\lambda(x) = \lim_{r\to 0} \frac{\mathbb{E}\Phi(b(x,r))}{|b(o,r)|}\,.$$

The denominator is the volume of the d-dimensional ball of radius r, which is $c_d r^d$, where

$$c_d = \frac{\pi^{d/2}}{\Gamma(d/2+1)}$$

is the volume of the unit ball.

While stationarity implies that the intensity function is constant, the converse is not true: A constant intensity function does *not* imply stationarity.

Theorem 2.3 (Conditional property) *Let* Φ *a general PPP* $\in \mathbb{R}^d$ *with intensity* λ. *Take* B *such that* $0 < \Lambda(B) < \infty$. *Conditioned on* $\Phi(B) = n$, *the* n *points in* B *are distributed with probability density*

$$\frac{\lambda(x)}{\Lambda(B)}\,.$$

Proof: In the homogeneous case (constant λ), we have seen that, conditioned on $\Phi(B) = n$, the n points form a binomial point process. This is a straightforward generalization to the inhomogeneous case. The intensity $\lambda(x)$ is a density that needs to be normalized by $\Lambda(B)$ to be a proper probability density. □

2.5 Properties of Point Processes

Let $\Phi_x = \{x_1 + x, x_2 + x, \ldots\}$ be the point process translated by $x \in \mathbb{R}^d$.

Stationarity. A PP on $\mathbb{R}^d$ is *stationary*, if its distribution is translation-invariant, *i.e.*, $\mathsf{P}(E) = \mathsf{P}(E + x)$, for all $E \in \mathcal{N}$ and $x \in \mathbb{R}^d$. $E + x$ or E_x is defined as

$$E + x = E_x = \{\varphi \in \mathsf{N} \colon \varphi_{-x} \in E\}\,, \quad E \in \mathcal{N}\,.$$

Equivalently, stationarity implies $\mathbb{P}(\Phi \in E) = \mathbb{P}(\Phi_x \in E)$, again for all E and x.

Isotropy. A PP on $\mathbb{R}^d$ is *isotropic*, if its distribution is rotationally invariant with respect to rotations about the origin o, *i.e.*, $\mathsf{P}(E) = \mathsf{P}(\mathbf{r}E)$, where $\mathbf{r}$ is an arbitrary rotation about o in $\mathbb{R}^d$,*i.e.*,

$$\mathbf{r}E = \{\varphi \in \mathsf{N} \colon \mathbf{r}^{-1}\varphi \in E\}\,, \quad E \in \mathcal{N}\,.$$

Motion-invariance. A stationary and isotropic PP is called *motion-invariant*.

Remarks. The class of motion-invariant PPs is by far the most important for our purposes. The one exception are inhomogeneous PPPs, which are analytically tractable due to their independence property.

For stationary PPs, the intensity measure is translation-invariant:

$$\Lambda(B) = \mathbb{E}\Phi(B) = \mathbb{E}\Phi_x(B) = \mathbb{E}\Phi(B_{-x}) = \Lambda(B_{-x})\,, \quad \forall x\,.$$

If follow that the intensity measure is then just a multiple of the Lebesgue measure, with λ being the proportionality constant:

$$\Lambda(B) = \lambda|B|\,.$$

Example 2.1

- **Isotropic, non-stationary PP:** *An inhomogeneous PPP over $\mathbb{R}^d$ with $\lambda(x) = f(\|x\|)$ (*i.e.*, radially symmetric intensity function). Also, a binomial point process of constant intensity defined on $b(o, r)$.*
- **Stationary, non-isotropic PP:** *A randomly translated 2D integer lattice: $\Psi \triangleq \mathbb{Z}^2 + (U, V)$, where U and V are uniform random variables over the unit interval $(0, 1)$. To demonstrate anisotropy, let $B_\square$ be a square of side length 1.1. Then $\mathbb{P}(\Psi(B_\square) = 0) = 0$, while for the rotated square (with the same area) by $\pi/4$, denoted by $B_\diamondsuit$, we have $\mathbb{P}(\Psi(B_\diamondsuit) = 0) > 0$, since the rotated square fits in the lattice without intersecting with any lattice point.*

2.6 Point Process Transformations

2.6.1 Shaping an inhomogeneous Poisson point process

Theorem 2.4 (Mapping theorem) *Let Φ be a non-homogeneous PPP on $\mathbb{R}^d$ with intensity function λ. Let $f : \mathbb{R}^d \to \mathbb{R}^s$ be measurable and $\Lambda(f^{-1}\{y\}) = 0, \forall y \in \mathbb{R}^s$ (f does not shrink a compact set to a point). Let*

$$\mu(B) \triangleq \Lambda(f^{-1}(B)) = \int_{f^{-1}(B)} \lambda(x)\mathrm{d}x\,,$$

where it satisfies $\mu(B) < \infty, \forall$ compact B.
Then

$$f(\Phi) = \bigcup_{x \in \Phi} f(x)$$

is a PPP with intensity measure μ.

Corollary 2.5 (Linear mapping) *Consider a stationary PPP Φ of intensity λ, and let $A\colon \mathbb{R}^d \mapsto \mathbb{R}^d$ be a non-singular linear mapping, given by a transformation matrix $A \in \mathbb{R}^{d\times d}$. Then $A(\Phi) = \{Ax\colon x \in \Phi\}$ is a stationary PPP with intensity $\lambda \det(A^{-1})$*

Example 2.2 *If $\Phi = \{x_i\}$ is a homogeneous PPP of intensity σ on $\mathbb{R}^2$, what is the intensity (function) $\lambda(x)$ of the one-dimensional PPP $\Psi = \{\|x_i\|\}$? What is the intensity function of $\Phi' = \{\|x_i\|^2\}$?*

In the first case, $f(x) = \|x\|$. Let $B = [0, r)$. Then $f^{-1}(B) = b(o, r)$, and we have

$$\mu([0, r)) = \Lambda(b(o, r)) = \int_{b(o,r)} \sigma \mathrm{d}x = \sigma\pi r^2 \,,$$

from which the density follows as

$$\lambda(x) = \frac{\mathrm{d}}{\mathrm{d}x}\mu([0, x)) = 2\sigma\pi x \,, \quad x \geq 0 \,.$$

Hence this is an inhomogeneous one-dimensional PPP with intensity function $\lambda(x) = 2\sigma\pi x \mathbf{1}\{x \geq 0\}$.

In the second case, $f(x) = \|x\|^2$. The preimage of the interval $[0, r)$ is the ball $b(o, \sqrt{r})$, and we find $\mu([0, r)) = \sigma\pi r$ and $\lambda(x) = \sigma\pi \mathbf{1}\{x \geq 0\}$.

FonThinning of a Poisson point process

Theorem 2.6 (Independent thinning) *Let $g\colon \mathbb{R}^d \to [0, 1]$ be a* thinning function*,* i.e.*, consider a realization of a stationary PPP Φ and delete each point with probability $1 - g(x)$, independently of all other points. This procedure, called* thinning*, generates an inhomogeneous PPP with density function $\lambda g(x)$.*

Proof: Let Φ be the original process and $\tilde{\Phi}$ the process after thinning. Independence follows from the construction. The distribution of $\tilde{\Phi}$ is computed as follows:

$$\mathbb{P}(\tilde{\Phi}(B) = k) = \sum_{i=k}^{\infty} \mathbb{P}(\Phi(B) = i)\mathbb{P}(\tilde{\Phi}(B) = k \mid \Phi(B) = i)$$

Given $\Phi(B) = i$, the i points of Φ in B are uniformly distributed. Thus the conditional probability of $\tilde{\Phi}(B) = 1$ given $\Phi(B) = 1$ is just

$$\mathbb{P}(\tilde{\Phi}(B) = 1 \mid \Phi(B) = 1) = |B|^{-1} \int_B g(x)\mathrm{d}x$$

and the complete conditional distribution is

$$\mathbb{P}(\tilde{\Phi}(B) = k \mid \Phi(B) = i) = \binom{i}{k} \left(|B|^{-1} \int_B g(x)\mathrm{d}x\right)^k \left(1 - |B|^{-1} \int_B g(x)\mathrm{d}x\right)^{i-k} .$$

Hence, with $G \triangleq \int_B g(x)\mathrm{d}x$,

$$\begin{aligned}
\mathbb{P}(\tilde{\Phi}(B) = k) &= \sum_{i=k}^{\infty} e^{-\lambda|B|} \frac{(\lambda|B|)^i}{i!} \binom{i}{k} \left(|B|^{-1}G\right)^k (1 - |B|^{-1}G)^{i-k} \\
&= e^{-\lambda|B|} \frac{(\lambda G)^k}{k!} \sum_{i=k}^{\infty} \frac{(\lambda|B|(1 - |B|^{-1}G))^{i-k}}{(i-k)!} \\
&= e^{-\lambda|B|} \frac{(\lambda G)^k}{k!} e^{\lambda|B|(1-|B|^{-1}G)} \\
&= \frac{(\lambda G)^k}{k!} e^{-\lambda G} \,.
\end{aligned}$$

□

2.7 Distances

Definition 2.7 (Contact distance) *The* contact distance *is* $\operatorname{dist}(u, \Phi)$*,* $u \in \mathbb{R}^d$*,* $\Phi \subset \mathbb{R}^d$ *the distance from point* u *to the nearest point of* Φ*:*

$$\operatorname{dist}(u, \Phi) \triangleq \min\{x \in \Phi \colon \|x - u\|\}, \quad u \in \mathbb{R}^d.$$

Example 2.3 *Let* Φ *be a homogeneous PPP of intensity* λ*.* $\mathbb{P}(\operatorname{dist}(u, \Phi) \leq r) = \mathbb{P}(N(b(u, r)) > 0) = 1 - e^{-\lambda c_d r^d}$*, where* $c_d = |b(o, 1)|$ *is the volume of the* d*-dimensional unit ball. So* $V = c_d \cdot \operatorname{dist}(u, \Phi)^d$ *is the largest ball that we can fit in before we hit a point in* Φ*. We have that* $\mathbb{P}(V < v) = 1 - e^{-\lambda \cdot v}$*.*

Definition 2.8 (Contact distribution function or empty space function) *The contact distribution function or empty space function* F *is the cdf of* $\operatorname{dist}(u, \Phi)$*:*

$$F^u(r) \triangleq \mathbb{P}(\operatorname{dist}(u, \Phi) \leq r) = \mathbb{P}(N(b(u, r)) > 0) = T(b(u, r))$$

Remarks. If Φ is stationary, F^u does not depend on u. $F^u(r) = T(b(u, r))$ is the special case of the capacity function where we use balls for the argument of the capacity function.

Definition 2.9 (Nearest-neighbor distance) *The* nearest-neighbor distance $\operatorname{dist}_{\text{NN}}(x, \Phi)$*,* $x \in \Phi$*, is the distance from a point in* Φ *to its nearest neighbor:*

$$\operatorname{dist}_{\text{NN}}(x, \Phi) = \min\{y \in \Phi \setminus \{x\} \colon \|x - y\|\}, \quad x \in \Phi.$$

The corresponding distribution function is the nearest-neighbor distance distribution function*.*

For the PPP, the contact distribution function and nearest-neighbor distance distribution function are identical in the following sense:

$$\operatorname{dist}(u, \Phi) \stackrel{\text{d}}{=} \operatorname{dist}_{\text{NN}}(x, \Phi \mid x \in \Phi).$$

This follows from the independence property of the PPP: Conditioning on the PPP having a point at x does not affect the distribution of the other points. A rigorous discussion of conditioning on such events of probability 0 follows in Chapter 5.

2.8 Marked point processes

In a *marked point process* a mark $m_i \in \mathbb{M}$, where $\mathbb{M}$ is the *mark space*, is added to each point x_i.

Example 2.4 (Marked point processes)

- *In a forest,* m *may denote the diameter of the trees.*
- *In a sensor network* m *may denoted the battery level or transmission power, etc.*

- *In ALOHA, $m_i \in \{0, 1\}$ could be the transmit or receive state of the node.*

Definition 2.10 (Marked point process:) *A marked point process in space $\mathbb{R}^d$ with marks in space $\mathbb{M}$ is a PP $\hat{\Phi}$ on $\mathbb{R}^d \times \mathbb{M}$ such that $\hat{\Phi}(K \times \mathbb{M}) < \infty, \quad \forall K \subset \mathbb{R}^d$, K compact. So $\hat{\Phi} = \{(x_i, m_i)\} \subset \mathbb{R}^d \times \mathbb{M}$.*

Example 2.5 *A PPP on $\mathbb{R}^3$ cannot be viewed as a* marked PP *in $\mathbb{R}^2 \times \mathbb{R}$ (with points in $\mathbb{R}^2$ and the third coordinate as the mark). The reason is the following: In a general marked PP, M need not be compact. It can be the whole mark space, so if we use the 3^{rd} dimension as the mark space then in a compact set $K \in \mathbb{R}^2$ there will be an infinite number of points such that $N_Y(K \times \mathbb{R}) \not< \infty$*

Example 2.6 *On the other hand, a PPP on $\mathbb{R}^2 \times [0, a]$ of intensity λ can be interpreted as a marked point process on $\mathbb{R}^2$ with marks from $\mathbb{M} = [0, a]$. The projected point process on $\mathbb{R}^2$ has intensity λa, since $\Lambda(B \times [0, a]) = \lambda a|B|$ for $B \subset \mathbb{R}^2$. The marks attached to each points are uniformly distributed on $[0, a]$.*

The mark space can be $\mathbb{R}$ or $\mathbb{Z}$, or a subset thereof, but it can also be more complicated, such as $\mathbb{R}^n$ a function space. For example, the fading states from a node to all other nodes in the process may be attached as a mark.

It is always possible to interpret the marked PP as an ordinary PP on the product space $\mathbb{R}^d \times \mathbb{M}$. So why do we need marked PPs? The reason is that Euclidean motions affect the positions of the points but leaves the marks unchanged. Thus it is often more intuitive to deal with marks than PPs on product spaces.

Theorem 2.7 (Marking theorem for Poisson point processes) *Let $\hat{\Phi}$ be a marked point process on $\mathbb{R}^d \times \mathbb{M}$. Let Φ be the projected point process (without marks). The the following two statements are equivalent:*

1. *Φ is a Poisson process on $\mathbb{R}^d$ with intensity measure Λ, and, given Φ, the marks $\{m_i\}$ are iid with distribution Q on $\mathbb{M}$.*
2. *$\hat{\Phi}$ is a Poisson process on $\mathbb{R}^d \times \mathbb{M}$ with intensity measure $\Lambda \otimes Q$.*

Proof: See [2, Sec. 5.2]. □

Chapter 3

Sums over Point Processes

3.1 Notation

Let $\Phi = \{x_1, x_2, \ldots\} = \{x_i\}$ be a point process. A sum of $f(x)$ over Φ can be alternatively written as

$$\sum_{x\in\Phi} f(x) = \int_{\mathbb{R}^d} f(x)\Phi(\mathrm{d}x) = \int_{\mathbb{R}^d} f(x)p(x)\mathrm{d}x\,,$$

where

$$p(x) = \sum_{y\in\Phi} \delta(x-y)\,.$$

$\delta(x)$ is the Dirac delta function.

The mean value of the sum can alternatively be written as

$$\mathbb{E}\left(\sum_{x\in\Phi} f(x)\right) = \int_{\mathsf{N}} \sum_{x\in\varphi} f(x)\mathsf{P}(\mathrm{d}\varphi) = \int_{\mathsf{N}} \int_{\mathbb{R}^d} f(x)\varphi(\mathrm{d}x)\mathsf{P}(\mathrm{d}\varphi)\,.$$

The different ways of writing these expressions reflect the variety of approaches to the theory.

