Mean value formulae of one-dimensional stationary line-segment processes

Albert K.L. Tsang
Census and Statistics Department, Hong Kong, China
e-mail: akl_tsang@yahoo.com

Abstract

Random processes of line segments embedded in $\mathbb{R}^2$ and $\mathbb{R}^3$ have been studied for a number of physical applications, e.g. for the fiber structure of papers and textiles. Pioneers such as Coleman, Parker and Cowan had considered a model for random line-segment process in $\mathbb{R}^2$ which makes very weak assumptions concerning the stationarity and isotropy only in the 1970's. This paper studies the parallel model for a line-segment process in $\mathbb{R}^1$. Mean value formulae for geometric variables are derived, relying only on very mild stationarity assumption. When interpreting $\mathbb{R}^1$ as the time axis, the one-dimensional line-segment process is actually a natural generalization of the stationary point process: for which an event occurs not just at a single time point, but may exist for a certain period. The model can thus be applied in modelling the occurrence and existence of some natural phenomena.

Keywords: Stochastic process, stationary point process, stochastic geometry

1 Introduction

Random processes of line segments embedded in $\mathbb{R}^2$ and $\mathbb{R}^3$ have been studied for a number of physical applications, e.g. for the fiber structure of paper and textiles. Coleman (1972, 1974) gives a mathematical model which is sufficiently general to incorporate the common applications. Parker and Cowan (1976) and Cowan (1979) consider a model for a random line-segment process in $\mathbb{R}^2$ which makes very weak assumptions concerning the stationarity and isotropy of the process only. In this paper we study the parallel model for a stationary line-segment process in $\mathbb{R}^1$.

Each line-segment in one dimension can be parametrised by the position of its left end-point $x \in \mathbb{R}^1$, and its length $l \in [0, \infty)$. So a one dimensional line-segment process is effectively a point process in the space $Z = \mathbb{R}^1 \times [0, \infty)$. We call this the *associated point process* (APP).

We commence with only the stationarity assumption of the APP in $\mathbb{R}^1$. That is, the point process of left end-points is stationary in $\mathbb{R}^1$ with intensity τ. Note that the choice of left end-points is arbitrary and only for convenience. One may choose any well-defined points (for example the mid-points or right end-points) as references of positions.

2 Integral geometric considerations

First of all, we will convert the arguments adopted in Cowan (1979) for 2-dimensional line-segment processes into the 1-dimensional setting. Consider the APP on Z. For any Borel subset $U \subseteq Z$ and any realization ω of this point process, we denote the number of points in U by $K(U, \omega)$. By the stationarity assumption, $\mathbb{E}K(\cdot)$ must be a measure on Z, invariant under the group of translations in $\mathbb{R}^1$. Using a product lemma of Cowan (1979), with the stationarity assumption alone, we have

$$\mathbb{E}K(U) = \int\int_{(x,l)\in U} \tau \, dx \, dG(l), \tag{1}$$

for some function G. For any $A \in \mathbb{R}^1$ with Lebesgue measure $\mathcal{L}$, $\tau\mathcal{L}G(l)$ can be interpreted as the expected number of line-segments having left-end points in A, length $\leq l$. However G cannot be interpreted as the length distribution of line-segments directly, unless we clearly define the meaning of a typical line-segment sampled through some special method, for example the ergodic method discussed in Cowan (1978, 1980), and prove that G has this interpretation. We shall prove that G is the length distribution of a typical line-segment defined in the ergodic way.

The calculation of $\mathbb{E}K(\cdot)$ by (1) for suitable Borel subsets in Z yields a number of interesting mean value formulae. Let A be an interval of length $\mathcal{L}$ in $\mathbb{R}^1$. Without loss of generality, assume $A = (0, \mathcal{L})$. Define (taking the notation # as *the number of*)

$$\begin{aligned}
L(A) &= \#\text{ left end-points within } A, \\
X(A) &= \text{total length of these } L(A) \text{ segments}, \\
N(A) &= \#\text{ segments that intersect with } A, \\
Y(A) &= \text{total length of these } N(A) \text{ segments}, \\
S(A) &= \text{total length within } A \text{ of these } N(A) \text{ segments}, \\
W(A) &= \#\text{ segments wholly contained in } A, \\
C(A) &= \#\text{ segments covering } A, \\
Z_R(A) &= \#\text{ segments covering the right end-point of } A\text{, but not the left end-point}, \\
Z_L(A) &= \#\text{ segments covering the left end-point of } A\text{, but not the right end-point.}
\end{aligned}$$

Since the APP is also a marked point process embedded in $\mathbb{R}^1$, recall the Campbell's theorem concerning marked point processes, which in our context means:

Theorem 1. (Campbell) *For any measurable function $f : Z \to \mathbb{R}^1$, and any Borel subset U of Z,*

$$\mathbb{E} \sum_{(x,l)\in U\cap\omega} f(x,l) = \int\int_U \tau f(x,l)\, dx\, dG(l),$$

where ω runs through all realisations of the APP.

Using theorem 1, as well as (1), expected values of the above random variables can be found easily when $A = (0, \mathcal{L})$.

Substituting $U = \{(x,l) : l \in (0,\infty), x \in (0,\mathcal{L})\}$ and $f(x,l) = 1$ into theorem 1 yields

$$\mathbb{E}L(A) = \int_0^\infty \int_0^{\mathcal{L}} \tau\, dx\, dG(l) = \tau\mathcal{L}. \tag{2}$$

Applying the same U, but taking $f(x,l) = l$, we have

$$\mathbb{E}X(A) = \int_0^\infty \int_0^{\mathcal{L}} \tau l\, dx\, dG(l) = \tau\mathcal{L}\nu, \tag{3}$$

where $\nu = \int_0^\infty l\, dG(l)$, a quantity which we shall be able to interpret as the mean length of the typical line-segment once we establish that G is the length distribution in the next section.

Taking $U = \{(x,l) : l \in (0,\infty), x \in (-l,\mathcal{L})\}$ and $f(x,l) = 1$ gives

$$\mathbb{E}N(A) = \int_0^\infty \int_{-l}^{\mathcal{L}} \tau\, dx\, dG(l) = \tau(\nu + \mathcal{L}). \tag{4}$$

Putting the same U but $f(x,l) = l$ gives

$$\begin{aligned} \mathbb{E}Y(A) &= \int_0^\infty \int_{-l}^{\mathcal{L}} \tau l \, dx \, dG(l) \\ &= \tau(\int_0^\infty l^2 \, dG(l) + \nu\mathcal{L}) \\ &= \tau(\sigma^2 + \nu^2 + \nu\mathcal{L}), \end{aligned} \tag{5}$$

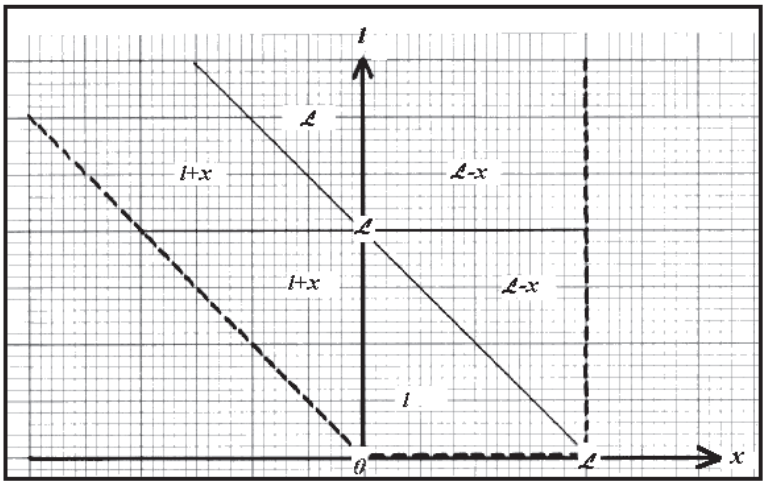
where $\sigma^2 = \int_0^\infty (l-\nu)^2 \, dG(l)$.

Considering $U = \{(x,l) : l \in (0,\infty), x \in (-l,\mathcal{L})\}$ as above, and $f(x,l) =$ *length of the line-segment* (x,l) *within* A, we obtain

$$\begin{aligned} \mathbb{E}S(A) = & \int_0^{\mathcal{L}} \int_{\mathcal{L}-l}^{\mathcal{L}} \tau(\mathcal{L}-x) \, dx \, dG(l) + \int_{\mathcal{L}}^\infty \int_0^{\mathcal{L}} \tau(\mathcal{L}-x) \, dx \, dG(l) \\ + & \int_{\mathcal{L}}^\infty \int_{\mathcal{L}-l}^0 \tau\mathcal{L} \, dx \, dG(l) + \int_0^{\mathcal{L}} \int_0^{\mathcal{L}-l} \tau l \, dx \, dG(l) \\ + & \int_0^{\mathcal{L}} \int_{-l}^0 \tau(l+x) \, dx \, dG(l) + \int_{\mathcal{L}}^\infty \int_{-l}^{\mathcal{L}-l} \tau(l+x) \, dx \, dG(l) \\ = & \tau\nu\mathcal{L}, \end{aligned} \tag{6}$$

(See Figure 1 for details of the integrals on the right-hand side).

Figure 1: Length of line-segments (x,l) within A.



Setting $U = \{(x,l) : l \in (0,\mathcal{L}), x \in (0,\mathcal{L}-l)\}$ and $f(x,l) = 1$ gives

$$\mathbb{E}W(A) = \int_0^{\mathcal{L}} \int_0^{\mathcal{L}-l} \tau \, dx \, dG(l) = \tau[\mathcal{L}\, G(\mathcal{L}) - \mathcal{L}\, G(0) - \int_0^{\mathcal{L}} l \, dG(l)]. \tag{7}$$

Putting $U = \{(x,l) : l \in (\mathcal{L},\infty), x \in (\mathcal{L}-l, 0)\}$ and $f(x,l) = 1$ yields

$$\mathbb{E}C(A) = \int_{\mathcal{L}}^\infty \int_{\mathcal{L}-l}^0 \tau \, dx \, dG(l) = \tau[\int_{\mathcal{L}}^\infty l \, dG(l) - \mathcal{L}(1 - G(\mathcal{L}))]. \tag{8}$$

For $U = \{(x,l) : x \in (0,\mathcal{L}), l \in (\mathcal{L}-x, \infty\}$ and $f(x,l) = 1$ the Campbell's theorem gives

$$\mathbb{E}Z_R(A) = \int_0^\infty \int_{\mathcal{L}-x}^\infty \tau \, dG(l) \, dx = \tau[\mathcal{L} - \int_0^{\mathcal{L}} G(x) \, dx]. \tag{9}$$

Note that by symmetry, $Z_L(A)$ has the same expected value as $Z_R(A)$.

We have shown that the above formulas hold for convex A (i.e. an interval). However it is clear that since $\mathbb{E}L(\cdot), \mathbb{E}X(\cdot)$ and $\mathbb{E}S(\cdot)$ are measures in $\mathbb{R}^1$, (2), (3) and (6) can be extended to any Borel set A immediately.

3 Ergodic definition and results

We shall now show that the mean length of line-segments defined by the ergodic approach equals ν, the constant obtained from the function G.

Intuitively, the ergodic method considers the objects within the observation window B_r (centred at origin o), including those truncated by the boundary of the window B_r. A 'typical object' is thus defined as the object sampled equally likely from those objects being observed (finitely many), and then taking the limit as $r \to \infty$. The distributions, moments etc. of the 'typical object' are therefore the limiting distributions, moments etc. of the sampled object when taking the limit as $r \to \infty$.

Mathematically, ergodicity of a line-segment process (or similarly other stochastic processes) is defined as follows:

Definition 1. *Let $M : (\Omega, \mathcal{S}, \mathcal{P}) \to (\mathbb{M}, \mathcal{M}, P)$ be a line segment process, and $T^{\boldsymbol{a}} : \Omega \to \Omega$ the translation defined as $T^{\boldsymbol{a}}\boldsymbol{x} = \boldsymbol{x} + \boldsymbol{a}$. for any $\boldsymbol{a} \in \mathbb{R}$ and $\boldsymbol{x} \in \Omega$. $T^{\boldsymbol{a}}$ is said to be ergodic if*

$$T^{\boldsymbol{a}}E = E \Rightarrow \mathcal{P}(E) = 0 \text{ or } 1.$$

Ergodicity is a very natural assumption, particularly if we recall that it is a consequence of the somewhat stronger mixing condition, which means intuitively that two events connected with the stochastic process tend to be independent if they refer to features which are at considerable spatial distance.

In our context, we assume that the line-segment process under consideration is ergodic. The most important reason for making the ergodic assumption is to allow the application of the Wiener ergodic theorem (Wiener (1939)):

Theorem 2. **(Wiener)** *If $T^{\boldsymbol{a}}$ is a measure-preserving translation from $\Omega \to \Omega$ for some $\boldsymbol{a} \in \mathbb{R}$, X is any random variable such that $\mathbb{E}|X| < \infty$, and V_d is the Lebesgue measure of the closed ball B_r of radius r, centre $\boldsymbol{o}$ in $\mathbb{R}^d$, then for almost all realisations ω of X,*

$$\lim_{r\to\infty} V_d^{-1} \int_{B_r} X(T^{-\boldsymbol{a}}\omega)\, d\boldsymbol{a}$$

exists. Furthermore, if $T^{\boldsymbol{a}}$ is ergodic then the limit equals $\mathbb{E}(X)$ almost surely.

Denote B_r the 1-dimensional ball with radius r and $T^t\omega$ the translation of the realization ω by an amount t. By standard argument it is easy to see that

$$\int_{B_{r-y}} S(B_y, T^t\omega)dt \le 2y\, S(B_r) \le \int_{B_{r+y}} S(B_y, T^t\omega)dt.$$

Thus by Wiener ergodic theorem,

$$\frac{S(B_r)}{2r} \xrightarrow{\text{a.s.}} \tau\nu \text{ as } r \to \infty, \tag{10}$$

provided that $\mathbb{E}S(B_y)$ is finite for all y.

Next we shall prove that

$$\frac{N(B_r) - W(B_r)}{2r} \xrightarrow{\text{a.s.}} 0, \tag{11}$$

that is, the normalized number of segments cutting the boundary converges almost surely to zero. This can be confirmed by the following inequality

$$N(B_r)-W(B_r) \leq y^{-1}[S(B_{r+y})-S(B_r)]+[L(B_{r+y})-L(B_r)]+[R(B_{r+y})-R(B_r)], \tag{12}$$

where $R(B_r) = \#$ right end-points within B_r. It is clear that both $L(B_r)$ and $R(B_r)$ converge to τ almost surely after normalization by $2r$ (simply consider a similar argument to arrive at 10). Applying (10) and this fact to (12), which when divided by $2r$, confirms (11).

Now

$$N(B_r) - W(B_r) = Z_L(B_r) + Z_R(B_r) + C(B_r) \tag{13}$$

Obviously by (11),

$$\frac{Z_L(B_r) + Z_R(B_r)}{2r} \xrightarrow{\text{a.s.}} 0, \tag{14}$$

$$\frac{C(B_r)}{2r} \xrightarrow{\text{a.s.}} 0. \tag{15}$$

Observe that

$$L(B_r) + R(B_r) = