Estimation of intersection intensity in a Poisson process of segments

Tomáš Mrkvička

Abstract. The minimum variance unbiased estimator of the intensity of intersections is found for stationary Poisson process of segments with parameterized distribution of primary grain with known and unknown parameters. The minimum variance unbiased estimators are compared with commonly used estimators.

Keywords: complete statistic, Poisson process, segment process, sufficient statistic, intensity of intersections

Classification: Primary 60D05, Secondary 62B05

1. Introduction

There are several ways in statistics how to find an appropriate estimator for a characteristic. One of such concepts is to find a minimum variance unbiased estimator with the help of the complete and sufficient statistic [2]. This work uses this concept in spatial statistics, especially for a Poisson process of segments. We will suppose that the model is parameterized since otherwise, the complete and sufficient statistic would be whole observation and no reduction, by the complete and sufficient statistic, would be achieved. The Poisson process of segments Φ is determined by its intensity $\alpha \in \mathbb{R}^+$ and by a parametric distribution of primary grain $\Lambda_0(\theta)$, where $\theta \in \Theta \subset \mathbb{R}^k$ is a k-dimensional parameter of the primary grain distribution (the distribution of the typical segment). For detailed introduction of the Poisson segment process we refer to [1], [5] or [3]. Denote a realization of Φ by ϕ.

There are two cases which can be distinguished. First when θ is known and second when θ is unknown. The complete and sufficient statistic for α is the number of segments (generally compact sets when the Poisson process of compact sets is considered) visible in the observation window W in the first case [3].

This work was partly supported by the Grant Agency of the Czech Republic, Project No. 201/03/D062.

In the second case, it is not possible to find a complete and sufficient statistic for α and θ when all segments are considered. But when we consider only segments which have a reference point (the point which is uniquely determined for every segment, for example the lexicographic minimum point) inside the observation window W then the model could be considered as a marked Poisson process, where points of the process are the reference points and marks are corresponding lengths and orientations of the segments. When we know the lengths of all segments which are considered then the complete and sufficient statistic for α and θ is the vector consisting of the number of reference points and the complete and sufficient statistic for θ [4]. The last statistic is dependent on the family of distributions to which the primary grain distribution belongs.

The intensity of intersection is of interest for example in forestry where the intersection intensity of the fallen trees specifies the danger of the treatment with the fallen trees. The intensity of intersections can be estimated completely non-parametrically or we can use the formula derived in Section 2 which expresses the intensity of intersections by the model parameters. In Section 2, we introduce an estimator which plugs the estimators of the model parameters in the formula. This estimator uses the Poisson assumption, independence between lengths and direction and isotropy of the directions.

The minimum variance unbiased estimator is found in Section 3 for the case of known parameter θ and it is found in Section 4 for the case of unknown parameter θ.

The disadvantage of the minimum variance unbiased estimator is that we need to know all the segment lengths, the uncensored segments but censored too. This can happen for example in forestry where we can measure the length of the fallen tree even if it falls outside the border of our observation window. But more often we do not observe complete segments. In Section 5 the question how to deal with unknown segment lengths is discussed. We propose an estimator which is based on the minimum variance unbiased estimator but the complete and sufficient statistics on which this estimator is based are estimated from the observed data only. This estimator is compared with the estimators introduced in Section 2.

2. Intensity of intersections

Let $\mathcal{S}$ denote the set of all segments in $\mathbb{R}^d$. For $S \in \mathcal{S}$, let $c(S) \in S$ denote the lexicographic minimum point, the 'reference point' of S. Each segment is completely determined by its reference point $c(S)$, length $0 \leq r < \infty$ and direction $\beta \in \mathcal{U}_d$ ($\mathcal{U}_d$ is the space of all linear one-dimensional subspaces in $\mathbb{R}^d$).

A stationary Poisson process of segments is a stationary Poisson process Φ on $\mathcal{S}$. The primary grain distribution Λ_0 is a probability measure on $\mathcal{S}_0 = \{S \in \mathcal{S} : c(S) = 0\}$. $\mathcal{S}_0$ is isomorphic to the space $\mathcal{U}_d \times \mathbb{R}^+$. Let $\mathcal{D}$ and ρ denote the distribution of length and direction, respectively, of a typical segment (i.e. segment with distribution Λ_0). The segment $S_0 \in \mathcal{S}_0$ is determined by its length

r and by the direction β. We will denote such segment $S_0(r,\beta)$. The length and direction of a typical segment need not be independent random variables. We observe the realization of the Poisson process through a bounded observation window W. Denote the Lebesgue measure on $\mathbb{R}^d$ by λ^d and for shorter expressions denote the volume of W by $|W|$.