Example 3.1 (Number of points) *The number of points in B is written as*

$$\Phi(B) = \sum_{x\in\Phi} \mathbf{1}_B(x) = \int_B \Phi(\mathrm{d}x)\,.$$

Its mean value can be written in the following ways:

$$\begin{aligned}
\mathbb{E}\Phi(B) &= \mathbb{E}\left(\sum_{x\in\Phi} \mathbf{1}_B(x)\right) \\
&= \int_{\mathsf{N}} \varphi(B)\mathsf{P}(\mathrm{d}\varphi) \\
&= \int_{\mathsf{N}} \sum_{x\in\varphi} \mathbf{1}_B(x)\mathsf{P}(\mathrm{d}\varphi) \\
&= \int_{\mathsf{N}} \int_{\mathbb{R}^d} \mathbf{1}_B(x)\varphi(\mathrm{d}x)\mathsf{P}(\mathrm{d}\varphi) \\
&= \int_{\mathsf{N}} \int_{B} \varphi(\mathrm{d}x)\mathsf{P}(\mathrm{d}\varphi)
\end{aligned}$$

3.2 Campbell's Theorem for the Mean

Theorem 3.1 (Campbell's formula for the mean) *Let Φ be a PP on $\mathbb{R}^d$ and let $f\colon \mathbb{R}^d \to \mathbb{R}$ be a measurable function. Then the random sum*

$$S = \sum_{x\in\Phi} f(x)$$

is a random variable with mean

$$\mathbb{E}S = \int_{\mathbb{R}^d} f(x)\Lambda(\mathrm{d}x)\,,$$

provided the RHS is finite. If Φ is a PP on $\mathbb{R}^d$ with intensity function λ,

$$\mathbb{E}S = \int_{\mathbb{R}^d} f(x)\lambda(x)\mathrm{d}x\,.$$

This formula also applies to non-simple PPs.

Proof: We prove that the theorem holds for simple f, *i.e.*, if f is a step function

$$f = \sum_{i=1}^{m} c_i \mathbf{1}_{B_i}$$

for compact $B_i \subset \mathbb{R}^d$ and $c_i \in \mathbb{R}$. We have

$$S = \sum_{x\in\Phi} f(x) = \sum_{x\in\Phi}\sum_{i=1}^{m} c_i \mathbf{1}_{B_i}(x) = \sum_{i=1}^{m} c_i \Phi(B_i)$$

and thus

$$\mathbb{E}S = \mathbb{E}\left(\sum_{i=1}^{n} c_i\Phi(B_i)\right) = \sum_{i=1}^{m} c_i \mathbb{E}\Phi(B_i) = \sum_{i=1}^{m} c_i\Lambda(B_i) = \int_{\mathbb{R}^d} f(x)\Lambda(\mathrm{d}x)\,.$$

The result for general f follows by monotone approximation. □

Written differently, Campbell's formula says

$$\int_{\mathsf{N}}\int_{\mathbb{R}^d} f(x)\varphi(\mathrm{d}x)\mathsf{P}(\mathrm{d}\varphi) = \int_{\mathbb{R}^d} f(x)\Lambda(\mathrm{d}x)\,.$$

The formula reminds of the expectation of a numerical random variable X with pdf $p_X(x)$:

$$\mathbb{E}f(X) = \int f(x)p_X(x)\mathrm{d}x$$

Focusing on a PP on a finite domain W with a fixed number of points n and averaging the contribution per point in the sum, we have

$$\frac{1}{n}\mathbb{E}\sum_{x\in\Phi} f(x) = \int_W f(x)\frac{\lambda(x)}{n}\mathrm{d}x\,.$$

In the case of a BPP, $\lambda(x)/n$ is indeed the pdf of the node distribution, which shows that there is an analogy between the mean contribution of a point in a sum over a PP and the expectation of a numerical random variable.

Corollary 3.2 (Campbell's formula for stationary point processes) *If $\Phi \subset \mathbb{R}^d$ is stationary with intensity λ, the sum $S = \sum_{x\in\Phi} f(x)$ is a random variable with mean*

$$\mathbb{E}S = \lambda \int_{\mathbb{R}^d} f(x)\mathrm{d}x\,.$$

3.3 The Probability Generating Functional

Definition 3.1 (Probability generating functional of a point process) *Let $v\colon \mathbb{R}^d \mapsto [0,1]$ a non-negative measurable function. The* generating functional *of the point process Φ is*

$$G[v] \triangleq \mathbb{E}\left(\prod_{x\in\Phi} v(x)\right) = \int_{\mathsf{N}} \prod_{x\in\varphi} v(x)\mathsf{P}(\mathrm{d}\varphi)\,.$$

$G[v]$ can also be written as

$$G[v] = \mathbb{E}\left[\exp\left(\sum_{x\in\Phi}\log v(x)\right)\right] = \mathbb{E}\left[\exp\left(\int_{\mathbb{R}^d}\log v(x)\Phi(\mathrm{d}x)\right)\right]\,.$$

Letting $f(x) \triangleq \log v(x)$ and $S = \sum_{x\in\Phi} f(x)$, the product $G[v]$ can, as a function of the sum S, be viewed as the *moment-generating function* of the sum S, *i.e.*,

$$M_S(t) \triangleq \mathbb{E}(e^{tS})$$

for $t = 1$, or its Laplace transform

$$\mathcal{L}_S(s) \triangleq \mathbb{E}(e^{-sS})$$

for $s = -1$, *i.e.*, $G[v] = M_S(1) = \mathcal{L}_S(-1)$.

In the Poisson case, there is a simple formula for these expressions, given by another theorem by Campbell.

3.3.1 The moment-generating function of the sum

Theorem 3.3 (Campbell's theorem for Poisson point processes) *Let Φ be a uniform PPP of intensity λ on $\mathbb{R}^d$ and $f\colon \mathbb{R}^d \mapsto \mathbb{R}$ be measurable. Then the sum*

$$S = \sum_{x\in\Phi} f(x)$$

is absolutely convergent with probability 1 *iff*

$$\int_{\mathbb{R}^d} \min(|f(x)|, 1)\mathrm{d}x < \infty\,. \tag{3.3.1}$$

If it is, then

$$\mathbb{E}(e^{tS}) = \exp\left(\lambda\int_{\mathbb{R}^d}(e^{tf(x)} - 1)\mathrm{d}x\right)$$

for any complex t for which the integral converges, and, in particular, when t is purely imaginary.

Proof: Consider (again) a *simple* function (step function) f that assumes finitely many non-zero values $f_1, \ldots, f_k$ and is zero outside some bounded region. Let

$$A_j = \{x\colon f(x) = f_j\}\,, \qquad j = 1, \ldots, k\,.$$

Since the A_j are disjoint, the RVs $N_j = \Phi(A_j)$ are iid Poisson with mean $\lambda|A_j|$. We also have that

$$S = \sum_{j=1}^{k} f_j N_j\,.$$

We know that for a Poisson RV X with mean μ

$$\mathbb{E}(e^{tX}) = \exp(\mu(e^t - 1))\,.$$

So the moment-generating function is

$$\begin{aligned}
\mathbb{E}(e^{tS}) &= \prod_{j=1}^{k} \mathbb{E}(e^{tf_j N_j}) \\
&= \prod_{j=1}^{k} \exp(\lambda|A_j|(e^{tf_j} - 1)) \\
&= \exp\left(\sum_{j=1}^{k} \int_{A_j} \lambda(e^{tf(x)} - 1)\mathrm{d}x\right) \\
&= \exp\left(\int_{\mathbb{R}^d} \lambda(e^{tf(x)} - 1)\mathrm{d}x\right)\,.
\end{aligned}$$

□

This Campbell theorem also holds for non-stationary processes. In this case, the convergence condition for the sum is

$$\int_{\mathbb{R}^d} \min(|f(x)|, 1)\Lambda(\mathrm{d}x) < \infty\,,$$

and the result is

$$\mathbb{E}(e^{tS}) = \exp\left(\int_{\mathbb{R}^d} (e^{tf(x)} - 1)\Lambda(\mathrm{d}x)\right)\,.$$

Since the moment-generating function defines the moments, we have the following corollary:

Corollary 3.4 (Mean and variance for PPPs) *Let $S = \sum_{x\in\Phi} f(x)$ for a Poisson point process Φ on $\mathbb{R}^d$. Then*

$$\mathbb{E}S = \int_{\mathbb{R}^d} f(x)\Lambda(\mathrm{d}x)$$

and

$$\mathrm{var}(S) = \int_{\mathbb{R}^d} f^2(x)\Lambda(\mathrm{d}x)\,.$$

Proof: Using the properties of the moment-generating function

$$\mathbb{E}S = \frac{\partial}{\partial t}\mathbb{E}\exp(tS)\Big|_{t=0}$$

and

$$\operatorname{var} S = \frac{\partial^2}{\partial t^2}\mathbb{E}\exp(tS)\Big|_{t=0} - (\mathbb{E}S)^2 .$$

□

The expression for the mean $\mathbb{E}S$ was obtained for general processes before, but the formula for the variance is new. It only holds for PPPs.

3.3.2 The characteristic or Laplace functional for the Poisson point process

Specializing Campbell's expression for the moment-generating function to the case $f \geqslant 0$ (to ensure convergence) and $t = -1$, we obtain

$$\mathbb{E}(e^{-S[f]}) = \exp\left(-\int_{\mathbb{R}^d}(1 - e^{-f(x)})\Lambda(\mathrm{d}x)\right) ,$$

where $S[f]$ is the sum of the function f over the point process. This is a functional since the argument is the function f. It is called the *characteristic functional* or the *Laplace functional*, sometimes denoted by $L[f]$. It is characteristic since if it holds for all step functions f, the process is Poisson.

3.3.3 The probability generating functional for the Poisson point process

Theorem 3.5 (Probability generating functional for the Poisson point process) *Let $v\colon \mathbb{R}^d \mapsto [0,1]$ be measurable and Φ a Poisson process with intensity measure Λ. Then*

$$G[v] \triangleq \mathbb{E}\left(\prod_{x\in\Phi} v(x)\right) = \exp\left(-\int_{\mathbb{R}^d}[1 - v(x)]\Lambda(\mathrm{d}x)\right) .$$

Proof: This follows from the characteristic functional by setting $v(x) = e^{-f(x)}$. An alternative, direct proof is as follows: Consider

$$v(x) = 1 - \sum_{i=1}^{n}(1 - z_i)\mathbf{1}_{C_i}(x)$$

for $z_i \in [0,1]$ and $C_1, \ldots, C_n$ pairwise disjoint and compact. Then

$$\begin{aligned} G[v] &= \mathbb{E}\left(\prod_{x\in\Phi} v(x)\right) \\ &= \int_{\mathsf{N}} z_1^{\varphi(C_1)} \cdot \ldots \cdot z_n^{\varphi(C_n)} P(\mathrm{d}\varphi) \\ &= \mathbb{E}\left(z_1^{\Phi(C_1)} \cdot \ldots \cdot z_n^{\Phi(C_n)}\right) \\ &= \mathbb{E}\left(z_1^{\Phi(C_1)}\right) \cdot \ldots \cdot \mathbb{E}\left(z_n^{\Phi(C_n)}\right) \end{aligned}$$

With $\mathbb{E}\left(z_1^{\Phi(C_1)}\right) = \exp(-\Lambda(C_1)(1-z_1))$ by the moment-generating (or probability-generating) function for the Poisson distribution,

$$\begin{aligned} G[v] &= \exp(-\Lambda(C_1)(1-z_1)) \cdot \ldots \cdot \exp(-\Lambda(C_n)(1-z_n)) \\ &= \exp\left(-\sum_{i=1}^{n} \int_{C_i} (1-z_i)\Lambda(\mathrm{d}x)\right) \\ &= \exp\left(-\int_{\mathbb{R}^d} (1-v(x))\Lambda(\mathrm{d}x)\right) \end{aligned}$$

This holds for piecewise constant v. The result for general v follows again by standard arguments from measure theory. □

Example 3.2 (Void probability) *Consider a uniform PPP of intensity λ. Let $v(x) = 1 - \mathbf{1}_B(x)$. Then the event $\Phi(B) = 0$ can be expressed as*

$$\prod_{x \in \Phi} v(x) = 1\,.$$

It occurs with probability

$$\mathbb{P}(\Phi(B) = 0) = \mathbb{E}\left(\prod_{x \in \Phi} v(x)\right) = G[v] = \exp(-\lambda|B|)\,,$$

as expected.

3.3.4 Relationship between moment-generating function and the functionals

Here, the relationships between the different expectations and functionals in this sections are explained. For notational simplicity, when discussing PPPs, we focus on the homogeneous case.

1. Let
$$S[f] = \sum_{x \in \Phi} f(x)\,; \qquad P[v] = \prod_{x \in \Phi} v(x)$$
be the sum of $f(x)$ over a point process Φ and P be the product of $v(x)$ over Φ, respectively. Then $S[f]$ and $P[v]$ are related as
$$P[v] = e^{S[\log v]}\,; \qquad S[f] = \log P[e^f]\,.$$
The same holds for their mean values, *e.g.*,
$$\mathbb{E}P[v] = \mathbb{E}(e^{S[\log v]})\,.$$

2. The expected value $\mathbb{E}(e^{S[f]})$ is the moment-generating function of S for $t = 1$ or the Laplace transform of S for $s = -1$:
$$\mathbb{E}(e^{S[f]}) = M_S(1) = \mathcal{L}_S(-1)$$
The moment-generating function for PPPs is given by Campbell's theorem.

3. The expected product $\mathbb{E}P[v]$ is called the *probability generating functional* (pgfl) of the point process:

$$G[v] \equiv \mathbb{E}P[v]$$

4. Campbell's theorem says that for PPPs,

$$\mathbb{E}(e^{S[f]}) = \exp\left(\lambda \int_{\mathbb{R}^d} (e^{f(x)} - 1)\mathrm{d}x\right)$$

if the integral converges. This is the mgf for $t = 1$.

5. If the integral diverges to $-\infty$, the sum $S = \infty$ a.s., and the mgf is 0. So the mgf is well defined unless the integral diverges to $+\infty$. A sufficient condition to avoid this is to focus on $f \leqslant 0$.

6. Equivalently, focus on $f \geqslant 0$ but consider the mgf for $t = -1$, *i.e.*, use

$$\mathbb{E}(e^{S[f]}) \equiv \mathbb{E}(e^{-S[-f]})\,.$$

This way, we obtain the *Laplace functional* or *characteristic functional*, defined as $L[f] \triangleq \mathbb{E}(e^{-S[f]})$ for $f \geqslant 0$. For the PPP,

$$L[f] = \mathbb{E}(e^{-S[f]}) = \exp\left(-\lambda \int_{\mathbb{R}^d} (1 - e^{-f(x)})\mathrm{d}x\right)\,, \quad f \geqslant 0\,.$$

7. The Laplace functional is related to the pgfl by setting $v = e^{-f}$, *i.e.*,

$$L[f] \equiv G[e^{-f}]\,.$$

This yields

$$G[v] = \mathbb{E}P[v] = \exp\left(-\lambda \int_{\mathbb{R}^d} (1 - v(x))\mathrm{d}x\right)\,,$$

where $v\colon \mathbb{R}^d \mapsto [0, 1]$.

3.4 Applications

3.4.1 Mean interference in stationary point process

Take a stationary point process $\Phi \subset \mathbb{R}^d$ of intensity λ. Assuming that each node (point) is transmitting at unit power and the path loss law is $\ell(x) = \|x\|^{-\alpha}$ for a path loss exponent $\alpha > 0$, we want to find.

$$\mathbb{E}I = \mathbb{E}\left(\sum_{x\in\Phi} \|x\|^{-\alpha}\right)\,.$$

By Campbell's formula for the mean,

$$\mathbb{E}I = \lambda \int_{\mathbb{R}^d} \|x\|^{-\alpha}\mathrm{d}x\,.$$

In two dimensions,

$$\lambda \int_{\mathbb{R}^2} \|x\|^{-\alpha} \mathrm{d}x = \lambda \int_0^{2\pi} \int_0^{\infty} r^{-\alpha} r \mathrm{d}r \mathrm{d}\beta = 2\pi \frac{\lambda}{2-\alpha} r^{2-\alpha} \Big|_0^{\infty}, \quad \alpha \neq 2.$$

For $\alpha = 2$,

$$\lambda \int_{\mathbb{R}^2} \|x\|^{-2} \mathrm{d}x = 2\pi\lambda \log r \Big|_0^{\infty}.$$

For this to be finite, we need $2 - \alpha < 0$ due to the upper integration bound and $2 - \alpha > 0$ due to the lower. And for $2 = \alpha$, it does not converge anyway. So the mean interference for the power path loss law is infinite in two dimensions for all values of the path loss exponent.

How about in three dimensions? We find

$$\mathbb{E}I = \lambda \int_{\mathbb{R}^3} \|x\|^{-\alpha} \mathrm{d}x = 4\pi \frac{\lambda}{3-\alpha} r^{3-\alpha} \Big|_0^{\infty}.$$

Again, the same problem. For larger number of dimensions d, these integrals may get tedious. We can use a mapping first as follows:

1. Map to one dimension. using $f(x) = \|x\|$. The new point process Φ' has intensity measure $\Lambda'([0, r)) = \lambda c_d r^d$ and intensity function $\lambda'(r) = \lambda c_d d r^{d-1}$.

2. Apply Campbell's theorem:

$$\begin{aligned} \mathbb{E}I = \mathbb{E}\left(\sum_{r \in \Phi'} r^{-\alpha} \right) &= \int_0^{\infty} r^{-\alpha} \lambda'(r) \mathrm{d}r \\ &= \lambda c_d \frac{d}{d-\alpha} r^{d-\alpha} \Big|_0^{\infty}, \quad \alpha \neq d. \end{aligned}$$

So for all d, there is no α for which the mean exists. If the upper bound is the culprit, that means that all the interferers far away contribute most of the interference. If $\alpha > d$, this problem is solved. The problem at $r = 0$ is due to the singularity of the path loss law near 0. This is clearly a modeling artefact, since no receiver ever gets more power than was transmitted. If the path loss function is replaced by a more accurate one, for example

$$g(x) = \min\{1, \|x\|^{-\alpha}\} \qquad \text{or} \qquad g(x) = (1 + \|x\|)^{-\alpha},$$

the issue at $r \to 0$ is solved, and the mean interference is finite as soon as $\alpha > d$.

Comparing the interference sum with condition (3.3.1), we notice that for $f(x) = \|x\|^{-\alpha}$, the condition is not violated, since the integrand is $\min\{|f(x)|, 1\}$. So the interference is an absolutely converging sum with a proper distribution, even though the mean may not be finite. If the condition is violated, then I has no proper distribution, since $I = \infty$ a.s.

Example 3.3 (Mean interference with bounded path loss)
Let the path loss law be $g(x) = \min\{1, \|x\|^{-\alpha}\}$. Then the mean interference in a PPP of intensity λ on the plane is

$$\mathbb{E}I = \lambda\pi + \frac{2\lambda\pi}{\alpha - 2}, \quad \alpha > 2.$$

If the diameter of the network is bounded to $R > 1$,

$$\mathbb{E}I = \lambda\pi + \frac{\lambda\pi}{\alpha - 2}(1 - R^{2-\alpha}), \quad \forall \alpha > 0.$$

Remarks.

- Why can we not use $\|x\|^{-\alpha}$ directly as the mapping function? Check the mean measure: In this case $\Lambda'([0,1]) = \Lambda([1,\infty]) = \infty$ (a.s.), which violates the finiteness of the mean measure.
- The mean is identical for all point process with the same intensity. Also, due to stationarity, the interference is the same no matter where on $\mathbb{R}^d$ we measure.

3.4.2 Variance of the interference in stationary Poisson point process

For general stationary point process, we are not in the position to find the variance. But in the Poisson case, we can apply Campbell's theorem. Let $g(x) = \min\{r_0^{-\alpha}, \|x\|^{-\alpha}\}$ for $r_0 > 0$. For a homogeneous PPP on $\mathbb{R}^d$,

$$\operatorname{var} I = \lambda \int_{\mathbb{R}^d} g^2(x)\mathrm{d}x = \lambda c_d r_0^{d-2\alpha} + \frac{\lambda c_d d}{d - 2\alpha} r^{d-2\alpha}\Big|_{r_0}^{\infty} = \lambda c_d r_0^{d-2\alpha}\left(\frac{2\alpha}{2\alpha - d}\right).$$

if $2\alpha > d$ (which is usually the case). For finite variance if $r_0 \to 0$ (and a finite upper integration bound), we would need $\alpha < d/2$—which would be highly unusual.

3.4.3 Interference from nearest transmitters

Let I_1 be the interference from the nearest transmitter. We find

$$\mathbb{P}(I_1 \leqslant x) = \mathbb{P}(R^{-\alpha} \leqslant x) = \mathbb{P}(R > x^{-1/\alpha}) = \exp(-\lambda c_d x^{-\delta}),$$

where $\delta = d/\alpha$. We get

$$\mathbb{E}I_1 = c_d^{1/\delta}\Gamma(1 - 1/\delta).$$

If $\delta < 1$ then this does not exist. So with the singular path loss law, the mean of only one interferer is already infinite.
The tail $\mathbb{P}(I_1 > x)$ of this distribution is

$$\mathbb{P}(I_1 > x) \sim \lambda c_d x^{-\delta} \qquad x \to \infty.$$

If $\delta < 1$, $\int \mathbb{P}(I_1 > x)\mathrm{d}x$ diverges and thus the mean does not exist. There is a heavy tail. Generally, $\mathbb{E}(I_1^p)$ exists for $p < \delta$.

Analogously, we have for the pdf

$$f_{I_1}(x) \sim \lambda c_d \delta x^{-\delta-1}, \quad x \to \infty.$$

Now, let I_n denote the interference from the n-th nearest interferer. What is the distribution of I_n for general n?

The ccdf of the distance to the n-th nearest neighbor R_n is

$$\mathbb{P}(R_n > r) = \frac{\Gamma_{\mathrm{ic}}(n, \lambda c_d r^d)}{\Gamma(n)}.$$

So for $n = 2$,

$$\mathbb{P}(I_2 < x) = \exp(-\lambda c_d x^{-\delta})(1 + \lambda c_d x^{-\delta})$$

and

$$\mathbb{P}(I_2 > x) \sim \frac{1}{2}(\lambda c_d)^2 x^{-2\delta}\,.$$

So we need $2\delta > 1$ for $\mathbb{E}I_2$ to exist.
For general n:

$$\mathbb{P}(I_n < x) = \exp(-\lambda c_d x^{-\delta}) \sum_{i=0}^{n-1} \frac{\left(\lambda c_d x^{-\delta}\right)^i}{i!}$$

For the tail probability we need to sum from n to ∞, so the dominant term will be the one for $i = n$ as $x \to \infty$. Therefore

$$\mathbb{P}(I_n > x) \sim \frac{1}{n!}(\lambda c_d)^n x^{-n\delta}\,.$$

This means that $\mathbb{E}(I_n^p)$ exists for $p < n\delta$.

So for $d = 2$ (two dimensions), we need to cancel $k > \alpha$ interferers to have a finite second moment.

Although we can find the distribution of I_n for all n, it is difficult to obtain the distribution of the total interference this way, since the I_n are neither independent nor identically distributed. We proceed differently.

3.4.4 Interference distribution without fading

In this subsection we focus on the case of two-dimensional networks and assume there is no fading, *i.e.*, $I = \sum_{x\in\Phi} g(x)$. Since $g(x)$ is assumed isotropic, we use $\ell\colon \mathbb{R}^+ \mapsto \mathbb{R}^+$, defined by $\ell(\|x\|) \equiv g(x)$. It is assumed that $\ell(x)$ strictly monotonically decreasing (invertible), and that $\lim_{x\to\infty} \ell(x) = 0$.

Our goal is to find the characteristic function of the interference and from there, if possible, the distribution.

We follow a basic yet powerful technique as it was used, for example, in [3]. It consists of two steps:

1. Consider first a finite network, say on a disk of radius a centered at the origin, and condition on having a fixed number of nodes in this finite area. The nodes' locations are then iid.

2. Then de-condition on the (Poisson) number of nodes and let the disk radius go to infinity.

Step 1. Consider the interference from the nodes located within distance a of the origin:

$$I_a = \sum_{x\in\Phi\cap b(o,a)} \ell(\|x\|) \tag{3.4.1}$$

In the limit $a \to \infty$, $I_a \to I$. Let $\mathcal{F}_{I_a}$ be the characteristic function (Fourier transform) of I_a, *i.e.*,

$$\mathcal{F}_{I_a}(\omega) \triangleq \mathbb{E}(e^{j\omega I_a})\,. \tag{3.4.2}$$

Conditioning on having k nodes in the disk of radius a,

$$\mathcal{F}_{I_a}(\omega) = \mathbb{E}\big(\mathbb{E}(e^{j\omega I_a} \mid \Phi(b(o,a)) = k)\big)\,. \tag{3.4.3}$$

Given that there are k points in $b(o,a)$, these points are iid uniformly distributed on the disk with radial density

$$f_R(r) = \begin{cases} \frac{2r}{a^2} & \text{if } 0 \leqslant r \leqslant a \\ 0 & \text{otherwise}\,, \end{cases} \tag{3.4.4}$$

and the characteristic function is the product of the k individual characteristic functions:

$$\mathbb{E}(e^{j\omega I_a} \mid \Phi(b(o,a)) = k) = \left(\int_0^a \frac{2r}{a^2} \exp(j\omega \ell(r)) \mathrm{d}r \right)^k \tag{3.4.5}$$

Step 2. The probability of finding k nodes in $b(o,a)$ is given by the Poisson distribution, hence:

$$\mathcal{F}_{I_a}(\omega) = \sum_{k=0}^{\infty} \frac{\exp(-\lambda\pi a^2)(\lambda\pi a^2)^k}{k!} \mathbb{E}\big(e^{j\omega I_a} \mid \Phi(b(o,a)) = k\big) \tag{3.4.6}$$

Inserting (3.4.5), summing over k, and interpreting the sum as the Taylor expansion of the exponential function, we obtain

$$\mathcal{F}_{I_a}(\omega) = \exp\left(\lambda\pi a^2 \left(-1 + \int_0^a \frac{2r}{a^2} \exp(j\omega\ell(r)) \mathrm{d}r \right) \right) . \tag{3.4.7}$$

Integration by parts, substituting $r \to \ell^{-1}(x)$, where ℓ^{-1} is the inverse of ℓ, and letting $a \to \infty$ yields

$$\lim_{a\to\infty} a^2 \left(-1 + \int_0^a \frac{2r}{a^2} \exp(j\omega\ell(r)) \mathrm{d}r \right) = \int_0^\infty (\ell^{-1}(x))^2 j\omega e^{j\omega x} \mathrm{d}x\,,$$

so that

$$\mathcal{F}_I(\omega) = \exp\left(j\lambda\pi\omega \int_0^\infty (\ell^{-1}(x))^2 e^{j\omega x} \mathrm{d}x \right) . \tag{3.4.8}$$

To get more concrete results, we need to specify the path loss law. For the standard power law $\ell(r) = r^{-\alpha}$, we obtain

$$\mathcal{F}_I(\omega) = \exp\left(j\lambda\pi\omega \int_0^\infty x^{-2/\alpha} e^{j\omega x} \mathrm{d}x \right) . \tag{3.4.9}$$

For $\alpha \leqslant 2$, the integral diverges, indicating that the interference is infinite almost surely. For $\alpha > 2$,

$$\mathcal{F}_I(\omega) = \exp\left(-\lambda\pi\Gamma(1-2/\alpha)\omega^{2/\alpha} e^{-j\pi/\alpha} \right) , \quad \omega \geqslant 0. \tag{3.4.10}$$

The values for negative ω are determined by the symmetry condition $\mathcal{F}_I^*(-\omega) = \mathcal{F}_I(\omega)$. For $\alpha = 4$,

$$\mathcal{F}_I(\omega) = \exp\left(-\lambda\pi^{3/2} \exp(-j\pi/4)\sqrt{\omega} \right) . \tag{3.4.11}$$

This case is of particular interest, since it is the only one where a closed-form expression for the density exists:

$$f_I(x) = \frac{\pi\lambda}{2x^{3/2}} \exp\left(-\frac{\pi^3\lambda^2}{4x} \right) . \tag{3.4.12}$$

This is the so-called Lévy distribution, which can also be viewed as an inverse gamma distribution, or as the inverse Gaussian distribution with infinite mean. For other values of α, the densities may be expressed

in an infinite series [3, Eqn. (22)].

The characteristic function (3.4.10) indicates that the interference distribution is a *stable distribution* with characteristic exponent $2/\alpha < 1$, drift 0, skew parameter $\beta = 1$, and dispersion $\lambda\pi\Gamma(1-2/\alpha)\cos(\pi/\alpha)$. Details on stable distributions are available in [4]. The corresponding Laplace transform is

$$\mathcal{L}_I(s) = \exp(-\lambda\pi\Gamma(1-2/\alpha)s^{2/\alpha})\,. \tag{3.4.13}$$

Stable distributions with characteristic exponents less than one do not have any finite moments. In particular, the mean interference diverges, which is due to the singularity of the path loss law at the origin. This also follows immediately from the fact that $\mathbb{E}(I) = -\frac{\mathrm{d}}{\mathrm{d}s}\log(\mathcal{L}_I(s))|_{s=0} = \lim_{s\to 0} c\, s^{2/\alpha-1} = \infty$.

The method of conditioning on a fixed number of nodes, using the iid property of the node locations, and de-conditioning with respect to the Poisson distribution is applicable to many other problems.

