FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

BACHELOR THESIS

Milan Pultar

Modelling of segment process in the plane

Department of Probability and Mathematical Statistics

Supervisor of the bachelor thesis: Prof. RNDr. Viktor Beneš, DrSc.
Study programme: Mathematics
Study branch: General mathematics

Prague 2017

I declare that I carried out this bachelor thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date .... signature of the author

Title: Modelling of segment process in the plane

Author: Milan Pultar

Department: Department of Probability and Mathematical Statistics

Supervisor: Prof. RNDr. Viktor Beneš, DrSc., Department of Probability and Mathematical Statistics

Abstract:

We consider a finite planar segment process in a circle, having a density with respect to the Poisson process. This density involves unknown parameters and a reference length distribution which is not observed. The aim is to estimate these quantities semiparametrically. The segment process is inhomogeneous, but it is isotropic. Combining the relation between the observed and reference length distribution and the maximum pseudolikelihood method we suggest an estimation procedure. Its properties (bias and variability) are investigated in a simulation study. In the last part we present two more complex models. The motivation is to model stress fibers observed in cultured stem cells.

Keywords: a point process density, a conditional intensity, Metropolis-Hastings algorithm, maximum pseudolikelihood, stem cells

I would like to thank my supervisor Prof. RNDr. Viktor Beneš, DrSc. for his valued advice and hours of consultations. I would also like to thank my family and my girlfriend for their support.

Contents

List of Abbreviations

μ_d ... Lebesgue measure

$\mathbf{X}$ ... a point process

$\mathbf{x}$ ... an observed realization of a process $\mathbf{X}$

N ... a counting measure

λ ... an intensity measure

ρ ... an intensity function

λ^* ... a conditional intensity

$f_{\mathbf{X}}^{(y)}$ ... a density of Palm mark distribution

$p(\mathbf{x})$ ... a probability density with respect to a Poisson process

$h(\mathbf{x}, u)$ ... Hastings ratio

f_1, g ... a reference length distribution

A ... a rotation operator in $\mathbb{R}^2$

Introduction

Random processes of geometrical objects (see [5]) have many applications in biomedicine, materials research and other sciences. E.g. projected stress fibres in stem cells form a segment process, see [7]. Two models of segment processes which can be used to fit the data of stress fibres are suggested in [3], where a parametric estimation is developed. In the present theoretical work we suggest a semiparametric estimation in this modeling which enables more freedom for the choice of the distribution of the random set.

In the first chapter we go through theoretical foundations we need for our work. We define a marked point process, an intensity measure and an intensity function. Then we cite the Campbell's formula which is later used. In this work we focus on finite point processes which are given by a probability density with respect to the Poisson process. Simulated realisations of processes in our models are computed using Metropolis-Hastings algorithm.

In the second chapter we define a segment point process on a bounded region. Parametric models involve reference distributions of marks of segments. These distributions generally do not coincide with corresponding observed distributions. We consider an inhomogeneous planar segment process with given probability density with respect to the Poisson segment process. From simulated realizations of the process using the maximum pseudolikelihood method we firstly estimate scalar parameters of the model. Then using the kernel density estimator of the Palm mark distribution we compute the estimate of the reference probability density on the set of all possible lengths of segments. We compare both the estimators and computed density with true parameters and reference density of the model.

In the third chapter we introduce two more complex models and show the conditions on the density w.r.t. Poisson process to be integrable. This part is motivated by modelling of the above mentioned stress fibres in stem cells.

1. Theoretical foundations

In stochastic geometry one may study random patterns of points. It may be a distribution of trees in a forest or occurrences of a disease in some area. In this chapter we formally define a spatial point process, define some key terms and introduce an algorithm to simulate a spatial point process.

1.1 Spatial point processes

Let $(\mathbb{R}^d, \mathcal{B}^d)$ be $d-$dimensional Euclidean space equipped with its Borel $\sigma-$field and μ_d be Lebesgue measure. Let $Y \in \mathcal{B}^d$ with $\mu_d(Y) > 0$, $\mathcal{B}_Y := \{D \cap Y, D \in \mathcal{B}^d\}$.

Definition 1. *A **counting measure** N on Y is a measure with values in $\{0, 1, 2, \dots\} \cup \{+\infty\}$. It is called **simple** if $N(\{x\}) \leq 1 \; \forall x \in Y$. A counting measure is* **locally finite** *if it is finite on every compact set.*

To formally define a point process on Y we consider $\mathbf{M}$ to be the set of all locally finite simple counting measures on $(Y, \mathcal{B}_Y)$. Let for $D \in \mathcal{B}_Y, k \in \mathbb{N}$

$$E_{D,k} = \{N \in \mathbf{M} : N(D) = k\},$$

where $N(D)$ is the number of points of the support of N in D. Let $\mathcal{M}$ be the σ-field of subsets of $\mathbf{M}$ generated by all sets of the form $E_{D,k}$.

Definition 2. *Let $\mathbf{M}$, $\mathcal{M}$ be as above. The pair $(\mathbf{M}, \mathcal{M})$ is called the **outcome space**.*

Definition 3. *A **point process** is a measurable map $\mathbf{X} : \Omega \to \mathbf{M}$ from a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ to the outcome space $(\mathbf{M}, \mathcal{M})$.*

Measurability of $\mathbf{X}$ ensures that for any set $E \in \mathcal{M}$, the event

$$\{\mathbf{X} \in E\} = \{\omega \in \Omega : \mathbf{X}(\omega) \in E\}$$

is measurable. Therefore the probability $\mathbb{P}(\mathbf{X} \in E)$ is well-defined. Also from the construction of $\mathcal{M}$ it comes that the variables $N(D)$ for any compact set D are random variables on $(\Omega, \mathcal{A}, \mathbb{P})$.

In the following we will also use $\mathbf{X}$ as a random measure or as a random locally finite set of points corresponding to the support of $\mathbf{X}$.

Definition 4. *The* **distribution of a point process X** *is the probability measure $\mathbf{P_X}$, on the outcome space $(\mathbf{M}, \mathcal{M})$, defined by*

$$\mathbf{P_X}(E) = \mathbb{P}(\mathbf{X} \in E), \quad E \in \mathcal{M}.$$

Definition 5. *A point process in $\mathbb{R}^2$ is called* **isotropic** *if its distribution is invariant under all rotations represented by the matrix*

$$A = \begin{bmatrix} cos\ \theta & -sin\ \theta \\ sin\ \theta & cos\ \theta \end{bmatrix},$$

i.e. $\mathbf{P_X}(E) = \mathbf{P_X}(AE), E \in \mathcal{M}$, where $AE = \{A\mathbf{x};\ \mathbf{x} \in E\}$ and $A\mathbf{x} = \{Ay;\ y \in supp\ \mathbf{x}\}$.