The intensity N of intersections in the plane is defined by the formula

$$N = \mathbb{E} \sum_{S \neq S', S, S' \in \Phi} \frac{I(S \cap S' \cap [0,1]^2 \neq \emptyset)}{2}.$$

Theorem 1. *Let Φ be a Poisson segment process in the plane with the directional distribution ρ independent of the length distribution $\mathcal{D}$. Then*

$$N = \frac{\alpha^2}{2} (\mathbb{E}r)^2 \int\int \sin(|\beta - \beta'|)\rho(d\beta)\rho(d\beta'), \tag{1}$$

where $\mathbb{E}r$ denotes the mean length of the typical segment. If Φ is isotropic then

$$N = \frac{\alpha^2}{\pi} (\mathbb{E}r)^2. \tag{2}$$

Proof: The point process of intersections can be defined by the following formula

$$\Xi(B) = \int\int_{S \neq S'} \frac{I(S \cap S' \cap B \neq \emptyset)}{2} \Phi(dS)\Phi(dS'),$$

where B is an arbitrary Borel set. Since Φ is a Poisson process, the expectation of $\Xi(B)$ can be expressed by

$$\mathbb{E}\Xi(B) = \int\int \frac{I(S \cap S' \cap B \neq \emptyset)}{2} \Lambda(dS)\Lambda(dS').$$

The measure $f_{S'}(B) = \int_{\mathcal{S}} \frac{I(S\cap S'\cap B\neq\emptyset)}{2}\Lambda(dS)$ has the property $f_{S'}(B) = f_{S'+x}(B+x)$ for $x \in \mathbb{R}^2$, thus we can write

$$\begin{aligned}
\mathbb{E}\Xi(B) &= \alpha \int\int f_{S_0'+z'}(B)\lambda^2(dz')\Lambda_0(dS_0') = \alpha \int\int f_{S_0'}(B - z')\lambda^2(dz')\Lambda_0(dS_0') \\
&= \alpha \int\int\int I_{B-z'}(y)\lambda^2(dz') f_{S_0'}(dy)\Lambda_0(dS_0') = \alpha\lambda^2(B) \int f_{S_0'}(\mathbb{R}^2)\Lambda_0(dS_0').
\end{aligned}$$

Since Ξ is stationary, the intensity of intersections can be expressed by

$$
\begin{aligned}
N &= \alpha \int\int \frac{I(S \cap S_0' \neq \emptyset)}{2} \Lambda(dS)\Lambda_0(dS_0') \\
&= \alpha^2 \int\int\int \frac{I((S_0+z) \cap S_0' \neq \emptyset)}{2} \lambda^2(dz)\Lambda_0(dS_0)\Lambda_0(dS_0') \\
&= \alpha^2 \int\int\int\int\int \frac{I((S_0(r,\beta)+z) \cap S_0'(r',\beta') \neq \emptyset)}{2} \lambda^2(dz)\rho(d\beta)\rho(d\beta')\mathcal{D}(dr)\mathcal{D}(dr') \\
&= \frac{\alpha^2}{2} \int\int\int\int rr' \sin(|\beta - \beta'|)\rho(d\beta)\rho(d\beta')\mathcal{D}(dr)\mathcal{D}(dr') \\
&= \frac{\alpha^2}{2} (\mathbb{E}r)^2 \int\int \sin(|\beta - \beta'|)\rho(d\beta)\rho(d\beta').
\end{aligned}
$$

Above we used the fact that the integral $\int I((S_0(r,\beta) + z) \cap S_0'(r',\beta') \neq \emptyset)\lambda^2(dz) = rr' \sin(|\beta - \beta'|)$.

If the process Φ is isotropic then the integral $\int\int \sin(|\beta - \beta'|)\rho(d\beta)\rho(d\beta') = \frac{2}{\pi}$. $\square$

It is possible to estimate N by a non-parametric estimator

$$
\widehat{N}_c(\phi) = \sum_{S \neq S', S, S' \in \phi} \frac{I(S \cap S' \cap W \neq \emptyset)}{2|W|}.
$$

The second possibility is to use the parametric approach. We can estimate the parameters and plug the estimates in Formula 1.