3.4.5 Interference distribution with fading

Here we consider the sum

$$I = \sum_{x\in\Phi} h_x g(x)\,,$$

where the h_x are iid. In fact, the random variables h_x can be viewed as the marks in a marked point process.

$$\mathcal{L}(s) = \mathbb{E}[e^{-sI}] = \mathbb{E}\left(\prod_{x\in\Phi} e^{-shg(x)}\right)$$

be the Laplace transform. Since the fading is iid,

$$\mathcal{L}(s) = \mathbb{E}_\Phi\left(\prod_{x\in\Phi} \mathbb{E}_h(e^{-shg(x)})\right)\,.$$

Mapping the PPP to one dimension, we know that $\lambda(r) = \lambda c_d d r^{d-1}$. Again let $\ell(\|x\| \equiv g(x)$. Now we use the pgfl for $v = \mathbb{E}_h(e^{-shg(x)})$ to obtain for the one-dimensional PPP

$$\mathcal{L}(s) = \exp\left\{-\int_0^\infty \mathbb{E}_h[1-e^{-sh\ell(r)}]\lambda(r)\mathrm{d}r\right\}.$$

Conditioned on h, we have

$$\begin{aligned}
\int_0^\infty \big(1-\exp(-sh\ell(r)\big)\lambda(r)\mathrm{d}r &= \lambda c_d \int_0^\infty \big(1-\exp(-shr^{-\alpha})\big) d r^{d-1}\mathrm{d}r \\
&= \lambda c_d \int_0^\infty \big(1-\exp(-sh/y)\big)\delta y^{\delta-1}\mathrm{d}y \qquad (\text{subst. } y \leftarrow r^{-1/\alpha}) \\
&= \lambda c_d \int_0^\infty \big(1-\exp(-shx)\big)\delta x^{-\delta-1}\mathrm{d}x \qquad (\text{subst. } x \leftarrow y^{-1}) \\
&= \lambda c_d \int_0^\infty x^{-\delta} sh\exp(-shx)\mathrm{d}x \qquad (\text{integration by parts}) \\
&= \lambda c_d (hs)^\delta \Gamma(1-\delta)\,, \quad 0<\delta<1\,.
\end{aligned}$$

With the expectation over h, we obtain

$$\mathcal{L}(s) = \exp\Big(- \lambda c_d \mathbb{E}[h^\delta]\Gamma(1-\delta)s^\delta \Big) .$$

So the interference has a stable distribution with characteristic exponent δ and dispersion $\lambda c_d \mathbb{E}[h^\delta]\Gamma(1-\delta)$.

In the Rayleigh fading case (exponential h), $\mathbb{E}[h^\delta] = \Gamma(1+\delta)$, and

$$\mathcal{L}(s) = \exp\Big(-\lambda c_d \Gamma(1+\delta)\Gamma(1-\delta)s^\delta\Big) = \exp\left(-\lambda c_d s^\delta \frac{\pi\delta}{\sin(\pi\delta)}\right) .$$

As $\delta \to 1$, $\sin(\pi\delta) \approx \pi(1-\delta)$, so at the limit we have

$$\mathcal{L}(s) \approx \exp\left(-\lambda c_d s^\delta \frac{\delta}{1-\delta}\right) .$$

As $\delta \to 1$, $\mathcal{L}(s) \equiv 0$, so $I = \infty$ a.s.

3.4.6 Outage in Poisson networks with Rayleigh fading

With Rayleigh fading, the desired signal strength at the receiver S is exponential with mean $r^{-\alpha}$. Let $r = 1$ first. What is the success probability $\mathbb{P}(\mathrm{SIR} > \theta)$ in a Poisson field of interferers of intensity λ that are subject to Rayleigh fading?

This is exactly the Laplace transform:

$$p_s \triangleq \mathbb{P}(\mathrm{SIR} \geqslant \theta) = \mathbb{P}(S > I\theta) = \mathbb{E}_I(\exp(-\theta I)) = \exp(-c_d\lambda\theta^\delta\Gamma(1+\delta)\Gamma(1-\delta))$$

If the desired link has distance r,

$$p_s(r) = \mathbb{E}_I(\exp(-\theta r^\alpha I)) = \exp(-c_d\lambda\theta^\delta r^d\Gamma(1+\delta)\Gamma(1-\delta)) .$$

So, in two dimensions, the success probability decays only with r^2, although the path loss is r^α.

The success probability is the cdf of the SIR— so we have a closed-form expression for the SIR distribution in Poisson networks with Rayleigh fading! The pdf is

$$f_{\mathrm{SIR}}(x) = c\delta x^{\delta-1}e^{-cx^\delta} ,$$

where $c = c_d\lambda r^d\Gamma(1+\delta)\Gamma(1-\delta))$. This is a Weibull distribution. Its mean is

$$\mathbb{E}(\mathrm{SIR}) = \frac{\Gamma(1/\delta)}{\delta c^{1/\delta}} = \frac{1}{r^\alpha}\frac{\Gamma(1+1/\delta)}{(\lambda C(\alpha))^{1/\delta}} , \quad C(\alpha) = c_d\Gamma(1+\delta)\Gamma(1-\delta) .$$

While the success probability is a function of r^d, the mean SIR is proportional to $r^{-\alpha}$, as expected, since the received signal power S decays with $r^{-\alpha}$, while the interference does not depend on r.

The outage expression is valid also when the interferers are not subject to Rayleigh fading. As long as the desired signal is exponentially distributed, we have

$$p_s = \mathbb{E}(\exp(-\theta I)) = \exp(-c_d\lambda\theta^\delta r^d\mathbb{E}(h^\delta)\Gamma(1-\delta)) ,$$

where the interferers' fading distribution F_h can be arbitrary (as long as it has a finite δ-th moment).

3.5 Stable Distributions

3.5.1 Definition

Definition 3.2 (Stable distribution) *A RV X is said to have a* stable *distribution if for all $a, b > 0$, there exists $c > 0$, d, such that*

$$aX_1 + bX_2 \stackrel{d}{=} cX + d$$

where X_1, X_2 are iid with the same distribution as X.

Theorem 3.6 *For any stable RV X, there is $0 < \delta \leqslant 2$ such that the number c in the definition satisfies $c^\delta = a^\delta + b^\delta$.*

δ is the characteristic exponent. For Gaussian RVs, $\delta = 2$ since

$$aX_1 + bX_2 \sim \mathcal{N}((a+b)\mu, (a^2+b^2)\sigma^2)$$

i.e., the definition holds with $c = \sqrt{a^2+b^2}$ and $d = (a+b-c)\mu$.

The Laplace transform of a stable RV is

$$\mathbb{E}(e^{-sX}) = \exp\left(-\frac{\sigma^\delta}{\cos\frac{\pi\delta}{2}} s^\delta\right) \qquad \text{if } \delta < 1\,.$$

Sometimes $\kappa \triangleq \frac{\sigma^\delta}{\cos\frac{\pi\delta}{2}}$ is called the *dispersion*.

3.5.2 LePage Series representation

Theorem 3.7 (Series representation) *Let τ_i be the arrival times of a PPP of intensity λ and let h_i be iid RVs, independent of τ_i. If the infinite sum*

$$\sum_{i=1}^{\infty} \tau_i^{-1/\delta} h_i$$

converges a.s., then it converges to a stable RV.

This is exactly what we need, since in our case, the distances r_i^d constitute a homogeneous PPP of intensity λc_d. So our sum is

$$\sum_{i=1}^{\infty} \left(r_i^d\right)^{-\alpha/d},$$

and $\delta = d/\alpha$ as expected.

3.5.3 Shot noise

The sum

$$I(t) = \sum_{\tau\in\Phi} g(t-\tau)$$

for Φ a PPP and g an impulse response is called *Poisson shot noise*. While shot noise was studied since 1909 for one-dimensional processes, it is easily generalized to d dimensions. If $g(x) = \|x\|^{-a}$, the shot noise is more specifically called *power-law Poisson shot noise* (with exponent a). So interference $I(x)$ in a Poisson field is a power-law Poisson shot noise process.

Chapter 4

Moment Measures of Point Processes

4.1 Introduction

Point process theory provides useful analogues to the mean, variance, and higher-order statistics of numerical random variables. Since point process are random elements in the space of locally finite counting measures, the moments will be replaced by *moment measures*. In this chapter, we will define such moment measures and study applications.

4.2 The First-Order Moment Measure

The first moment of a point process is its intensity measure $\Lambda(B) = \mathbb{E}\Phi(B)$ that we are already familiar with. It corresponds to the mean of a numerical random variable.

Example 4.1 (Stationary lattice) *Let*

$$\Phi = \{m, n \in \mathbb{Z} \,:\, (U_1 + ms, U_2 + ns)\}\,, \quad U_1, U_2 \sim \mathcal{U}[0, s)\,.$$

We find

$$\Lambda(B) = \mathbb{E}\Phi(B) = \frac{1}{s^2}|B|\,.$$

If Φ is stationary on $\mathbb{R}^d$, then $\Lambda(B) \equiv \Lambda(B + v)$ for all $v \in \mathbb{R}^d$, indicating that the intensity measure is invariant under translations. But only multiples of the Lebesgue measure have this property, so:

Theorem 4.1 *If ν is a translation-invariant measure on $\mathbb{R}^d$, then $\nu(B) = c|B|$ for some $c > 0$.*

Corollary 4.2 *If Φ is a stationary PP on $\mathbb{R}^d$, then its intensity measure Λ is a constant multiple of the Lebesgue measure. The constant is called the* intensity *of Φ.*

If the intensity measure Λ satisfies

$$\Lambda(B) = \int_B \lambda(x)\mathrm{d}x$$

for some $\lambda(x)$, we call λ the *intensity function*. It has the following interpretation: For some small region $\mathrm{d}x \subset \mathbb{R}^d$,

$$\mathbb{P}(\Phi(\mathrm{d}x) > 0) \sim \mathbb{E}\Phi(\mathrm{d}x) = \Lambda(\mathrm{d}x) = \lambda(x)\mathrm{d}x\,.$$

4.2.1 Constant density vs. stationarity

We know that the density is constant for stationary point processes, *i.e.*, $\mathbb{E}\Phi(B) = c|B|$. Does the converse hold, *i.e.*, does $\mathbb{E}\Phi(B) = c|B|$ imply stationarity?

No, we can construct a counterexample. A Poisson cluster process with parent intensity $\lambda(r) = \min\{1, r^{-1}\}$ and average number of nodes per cluster $\mu(r) = \max\{1, r\}$. Another counterexample is a mixed PPP/BPP, with a BPP with one node on $[0,1]^2$ and a PPP on $\mathbb{R}^2 \setminus [0,1]^2$ of intensity 1.

4.3 Second Moment Measures

(See also the Baddeley handout, p. 32ff.)
The variance of the point counts $\Phi(B)$ is

$$\operatorname{var}\Phi(B) = \mathbb{E}(\Phi(B)^2) - (\mathbb{E}\Phi(B))^2$$

and the covariance is

$$\operatorname{cov}(\Phi(A), \Phi(B)) = \mathbb{E}(\Phi(A)\Phi(B)) - \mathbb{E}\Phi(A)\mathbb{E}\Phi(B)\,.$$

Definition 4.1 (Second moment measure) *Let Φ be a PP on $\mathbb{R}^d$. Then $\Phi^{(2)} \triangleq \Phi \times \Phi$ is a point process on $\mathbb{R}^d \times \mathbb{R}^d$ consisting of all ordered pairs (x, x') of points $x, x' \in \Phi$. The intensity measure $\mu^{(2)}$ of $\Phi \times \Phi$ is a measure on $\mathbb{R}^{2d}$ defined as*

$$\mu^{(2)}(A \times B) = \mathbb{E}(\Phi(A)\Phi(B))\,.$$

If $A = B$, then $\mu^{(2)}(A^2) = \mathbb{E}(\Phi(A)^2)$, so this contains all information about variances and covariances.

So we can write

$$\operatorname{cov}(\Phi(A), \Phi(B)) = \mu^{(2)}(A \times B) - \Lambda(A)\Lambda(B)$$

and

$$\operatorname{var}(\Phi(A)) = \mu^{(2)}(A^2) - (\Lambda(A))^2\,.$$

Campbell's formula for the mean also applies to $\Phi \times \Phi$, hence

$$\mathbb{E}\left(\sum_{x\in\Phi}\sum_{y\in\Phi} f(x,y)\right) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x,y)\mu^{(2)}(\mathrm{d}x, \mathrm{d}y)\,.$$

Example 4.2 (Uniform PPP) *For the uniform PPP of intensity λ in $\mathbb{R}^d$, the second moment measure satisfies*

$$\mu^{(2)}(A \times B) = \lambda^2|A||B| + \lambda|A \cap B|\,.$$

To see this, let's write A, B as

$$A = (A \cap B) \cup (A \setminus B)\,; \qquad B = (A \cap B) \cup (B \setminus A)\,.$$

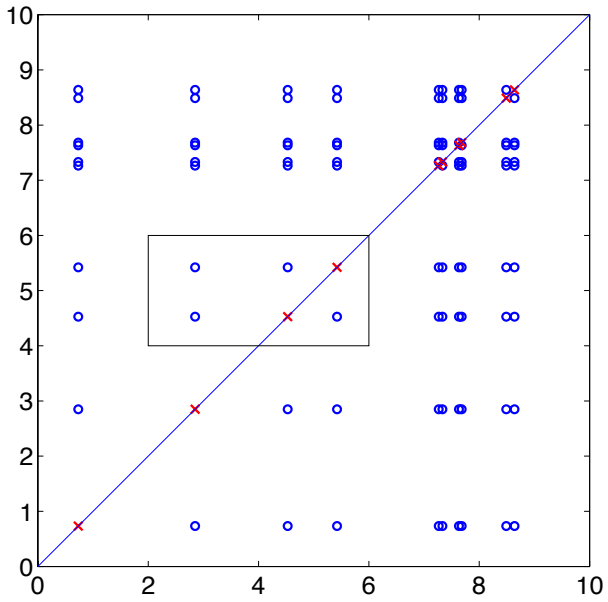

Figure 4.1: Illustration for second moment measure. Here $\Phi \subset \mathbb{R}^+$ is a PPP of intensity 1 on $\mathbb{R}^+$. Shown is the product PP $\Phi^{(2)} = \Phi \times \Phi$ on $(\mathbb{R}^+)^2$. The points on the diagonal line are the points (x, x) for $x \in \Phi$. The box is the rectangle $C = A \times B$ with $A = [2, 6]$ and $B = [4, 6]$. For this realization, $\Phi^{(2)}(C) = 6$. If $A \cap B = \emptyset$, then no points on the diagonal lie in C. The expected number of points in C is $\mu^{(2)}(C) = 4 \cdot 2 + 2 = 10$.

Then

$$\begin{aligned}
\mu^{(2)}(A \times B) &= \mathbb{E}(\Phi(A)\Phi(B)) \\
&= \mathbb{E}\left[(\Phi(A \cap B) + \Phi(A \setminus B)) \cdot (\Phi(A \cap B) + \Phi(B \setminus A))\right] \\
&= \mathbb{E}(\Phi(A \setminus B))\mathbb{E}(\Phi(B \setminus A)) + \mathbb{E}(\Phi(A \cap B))\mathbb{E}(\Phi(A \setminus B)) \\
&\quad + \mathbb{E}(\Phi(A \cap B)\mathbb{E}(\Phi(B \setminus A)) + \mathbb{E}((\Phi(A \cap B))^2) \\
&= \mathbb{E}(\Phi(A))\mathbb{E}(\Phi(B)) + \underbrace{\mathbb{E}((\Phi(A \cap B))^2) - (\mathbb{E}(\Phi(A \cap B)))^2}_{\operatorname{var} \Phi(A \cap B)} \\
&= \Lambda(A)\Lambda(B) + \Lambda(A \cap B) \\
&= \lambda^2 |A||B| + \lambda |A \cap B| \,.
\end{aligned}$$

So there is a constant density λ^2 on all of $\mathbb{R}^d \times \mathbb{R}^d$, plus a positive mass on the diagonal $\{(x, x) \colon x \in \mathbb{R}^d\}$.

The fact that points inside $A \cap B$ are of special importance can also be explained as follows: Consider what happens if a point is added. If the additional point falls in $A \setminus B$ or $B \setminus A$, the change is only linear in $|A|$ or $|B|$, respectively. If it falls into $A \cap B$, however, the change is quadratic.

Fig. 4.1 shows a realization of a PP $\Phi^{(2)}$, where Φ is a PPP on $\mathbb{R}^+$.