Let D be a compact set. Since the sets

$$\{\mathbf{X}(D) = k\} = \{\omega \in \Omega; \mathbf{X}(\omega, D) = k\},$$

$$\{\mathbf{X} \equiv 0\} = \{\omega \in \Omega; \mathbf{X}(\omega, Y) = 0\}$$

are measurable, we can write

$$\mathbb{P}(\mathbf{X}(D) = k); \quad \mathbb{P}_{\mathbf{X}}(N \equiv 0) = \mathbb{P}(\mathbf{X} = \emptyset).$$

Definition 6. *Let* $\mathbf{X}$ *be a point process on* Y*. Then*

$$\lambda(D) := \mathbb{E}[\mathbf{X}(D)], \quad D \in \mathcal{B}_Y,$$

defines a measure on Y *which is called the* **intensity measure of X.**

Definition 7 (General Poisson process)**.** *Let* Y *be as above and* Λ *a measure which is finite on every compact set and which has no atoms.*

The ***Poisson process*** *on* Y *with intensity measure* Λ *is a point process on* Y *such that:*

- *If* $D_1, \dots D_m$ *are disjoint, then* $N(D_1), \dots N(D_m)$ *are independent random variables.*
- *For every bounded closed set* D*, the count* $N(D)$ *of points in* D *has a Poisson distribution with mean* $\Lambda(D)$*.*

Definition 8. *The* **homogeneous Poisson process** *with intensity* $\rho > 0$ *is a Poisson process with measure* $\Lambda(D) = \rho\mu_d(D)$*. Otherwise we call it the* **non-homogeneous Poisson process***.*

Definition 9. *Suppose the intensity measure* λ *of a point process* $\mathbf{X}$ *in* $\mathbb{R}^d$ *satisfies*

$$\lambda(D) = \int_D \rho(u)\mathrm{d}u$$

for some function ρ*. Then* ρ *is called the* **intensity function of X.**

Example 1. *For the homogeneous Poisson process the intensity function* $\rho(u) \equiv \rho$*.*

Theorem 1 (Campbell's formula, [1], Theorem 2.2, p. 28)**.** *Let* $\mathbf{X}$ *be a point process on* Y *with intensity measure* λ *and let* $f : Y \to \mathbb{R}$ *be a measurable function. Then the random sum*

$$T = \sum_{x \in \mathbf{X}} f(x)$$

is a random variable with expected value

$$\mathbb{E}\,T = \int_Y f(x)\lambda(\mathrm{d}x).$$

If $\mathbf{X}$ *is a point process with intensity function* ρ*, then*

$$\mathbb{E}\,T = \int_Y f(x)\rho(x)\mathrm{d}x.$$

In many applications we need to study spatial point processes with some additional information attached to each point of the process. A common example is the map of emergency calls where the extra information is the time of the call and the nature of the emergency. In the process of trees growing in a forest the extra information may be the height of each tree. Another example is the process of stress fibers observed during the life of a stem cell.

Definition 10. *Let $(R, \mathcal{R})$ be a measurable space, $\mathcal{R}$ $\sigma-$algebra. A* ***marked point process*** *in a space Y with marks in a space R is a point process $\mathbf{X}$ on $Y \times R$ such that $\mathbf{X}(D \times R) < \infty$ a.s. for all compact $D \subset Y$.*

In other words, the corresponding projection of a marked point process (a process of points without marks) must be locally finite. Since there are no further general requirements on the space of marks R, it can be very diverse. An example is a set of real numbers, an interval or a set of geometric figures.

Definition 11. *For a marked point process $\mathbf{X}$ the intensity measure λ is defined by*

$$\lambda(U) = \mathbb{E}\,\mathbf{X}(U), \quad U \subset Y \times R,$$

provided $\lambda(U) < \infty$ $\forall U \subset D \times R$, D compact.

Theorem 2 ([1], Theorem 1.3, p. 21)**.** *Let $\mathbf{Y}$ be a marked point process on Y with marks in R. Let $\mathbf{X}$ be the projected process in Y of points without marks. Then the following are equivalent:*

- *$\mathbf{X}$ is a Poisson process with intensity measure λ. The marks attached to the points of $\mathbf{X}$ are independent and identically distributed with common distribution Q on λ.*
- *$\mathbf{Y}$ is a Poisson process in $Y \times R$ with intensity measure $\lambda \otimes Q$.*

1.2 Finite point processes

Definition 12. *Let $\mathbf{X}$ be a point process on a space Y. If $N(S) < \infty$ a.s., then $\mathbf{X}$ is called a* ***finite point process****.*

Realisations of a finite point process $\mathbf{X}$ belong to the space

$$\mathbf{M}^f = \{N \in \mathbf{M} : N(Y) < \infty\}$$

of simple finite counting measures on Y. Let

$$\mathbf{M}_k = \{N \in \mathbf{M} : N(Y) = k\}, \quad k \in \mathbb{N}_0.$$

Then

$$\mathbf{M}^f = \bigcup_{k=0}^{\infty} \mathbf{M}_k$$

is the set of all counting measures with total mass k or equivalently the set of all configurations of k points. This can be also represented by introducing the space of ordered k-tuples

$$Y^{!k} = \{(x_1, \dots, x_k) : x_i \in Y, x_i \neq x_j \;\forall\; i \neq j\}$$

and defining a mapping $I_k : Y^{!k} \to \mathbf{M}_k$ by

$$I_k(x_1, \dots, x_k) = \delta_{x_1} + \cdots + \delta_{x_k}.$$

This gives $\mathbf{M}_k \equiv Y^{!k}/\sim$,where $\sim$ is the equivalence relation under permutation, i.e.

$$(x_1, \dots, x_k) \sim (y_1, \dots, y_k) \Leftrightarrow \{x_1, \dots, x_k\} = \{y_1, \dots, y_k\}$$

1.3 Point process density

In the following let $\mathbf{P}_\eta$ be the distribution of Poisson point process on a space Y with intensity measure λ, where $0 < \lambda(Y) < \infty$.

Definition 13. *Let* p*:* $\mathbf{M}^f \to \mathbb{R}_+$ *be a measurable function satisfying* $\int_{\mathbf{M}} p(\mathbf{x})\mathbf{P}_\eta(\mathrm{d}\mathbf{x}) = 1$*. Define*

$$\mathbf{P}(E) = \int_E p(\mathbf{x})\mathbf{P}_\eta(\mathrm{d}\mathbf{x})$$

for any event $E \in \mathcal{M}$*. Then* $\mathbf{P}$ *is a distribution of a point process. The function* p *is called the* **probability density with respect to the Poisson process** η *of the point process with distribution* $\mathbf{P}$*.*

Lemma 3 ([1], Lemma 4.1)**.** *For a point process* $\mathbf{X}$ *with probability density* p *we have*

$$\mathbb{P}(\mathbf{X} \in E) = \tag{1.1}$$

$$e^{-\lambda(Y)} \sum_{n=0}^{\infty} \frac{1}{n!} \int_Y \cdots \int_Y 1\{I_n(x_1, \dots, x_n) \in E\} p(I_n(x_1, \dots, x_n))\lambda(\mathrm{d}x_1) \dots \lambda(\mathrm{d}x_n),$$

$E \in \mathcal{M}$.

Definition 14. *Probability density* p *is in the* **exponential form** *if*

$$p(\mathbf{x}) = c\exp(\mathbf{h} \cdot V(\mathbf{x})), \quad \mathbf{h} \in \mathbb{R}^m,$$

with respect to some Poisson process η*. Here* $V(\mathbf{x})$ *is a vector of some geometrical characteristics of* $\mathbf{x}$*,* $\mathbf{h} \cdot V(\mathbf{x})$ *is an inner product and* c *is a normalizing constant.*

We have the identity

$$\mathbb{E}[\exp(\mathbf{h} \cdot V(\eta))] = \frac{1}{c},$$

therefore a point process with probability density p in the exponential form is well defined, i.e. p is integrable, if and only if

$$\mathbb{E}[\exp(\mathbf{h} \cdot V(\eta))] < \infty.$$

Definition 15 (Conditional intensity)**.** *Let* p *be the density of a point process* $\mathbf{X}$ *on a bounded set* $Y \subset \mathbb{R}^d$*. If* p *is* **hereditary***, that is*

$$p(\mathbf{x}) > 0 \Rightarrow p(\mathbf{y}) > 0, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbf{M}^f, \mathbf{x} \subset \mathbf{y},$$

then we define the **conditional intensity** *of the process* $\mathbf{X}$ *as*

$$\lambda^*(\mathbf{x}, u) = \frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})}, \quad \mathbf{x} \in \mathbf{M}^f,\, u \in Y.$$

1.4 Metropolis-Hastings algorithm

To simulate realizations of the point process $\mathbf{X}$ we use Markov Chain Monte Carlo method ([6]).