Suppose for the simplicity that Φ is isotropic and ρ and $\mathcal{D}$ are independent throughout the rest of this note. Then we can use Formula 2 and construct the estimator

$$
\widehat{N}_l(\phi) = \left(\frac{1}{|W|} \sum_{S \in \phi} L(S \cap W) \right)^2 / \pi,
$$

where $L(S \cap W)$ is the segment length visible in the observation window. This estimator uses only the information that the segment process is Poisson. This estimator is biased and asymptotically unbiased as $|W| \to \infty$.

The third possibility is to construct a minimum variance unbiased estimator which uses the information that the segment process is a Poisson process and the knowledge of a family of length distribution. The minimum variance unbiased estimator is found in Section 3 for the case of known parameters of the length distribution and it is found in Section 4 for the case of unknown parameters of the length distribution.

3. Known distribution of primary grain

We will consider the stationary Poisson segment process Φ in $\mathbb{R}^d$ with known primary grain distribution Λ_0 in this section. The problem of estimation of a real function N_α of the intensity parameter α is considered. The data are observed through a bounded measurable observation window $W \subseteq \mathbb{R}^d$. Let $\mathcal{W}$ be a measurable subset of $\mathcal{S}$. Let $\mathcal{E}_{\mathcal{W}}$ be the set of all estimators which depend only on segments from $\mathcal{W}$ (i.e., $e(\phi) = e(\phi \,|\, \mathcal{W})$).

Denote $C = \int_{\mathcal{U}_d \times \mathbb{R}^+} \lambda^d(W(r,\beta))\, \Lambda_0(d(r,\beta))$, where $W(r,\beta) = \{z : z + S(r,\beta) \in \mathcal{W}\}$ is the set of all reference points of shifts of the segment $S(r,\beta)$ which belong to $\mathcal{W}$, $S(r,\beta) \in \mathcal{S}_0$ denotes the segment of length r which starts from the origin under the direction β.

The following lemma is of value for computing the second moment of $\Phi(\mathcal{W})$, the proof can be found in [3].

Lemma 1. *Let Ψ denote a stationary Poisson point process on a Polish space $\mathcal{X}$. Let F, G denote the functions $F(\psi) = \int f(x)\ \psi(dx)$, $G(\psi) = \int g(x)\ \psi(dx)$, respectively, where f, g are nonnegative measurable functions on $\mathcal{X}$. Then*

$$\operatorname{cov}[F(\Psi), G(\Psi)] = \int f(x) g(x) \Lambda(dx).$$

Theorem 2. *Let Φ be an isotropic Poisson segment process in the plane with ρ and $\mathcal{D}$ independent. Then the estimator*

$$\widehat{N} = \frac{\Phi(\mathcal{W})^2 - \Phi(\mathcal{W})}{C^2 \pi} (\mathbb{E} r)^2$$

is the minimum variance unbiased estimator of the intensity N of the segments intersections among all estimators from $\mathcal{E}_{\mathcal{W}}$.

Proof: First compute the expectation of $\Phi(\mathcal{W})$

$$\begin{aligned}\mathbb{E}\Phi(\mathcal{W}) &= \int_{\mathcal{S}} I_{\mathcal{W}}(S)\ \Lambda(dS) = \alpha \int_{\mathcal{S}_0} \int_{\mathbb{R}^2} I_{\mathcal{W}}(x + S_0) dx \Lambda_0(dS_0) \\ &= \alpha \int_{\mathcal{U}_d \times \mathbb{R}^+} \lambda^d(W(r,\beta))\ \Lambda_0(d(r,\beta)) = \alpha C.\end{aligned}$$

To compute the expectation of $\Phi(\mathcal{W})^2$ we set $F(\phi) = \int I_{\mathcal{W}}(S)\ \phi(dS) = \phi(\mathcal{W})$. Using Lemma 1 we have

$$\operatorname{Var}(F(\Phi)) = \int_{\mathcal{S}} (I_{\mathcal{W}}(S))^2\ \Lambda(dS) = \alpha C.$$

Thus

$$\mathbb{E}\Phi(\mathcal{W})^2 = \mathrm{Var}(F(\Phi)) + (\mathbb{E}\Phi(\mathcal{W}))^2 = \alpha C + \alpha^2 C^2.$$

Now it is easy to show that $\widehat{N}$ is an unbiased estimator of N. Then Rao-Blackwell theorem with the fact that $\Phi(\mathcal{W})$ is the complete and sufficient statistic finishes the proof. $\square$

Lemma 1 can be easily extended for higher orders. The extensions for the third and fourth order help us to compute the variance of $\widehat{N}$

$$\mathrm{Var}(\widehat{N}) = \frac{4\alpha^3 C + 2\alpha^2}{C^2\pi^2}(\mathbb{E}r)^4.$$

There are two important examples of $\mathcal{W}$.