If $A = B$: $\mu^{(2)}(A^2) = \mathbb{E}(\Phi^2(A)) \neq \mathbb{E}^2(\Phi(A))$. The difference is the mass on the diagonal. But the difference is exactly the variance. So

$$\operatorname{var}(\Phi(A)) = \lambda |A| \,.$$

We already know that—$\Phi(A)$ is Poisson after all.

In many instances, it is needed or makes sense to remove the mass on the diagonal. This leads to the *second factorial moment measure*.

Definition 4.2 (Second factorial moment measure) *The second factorial moment measure is the intensity measure of the process $\Phi \star \Phi$ of all ordered pairs of* distinct *points of Φ:*

$$\alpha^{(2)}(A \times B) = \mathbb{E}(\Phi(A)\Phi(B)) - \mathbb{E}(\Phi(A \cap B))\,.$$

The set of points $\Phi \star \Phi$ can be written as

$$\Phi \star \Phi \triangleq \{(x, x') \in \Phi \times \Phi \colon x \neq x'\}\,.$$

Again from Campbell, we know that the second factorial moment measure satisfies

$$\mathbb{E}\left(\sum_{x,y\in\Phi}^{\neq} f(x,y)\right) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x,y)\alpha^{(2)}(\mathrm{d}x, \mathrm{d}y)\,.$$

The notation $\sum^{\neq}$ indicates a multi-sum over a set, where none of the elements of the set may be taken more than once. For example,

$$\sum_{m,n\in[5]} 1 = 25\,, \qquad \text{whereas} \qquad \sum_{m,n\in[5]}^{\neq} 1 = 20\,.$$

If A and B are disjoint, then $\mu^{(2)}(A \times B) = \alpha^{(2)}(A \times B)$. Generally we have

$$\mu^{(2)}(A \times B) = \alpha^{(2)}(A \times B) + \Lambda(A \cap B)\,.$$

The circles in Fig. 4.1 show the points of the product $\Phi \star \Phi$.

We can express it as follows:

$$\begin{aligned}\alpha^{(2)}(A \times B) &= \mathbb{E}(\#\{(x,y)\colon x \in \Phi \cap A,\ y \in \Phi \cap B,\ x \neq y\}) \\ &= \mathbb{E}\left(\sum_{x,y\in\Phi}^{\neq} \mathbf{1}_A(x)\mathbf{1}_B(y)\right)\,.\end{aligned}$$

The difference between $\mu^{(2)}(A \times B)$ and $\alpha^{(2)}(A \times B)$ lies in the expectation of the sum

$$\sum_{x,y\in\Phi\colon\, x=y} \mathbf{1}_A(x)\mathbf{1}_B(y) = \sum_{x\in\Phi} \mathbf{1}_{A\cap B}(x)\,.$$

The name "factorial" comes from the fact that

$$\begin{aligned}\alpha^{(2)}(A \times A) &= \mathbb{E}(\Phi(A)^2) - \mathbb{E}(\Phi(A)) \\ &= \mathbb{E}\big(\Phi(A)(\Phi(A) - 1))\big)\,.\end{aligned}$$

Also:

$$\operatorname{var}\Phi(A) = \alpha^{(2)}(A \times A) + \Lambda(A) - (\Lambda(A))^2\,.$$

Example 4.3 (PPP) *For a Poisson point process,*

$$\alpha^{(2)}(A \times B) = \Lambda(A)\Lambda(B)\,.$$

In the uniform case,

$$\alpha^{(2)}(A \times B) = \lambda^2 |A||B|\,.$$

Usually there exists a density (with respect to the Lebesgue measure) of the second factorial moment measure, the *second moment density*, discussed in the next section. In contrast, the second moment measure $\mu^{(2)}$ does not have a density. This is another principal motivation for using $\varrho^{(2)}$.

4.4 Second Moment Density

Definition 4.3 (Second moment density (or second-order product density)) *A point process Φ on $\mathbb{R}^d$ is said to have* second moment density $\varrho^{(2)}$ *if*

$$\alpha^{(2)}(C) = \alpha^{(2)}(A \times B) = \int_A \int_B \varrho^{(2)}(x,y)\mathrm{d}y\mathrm{d}x = \int_C \varrho^{(2)}(x,y)\mathrm{d}x\mathrm{d}y$$

for any compact $C = A \times B \subset \mathbb{R}^d \times \mathbb{R}^d$.

In differential form, $\alpha^{(2)}(\mathrm{d}x, \mathrm{d}y) = \varrho^{(2)}(x,y)\mathrm{d}x\mathrm{d}y$. Informally, $\varrho^{(2)}(x,y)$ is the joint probability that there are two points of Φ at two specified locations x and y:

$$\mathbb{P}(\Phi(\mathrm{d}x) > 0, \Phi(\mathrm{d}y) > 0) \sim \mathbb{E}(\Phi(\mathrm{d}x)\Phi(\mathrm{d}y)) = \varrho^{(2)}(x,y)\mathrm{d}x\mathrm{d}y\,.$$

Another interpretation is the following: If C_1 and C_2 are disjoint spheres with centers x_1 and x_2 and infinitesimal volumes $\mathrm{d}V_1$ and $\mathrm{d}V_2$. Then $\varrho^{(2)}(x_1,x_2)\mathrm{d}V_1\mathrm{d}V_2$ is the probability that there is a point of Φ in both C_1 and C_2.

Example 4.4 (PPP) *The uniform PPP of intensity λ has $\varrho^{(2)}(x,y) = \lambda^2$. In the non-uniform case, $\varrho^{(2)}(x,y) = \lambda(x)\lambda(y)$.*

If Φ is stationary then $\varrho^{(1)} = \lambda$ and $\varrho^{(2)}$ depends only on the difference of its arguments, *i.e.*, there exists a $\varrho^{(2)}_{\mathrm{st}} \colon \mathbb{R}^d \to \mathbb{R}^+$ such that

$$\varrho^{(2)}(x,y) \equiv \varrho^{(2)}_{\mathrm{st}}(x-y) \quad \forall x, y \in \mathbb{R}^d\,.$$

If Φ is motion-invariant then $\varrho^{(2)}_{\mathrm{st}}$ depends only on the distance $r = \|x-y\|$, *i.e.*, there exists a $\varrho^{(2)}_{\mathrm{mi}} \colon \mathbb{R}^+ \to \mathbb{R}^+$ such that

$$\varrho^{(2)}(x,y) \equiv \varrho^{(2)}_{\mathrm{st}}(x-y) \equiv \varrho^{(2)}_{\mathrm{mi}}(\|x-y\|) = \varrho^{(2)}_{\mathrm{mi}}(r)\,, \quad \forall x, y \in \mathbb{R}^d\,.$$

In these cases, $\varrho^{(2)}_{\mathrm{st}}$ and $\varrho^{(2)}_{\mathrm{mi}}$ are also called second moment density or second-order product density. For example, in the uniform Poisson case, $\varrho^{(2)}_{\mathrm{st}}(x) = \varrho^{(2)}_{\mathrm{mi}}(\|x\|) = \lambda^2$.

Example 4.5 (Binomial point process) *The second moment density for a (uniform) binomial point process with n points in W is*

$$\varrho^{(2)}(x,y) = \frac{n(n-1)}{|W|^2}\,.$$

Calculation. *Let $N_A = \Phi(A\backslash B)$, $N_B = \Phi(B\backslash A)$, $N_C = \Phi(A\cap B)$; $p_A = |A\backslash B|/|W|$, $p_B = |B\backslash A|/|W|$, $p_C = |A\cap B|/|W|$. Note that N_A, N_B, N_C follow a multinomial distribution with probabilities p_A, p_B, and p_C, respectively.*

$$\begin{aligned}\alpha^{(2)}(A\times B) &= \mathbb{E}\left[(N_A+N_C)(N_B+N_C)\right] - \mathbb{E}N_C \\ &= \mathbb{E}(N_C^2) + \mathbb{E}(N_A N_C) + \mathbb{E}(N_B N_C) + \mathbb{E}(N_A N_B) - \mathbb{E}N_C \\ &= n(n-1)\left[p_C^2 + p_A p_C + p_B p_C + p_A p_B\right] \\ &= n(n-1)\frac{|A||B|}{|W|^2}\,.\end{aligned}$$

The mean and variance of N_C are np_C and $np_C(1-p_C)$, respectively, and $\operatorname{cov}(N_A, N_B) = -np_Ap_B$ so that $\mathbb{E}(N_C^2) - \mathbb{E}N_C = n(n-1)p_C^2$ and $\mathbb{E}(N_AN_B) = \operatorname{cov}(N_A,N_B) + \mathbb{E}(N_A)\mathbb{E}(N_B) = n(n-1)p_Ap_B$.

Definition 4.4 (Pair correlation function) *For a PP $\Phi \subset \mathbb{R}^d$ with intensity function $\lambda(x)$ and second moment density $\varrho^{(2)}(x,y)$, the pair correlation function is defined as*

$$g(x,y) \triangleq \frac{\varrho^{(2)}(x,y)}{\lambda(x)\lambda(y)}\,.$$

For motion-invariant processes, it is

$$g_{\mathrm{mi}}(r) \triangleq \frac{\varrho^{(2)}_{\mathrm{mi}}(r)}{\lambda^2}\,.$$

The pair correlation function is identical to 1 in the uniform Poisson case, and it is smaller than 1 (for small r) if there is repulsion and larger than 1 if there is clustering.
For a BPP of n points on W, it is (from Example 4.5)

$$g(x,y) = \frac{n(n-1)}{|W|^2}\frac{|W|^2}{n^2} = 1 - \frac{1}{n}\,.$$

The pair correlation function measures the degree of correlation between the RVs $\Phi(A)$ and $\Phi(B)$ in a non-centered way. If $g(x,y)\equiv 1$, then $\operatorname{cov}(\Phi(A),\Phi(B)) = 0$ for disjoint A, B.

Example 4.6 (PP with fixed finite number of points) *Let $\Phi \subset \mathbb{R}^d$ consist of n random points $\{x_1,\ldots,x_n\}$, where x_i has marginal probability density $f_i(u)$, $u\in\mathbb{R}^d$. Then Φ has intensity*

$$\lambda(x) = \sum_{i=1}^{n} f_i(x)\,.$$

Let $f_{ij}(u,v)$ be the marginal joint *density of (x_i,x_j). Then Φ has second moment density*

$$\varrho^{(2)}(x,y) = \sum_{i,j\in[n]}^{\neq} f_{ij}(x,y) \tag{4.4.1}$$

and pair correlation function

$$g(x,y) = \frac{\sum_{i,j\in[n]}^{\neq} f_{ij}(x,y)}{\left(\sum_{i\in[n]} f_i(x)\right)\left(\sum_{j\in[n]} f_j(y)\right)}\,.$$

For the simplest case $n = 2$, the joint probability having one point in A and the other in B is

$$\alpha^{(2)}(A \times B) = \int_A \int_B (f_{12}(x, y) + f_{21}(x, y))\mathrm{d}x\mathrm{d}y\,,$$

consistent with (4.4.1)*.*

The BPP is a special case with $f_i(x) = |W|^{-1}\mathbf{1}\{x \in W\}$ (if it is uniform) and $f_{ij}(x, y) = f_i(x)f_j(y) = |W|^{-2}\mathbf{1}\{x \in W\}$. Its intensity is $n/|W|$, and the second moment density follows from (4.4.1) *to be $n(n - 1)/|W|^2$, as expected.*

Example 4.7 (Poisson cluster process) *Consider a Poisson cluster process Φ, formed from a uniform PPPP (Poisson parent point process) Φ_p with intensity λ_p by replacing each $x \in \Phi_p$ by a random cluster Z_x, which is a finite point process. The clusters Z_x for different $x \in \Phi_p$ are independent processes. Let Z_x have intensity function $g(u \mid x)$ and second moment density $h(u, v \mid x)$. Conditioned on Φ_p, the cluster process has intensity*

$$\lambda(u \mid \Phi_p) = \sum_{x \in \Phi_p} g(u \mid x)\,.$$

The (unconditioned) intensity of the cluster process Φ is thus, by Campbell's formula,

$$\begin{aligned}\lambda(u) &= \mathbb{E}(g(u \mid \Phi_p)) \\ &= \mathbb{E}\sum_{x \in \Phi_p} g(u \mid x) \\ &= \lambda_p \int_{\mathbb{R}^d} g(u \mid x)\mathrm{d}x\,.\end{aligned}$$

If all clusters have the same intensity function $\lambda_c(x)$, i.e.*, $g(u \mid x) \equiv \lambda_c(u - x)$, then $\lambda = \lambda_p\mu$, where $\mu = \Lambda_c(\mathbb{R}^d)$ is the mean number of points per cluster. The second moment density is*

$$\varrho^{(2)}(u, v) = \lambda^2 + \lambda_p \int_{\mathbb{R}^d} h(u, v \mid x)\mathrm{d}x\,.$$

There are two contributions to this second moment density, the one from the overall process and the one from pairs of points in the same cluster. The first contribution is λ^2 since it arises from pairs of points from different clusters, and clusters are independent. The second one is

$$\mathbb{E}\left(\sum_{x \in \Phi_p} h(u, v \mid x)\right) = \lambda_p \int_{\mathbb{R}^d} h(u, v \mid x)\mathrm{d}x\,.$$

Sanity check: Assume each parent had a daughter point at exactly the location of the parent. Then $g(u \mid x) = \delta_x(u)$ and $h(u, v \mid x) = \delta_x(u)\delta_x(v) = 0$ since $u \neq v$. So $\lambda(u) = \lambda_p \int \delta_x(u)\mathrm{d}x = \lambda_p$, and we obtain the desired result.

In the Matérn cluster process*, the points in each cluster are distributed uniformly at random in a ball of radius R around their parent points, and each cluster contains a Poisson number of points with mean μ. Accordingly, its intensity function is*

$$\lambda_c(x) = \frac{\mu}{c_d R^d}\mathbf{1}_{b(o,R)}(x)\,.$$

In the two-dimensional case, the second moment density of a cluster Z_x is $h(u, v \mid x) = \mu^2/(\pi^2 R^4)$ if $u, v \in b(x, R)$ and 0 otherwise. We have

$$\begin{aligned}\int_{\mathbb{R}^d} \mathbf{1}\{u, v \in b(x, R)\}\mathrm{d}x &= \int_{\mathbb{R}^d} \mathbf{1}_{b(x,R)}(u)\mathbf{1}_{b(x,R)}(v)\mathrm{d}x \\ &= \int_{\mathbb{R}^d} \mathbf{1}_{b(u,R)}(x)\mathbf{1}_{b(v,R)}(x)\mathrm{d}x \\ &= |b(u, R) \cap b(v, R)|\,.\end{aligned}$$

Hence the second moment density of the Matérn cluster process is

$$\varrho^{(2)}(u, v) = \lambda_p^2\mu^2 + \lambda_p \frac{\mu^2}{\pi^2 R^4}|b(u, R) \cap b(v, R)|$$

or, since this cluster process is motion-invariant, with $r = \|u - v\|$,

$$\varrho^{(2)}_{\mathrm{mi}}(r) = \lambda_p^2\mu^2 + \lambda_p \frac{\mu^2}{\pi^2 R^4}|b(o, R) \cap b(r, R)|\,.$$

The pair correlation function follows as

$$g_{\mathrm{mi}}(r) = \frac{\varrho^{(2)}(r)}{\lambda_p^2\mu^2} = 1 + \frac{1}{\lambda_p}\frac{|b(o, R) \cap b(r, R)|}{|b(o, R)|^2}\,.$$

Interpretation: For $r \geqslant 2R$, this is just the pair correlation function of the PPP. Also, as $\lambda_p \to \infty$, this approaches the PPP's pair correlation function. For smaller r, $g(r) > 1$, so the PP exhibits clustering — of course.