Algorithm 1 (Metropolis-Hastings birth-death algorithm)**.**

Let $\mathbf{X}$ *be a point process on* $Y \in \mathcal{B}^d$ *with given probability density* p *with respect to the unit (uniform with* $\lambda = \mu_d$*) Poisson process.* $\lambda^*(\mathbf{x}, u), u \in Y$*, is the corresponding conditional intensity.*

Define Hastings ratio

$$h(\mathbf{x}, u) = \lambda^*(\mathbf{x}, u) \frac{\mu_d(Y)}{n(\mathbf{x}) + 1},$$

where $n(\mathbf{x})$ *denotes number of points in* $\mathbf{x}$*.*

1. *Let* I *be number of iterations,* $\mathbf{x}^i$ *be the configuration of segments in the* i*-th iteration.* $\mathbf{x}^0$ *is an initial configuration of an arbitrary number of points in* Y *under the condition* $p(\mathbf{x}) > 0$*.*

2. (a) *With probability* $\frac{1}{2}$ *we consider adding a point* u *chosen uniformly randomly in* Y *under the condition* $p(u) > 0$*. Then add the point (formally* $\mathbf{x}^{i+1} = \mathbf{x}^i \cup \{u\}$*) with probability*

   $$\min(1, h(\mathbf{x}, u)),$$

   else we set $\mathbf{x}^{i+1} = \mathbf{x}^i$*.*

   (b) *Else we consider removing one point from* $\mathbf{x}$*. We choose a point uniformly randomly from existing points and remove it (i.e.* $\mathbf{x}^{i+1} = \mathbf{x}^i \setminus \{u\}$*) with probability*

   $$\min\left(1, \frac{1}{h(\mathbf{x}, u)}\right),$$

   else we set $\mathbf{x}^{i+1} = \mathbf{x}^i$*.*

3. *Repeat* 2. *until we reach the* I*-th iteration. Then set* $\mathbf{x} = \mathbf{x}^I$ *to be the simulated realization of the process* $\mathbf{X}$*.*

This is an approximation in the sense that the distribution of configurations tends to the target distribution of $\mathbf{X}$ for the number of iterations approaching infinity.

2. Segment process with reference mark distribution

A **segment process** in $\mathbb{R}^2$ can be considered as a marked point process with two marks corresponding to the length and direction of a segment. Let $B \subset \mathbb{R}^2$ be bounded measurable,

$$Y = B \times S, \quad S = (0, e_a] \times [0, \pi), \tag{2.1}$$

where $e_a > 0$ is an upper bound for the segment length, $[0, \pi)$ is the manifold of axial directions. Let the Poisson process η on Y have intensity measure λ,

$$\lambda(\mathrm{d}(y, r, \varphi)) = \mathrm{d}y \frac{1}{e_a} \mathrm{d}r \frac{1}{\pi} \mathrm{d}\varphi, \tag{2.2}$$

where for a segment $u = (y, r, \varphi)$ we denote y the location of the centre, r the length and φ the direction. Let the segment process $\mathbf{X}$ have a hereditary density p with respect to η and conditional intensity λ^*. Let ρ be the intensity function of the process $\mathbf{X}$.

For $D \subset B$ measurable and the point process ν of segment centres, denote

$$\kappa(D) = \mathbb{E} \sum_{y \in \nu} 1_D(y).$$

Let $G \subset S$,

$$C(D \times G) = \mathbb{E} \sum_{(y,\xi) \in \mathbf{X}} 1_D(y) 1_G(\xi).$$

The measure ω_G, $\omega_G(D) = C(D \times G)$, is absolutely continuous with respect to κ. Let $P^y(G)$ be the corresponding Radon-Nikodym density. It satisfies

$$\omega_G(D) = \int_D P^y(G) \kappa(\mathrm{d}y). \tag{2.3}$$

Definition 16. *$P^y(G)$ is called the* **Palm mark distribution** *of* $\mathbf{X}$ *at y.*

Let exist $f_{\mathbf{X}}^{(y)}$, the density with respect to Lebesgue measure of the Palm mark distribution of length and direction of a typical segment of the process $\mathbf{X}$ at the location $y \in B$.

Proposition 1. *For each $y \in B$ we have*

$$f_{\mathbf{X}}^{(y)}(\xi) = \frac{\rho(y, \xi)}{\int_S \rho(y, \xi) \mathrm{d}\xi}. \tag{2.4}$$

Proof: For Borel sets $D \subset B$, $G \subset S$ it holds using the Campbell theorem

$$C(D \times G) = \int_Y 1_D(y) 1_G(\xi) \rho(y, \xi) \mathrm{d}\xi \mathrm{d}y.$$

Specially for $G = S$ we have

$$\kappa(D) = \int_D \int_S \rho(y, \xi) \mathrm{d}\xi \mathrm{d}y.$$

Then from (2.3) we have

$$\int_D P^y(G) \int_S \rho(y,\xi)\mathrm{d}\xi\mathrm{d}y = \int_D \int_G \rho(y,\xi)\mathrm{d}\xi\mathrm{d}y$$

for all Borel sets $D \subset B$, finally

$$P^y(G) = \frac{\int_G \rho(y,\xi)\mathrm{d}\xi}{\int_S \rho(y,\xi)\mathrm{d}\xi}$$

and (2.4) follows. □

Using formula (1.1) we can alternatively transfer the reference mark density f and parameter τ into p and consider a modified density p' (of the same process) with respect to unit Poisson process. In Lemma 3 we have

$$\int_Y \cdots \int_Y 1\{I_n(u_1,\ldots,u_n) \in Z\} p(I_n(u_1,\ldots,u_n))\lambda(\mathrm{d}u_1)\ldots\lambda(\mathrm{d}u_n) =$$
$$\int_Y \cdots \int_Y 1\{I_n(u_1,\ldots,u_n) \in Z\} p(I_n(u_1,\ldots,u_n)) e_a^{-n}\pi^{-n}\mathrm{d}u_1\ldots\mathrm{d}u_n =$$
$$\int_Y \cdots \int_Y 1\{I_n(u_1,\ldots,u_n) \in Z\} p^{'}(I_n(u_1,\ldots,u_n))\mathrm{d}u_1\ldots\mathrm{d}u_n.$$

This approach is applied in the following.

2.1 The segment process with reference length distribution

Consider a circle $B \subset \mathbb{R}^2$ centered in the origin with diameter $e_a > 0$. Let $L_o = [0, e_a]$ be the interval of segment lengths. We demand that segments lie completely in the window B. Let

$$Y = B \times L_o \times [0,\pi) \tag{2.5}$$

be the space of segments $u = (y, r, \varphi) \in Y$ which have centre y, length r and axial direction φ.

We consider the Poisson segment process η with the intensity measure λ on Y. Let the segment process $\mathbf{X}$ have a density p w.r.t. η :

$$p(\mathbf{x}) = c 1_{[\mathbf{x}\subset B]} \exp(b\, D(\mathbf{x}))\tau^{n(\mathbf{x})} \prod_{u_i \in \mathbf{x}} f_1\left(\frac{r_i}{e_a}\right), \tag{2.6}$$

$b \in \mathbb{R}$, $\tau > 0$, $n(\mathbf{x})$ is the total number of segments in $\mathbf{x}$, c is the normalizing constant, r_i is the length of $i-$th segment u_i, f_1 is a reference length probability density on L_o and

$$D(\mathbf{x}) = \sum_{u\in\mathbf{x}} d(u), \quad d(u) = \max_{z\in u} \frac{||z||}{e_a}.$$

For $u \subset B$ the quantity $d(u) \in (0, \frac{1}{2}]$ is the normalized distance of the most distant point of u from the centre of B. For positive, negative values of parameter b more, less distant segments from the origin prevail, respectively, cf. Fig. 2.1.