1. $\mathcal{W}_1 = \{S \in \mathcal{S} : c(S) \in W\}$, the set of all segments which have the reference point within the window W (in this case the constant $C = |W|$).
2. $\mathcal{W}_2 = \{S \in \mathcal{S} : S \cap W \neq \emptyset\}$, the set of all segments which hit the observation window W.

The estimator $\widehat{N}_2$ (i.e. the estimator $\widehat{N}$ in the case of $\mathcal{W}_2$) is the minimum variance unbiased estimator of all possible estimators. Under the assumption of Theorem 2 and for W being a square window with the side length a we can write

$$C = \int \lambda^2(W \oplus \check{S}_0)\Lambda_0(dS_0) = a^2 + \frac{4a}{\pi}\mathbb{E}r.$$

When it is not possible to compute C, we can use $\widehat{N}$ in the case of $\mathcal{W}_1$ for two different reference points. First we use $\Phi_1(W)$ the number of the lexicographic minimum points in W and second we use $\Phi_2(W)$ the number of the lexicographic maximum points in W. The resulting estimator will be the average of these two and it will be denoted by $\widetilde{N} = \frac{\Phi_1^2(W)-\Phi_1(W)+\Phi_2^2(W)-\Phi_2(W)}{2|W|^2\pi}(\mathbb{E}r)^2$. This estimator does not use only segments which have both end points outside W. Using Lemma 1, it is possible to compute

$$\mathrm{cov}(\Phi_1(W), \Phi_2(W)) = \Lambda(\{S \in \Phi : S \subseteq W\}) = \alpha \int \lambda^2(W \ominus \check{S}_0)\Lambda_0(dS_0) = \alpha D.$$

Now the variance of $\widetilde{N}$ can be computed by the extensions of third and fourth order of Lemma 1

$$\mathrm{Var}(\widetilde{N}) = \frac{2\alpha^3(|W|^3 + |W|^2 D) + \alpha^2(|W|^2 + D^2)}{|W|^4\pi^2}(\mathbb{E}r)^4.$$

Under the assumption of Theorem 2 and for W being a square window with the side length a we can write

$$D = \int \lambda^2(W \ominus \check{S}_0)\Lambda_0(dS_0) = a^2 - \frac{4a}{\pi}\mathbb{E}r + \frac{\mathbb{E}r^2}{\pi}, \quad \text{when} \quad r \le a \quad \text{a.s.}$$

and we can compute the ratio of the variances of the estimators $\widehat{N}_2$ and $\widetilde{N}$. The ratio is shown in Figure 1 for the uniform length distribution U(0,0.1).

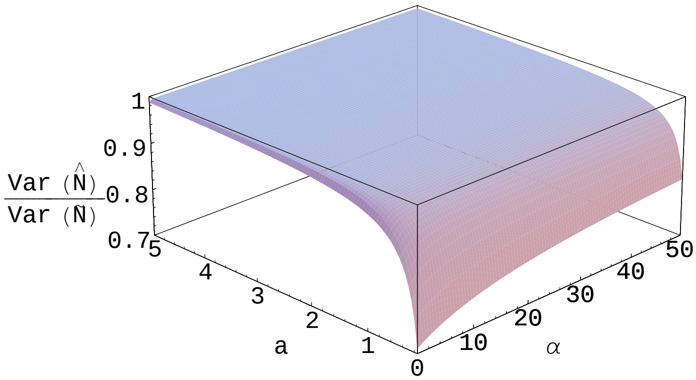

FIGURE 1: The ratio $\mathrm{Var}(\widehat{N}_2)/\mathrm{Var}(\widetilde{N})$ for the uniform length distribution with the dependence on the intensity α and the side length a of the square window W (a goes from 0.15 to 5).

Since we were not able to compute a formula for $\mathrm{Var}(\widehat{N}_c)$, we compared the estimators $\widehat{N}_c$ and $\widetilde{N}$ by simulations. We simulated the realizations of the Poisson stationary segment process with intensity α and with exponential length distributions with expected length $\mathbb{E}r$ in the observation window W which is a square with side length a. We chose 17 different parameters of the model (α, a) and for each parameter we did 1000 simulations. We fixed $\mathbb{E}r = 0.05$ in all cases to prevent same realizations with different scaling only. The resulted ratios $\mathrm{Var}(\widetilde{N})/\mathrm{Var}(\widehat{N}_c)$ are given in Table 1.

$\alpha \backslash a$	0.05	0.1	0.25	0.5	1	2
20			0.105	0.054	0.059	0.057
50			0.124	0.138	0.129	0.120
100			0.209	0.209	0.189	
200		0.334	0.277	0.262	0.274	
500	0.615	0.497				

TABLE 1: The ratio $\mathrm{Var}(\widetilde{N})/\mathrm{Var}(\widehat{N}_c)$ for the exponential length distribution with the dependence on the intensity α and window side length a.

The simulation shows that the ratio $\mathrm{Var}(\widetilde{N})/\mathrm{Var}(\widehat{N}_c)$ depends mostly on the intensity α when the expectation of lengths is fixed. The similar results were achieved for uniform length distribution.

Under the assumption of Theorem 2 and for W being a square window with the side length a we can compute the expectation of $\widehat{N}_l$. Using the formula for variance of $\frac{1}{|W|}\sum_{S\in\phi} L(S\cap W)$ computed in [3] we have that

$$\frac{\mathbb{E}\widehat{N}_l}{N} = 1 + \frac{a^2\mathbb{E}r^2 - \frac{4}{3}a\mathbb{E}r^3 + \frac{1}{6}\mathbb{E}r^4}{a^4\alpha(\mathbb{E}r)^2}\,.$$

The bias for exponential length distribution is presented in Figure 2.

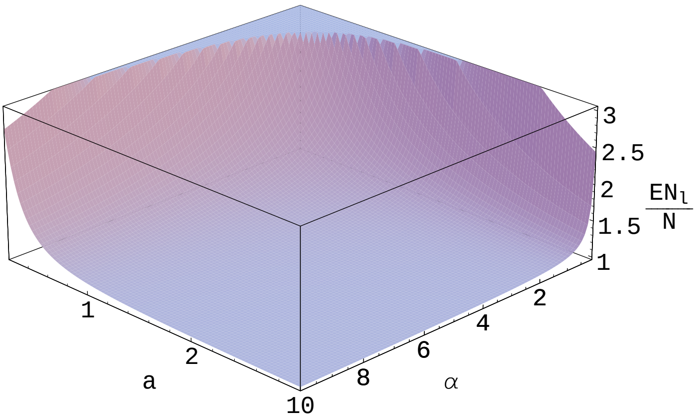

FIGURE 2: The ratio $\mathbb{E}\widehat{N}_l/N$ for the exponential length distribution with the dependence on the intensity α and the side length a of the square window W.

The simulations also show that the ratio $\operatorname{Var}(\widetilde{N})/\operatorname{Var}(\widehat{N}_l)$ depends mainly on the mean number of segments $\alpha|W|$ in the observation window when the expectation of lengths is fixed. The graph of the ratio $\operatorname{Var}(\widetilde{N})/\operatorname{Var}(\widehat{N}_l)$ has approximately the same shape for different distributions but the limit of the ratio as the mean number of segments goes to the infinity is different and it is equal to

$$\lim_{\alpha|W|\to\infty}\frac{\operatorname{Var}(\widetilde{N})}{\operatorname{Var}(\widehat{N}_l)} = \frac{(\mathbb{E}r)^2}{\mathbb{E}r^2}\,.$$

Here W is a convex set. The same ratio is reached by the ratio of the variance of the minimum variance unbiased estimator of length density L_A and the variance of the estimator of L_A which sums all lengths visible in W [3]. For example: exponential distribution gives ratio $1/2$, uniform distribution with parameters (A,B) gives ratio from the interval $(0,1]$ or lognormal distribution with parameters (μ, σ) gives ratio equal to $1/\exp(\sigma^2)$. The graph of the dependence for exponential length distribution fitted from the simulations is presented in Figure 3.

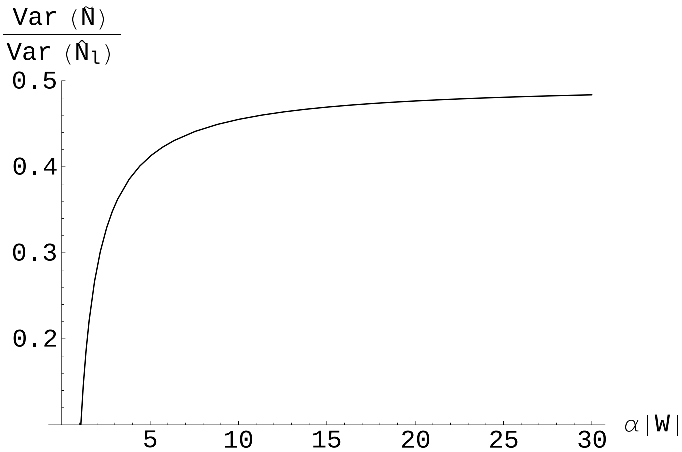

Figure 3: The ratio $\mathrm{Var}(\widetilde{N})/\,\mathrm{Var}(\widehat{N}_l)$ for the exponential length distribution with the dependence on the mean number of segments $\alpha|W|$ in the observation window.

4. Unknown parameters of distribution of primary grain

We will consider the stationary Poisson segment process Φ in $\mathbb{R}^d$ with known primary grain distribution $\Lambda_0(\theta)$ with unknown parameter θ observed through a bounded window W in this section. Generally some of the segments are censored by the border of the window W. Assume that we know the censored segments lengths in this section. Then we observe the marked Poisson point process where points of the process are the reference points and marks are corresponding lengths and orientations of the segments [4]. Then a statistic $S_T = \binom{\Phi(W)}{T_{\Phi(W)}}$ is a complete and sufficient statistic for α, θ [4]. Here $\Phi(W)$ is the number of the reference points in the observation window and $T_{\Phi(W)}$ is the complete and sufficient statistic for θ. This statistic is different for different primary grain distributions. Like in the previous section if we find a function of S_T which expectation is $\tau(\alpha, \theta)$ then such estimator will be the minimum variance unbiased estimator of $\tau(\alpha, \theta)$ among all estimators from $\mathcal{E}_{\mathcal{W}_1}$.

Theorem 3. *Let Φ be an isotropic Poisson segment process in the plane with ρ and $\mathcal{D}$ independent. Then the estimator*

$$\widehat{N}^{MVUE} = \frac{\Phi(W)^2 - \Phi(W)}{|W|^2\pi} \cdot e_{2,\Phi(W)}$$

is the minimum variance unbiased estimator of the intensity of the segments intersections N among all estimators from $\mathcal{E}_{\mathcal{W}_1}$ under the condition that we know the censored segments lengths. Here $e_{2,\Phi(W)}$ is the minimum variance unbiased estimator of $(\mathbb{E}_\theta r)^2$.

Proof: It is easy to show that $\widehat{N}^{MVUE}$ is an unbiased estimator of N. Thus Rao-Blackwell theorem finishes the proof. □

5. Unknown length of censored segments

We do not know very often all the segment lengths. Thus it is a question how to take advantages of these estimators in practice. It was shown that when we use minus sampling (we take into account only segments which are whole observable) the minimum variance unbiased estimator could have lower variance then the common estimator $\widehat{L}_A = \frac{1}{|W|}\sum_{S\in\phi} L(S\cap W)$ in the case of estimating the length density [4]. But this approach works only when the segment length distribution is bounded, like in the case of the uniform length distribution. When the distribution is not bounded then there may appear very long segments and we lose too much information. Therefore we found a different approach which works for both types of distributions and even more it gives better results than the minus sampling approach for bounded distributions.

The minimum variance unbiased estimator is based on measuring of the complete, sufficient statistic S_T then the statistic is plugged in the formula of estimator. Our approach takes the most efficient unbiased estimator of S_T, which is available from the data, and then it is plugged in the formula of minimum variance unbiased estimator pretending that S_T was measured from whole observation and that no censoring of segments appeared.

The number of reference points $\Phi(W)$ which is the first part of S_T does not change in this approach. We use first the lexicographic minimum of the segment as the reference point and compute the estimator with these reference points, then we use the lexicographic maximum of the segment as the reference point and compute the estimator with these reference points. At the end these two estimators are averaged. We will denote this kind of estimators by $\overline{N}$.