In the (two-dimensional) Thomas cluster process, each cluster is a PPP with Gaussian intensity function

$$\lambda_c(x) = \frac{\mu}{2\pi\sigma^2}\exp\left(-\frac{\|x\|^2}{2\sigma^2}\right)\,, \quad x \in \mathbb{R}^2\,.$$

In this case, the pair correlation function is

$$g_{\mathrm{mi}}(r) = 1 + \frac{1}{4\pi\lambda_p\sigma^2}\exp\left(-\frac{r^2}{4\sigma^2}\right)\,.$$

Again we note that the pair correlation function approaches 1 from above as $\lambda_p \to \infty$.

Generally, for Poisson cluster processes with A Poisson number of nodes per cluster and symmetric probability density function $f = f(-x)$ for the points in each cluster,

$$\varrho^{(2)}_{\mathrm{st}}(x) = \lambda^2\left[1 + \frac{(f * f)(x)}{\lambda_p}\right]\,, \tag{4.4.2}$$

where $\lambda = \lambda_p\mu$ is the density of the cluster process. To show this, we focus on a single cluster and go back to the case of the finite point process. Given the number of nodes n, they are independently placed. Hence, given n, $\varrho^{(2)}(x, y \mid n) = n(n-1)f(x)f(y)$ and

$$\varrho^{(2)}(x, y) = \mu^2 f(x)f(y)\,.$$

Still considering the case where x and y belong to the same cluster, but now averaging over the parent process Φ_p yields

$$\begin{aligned}\varrho_{\mathrm{cl}}^{(2)}(u,v) &= \mathbb{E}\sum_{x\in\Phi_p}\mu^2 f(u-x)f(v-x)\\ &= \lambda_p\mu^2\int_{\mathbb{R}^d} f(u-x)f(v-x)\mathrm{d}x\\ &= \lambda_p\mu^2\int_{\mathbb{R}^d} f(z)f(v-u+z)\mathrm{d}z\\ &= \lambda_p\mu^2(f*f)(u-v)\,.\end{aligned}$$

To this we need to add the λ^2 for the case when x and y belong to different clusters to obtain (4.4.2).

The second moment density is invariant under permutation of its arguments.

4.5 Second Moments for Stationary Processes

For a PP $\Phi\subset\mathbb{R}^d$, stationarity implies

$$\mathbb{E}[\Phi(A+v)\Phi(B+v)] = \mathbb{E}[\Phi(A)\Phi(B)]$$

for all $v\in\mathbb{R}^d$. Thus $\mu^{(2)}$ and $\alpha^{(2)}$ are invariant under *simultaneous* shifts.
Now apply the transform $T(x,y)=(x,y-x)$ from $\mathbb{R}^d\times\mathbb{R}^d$ onto itself. Under this transformation, the simultaneous shift becomes a shift in the first coordinate, *i.e.*,

$$(s,t)\longrightarrow(s+v,t)\,.$$

The image of $\alpha^{(2)}$ under the transformation is a measure μ which is invariant under translations of the first coordinate. This process is illustrated in Fig. 4.2. Since translation-invariant measure are multiples of the Lebesgue measure, it follows that

$$\mu = \lambda\nu_d\otimes\mathcal{K}\,,$$

where λ is the intensity of the process, ν_d the d-dimensional Lebesgue measure, and $\mathcal{K}$ is a measure on $\mathbb{R}^d$ called the *reduced second moment measure*. So we have the product of two measures on $\mathbb{R}^d$ rather than one measure on $\mathbb{R}^d\times\mathbb{R}^d$ — there is a *disintegration* or *separation* in the second factorial moment measure: The second factorial moment measure is expressible as the product of a Lebesgue component $\Lambda(\mathrm{d}x)$ along the diagonal $x=y$ and a reduced measure along $u=y-x$. Using Campbell,

$$\begin{aligned}\mathbb{E}\left[\sum_{x,y\in\Phi}^{\neq} f(x,y)\right] &= \iint f(x,y)\alpha^{(2)}(\mathrm{d}x,\mathrm{d}y)\\ &= \iint f(x,x+u)\mu(\mathrm{d}x,\mathrm{d}u)\\ &= \lambda\iint f(x,x+u)\mathrm{d}x\,\mathcal{K}(\mathrm{d}u)\,.\end{aligned}$$

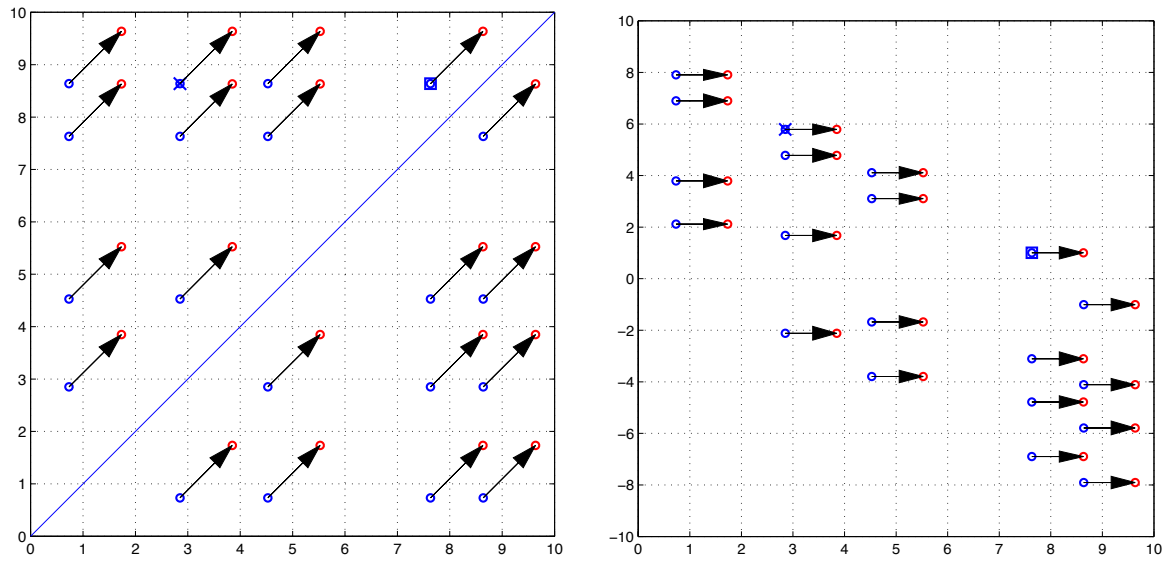

Figure 4.2: Illustration of coordinate transform $(x, y) \to (x, y - x)$. Left: Original product point process with translation $v = 1$. Right: Transformed product point process with same translation.

Theorem 4.3 (Reduced second moment measure) *Let Φ be a stationary PP on $\mathbb{R}^d$ with intensity λ. Then there is a measure $\mathcal{K}$ on $\mathbb{R}^d$ such that for general measurable f,*

$$\mathbb{E}\left[\sum_{x,y\in\Phi}^{\neq} f(x,y)\right] = \lambda \iint f(x, x+u)\mathrm{d}x\, \mathcal{K}(\mathrm{d}u)\,.$$

$\mathcal{K}$ is called the reduced second moment measure *of* Φ.

Choosing $f(x, y) = \mathbf{1}_A(x)\mathbf{1}_B(y - x)$,

$$\mathbb{E}\left[\sum_{x,y\in\Phi}^{\neq} \mathbf{1}_A(x)\mathbf{1}_B(y-x)\right] = \lambda \iint \mathbf{1}_A(x)\mathbf{1}_B(u)\mathrm{d}x\, \mathcal{K}(\mathrm{d}u) = \lambda|A|\mathcal{K}(B)\,. \tag{4.5.1}$$

Since $\lambda|A| = \mathbb{E}\Phi(A)$ and $\mathbf{1}_B(y - x) = \mathbf{1}_{B+x}(y)$,

$$\mathcal{K}(B) = \frac{\mathbb{E}\sum_{x\in\Phi\cap A} \Phi((B + x) \setminus \{x\})}{\mathbb{E}\Phi(A)}\,.$$

If the second moment density $\varrho^{(2)}(x, y)$ exists, it depends only on the difference $x - y$, and we can write

$$\mathcal{K}(B) = \frac{1}{\lambda}\int_B \varrho^{(2)}_{\mathrm{st}}(u)\mathrm{d}u\,. \tag{4.5.2}$$

This can be seen by expanding the LHS of (4.5.1) as follows:

$$\begin{aligned}\mathbb{E}\left[\sum_{x,y\in\Phi}^{\neq}\mathbf{1}_A(x)\mathbf{1}_B(y-x)\right] &= \iint \mathbf{1}_A(x)\mathbf{1}_B(y-x)\varrho^{(2)}(x,y)\mathrm{d}x\mathrm{d}y\\ &= \int_A\int_{\mathbb{R}^d}\mathbf{1}_B(y-x)\varrho^{(2)}(x,y)\mathrm{d}x\mathrm{d}y\\ &= \int_A\int_{\mathbb{R}^d}\mathbf{1}_B(u)\varrho^{(2)}_{\mathrm{st}}(u)\mathrm{d}u\mathrm{d}x\\ &= \int_A\int_B\varrho^{(2)}_{\mathrm{st}}(u)\mathrm{d}u\mathrm{d}x\\ &= |A|\int_B\varrho^{(2)}_{\mathrm{st}}(u)\mathrm{d}u\,,\end{aligned}$$

which, by (4.5.1), is equal to $\lambda|A|\mathcal{K}(B)$. Hence (4.5.2) follows.

In particular for motion-invariant processes, a simpler function is often useful, namely *Ripley's K function*.

Definition 4.5 (Ripley's K function) *Ripley's K function is defined as*

$$K(r) = \frac{1}{\lambda}\mathcal{K}(b(o,r)) \qquad r\geqslant 0\,.$$

So $\lambda K(r)$ is the mean number of points y of the process that satisfy $0<\|y-x\|\leqslant r$ for a given point x of the process.

Example 4.8 *For the uniform PPP on $\mathbb{R}^d$,*

$$K(r) = c_d r^d\,, \qquad r\geqslant 0\,.$$

For a stationary PP on $\mathbb{R}^d$ which has a second moment density,

$$K(r) = \frac{1}{\lambda^2}\int_{b(o,r)}\varrho^{(2)}_{\mathrm{st}}(x)\mathrm{d}x = \int_{b(o,r)} g_{\mathrm{st}}(x)\mathrm{d}x\,.$$

where g is the pair correlation function. If the process is on $\mathbb{R}^2$ and motion-invariant,

$$K(r) = 2\pi\int_0^r g(r)r\mathrm{d}r\,.$$

Lemma 4.4 *For a motion-invariant PP on $\mathbb{R}^2$, the pair correlation function is given by*

$$g_{\mathrm{mi}}(r) = \frac{1}{2\pi r}\frac{\mathrm{d}}{\mathrm{d}r}K(r)\,.$$

Lemma 4.5 (Invariance of K under thinning) *If Φ' is obtained by random (stationary) thinning from a stationary PP Φ, then the K functions of Φ and Φ' are identical.*

The definition of the reduced second moment measure varies slightly throughput the literature. Here we adopted the definition in [1]. Some authors scale their $\mathcal{K}(B)$ by λ^{-1} [5], such that $K(r)=\mathcal{K}(b(o,r))$, or λ [6].

A close relative of the K function is the L function.

Definition 4.6 (The L function)

$$L(r) \triangleq \left(\frac{K(r)}{c_d} \right)^{1/d}$$

The L function is sometimes preferred to the K function since it is simply $L(r) = r$ for the uniform PPP.

Chapter 5

Conditioning and Palm Theory

5.1 Introduction and Basic Concepts for Stationary Point Processes

The *Palm probability* in point process theory is the probability of an event given that the point process contains a point at some location. It formalizes the notion of a "typical" point of the process. Informally, the typical point results from a selection procedure in which every point has the same chance of being selected. This idea, while heuristically clear, needs to be made mathematically precise. For example, points chosen according to some sampling procedure, such as choosing the point closest to the origin are *not* typical, because they have been sampled in a specific way. Intuitively, the Palm distribution is the conditional point process distribution given that a point (the typical point) is located at a specific location.

In this section we discuss two heuristic approaches to Palm theory, a local and a global approach, applied to stationary point processes.

5.1.1 The local approach

We first introduce some new notation. For a point process property $Y \in \mathcal{N}$,

$$\begin{aligned}\mathbb{P}(\Phi \text{ has property } Y \parallel x) &\triangleq \mathbb{P}(\Phi \text{ has property } Y \mid x \in \Phi)\\ &= \mathbb{P}(\Phi \in Y \mid x \in \Phi)\\ &= \mathsf{P}(Y \mid x \in \Phi)\\ &= \mathsf{P}(Y \parallel x)\,.\end{aligned}$$

By stationarity

$$\mathbb{P}(\Phi \in Y \parallel x) = \mathbb{P}(\Phi_x \in Y \parallel o)\,,$$

where $\Phi_x \triangleq \{x_1 + x,\, x_2 + x, \ldots\}$. Note that $\mathbb{P}(\Phi_x \in Y \parallel o)$ is to be understood as $\mathbb{P}(\Phi_x \in Y \mid o \in \Phi)$, not $\mathbb{P}(\Phi_x \in Y \mid o \in \Phi_x)$. So for stationary PPs, conditioning may be restricted to $o \in \Phi$, without loss of generality.

The nearest-neighbor distance distribution, for example, may be expressed as

$$D(r) = \mathbb{P}(\Phi(b(o,r)) > 1 \parallel o) = 1 - \mathbb{P}(\Phi(b(o,r)) = 1 \parallel o)\,.$$

The conditioning event $o \in \Phi$ has probability zero. For the uniform PPP, the conditional distribution can be calculated by a limit procedure: The conditional probability

$$D_\epsilon(r) = 1 - \mathbb{P}(\Phi(b(o,r) \setminus b(o,\epsilon)) = 0 \mid \Phi(b(o,\epsilon)) = 1)$$

is well-defined for $0 < \epsilon < r$, since

$$\mathbb{P}(\Phi(b(o,\epsilon)) = 1) = \lambda c_d \epsilon^d \exp(-\lambda c_d \epsilon^d) > 0\,.$$

We have

$$\begin{aligned} D_\epsilon(r) &= 1 - \frac{\mathbb{P}(\Phi(b(o,r) \setminus b(o,\epsilon))) = 0)\mathbb{P}(\Phi(b(o,\epsilon)) = 1)}{\mathbb{P}(\Phi(b(o,\epsilon)) = 1)} \\ &= 1 - \mathbb{P}(\Phi(b(o,r) \setminus b(o,\epsilon)) = 0) \\ &= 1 - \exp(-\lambda c_d (r^d - \epsilon^d))\,. \end{aligned}$$

It seems reasonable to define $D(r) \triangleq \lim_{\epsilon \to 0} D_\epsilon(r)$, so

$$D(r) = 1 - \exp(-\lambda c_d r^d)\,, \qquad r \geqslant 0\,.$$

So the spherical contact distribution function (or empty space function) and nearest-neighbor distance distribution for the stationary PPP are identical. And for any compact B with $o \notin B$,

$$\mathbb{P}(\Phi(B) = n \parallel o) = \mathbb{P}(\Phi(B) = n) \qquad \text{for } n = 0, 1, 2, \ldots\,.$$

This suggests that the Palm distribution of the stationary PPP is identical to the distribution of the original PPP with a point added at the origin. Slivnyak's theorem formalizes this statement:

Theorem 5.1 (Slivnyak's theorem) *For the PPP,*

$$\mathbb{P}(\Phi \in Y \parallel o) = \mathbb{P}(\Phi \cup \{o\} \in Y)\,.$$

So conditioning on $o \in \Phi$ is the same as adding a point at o. This is not restricted to the uniform case. We will provide a proof later.