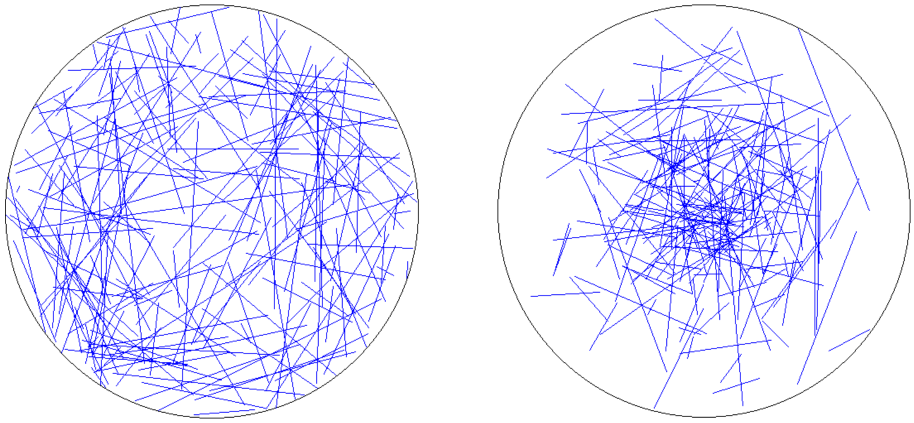


Figure 2.1: Simulated realizations of the process $\mathbf{X}$ using Algorithm 1 ($3 \cdot 10^6$ iterations) with $e_a = 1$ and the density (2.6) having parameters $b = 10,\ \tau = 3$ (left) and $b = -10,\ \tau = 4000$ (right), f_1 corresponds to beta distribution with parameters $\alpha = 2,\ \beta = 4$.

The corresponding conditional intensity is

$$\lambda^*(\mathbf{x}, u) = 1_{[\mathbf{x} \cup \{u\} \subset B]} \tau f_1 \left(\frac{r}{e_a} \right) \exp(bd(u)).$$

For the intensity function it holds

$$\begin{aligned} &\rho(u) = \mathbb{E}\, \lambda^*(\mathbf{X}, u) = \tau f_1 \left(\tfrac{r}{e_a} \right) \exp(bd(u)), && u \subset B, \\ &\rho(u) = 0, && u \cap (\mathbb{R}^2 \setminus B) \neq \emptyset. \end{aligned}$$

There are no interactions among the segments in the model (2.6) with statistics $D(\mathbf{x})$, $\mathbf{X}$ is in fact an inhomogeneous Poisson process with unknown reference density f_1 and a condition $\mathbf{X} \subset B$. The process $\mathbf{X}$ is isotropic since both the reference Poisson process and the density p are invariant with respect to rotations around the origin. This comes from Lemma 3 where

$$\int_Y \cdots \int_Y 1\{I_n(u_1, \dots, u_n) \in Z\} p(I_n(u_1, \dots, u_n)) \lambda(\mathrm{d}u_1) \dots \lambda(\mathrm{d}u_n) = \\ \int_Y \cdots \int_Y 1\{I_n(v_1, \dots, v_n) \in Z\} p(I_n(v_1, \dots, v_n)) \lambda(\mathrm{d}v_1) \dots \lambda(\mathrm{d}v_n),$$

$Z \in \mathcal{M}, v_i = Au_i$ is rotation of the segment u_i. The identity holds because we integrate over all configurations of $n \in \mathbb{N}$ segments where all segments lie (completely) in B.

We use the symbol $A\varphi$ for the corresponding direction of the segment $u = (y, r, \varphi)$ after rotation Au. Then also the intensity function is invariant with respect to rotations, i.e.

$$\rho(u) = \rho(Au)$$

for all $u \in Y$ and all rotations A.

Let $f_{\mathbf{X}}^{(y)}$ be the bivariate (Palm) density of the distribution of length and direction of the segment centered at y, from Proposition 1 we have for $y \in B$ a normalizing constant $C_y > 0$ such that

$$f_{\mathbf{X}}^{(y)}(r,\varphi) = \frac{\rho(y,r,\varphi)}{\int\limits_{L_0\times[0,\pi)} \rho(y,r,\phi)\mathrm{d}r\mathrm{d}\phi} = \frac{\tau f_1\left(\frac{r}{e_a}\right)\exp(bd(u))}{\tau \int\limits_{L_0\times[0,\pi)} f_1\left(\frac{r}{e_a}\right)\exp(bd(u))\mathrm{d}r\mathrm{d}\varphi} = \frac{f_1\left(\frac{r}{e_a}\right)\exp(bd(u))}{C_y}.$$

This gives the relation between the reference length density and the Palm distribution.

$$f_1\left(\frac{r}{e_a}\right) = \frac{C_y f_{\mathbf{X}}^{(y)}(r,\varphi)}{e^{bd(u)}}. \tag{2.7}$$

From the isotropy of $\mathbf{X}$ it holds

$$f_{\mathbf{X}}^{(y)}(r,\varphi) = f_{\mathbf{X}}^{(Ay)}(r,A\varphi) \tag{2.8}$$

and therefore

$$C_y = C_{Ay}$$

for all $y \in B$ and all rotations A.

2.2 Semiparametric estimation using maximum pseudolikelihood

The aim is to estimate parameters b, τ and density f_1 from a sample of realizations of $\mathbf{X}$ where we observe $f_{\mathbf{X}}^{(y)}$ instead of f_1. We are using a semiparametric approach so that f_1 is not parametrized. Because of inhomogeneity of $\mathbf{X}$ we have to use discretization during the solution, but we make use of isotropy. In the parametric part we use maximum pseudolikelihood method for parameter estimation. For an observed realization $\mathbf{x}$ the pseudolikelihood ([1]) is defined as

$$\mathcal{L} = \prod_{u \in \mathbf{x}} \lambda^*(\mathbf{x} \setminus u, u) \exp\left(- \int_{v \subset B} \lambda^*(\mathbf{x}, v) \mathrm{d}v \right).$$

The logarithmic pseudolikelihood

$$\begin{aligned} \log \mathcal{L} &= \sum_{u \in \mathbf{x}} \lambda^*(\mathbf{x} \backslash u, u) - \int_{u \subset B} \lambda^*(\mathbf{x}, u) \mathrm{d}u \\ &= \log(\tau) n(\mathbf{x}) + bD(\mathbf{x}) + \sum_{u \in \mathbf{x}} \log f_1\left(\frac{r}{e_a}\right) - \tau \int_{u \subset B} f_1\left(\frac{r}{e_a}\right) e^{bd(u)} \mathrm{d}u \end{aligned}$$

has to be maximized with respect to τ, b. We have

$$\frac{\partial \log \mathcal{L}}{\partial \tau} = \frac{n(\mathbf{x})}{\tau} - \int_{u \subset B} f_1\left(\frac{r}{e_a}\right) e^{bd(u)} \mathrm{d}u,$$

$$\frac{\partial \log \mathcal{L}}{\partial b} = D(\mathbf{x}) - \tau \int_{u \subset B} d(u) f_1\left(\frac{r}{e_a}\right) e^{bd(u)} \mathrm{d}u.$$

Using (2.7) we obtain equations

$$n(\mathbf{x}) = \tau \int_{u \subset B} C_y f_{\mathbf{X}}^{(y)}(r, \varphi) \mathrm{d}u, \tag{2.9}$$

$$D(\mathbf{x}) = \tau \int_{u \subset B} C_y d(u) f_{\mathbf{X}}^{(y)}(r, \varphi) \mathrm{d}u.$$

In the estimation procedure we proceed in several steps:
(**i**) Consider discrete levels of $||y||$ like $0 < w_1 < \cdots < w_k < e_a/2$ and put $y_j = (0, w_j)^T \in B$. For

$$\Delta_f = \frac{e_a}{2k}$$

let

$$w_j = (j-1)\Delta_f + \frac{\Delta_f}{2},\ j = 1, \ldots, k,$$

$$Y_j = B_j \times L_o \times [0, \pi), \quad B_j = \{y \in B,\ (j-1)\Delta_f < ||y|| \leq j\Delta_f\},\ j = 1, \ldots, k.$$