The statistic $T_{\Phi(W)}$ which is the second part of S_T is different for different families of distributions. Suppose now that the length distribution and distribution of orientation are independent and the process is isotropic. Thus the unknown parameter θ is reduced to the parameter of length distribution. We investigated two families of length distributions. We compare the estimators $\overline{N}$ with $\widehat{N}_c$ and $\widehat{N}_l$ by simulations in the following subsections. We simulate the realizations of the Poisson stationary, isotropic segment process with intensity α in the observation window W which is a square with side length a. We chose 16 different parameters of the model and for each parameter we did 1000 simulations. We fixed $\mathbb{E}r = 0.05$ in all cases to prevent same realizations with different scaling only. The parameters of the length distribution were chosen to fit the condition $\mathbb{E}r = 0.05$.

The simulations show that the ratio $\operatorname{Var}(\overline{N})/\operatorname{Var}(\widehat{N}_l)$ (Figures 4, 5) and the bias of $\widehat{N}_l$ depend mainly on the mean number of segments $\alpha|W|$ in the observation window when the expectation of lengths is fixed. The bias of $\widehat{N}_l$ is very similar for different length distribution. The bias for exponential length distribution is presented in Figure 2.

The simulations also show that the ratio $\operatorname{Var}(\overline{N})/\operatorname{Var}(\widehat{N}_c)$ (Tables 2, 3) depends mainly on the intensity α when the length distribution is fixed (similarly as in Table 1).

All graphs are nonlinear regression fits of the simulated variances ratios.

5.1 Uniform (0,A) length distribution

First we compare the estimators for the uniform distribution which is bounded and the complete and sufficient statistic is different than $\sum r_i$ (the base of the estimator $\widehat{N}_l$).

The complete and sufficient statistic for A is $T = \max_{i=1,\dots,\Phi(W)} r_i$. This statistic is clearly estimated by the longest visible part of the segment. The minimum variance unbiased estimator of $\mathbb{E}_\theta r$ is $e_{\Phi(W)} = \frac{\Phi(W)+1}{2\Phi(W)}T$ and the minimum variance unbiased estimator of $(\mathbb{E}_\theta r)^2$ is $e_{2,\Phi(W)} = \frac{\Phi(W)+2}{4\Phi(W)}T^2$.

$\alpha \backslash a$	0.25	0.5	1	2	3
20		0.072	0.063	0.055	0.053
50	0.130	0.141	0.12725	0.126	0.130
100	0.291	0.265	0.238		
200	0.386	0.363	0.431		
500	0.543				

TABLE 2: The ratio $\operatorname{Var}(\overline{N})/\operatorname{Var}(\widehat{N}_c)$ with the dependence on the intensity α and window side length a for uniform distribution and unknown parameter A.

5.2 Exponential (μ) length distribution

Now we compare the estimators for the exponential distribution which is unbounded and the complete and sufficient statistic is $\sum r_i$ (the base of the estimator $\widehat{N}_l$). By a minus sampling method no improvement would be achieved in this case. The similar results were achieved for lognormal distribution too.

The complete and sufficient statistic for μ is $T = \sum_{i=1}^{\Phi(W)} r_i$. This statistic is estimated by the sum of all segment lengths but when the segment is censored

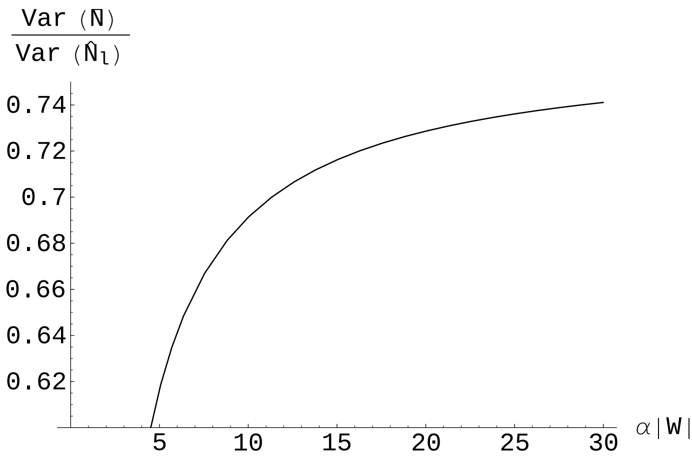

Figure 4: The ratio $\operatorname{Var}(\overline{N})/\operatorname{Var}(\widehat{N}_l)$ with the dependence on the mean number of segments $\alpha|W|$ in the observation window for uniform distribution and unknown parameter A.