If the point process is finite (and thus non-stationary), we may arrive at a definition of the nearest-neighbor distance distribution by considering all points (and averaging, if needed). In particular, for the BPP, we may proceed as follows. Let $\Phi = \{x_1, \ldots, x_n\}$ be a uniform BPP on $W \subset \mathbb{R}^2$, and let R_x, $x \in \Phi$, be the distance to x's nearest neighbor in $\Phi \setminus \{x\}$. We would like to find $\mathbb{P}(R_x \leq r \parallel x)$. For each $i \in [n]$,

$$\begin{aligned} \mathbb{P}(R_x \leq x \mid x_i = x) &= 1 - \mathbb{P}(R_x > r \mid x_i = x) \\ &= 1 - \mathbb{P}(\Phi^!(b(x,r)) = 0)\,, \end{aligned}$$

where $\Phi^! = \Phi \setminus \{x_i\}$ is a BPP with $n-1$ points. Hence

$$\mathbb{P}(R_x \leq r \mid x_i = x) = 1 - \left(\frac{|b(x,r) \cap W|}{|W|}\right)^{n-1}\,.$$

This is independent of the index i, as expected. Thus it is plausible to interpret it as the conditional probability $\mathbb{P}(R_x \leq r \parallel x)$.

5.1.2 The global approach

Consider all points falling into a test set B with $|B| > 0$. Denote by $\Phi_Y(B)$ the number of points x in B such that $\Phi_{-x} \in Y$:

$$\Phi_Y(B) \triangleq \#\{x \in \Phi \cap B \colon \mathbf{1}_Y(\Phi_{-x})\}$$

Then

$$\mathbb{P}(\Phi \in Y \parallel o) = \frac{\mathbb{E}(\Phi_Y(B))}{\lambda|B|}\,.$$

Since $\lambda|B|$ is the expected number of points in B, this expression defines the Palm probability for the property Y as the fraction of points x expected to fall in B such that $\Phi_{-x} \in Y$. By stationarity this does not depend on B if $|B|$ is fixed.

5.2 The Palm Distribution

5.2.1 Heuristic introduction

In the stationary case, we can define a distribution function

$$D(r) = 1 - \mathbb{P}(\Phi(b(o,r)) = 1 \parallel o)\,,$$

which, for a "typical point", would be the distribution function of the distance to its nearest neighbor. Due to the conditioning, $\Phi(b(o,r)) \parallel o \geq 1$. So this definition of the nearest neighbor distance adds a point at the origin, but then ignores it in the void probability. This approach leads to the Palm distribution, but needs to be made mathematically precise.

For $Y \in \mathcal{N}$, let

$$Y_x = \{\varphi \in \mathsf{N} \colon \varphi_{-x} \in Y\}\,.$$

Then

$$\mathbb{P}(\Phi_x \in Y \parallel o) = \mathbb{P}(\Phi \in Y_{-x} \parallel o)\,.$$

Now let Φ be a general PP with distribution P and locally finite intensity measure Λ. We consider a concrete example for Y: Let

$$Y = \{\varphi \in \mathsf{N} \colon \varphi(b(o,r)) = 1\} \in \mathcal{N}$$

and consider the function $h \colon \mathbb{R}^d \times \mathsf{N} \to \{0,1\}$ given by

$$h(x,\varphi) \triangleq \mathbf{1}_B(x)\mathbf{1}_Y(\varphi_{-x}) = \mathbf{1}_B(x)\mathbf{1}_{Y_x}(\varphi)$$

for some bounded Borel set B. When is $\mathbf{1}_Y(\varphi - x) = 1$? This is the case if the point x does not have any neighbor within distance r. So we may write the mean number of points in B whose neighbors are all at distance at least r as

$$\mathbb{E}\left(\sum_{x\in\Phi} h(x,\Phi)\right) = \int_{\mathsf{N}} \sum_{x\in\varphi} h(x,\varphi)\mathsf{P}(\mathrm{d}\varphi)\,.$$

Note: This is *not* an application of Campbell's formula, this is simply writing out the expectation.

More generally we wish to evaluate this expression for arbitrary non-negative measurable functions $h\colon \mathbb{R}^d \times \mathsf{N} \to \mathbb{R}^+$. If $\mathbb{R}^d$ is partitioned into domains $D_1, D_2, \ldots$, each with non-zero volume, then we can write

$$\mathbb{E}\left(\sum_{x\in\Phi} h(x,\Phi)\right) = \sum_k \mathbb{E}\left(\sum_{x\in\Phi\cap D_k} h(x,\Phi) \mid \Phi(D_k) > 0\right) \cdot \mathbb{P}(\Phi(D_k) > 0)\,.$$

Now suppose $D_k \to \mathrm{d}x$. Then $\mathbb{P}(\Phi(D_k) > 0) \to \Lambda(\mathrm{d}x)$, and the conditional mean should converge to the mean $\mathbb{E}(h(x,\Phi) \parallel x)$ of $h(x,\varphi)$ with respect to $\mathbb{P}(\cdot \parallel x)$. Hence we obtain

$$\mathbb{E}\left(\sum_{x\in\Phi} h(x,\Phi)\right) = \int_{\mathbb{R}^d} \mathbb{E}(h(x,\Phi) \parallel x)\Lambda(\mathrm{d}x)\,.$$

For stationary Φ, $\Lambda = \lambda\nu_d$, so we have

$$\mathbb{E}\left(\sum_{x\in\Phi} h(x,\Phi)\right) = \lambda \int_{\mathbb{R}^d} \mathbb{E}(h(x,\Phi_x) \parallel o)\mathrm{d}x\,,$$

since Φ_x has a point at x if Φ has a point at o.

Using the notation $\mathsf{P}_o(Y) \triangleq \mathbb{P}(\Phi \in Y \parallel o)$ for $Y \in \mathcal{N}$, this takes the form

$$\int_{\mathsf{N}} \sum_{x\in\varphi} h(x,\varphi)\mathsf{P}(\mathrm{d}\varphi) = \lambda \int_{\mathbb{R}^d}\int_{\mathsf{N}} h(x,\varphi_x)\mathsf{P}_o(\mathrm{d}\varphi)\mathrm{d}x \tag{5.2.1}$$

and so, coming back to our example, for $h(x,\varphi) = \mathbf{1}_B(x)\mathbf{1}_Y(\varphi_{-x})$, we obtain

$$\int_{\mathsf{N}} \sum_{x\in\varphi\cap B} \mathbf{1}_Y(\varphi_{-x})\mathsf{P}(\mathrm{d}\varphi) = \lambda|B|\mathsf{P}_o(Y)$$

for Borel B and $Y \in \mathcal{N}$. Now we need to construct a distribution on $[\mathsf{N},\mathcal{N}]$ with the desired behavior of P_o.

5.2.2 A first definition of the Palm distribution (stationary point processes)

Take Φ to be a stationary PP with finite non-zero intensity λ.

Definition 5.1 (Palm distribution for stationary point processes) *The* Palm distribution *(at o) of P is a distribution (probability measure) defined on $[\mathsf{N},\mathcal{N}]$ by*

$$\mathsf{P}_o(Y) \triangleq \int_{\mathsf{N}} \sum_{x\in\varphi\cap B} \frac{\mathbf{1}_Y(\varphi_{-x})\mathsf{P}(\mathrm{d}\varphi)}{\lambda|B|} \qquad \textit{for } Y \in \mathcal{N}\,.$$

This definition does not depend on B — it can be an arbitrary Borel set of positive volume.

We can give an intuitive explanation for this definition using marked point processes. To each point $x \in \Phi$ give a mark 1 or 0 depending on whether the shifted process Φ_{-x} belongs to Y or not. For example, consider (again) $Y = \{\varphi \in \mathsf{N} \mid \varphi(b(o,r)) = 1\}$. Then x has mark 1 precisely when its nearest neighbor is further than r away. The result is a stationary marked PP with a mark distribution M on $\{0,1\}$, and we can write

$$\mathsf{P}_o(Y) = M(\{1\}) = \frac{\lambda_{[1]}}{\lambda}\,,$$

where $\lambda_{[1]}$ is the intensity of the points whose mark is 1. The mean number of points of Φ with mark 1 in B is $\lambda_{[1]}|B|$, so the above definition makes $P_o(Y)$ independent of the test set B.

Alternative notation: Frequently, $\lambda(Y)$ is used instead of $\lambda_{[1]}$. So

$$\mathsf{P}_o(Y) = \frac{\lambda(Y)}{\lambda}\,.$$

In the derivation above, (5.2.1) suggests that the following result, the so-called *refined Campbell theorem* (or Campbell-Mecke theorem) holds for stationary PPs:

Theorem 5.2 (Campbell-Mecke) *For any non-negative measurable function h on $\mathbb{R}^d \times \mathsf{N}$,*

$$\begin{aligned}\mathbb{E}\left(\sum_{x\in\Phi} h(x,\Phi)\right) &= \int_{\mathsf{N}} \sum_{x\in\varphi} h(x,\varphi)\mathsf{P}(\mathrm{d}\varphi)\\ &= \lambda \int_{\mathbb{R}^d}\int_{\mathsf{N}} h(x,\varphi_x)\mathsf{P}_o(\mathrm{d}\varphi)\mathrm{d}x\,.\end{aligned}$$

Taking $h(x,\varphi) = f(x)$ reproduces the standard Campbell theorem.

5.2.3 A second definition of the Palm distribution (general point processes)

This approach uses the Campbell measure $\mathcal{C}$ and the Radon-Nikodým theorem.

Definition 5.2 (Campbell measure) *The* Campbell measure *$\mathcal{C}$ is defined as a measure on $[\mathbb{R}^d\times\mathsf{N}, \mathcal{B}^d\times\mathcal{N}]$ by the relationship*

$$\int_{\mathsf{N}} \sum_{x\in\varphi} f(x,\varphi)\mathsf{P}(\mathrm{d}\varphi) = \int_{\mathbb{R}^d\times\mathsf{N}} f(x,\varphi)\mathcal{C}(\mathrm{d}(x,\varphi))\,,$$

where f is any non-negative measurable function on $\mathbb{R}^d \times \mathsf{N}$. Equivalently, in integral form,

$$\mathcal{C}(B\times Y) = \mathbb{E}(\Phi(B)\mathbf{1}_Y(\Phi))\,.$$

The two definitions are equivalent, since for $f(x,\phi) = \mathbf{1}_B(x)\mathbf{1}_Y(\phi)$,

$$\int_{\mathsf{N}} \sum_{x\in\varphi} f(x,\varphi)\mathsf{P}(\mathrm{d}\varphi) = \int_Y \phi(B)\mathsf{P}(\mathrm{d}\phi) = \mathbb{E}(\Phi(B)\mathbf{1}_Y(\Phi)) = \mathcal{C}(B\times Y)\,.$$

Note that $\mathcal{C}(B\times Y) \leqslant \mathbb{E}\Phi(B) = \Lambda(B)$. For fixed Y, let $\mu_Y(B) \triangleq \mathcal{C}(B\times Y)$. So μ_Y is a measure and $\mu_Y \leqslant \Lambda$, so certainly μ_Y is absolutely continuous with respect to Λ, *i.e.*, $\mu_Y \ll \Lambda$. By the Radon-Nikodým theorem, there exists a density f_Y such that

$$\mu_Y(B) = \int_B f_Y(x)\Lambda(\mathrm{d}x)$$

for Borel B where $f_Y\colon \mathbb{R}^d \to \mathbb{R}^+$ is measurable (and unique up to equality Λ-a.s.). This density f_Y is the *Radon-Nikodým derivative* $\mathrm{d}\mu_Y/\mathrm{d}\Lambda$. As a consequence, we define the Palm distribution as follows:

Definition 5.3 (Palm distribution for general point processes) *The Palm distribution $P_x(Y)$ is defined as*

$$P_x(Y) \triangleq f_Y(x)\,,$$

where $f_Y(x)$ is the density pertaining to $\mathcal{C}(\cdot \times Y)$, i.e.*,*

$$\mathcal{C}(B \times Y) = \int_B f_Y(x)\Lambda(\mathrm{d}x) = \int_B \mathsf{P}_x(Y)\Lambda(\mathrm{d}x)\,. \tag{5.2.2}$$

For fixed x, $\mathsf{P}_x(\cdot)$ is indeed a distribution (probability measure) on $[\mathsf{N}, \mathcal{N}]$. We may write the Palm distribution compactly as the Radon-Nikodým derivative

$$\mathsf{P}_x(Y) \triangleq \frac{\mathrm{d}\mathcal{C}(\cdot \times Y)}{\mathrm{d}\Lambda}\,.$$

This is consistent with the intuition we initially developed: As $\epsilon \to 0$, we have

$$\begin{aligned}
\mathbb{P}(\Phi \in Y \mid \Phi(b(x,\epsilon)) > 0) &= \frac{\mathbb{P}(\Phi \in Y,\ \Phi(b(x,\epsilon)) > 0)}{\mathbb{P}(\Phi(b(x,\epsilon)) > 0)} \\
&\sim \frac{\mathbb{E}(\Phi(b(x,\epsilon))\mathbf{1}_Y(\Phi))}{\mathbb{E}(\Phi(b(x,\epsilon)))} \\
&= \frac{\mathcal{C}(b(x,\epsilon) \times Y)}{\Lambda(b(x,\epsilon))} \\
&\sim \mathsf{P}_x(Y)\,.
\end{aligned}$$

In the stationary case we have

$$\mathsf{P}_o(Y) = \mathsf{P}_x(Y_x) \qquad \text{for } Y \in \mathcal{N}\,,$$

since

$$\begin{aligned}
\lambda \int_B \mathsf{P}_z(Y)\mathrm{d}z = \mathcal{C}(B \times Y) &= \mathcal{C}(B_x \times Y_x) \\
&= \lambda \int_{B_x} \mathsf{P}_y(Y_x)\mathrm{d}y \\
&= \lambda \int_B \mathsf{P}_{x+z}(Y_x)\mathrm{d}z
\end{aligned}$$

for all B, x, and Y. This implies $P_z(Y) = P_{x+z}(Y_x)$. This second definition is applicable to general point processes. In the stationary case, we can retrieve the refined Campbell theorem as follows:

$$\begin{aligned}
\mathbb{E}\left(\sum_{x\in\Phi} h(x,\Phi)\right) &= \int_{\mathbb{R}^d \times \mathsf{N}} h(x,\varphi)\mathcal{C}(\mathrm{d}(x,\varphi)) \\
&= \lambda \int_{\mathbb{R}^d} \int_{\mathsf{N}} h(x,\varphi)\mathsf{P}_x(\mathrm{d}\varphi)\mathrm{d}x \\
&= \lambda \int_{\mathbb{R}^d} \int_{\mathsf{N}} h(x,\varphi_x)\mathsf{P}_o(\mathrm{d}\varphi)\mathrm{d}x
\end{aligned}$$

Example 5.1 (An elementary property of the Palm distribution in the stationary case) *Let* $Y = \{\varphi \in \mathsf{N}\colon \varphi(\{o\}) > 0\}$ *be the set of simple sequences that have a point at o. What is* $\mathsf{P}_o(Y)$*? We have*

$$\begin{aligned}\mathsf{P}_o(Y) &= \frac{\lambda(Y)}{\lambda} = \frac{\mathbb{E}(\Phi_Y([0,1]^d))}{\lambda} \\ &= \frac{\mathbb{E}\#\{x \in \Phi \cap [0,1]^d\colon \mathbf{1}_Y(\Phi_{-x})\}}{\lambda} \\ &= \frac{\mathbb{E}\Phi([0,1]^d)}{\lambda} \\ &= 1\,.\end{aligned}$$

This is of course to be expected, since we are conditioning on $o \in \Phi$*.*

Remarks.

- If Φ is ergodic, we have for all $Y \in \mathcal{N}$

$$\mathsf{P}_o(Y) = \lim_{m\to\infty} \frac{\Phi_Y([-m,m]^d)}{\Phi([-m,m]^d)}\,.$$

- The Campbell can also be defined on the original probability space, *i.e.*, as a measure on $[\mathbb{R}^d \times \Omega, \mathcal{B} \times \mathcal{A}]$, where $(\Omega, \mathcal{A}, \mathbb{P})$ is the original probability space. This is the definition that Baddeley uses.

Example 5.2 (Mixed Poisson point process) *Let* L *be a non-negative random variable defined on* Ω*. Given* $L = \lambda$*, let* Φ *be a homogeneous PPP of intensity* λ*. The intensity of this (non-ergodic)* mixed Poisson point process *is*

$$\mathbb{E}\Phi(B) = \mathbb{E}(\mathbb{E}\Phi(B) \mid L) = \mathbb{E}(L)|B|\,.$$

Let $A = \{L \geq \gamma\}$ *for some* $\gamma \leq 0$*. For* $B \subset \mathbb{R}^d$*, we have*

$$\begin{aligned}\mathcal{C}(B \times A) &= \mathbb{E}(\Phi(B)\mathbf{1}_A) \\ &= \mathbb{E}(\mathbb{E}(\Phi(B)\mathbf{1}_A \mid L)) \\ &= \mathbb{E}(L|B|\mathbf{1}\{L \leq \gamma\}) \\ &= \mathbb{E}(L\mathbf{1}\{L \leq \gamma\})|B|\,,\end{aligned}$$

and it follows that

$$\mathbb{P}_x(L \leq \gamma) = \frac{\mathbb{E}(L\mathbf{1}\{L \leq \gamma\})}{\mathbb{E}L}\,,$$

since

$$C(B \times A) = \int_B \mathbb{P}_x(A)\Lambda(\mathrm{d}x)$$

and $\int_B \Lambda(\mathrm{d}x) = \mathbb{E}(L)|B|$*. Hence the distribution of* L *under* $\mathbb{P}_x$ *is skewed compared to its original distribution.* The event $L \leq \gamma$ is less likely to have occurred if Φ has a point at x*. Say for* $\mathbb{P}(L \leq \gamma) = 1 - \exp(-\gamma)$*, we obtain*

$$\mathbb{P}_x(L \leq \gamma) = 1 - \exp(-\gamma) - \gamma\exp(-\gamma)\,.$$

This is an example where the event A *in the Campbell measure* $\mathcal{C}(\cdot \times A)$ *is an event in the original probability space.*

5.2.4 Alternative interpretation and conditional intensity

We have interpreted $\mathbb{P}_x(\Phi \in Y)$ as the limit of the probability $\mathbb{P}(\Phi \in Y \mid \Phi(b(x,\epsilon)) > 0)$ as $\epsilon \downarrow 0$. Using Bayes' theorem,

$$\mathbb{P}(\Phi(b(x,\epsilon)) > 0 \mid \Phi \in Y) = \frac{\mathbb{P}(\Phi(b(x,\epsilon)) > 0)}{\mathbb{P}(\Phi \in Y)} \mathbb{P}(\Phi \in Y \mid \Phi(b(x,\epsilon)) > 0)$$

and, as $\epsilon \downarrow 0$,

$$\frac{\mathbb{P}(\Phi(b(x,\epsilon)) > 0 \mid \Phi \in Y)}{\mathbb{P}(\Phi(b(x,\epsilon)) > 0)} \to \frac{\mathbb{P}_x(\Phi \in Y)}{\mathbb{P}(\Phi \in Y)} .$$

If Φ has an intensity function $\lambda(x)$, $\mathbb{P}(\Phi(b(x,\epsilon)) > 0) \sim \lambda(x)|b(x,\epsilon)|$, and

$$\frac{\mathbb{P}(\Phi(b(x,\epsilon)) > 0 \mid \Phi \in Y)}{|b(x,\epsilon)|} \to \lambda(x) \frac{\mathbb{P}_x(\Phi \in Y)}{\mathbb{P}(\Phi \in Y)} .$$

So the RHS can be interpreted as the *conditional intensity* of the point process given Y.

5.3 The Reduced Palm Distribution

5.3.1 Definition

In the reduced Palm distribution $\mathsf{P}^!_x$, the point at x on which we condition is not included in the distribution:

$$\mathsf{P}^!_x(Y) \triangleq \mathbb{P}(\Phi \setminus \{x\} \in Y \parallel x) \qquad \text{for } Y \in \mathcal{N}$$

In particular,

$$\mathsf{P}^!_o(Y) = \mathbb{P}(\Phi \setminus \{o\} \in Y \parallel o) \qquad \text{for } Y \in \mathcal{N}$$

To make precise the conditioning on the event with probability zero $x \in \Phi$, we may write in the stationary case:

$$\mathsf{P}_o(Y) = \int_{\mathsf{N}} \sum_{x \in \varphi \cap B} \frac{\mathbf{1}_Y(\varphi_{-x} \setminus \{o\}) \mathsf{P}(\mathrm{d}\varphi)}{\lambda |B|} .$$

In the general case, we define first the so-called *reduced* version of the Campbell measure. This is the

Definition 5.4 (Reduced Campbell measure) *The reduced Campbell measure $\mathcal{C}^!$ is defined as*

$$\begin{aligned} \int_{\mathsf{N}} \sum_{x \in \varphi} f(x, \varphi \setminus \{x\}) \mathsf{P}(\mathrm{d}\varphi) &= \int_{\mathsf{N}} \sum_{x \in \varphi} f(x, \varphi - \delta_x) \mathsf{P}(\mathrm{d}\varphi) \\ &= \int_{\mathsf{N}} f(x,\varphi) \mathcal{C}^!(\mathrm{d}(x,\varphi)) . \end{aligned}$$

$\varphi \setminus \{x\}$ and $\varphi - \delta_x$ are two different notations, a set-theoretic notation and a measure-theoretic notation, for the point pattern φ with the point $x \in \varphi$ deleted. Replacing $\mathcal{C}$ by $\mathcal{C}^!$ in (5.2.2), we arrive at the definition of the *reduced Palm distribution* $P^!_x$:

Definition 5.5 (Reduced Palm distribution) *The reduced Palm distribution* $\mathsf{P}_x^!$ *is defined by the relationship*

$$\mathcal{C}^!(B \times Y) = \int_B \mathsf{P}_x^!(Y)\Lambda(\mathrm{d}x)\,.$$

or, equivalently, as the Radon-Nikodým derivative

$$\mathsf{P}_x^!(Y) \triangleq \frac{\mathrm{d}\mathcal{C}^!(\cdot \times Y)}{\mathrm{d}\Lambda}$$

of the reduced Campbell measure with respect to the Lebesgue measure.

The nearest-neighbor distribution function can be expressed via P_o or $\mathsf{P}_o^!$ by

$$\begin{aligned} D(r) &= 1 - \mathsf{P}_o(\{\varphi \in \mathsf{N} \colon \varphi(b(o,r)) = 1\}) \\ &= 1 - \mathsf{P}_o^!(\{\varphi \in \mathsf{N} \colon \varphi(b(o,r)) = 0\})\,. \end{aligned}$$

D is an important characteristic of PPs. The ratio of the ccdf of the NN distance and the spherical contact distribution function is the so-called *J-function* defined as

$$J(r) \triangleq \frac{1 - D(r)}{1 - F(r)} \qquad \text{for } r \geqslant 0\,,$$

where $F(r)$ is the spherical contact distribution function (or empty space function).

Example 5.3 (Reduced Palm distribution for BPP) *Let Φ_n be a BPP with n nodes. Then the reduced Palm distribution of Φ_n is the distribution of Φ_{n-1}.*

5.3.2 Palm distribution for PPPs and proof of Slivnyak's theorem

The theorem of Slivnyak gives the Palm distribution P_x of a PPP of intensity measure Λ and distribution P:

$$\mathsf{P}_x = \mathsf{P} * \delta_{\delta_x} \qquad \text{for all } x\,.$$

δ_{δ_x} denotes the distribution of the degenerate point process that consists solely of the (non-random) point x, and "$*$" denotes the convolution of distributions, which corresponds to the superposition of point processes. This equation can be interpreted as

$$\mathsf{P}_x(Y) = \mathbb{P}(\Phi \in Y \parallel x) = \mathbb{P}(\Phi \cup \{x\} \in Y) \qquad \text{for } Y \in \mathcal{N}$$

or

$$\int_{\mathsf{N}} f(\varphi)\mathsf{P}_x(\mathrm{d}\varphi) = \int_{\mathsf{N}} f(\varphi \cup \{x\})\mathsf{P}(\mathrm{d}\varphi)$$

for all measurable non-negative functions f. If the reduced Palm distribution is used it takes on a more elegant form:

$$\mathsf{P}_x^! \equiv \mathsf{P}$$

This is a characterization of the PPP.

Proof of Slivnyak's Theorem. Since P_x and $\mathsf{P} * \delta_{\delta_x}$ are the distributions of simple processes, their equality is established if the corresponding void probabilities are equal, *i.e.*,

$$\mathsf{P} * \delta_{\delta_x}(V_K) = \mathsf{P}_x(V_K)$$

for all compact $K \subset \mathbb{R}^d$ where $V_K = \{\varphi \in \mathsf{N} \colon \varphi(K) = 0\}$.

Let A be any bounded Borel set. Then

$$\begin{aligned}\int_A \mathsf{P} * \delta_{\delta_x}(V_K)\Lambda(\mathrm{d}x) &= \int_{A\setminus K} \mathsf{P}(V_K)\Lambda(\mathrm{d}x)\\ &= \mathsf{P}(V_K)\Lambda(A \setminus K)\\ &= \mathbb{E}(\mathbf{1}\{\Phi(K) = 0\}) \cdot \mathbb{E}(\Phi(A \setminus K))\\ &= \mathbb{E}\left(\Phi(A \setminus K)\mathbf{1}\{\Phi(K) = 0\}\right)\\ &= \mathcal{C}((A \setminus K) \times V_K)\,.\end{aligned}$$

Clearly $\mathcal{C}((A \cap K) \times V_K) = \mathbb{E}(\Phi(A \cap K)\mathbf{1}\{\Phi(K) = 0\}) = 0$. (If there was a point in the intersection the void indicator would be zero.) Hence

$$\int_A \mathsf{P} * \delta_{\delta_x}(V_K)\Lambda(\mathrm{d}x) = \mathcal{C}(A \times V_K) = \int_A \mathsf{P}_x(V_K)\Lambda(\mathrm{d}x)$$

using $\mathcal{C}(B \times Y) = \int_B \mathsf{P}_x(Y)\Lambda(\mathrm{d}x)$.

5.3.3 Isotropy of Palm distributions

Clearly the Palm distribution P_o is never a stationary distribution since under P_o a PP must always contain o. But the example of the PPP shows that $\mathsf{P}_o^!$ can be stationary! However, if the PP Φ is motion-invariant then its Palm distribution is isotropic, *i.e.*,

$$\mathsf{P}_o(Y) = \mathsf{P}_o(\mathbf{r}Y) \qquad \text{for } Y \in \mathcal{N}$$

for every rotation $\mathbf{r}$ about the origin.

Example 5.4 (Palm distribution for motion-invariant lattice) *Consider a randomly translated and rotated lattice Φ. Conditioned on $o \in \Phi$, the lattice is no longer stationary, but it is still isotropic. One of the four nearest neighbors of the origin is bound to lie at an angle between 0 and $\pi/2$, with uniform distribution. If a second point is added in the conditioning, say $(0, 1) \in \Phi$ (if the underlying lattice is $\mathbb{Z}^2$), then the Palm distribution equals the (degenerate) distribution of $\mathbb{Z}^2$. In general, such two-fold Palm distributions can be derived from higher-order Campbell measures.*

5.3.4 Palm expectations

The Palm distribution translates to the probability measure in the original probability space naturally, *i.e.*,

$$P_x(Y) \equiv \mathbb{P}_x(\Phi \in Y)\,; \qquad P_x^!(Y) \equiv \mathbb{P}_x^!(\Phi \in Y)\,.$$

Expectations with respect to the (reduced) Palm distribution are often denoted as $\mathbb{E}_x$ and $\mathbb{E}_x^!$, respectively.

5.4 Second Moments Measures and Palm Distributions for Stationary Processes

In integral form, the reduced Campbell measure can be expressed as

$$\mathcal{C}^!(B \times Y) = \mathbb{E}_o^!(\Phi(B)\mathbf{1}_Y(\Phi))\,.$$

For a stationary point process Φ, the distribution of Φ given that $x \in \Phi$ is the same as the distribution of Φ_x given that $o \in \Phi$:

$$\Phi \parallel x \stackrel{\mathrm{d}}{=} (\Phi \parallel o) + x = \Phi_x \parallel o\,.$$

Using refined Campbell,

$$\begin{aligned}\alpha^{(2)}(A \times B) &= \mathbb{E}\left(\sum_{x_1,x_2\in\Phi}^{\neq} \mathbf{1}_A(x_1)\mathbf{1}_B(x_2)\right)\\ &= \int_{\mathsf{N}} \sum_{x\in\varphi} \mathbf{1}_A(x)\varphi(B\setminus\{x\})\mathsf{P}(\mathrm{d}\varphi)\\ &= \lambda \int_{\mathbb{R}^d}\int_{\mathsf{N}} \mathbf{1}_A(x)\varphi((B-x)\setminus\{o\})\mathsf{P}_o(\mathrm{d}\varphi)\mathrm{d}x\,.\end{aligned}$$

We want to express this as

$$\alpha^{(2)}(A \times B) = \lambda \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbf{1}_A(x)\mathbf{1}_B(x+u)\mathcal{K}(\mathrm{d}u)\mathrm{d}x \tag{5.4.1}$$

$$= \lambda \int_A \mathcal{K}(B-x)\mathrm{d}x\,. \tag{5.4.2}$$

So we may define the reduced second moment measure $\mathcal{K}$ as follows:

Definition 5.6 (Reduced second moment measure based on Palm distribution)

$$\mathcal{K}(B) \triangleq \int_{\mathsf{N}} \varphi(B\setminus\{o\})\mathsf{P}_o(\mathrm{d}\varphi) = \int_{\mathsf{N}} \varphi(B)\mathsf{P}_o^!(\mathrm{d}\varphi) = \mathbb{E}_o^!\Phi(B)\,.$$

Hence, $\mathcal{K}(B)$ is the mean number of points in $\Phi \cap B \setminus \{o\}$ under the condition that in o there is a point of Φ. In other words, $\mathcal{K}$ *is the intensity measure of the reduced Palm distribution.*

Accordingly, the K function can be defined as

$$K(r) \triangleq \frac{1}{\lambda}\mathbb{E}(\Phi(b(o,r)) - 1 \parallel o) = \frac{1}{\lambda}\mathbb{E}_o^!\Phi(b(o,r))\,.$$

The K function is an important second-order statistic of motion-invariant processes, but it does not fully characterize it. Different point processes may have the same K function — in much the same way as different random variables may share the same mean and variance.