The kernel estimator of the bivariate densities $f_{\mathbf{X}}^{(y_j)}$, $j = 1, \ldots, k$, is evaluated from the observed data in each class, i.e. from the sample of segments $u_j^{(i)} = (y_j^{(i)}, r_j^{(i)}, \varphi_j^{(i)})$, $i = 1, \ldots, m_j$ centered in B_j, $j = 1, \ldots, k$ (see Fig. 2.2). Because of (2.8) these segments are first rotated by such $A_j^{(i)}$ that

$$A_j^{(i)} y_j^{(i)} = q y_j$$

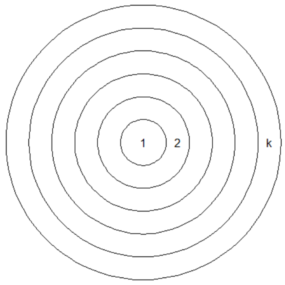


Figure 2.2: The circle B is divided in k annuli (classes) $B_j, j = 1, \ldots, k$.

for some $q > 0$. Then we apply kernel estimation to each sample

$$A_j^{(i)} u_j^{(i)}, i = 1, \ldots, m_j,$$

to estimate the joint length-direction density $f_{\mathbf{X}}^{(y_j)}$, $j = 1, \ldots, k$. Since the length component has values on a compact L_0, we use here a system of beta kernels ([2]). The angular component is a circular variable, cf. [6], jointly we use a bivariate product kernel. Let a_1, a_2 be positive constants, we use the following kernel density estimator $\hat{f}_j$ of the bivariate densities $f_{\mathbf{X}}^{(y_j)}$.

$$\hat{f}_j(\vec{x}) = \frac{1}{n} \sum_{i=1}^{n} K^{\star}_{x,a_1}(r_j^{(i)}) K_{a_2}(x_2 - \tilde{\varphi}_j^{(i)}),$$
$$\vec{x} = (x_1, x_2), \quad A^{(i)} u_j^{(i)} = (\tilde{y}_j^{(i)}, r_j^{(i)}, \tilde{\varphi}_j^{(i)})$$

$j = 1, \ldots, k$. Here the angular component is a circular variable, so

$$K_{a_2}(x) = \frac{1}{\sqrt{a_2}} K\left(\frac{min(|x|, \pi - |x|)}{\sqrt{a_2}}\right), \quad x \in [0, \pi),$$

where K is Epanechnikov kernel function

$$K(x) = \frac{3}{4}(1 - x^2)\mathbb{1}_{(|x| \leq 1)}, \quad x \in \mathbb{R}.$$

Further

$$K^{\star}_{x,a_1}(t) = \begin{cases} K_{x/a_1,(1-x)/a_1}(t) & \text{if } x \in [2a_1, 1 - 2a_1] \\ K_{\rho(x),(1-x)/a_1}(t) & \text{if } x \in [0, 2a_1) \\ K_{x/a_1,\rho(1-x)}(t) & \text{if } x \in (1 - 2a_1, 1] \end{cases}$$

where $K_{a,b}$ is the density function of distribution $Beta(a, b)$ and

$$\rho(x) = 2a_1^2 + 2.5 - \sqrt{4a_1^4 + 6a_1^2 + 2.25 - x^2 - \frac{x}{a_1}}.$$

(**ii**) From (2.7) we have

$$C_y^{-1} = \int_{\mathbb{R}} \frac{f_{\mathbf{X}}^{(y)}(r, \varphi)}{e^{bd(u)}} \mathrm{d}r \tag{2.10}$$

We choose fixed $\varphi = 0$. Since the length of the longest possible segment centered in the $j-$th class is

$$l_j = 2\sqrt{\left(\frac{e_a}{2}\right)^2 - (\Delta_f(j-1))^2},$$

we use the step

$$\Delta_c^{(j)} = \frac{l_j}{m}, \quad m \in \mathbb{N},$$

for numerical integration of (2.10) with $y = y_j$ using Simpson rule:

$$\int_a^b f(x)dx \approx \frac{f(a) + 4f(\frac{a+b}{2}) + f(b)}{6}$$

for some function f. We split the interval $(0, l_j)$ in m subintervals $(c_0, c_1), (c_1, c_2), \ldots (c_{m-1}, c_m)$, where

$$0 = c_0 < c_1 < \cdots < c_{m-1} < c_m = l_j$$

are equally spaced points.

We put

$$r_{l,1} = (l-1)\Delta_c^{(j)}$$
$$r_{l,2} = (l-1)\Delta_c^{(j)} + \frac{\Delta_c^{(j)}}{2}$$
$$r_{l,3} = l \cdot \Delta_c^{(j)}$$

Then we approximate the integral (2.10) with

$$C_{y_j}^{-1} = \Delta_c^{(j)} \sum_{l=1}^{m} \frac{F_j(r_{l,1}) + 4F_j(r_{l,2}) + F_j(r_{l,3})}{6},$$

where

$$F_j(x) = \frac{f_{\mathbf{X}}^{(y_j)}(x, \varphi)}{\exp(bd(y_j, x, \varphi))}$$

Now we have unknown constants $C_{y_j} = C_j$ expressed as functions of b.

(**iii**) Integrals in the equations (2.9) are evaluated by Monte Carlo simulation of M segments in Y uniformly randomly. Let n of them lie completely in B, denoted $\bar{u}_j^{(i)} = (\bar{y}_j^{(i)}, \bar{r}_j^{(i)}, \bar{\varphi}_j^{(i)})$, $i = 1, \ldots, n_j$ where n_j of them are centered in B_j, $\sum_{j=1}^k n_j = n$. Then

$$n(\mathbf{x}) = \frac{\tau\pi^2 e_a^3}{4M} \sum_{j=1}^{k} C_j \sum_{i=1}^{n_j} f_X^{(\bar{y}_j^{(i)})}(\bar{r}_j^{(i)}, \bar{\varphi}_j^{(i)})$$
$$D(\mathbf{x}) = \frac{\tau\pi^2 e_a^3}{4M} \sum_{j=1}^{k} C_j \sum_{i=1}^{n_j} d(\bar{u}_j^{(i)}) f_X^{(\bar{y}_j^{(i)})}(\bar{r}_j^{(i)}, \bar{\varphi}_j^{(i)})$$

(since $\frac{\pi e_a^2}{4} \cdot \pi \cdot e_a$ is the volume of Y). Thus we obtain

$$\frac{D(\mathbf{x})}{n(\mathbf{x})} = \frac{\sum_{j=1}^k C_j \sum_{i=1}^{n_j} d(\bar{u}_j^{(i)}) f_X^{(\bar{y}_j^{(i)})}(\bar{r}_j^{(i)}, \bar{\varphi}_j^{(i)})}{\sum_{j=1}^k C_j \sum_{i=1}^{n_j} f_X^{(\bar{y}_j^{(i)})}(\bar{r}_j^{(i)}, \bar{\varphi}_j^{(i)})}. \tag{2.11}$$

	true	mean	sd
b	3	3.059	0.484
τ	900	938.8	189.9
	true	mean	sd
b	3	3.048	0.486
τ	900	954	186

Table 2.1: Means and standard deviations (sd) of the maximum pseudolikelihood estimates of scalar parameters in the model (2.6) with reference length density. Upper part - sample I (20 simulations), lower part - sample II (40 simulations)

It is an equation with a single variable b (using (2.10)) which is solved numerically.
(**iv**) Having estimated b we obtain C_j from step (ii) and then τ from any of the equations

$$\tau = \frac{4M\, n(\mathbf{x})}{\pi^2 e_a^3 \sum_{j=1}^{k} C_j \sum_{i=1}^{n_j} f_X^{(\bar{y}_j^{(i)})}(\bar{r}_j^{(i)}, \bar{\varphi}_j^{(i)})},$$

$$\tau = \frac{4M\, D(\mathbf{x})}{\pi^2 e_a^3 \sum_{j=1}^{k} C_j \sum_{i=1}^{n_j} d(\bar{u}_j^{(i)}) f_X^{(\bar{y}_j^{(i)})}(\bar{r}_j^{(i)}, \bar{\varphi}_j^{(i)})}$$

which are Monte Carlo analogues of the equations in (2.9).
(**v**) Finally the estimator of the reference length density f_1 is obtained by plugging these estimators of b, τ and constants $C_j, j = 1, \ldots, k$ in (2.7) and renormalizing. We can average results of several levels of y.