we add to the visible part of the segment the expectation of the censored length with respect to the exponential distribution with parameter μ. The parameter μ was first estimated by $\widehat{L}_A/\Phi(W)$.

$$e_{\Phi(W)} = \frac{T}{\Phi(W)}, \quad e_{2,\Phi(W)} = \frac{T^2}{\Phi(W)^2 + \Phi(W)}.$$

$\alpha \backslash a$	0.25	0.5	1	2	3
20		0.099	0.109	0.104	0.128
50	0.457	0.276	0.243	0.243	0.246
100	0.518	0.385	0.357		
200	0.615	0.557	0.521		
500	0.773				

Table 3: The ratio $\operatorname{Var}(\overline{N})/\operatorname{Var}(\widehat{N}_c)$ with the dependence on the intensity α and window side length a for exponential length distribution with unknown parameter μ.

6. Conclusion

We can see from the simulations that it is much better to estimate N by using Formulas derived in Section 2 than to use the complete non-parametric

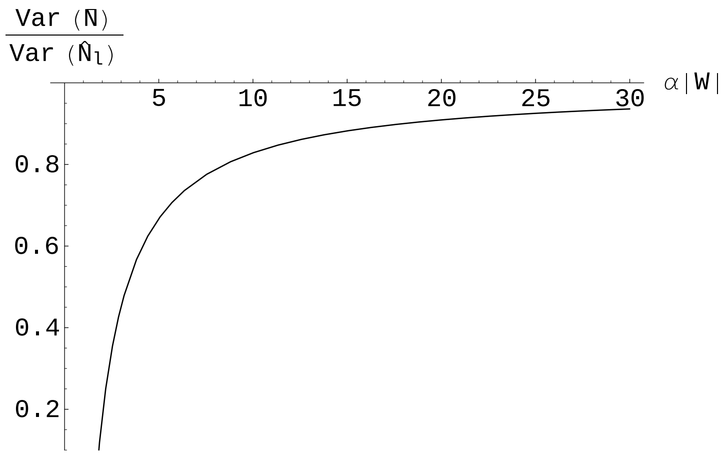

Figure 5: The ratio $\mathrm{Var}(\overline{N}) / \mathrm{Var}(\widehat{N}_l)$ with the dependence on the mean number of segments $\alpha|W|$ in the observation window for exponential distribution with unknown parameter μ.

estimator $\widehat{N}_c$ when the Poisson assumption, independence between lengths and direction and isotropy of the directions are satisfied. Furthermore when we know the distribution of the primary grain, it is better to use the estimator $\widetilde{N}$ because it is unbiased and it has lower variance than the estimator $\widehat{N}_l$. The estimators $\widehat{N}_l$ and $\overline{N}$ are asymptotically equivalent in the case of exponential distribution.

The simulations show that the profit gained by using complete and sufficient statistic is bigger than the loss caused by avoiding some information visible in the observation window. It also shows that using information about the primary grain distribution can bring a big improvement on the efficiency of the characteristics estimators.

Acknowledgments. I would like to thank the referee for his helpful comments.

References

[1] Daley D. J., Vere-Jones D., *An Introduction to the Theory of Point Processes*, Springer, New York, 1998.

[2] Lehmann E.L., *Theory of Point Estimation*, Wadsworth & Brooks, California, 1991.

[3] Mrkvička T., *Estimation variances for Poisson process of compact sets*, Adv. Appl. Prob. (SGSA) **33** (2001), 765–772.

[4] Mrkvička T., *Estimation variances for parameterized marked point processes and for parameterized Poisson segment processes*, Comment. Math. Univ. Carolin. **45** (2004), no. 1, 109–117.

[5] Stoyan D., Kendall W.S., Mecke J., *Stochastic Geometry and its Applications* (2nd Ed.), John Wiley & Sons, Chichester, 1995.

Mathematical Department, University of South Bohemia, Jeronýmova 10,
371 15 České Budějovice, Czech Republic

E-mail: mrkvicka@pf.jcu.cz

(*Received* June 26, 2005, *revised* August 18, 2006)