A numerical study is based on samples I, II of 20, 40 simulated independent realizations, respectively, of segment process with parameters $b = 3, \tau = 900, \alpha = 2, \beta = 4, e_a = 1$. In step (**i**) $k = 6$ classes were considered. In Fig. 2.3 there are kernel estimates of the observed length density in all six classes separately. The results of estimation of parameters b, τ (step (**iii**), (**iv**)) are in Table 2.1. Then in Fig. 2.4 there are semiparametric estimators of reference length density f_1 (step (**v**)) in all classes separately. Finally the resulting semiparametric estimator of reference length density f_1 is in Fig. 2.5, here we do not consider two outer classes where the inhomogeneity is the highest.

In all graphs the envelopes are created using the following procedure. In each point x of the 1000 points on horizontal axis we have $\{d_i\}_{i=1}^m$, the function values of m independent estimators. Denote $d_{(0)} \leq \cdots \leq d_{(m-1)}$ the corresponding ordered set. We take $i_1 = \left\lfloor \frac{1}{20}(m-1) \right\rfloor$ and $i_2 = \left\lceil \frac{19}{20}(m-1) \right\rceil$. Then $d_{(i_1)}, d_{(i_2)}$ are the lower and upper values of the envelopes at x. They correspond to empirical 90% confidence interval.

If we look at the results of our simulation study, there is only a small deviation of the estimations of both scalar parameters and the reference length density (if we omit the last two classes). We still observe large variance which may be due to the very small number of segments in the first class (the most segments are in classes 3,4,5, where variance is the smallest). It may be studied in later work.

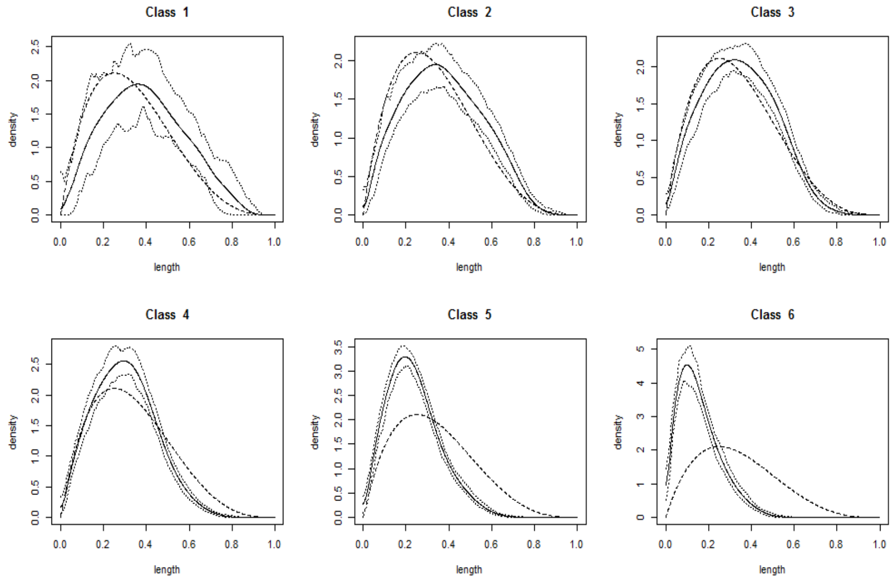


Figure 2.3: Kernel estimation of the observed length density based on 40 simulations of the inhomogeneous segment process $\mathbf{X}$, $b = 3, \tau = 900$. In each of six classes the average kernel estimator of the observed length density (full line) is compared to the true reference density (dashed line) of beta distribution with parameters $\alpha = 2$, $\beta = 4$. The envelopes (dotted lines) correspond to empirical 90% confidence interval for the kernel estimator, pointwise in 1000 points on horizontal axis.

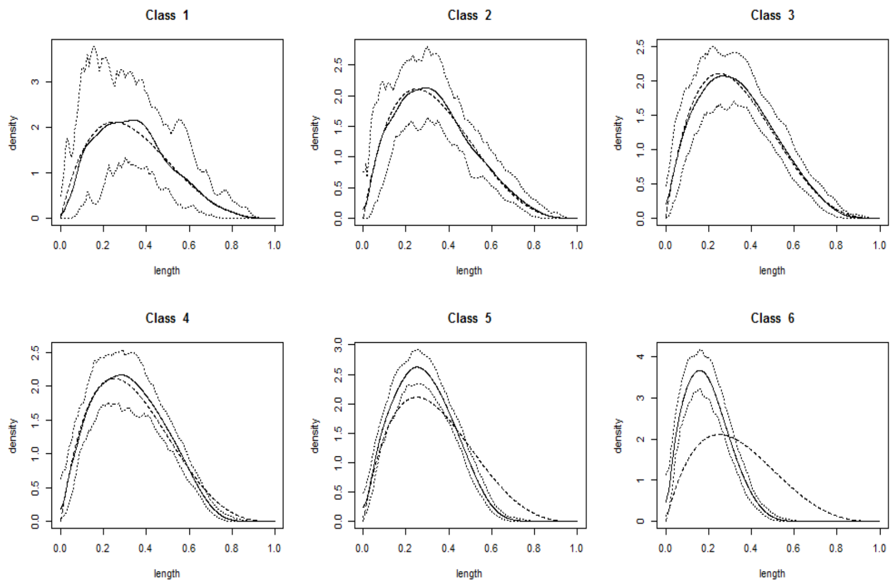


Figure 2.4: Estimated reference length density (sample II) of the inhomogeneous segment process $\mathbf{X}$, $b = 3, \tau = 900$. In each of six classes the average estimator of the observed length density (full line) is compared to the true reference density (dashed line) of beta distribution with parameters $\alpha = 2$, $\beta = 4$. The envelopes (dotted lines) correspond to empirical 90% confidence interval for the kernel estimator, pointwise in 1000 points on horizontal axis.

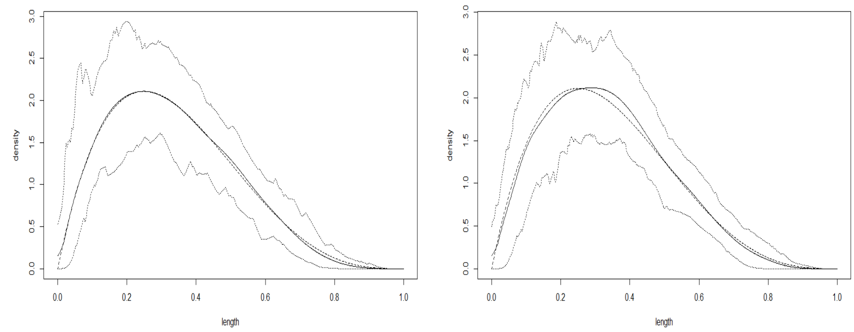


Figure 2.5: Semiparametric estimation of reference length density based on sample I (left), II (right) of 20, 40 simulations, respectively, of the segment process $\mathbf{X}$, $b = 3, \tau = 900$. The average (from first four classes) estimator of the reference density (full line) compared to the true reference density (dashed line) of beta distribution with parameters $\alpha = 2$, $\beta = 4$. The envelopes (dotted lines) correspond to empirical 90% confidence interval for the estimated reference density, pointwise in 1000 points on horizontal axis.

3. More complex models

In this chapter we introduce generalizations of the segment process model from Chapter 2. Models with some parameters are aimed for applications in biology. In a further research they should fit systems of stress fibers in stem cells. Here we present simulations of the models and find the set of parameters for which the density p is integrable.

3.1 Model I

Let $B, Y, L_0, \eta, n(\mathbf{x}), D(\mathbf{x}), d(u)$ and other labeling be as in previous chapter. Let the segment process $\mathbf{X}$ have density p with respect to η,

$$p(\mathbf{x}) = c1_{[\mathbf{x}\subset B]} \exp(bD(\mathbf{x}) + qH(\mathbf{x}))\tau^{n(\mathbf{x})} \prod_{x_i\in\mathbf{x}} g\left(\frac{r_i}{e_a}\right), \tag{3.1}$$

with parameters $b, q \in \mathbb{R}, \tau > 0, c$ is the normalizing constant, r_i is the length of i-th segment $u_i = (y_i, r_i, \varphi_i)$ and g is a reference probability density on L_0. Further

$$H(\mathbf{x}) = \sum_{u\in\mathbf{x}} \sin\gamma(u),$$

where $\gamma(u) \in [0, \frac{\pi}{2}]$, is the angle between the segment u and the tangent τ_u to the boundary ∂B of B at the point where the radius coming from the origin through the center y of u intersects ∂B.

To verify that p is integrable we use

$$p(\mathbf{x}) \leq p^{'}(\mathbf{x}) = c\exp\left(bD(\mathbf{x}) + qH(\mathbf{x}) + n(\mathbf{x})\log\tau + \sum_{x_i\in\mathbf{x}} \log g\left(\frac{r_i}{e_a}\right)\right),$$

where $p'(\mathbf{x})$ is in the exponential form. We apply the remark after Definiton 14, where

$$V(\mathbf{x}) = \left(D(\mathbf{x}), H(\mathbf{x}), n(\mathbf{x}), \sum_{x_i\in\mathbf{x}} \log g\left(\frac{r_i}{e_a}\right)\right), \quad \mathbf{h} = (b, q, \log\tau, 1).$$

Since coordinatewise

$$V(\mathbf{x}) \leq (n, n, n, c_2 n), \text{ for some } c_2 \in \mathbb{R},\ n = n(\mathbf{x}),$$

we have

$$\mathbb{E}[\exp(\mathbf{h}\cdot V(\eta))] =$$

$$\sum_{n=0}^{\infty} e^{-\lambda(Y)} \frac{\lambda^n(Y)}{n!} \int_Y \cdots \int_Y \exp\left(\mathbf{h}\cdot V(\mathbf{x})\right) \lambda(\mathrm{d}(u_1,\ldots,u_n)) \leq$$

$$\sum_{n=0}^{\infty} e^{-\lambda(Y)} \frac{\lambda^n(Y)}{n!} \exp\left(bn + qn + n\log\tau + c_2 n\right).$$

According to D'Alembert ratio criterion this series converges for $b, q \in \mathbb{R}, \tau > 0$.

The corresponding conditional intensity of the process $\mathbf{X}$ is

$$\lambda^*(\mathbf{x}, u) = 1_{[\mathbf{x} \cup u \subset B]} z g\left(\frac{r}{e_a}\right) \exp(bd(u) + q \sin\gamma(u)).$$

Using the conditional intensity of the process $\mathbf{X}$ with density with respect to unit Poisson process we simulate realizations the process using Algorithm 1 on page 8, see Fig. 3.1.

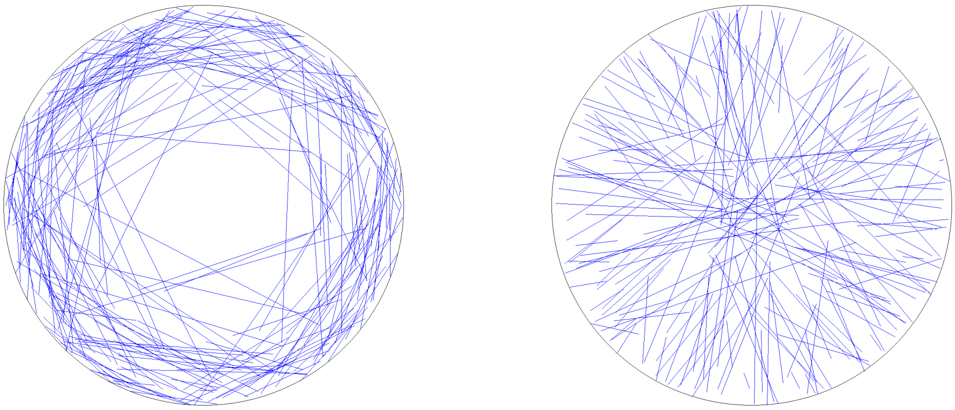


Figure 3.1: Simulated realisations of the process $\mathbf{X}$, model I, using Algorithm 1 ($I = 2 \cdot 10^7$ iterations) with $e_a = 1$ and the density (3.1) having parameters $(\tau, b, q) = (0.6, 20, 10)$ (left) and $(\tau, b, q) = (2 \cdot 10^{-6}, 27, -8)$ (right), g corresponds to beta distribution with parameters $\alpha = 2,\ \beta = 4$.

3.2 Model II

For the simulation of real data of stress fibres in stem cells ([7]) the following model is even more complex and comprises more variables. We no longer demand segments to be completely in the circle B. Let $\mathbf{X}$ be a segment process with a density

$$p(\mathbf{x}) = c_{\mathbf{x}} \exp(q_1 H_1(\mathbf{x}) + q_2 H_2(\mathbf{x}) + vS(\mathbf{x}) + aN(\mathbf{x}))\tau^{n(\mathbf{x})} \prod_{u \in \mathbf{x}} g\left(\frac{r}{e_a}\right) \quad (3.2)$$

with respect to the Poisson segment process η. Here $\tau > 0, a, v < 0, q_1, q_2 \in \mathbb{R}$ are parameters, $c_{\mathbf{x}}$ is a normalizing constant. Further let $\delta, \epsilon, \psi \geq 0$ be fixed constants. Denote

$$\text{dist}(u_1, u_2) =$$
$$\inf\{\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2};\ \vec{x} = (x_1, x_2) \in u_1, \vec{y} = (y_1, y_2) \in u_2\},$$

$u_1, u_2 \subset \mathbb{R}^2$, the distance between two sets in Euclidean space. Then we put

$$S(\mathbf{x}) = N(\{u_1,\ u_2 \in \mathbf{x};\ u_1 \neq u_2, \measuredangle(u_1, u_2) < \psi, \text{dist}(u_1, u_2) < \epsilon\})$$

$$N(\mathbf{x}) = N(\{u_1,\ u_2 \in \mathbf{x};\ u_1 \neq u_2, u_1 \cap u_2 \neq \emptyset\}),$$

where $N(D)$ is the number of elements in a finite set D and $\measuredangle(u_1, u_2)$ is the angle in $[0, \pi/2]$ between the lines we obtain if we extend the segments u_1, u_2.

Furthermore let the circle B be divided in an even number of sectors $B_i, i = 1, \ldots, 2m$, indexed in a clockwise direction. This means $B = C_1 \cup C_2$, where $C_1 = \bigcup_{i=1}^{m} B_{2i-1}$ is the union of all odd sectors and $C_2 = \bigcup_{i=1}^{m} B_{2i}$ is the union of all even sectors. Then we define

$$\Gamma(\mathbf{x}) = \{u \in \mathbf{x};\ \mathrm{dist}(\{y\}, \partial B) < \delta, u = (y, r, \varphi)\},$$

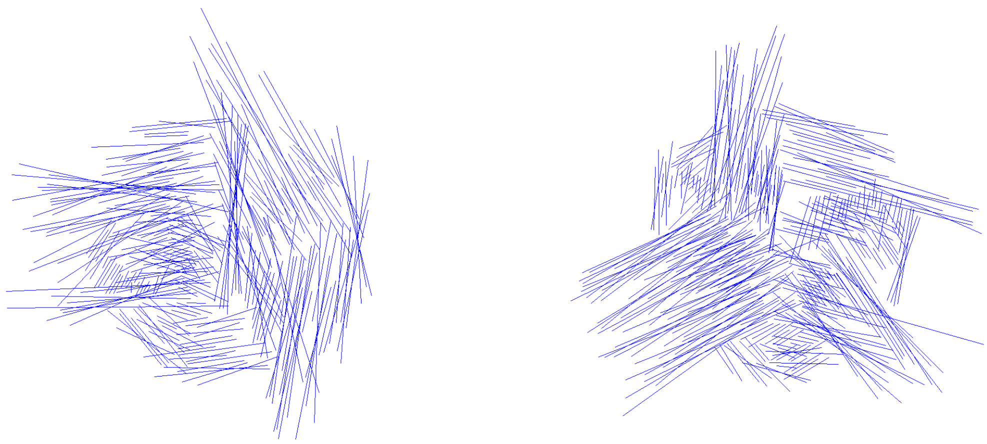
a subset of $\mathbf{x}$, and

$$H_1(\mathbf{x}) = \sum_{u \in \Gamma(\mathbf{x})} \sin(\gamma(u)) 1_{[y \in C_2]},$$
$$H_2(\mathbf{x}) = \sum_{u \in \Gamma(\mathbf{x})} \cos(\gamma(u)) 1_{[y \in C_1]}, \quad u = (y, r, \varphi).$$

Here $q_1 > 0$ applies only to even sectors where it favors segments u orthogonal to the tangent τ_u, $q_2 > 0$ applies to all odd sectors and favors segments u parallel to the tangent τ_u. Both $q_1, q_2 > 0$ favor segments $u \in Y$ for which $\mathrm{dist}(\{y\}, \partial B) < \delta$. Further $v < 0, a < 0$ inhibits the occurrence of segments which satisfy the conditions given in $N(\mathbf{x}), S(\mathbf{x})$, respectively. The sectors B_i may be set in the following way. We denote $w(\phi_1, \phi_2)$ the sector between the angles ϕ_1, ϕ_2, in a counterclockwise direction. Then in Fig. 3.2 we have

(a): $2m = 6$ sectors, $C_1 = w\left(-\frac{\pi}{6}, \frac{\pi}{3}\right) \cup w\left(\frac{\pi}{2}, \frac{5\pi}{6}\right) \cup w\left(\frac{7\pi}{6}, \frac{5\pi}{3}\right)$,

(b): $2m = 8$ sectors, $C_1 = w\left(-\frac{\pi}{6}, \frac{\pi}{12}\right) \cup w\left(\frac{\pi}{6}, \frac{\pi}{2}\right) \cup w\left(\frac{2\pi}{3}, \pi\right) \cup w\left(\frac{4\pi}{3}, \frac{5\pi}{3}\right)$



(a) (b)

Figure 3.2: Simulated realisations of the process $\mathbf{X}$ using Algorithm 1 ($I = 2 \cdot 10^7$ iterations) with $e_a = 1$ and the density (3.2) having parameters $(\tau, q_1, q_2, v, a) = (0.2, 20, 40, -20, -1)$ (left) and $(\tau, q_1, q_2, v, a) = (2, 20, 40, -20, -1)$ (right), g corresponds to beta distribution with parameters $\alpha = 2,\ \beta = 4$. The constants are $(\delta, \epsilon, \psi) = (0.4, 0.01, \pi/6)$.

Anallogically to Section 3.1 we can verify that the density p is well defined. In the exponential form of $p(\mathbf{x})$ we have

$$V(\mathbf{x}) = \left(H_1(\mathbf{x}), H_2(\mathbf{x}), S(\mathbf{x}), N(\mathbf{x}), n(\mathbf{x}), \sum_{x_i \in \mathbf{x}} \log g\left(\frac{r_i}{e_a}\right)\right),$$

$$\mathbf{h} = (q_1, q_2, v, a, \log\tau, 1).$$

It holds coordinatewise

$$V(\mathbf{x}) \leq (n, n, n^2, n^2, n, c_2 n), \text{ for some } c_2 \in \mathbb{R},\ n = n(\mathbf{x}),$$

therefore

$$\mathbb{E}[\exp(\mathbf{h} \cdot V(\eta))] \leq$$

$$\sum_{n=0}^{\infty} e^{-\lambda(Y)} \frac{\lambda^n(Y)}{n!} \exp\left(q_1 n + q_2 n + v n^2 + a n^2 + n \log\tau + c_2 n\right).$$

According to D'Alembert ratio criterion the series converges for $q_1, q_2 \in \mathbb{R}$, $v, a \leq 0, \tau > 0$.

The corresponding conditional intensity of the process $\mathbf{X}$ is

$$\lambda^*(\mathbf{x}, u) = \tau g\left(\frac{r}{e_a}\right) \exp(q_1 h_1(\mathbf{x}) + q_2 h_2(\mathbf{x}) + v s(\mathbf{x}) + a N'(\mathbf{x})), \tag{3.3}$$

where $h_1(\mathbf{x}), h_2(\mathbf{x}), s(\mathbf{x}), N'(\mathbf{x})$ are sums of those summands in $H_1(\mathbf{x}), H_2(\mathbf{x})$, $S(\mathbf{x}), N(\mathbf{x})$, where one of u_1, u_2 equals u. We use Algorithm 1 to simulate realizations of the process, see Fig. 3.2.

Conclusion

In this work, based on simulated data, we introduce a new approach to estimate the parameters and the density of reference length distribution of a segment process with a given probability density with respect to the Poisson process. The estimation procedure follows. We use characteristics of the process and discretize the model in several classes. In each class we evaluate a kernel density estimator of the Palm mark distribution. Combining the relation between the observed and reference length distribution and the maximum pseudolikelihood we compute the estimators of the scalar parameters. Then using these estimates we evaluate the estimator of the reference length density. The resulting density estimator is an average over some of the classes. We compare results of the procedure with true parameters and reference density of the model. In this work we also briefly introduce two more complex models. The motivation is to model stress fibers observed in cultured stem cells ([7]).

Bibliography

[1] Baddeley A (2007) Spatial point processes and their applications. In: Stochastic geometry. Ed by Weil W, Lecture Notes in Math, vol 1892, Springer, Berlin, 1–75.

[2] Chen SX (1999) Beta kernel estimators for density functions. Comput Statist Data Anal 31:131–145.

[3] Beneš V, Večeřa J, Eltzner B, Wollnik C, Rehfeldt F, Králová V,and Huckemann S (2017). Estimation of parameters in a planar segment process with a biological application. Image Analysis & Stereology, 36(1), 25-33.

[4] Beneš V, Večeřa J, and Pultar M (2017). Planar segment processes with reference mark distributions, modelling and estimation. arXiv:1701.01893v1 [math.ST], submitted.

[5] Stoyan D, Kendall W.S., Mecke J: Stochastic Geometry and its Applications. 2nd Ed., Wiley, New York, 1995.

[6] Moller, Jesper, and Rasmus Plenge Waagepetersen. Statistical inference and simulation for spatial point processes. CRC Press, 2003.

[7] Eltzner B, Wollnik C, Gottschlich C, Huckemann S, and Rehfeldt F (2016). The filament sensor for near real-time detection of cytoskeletal fiber structures. PLOS ONE 10(5):e0126346.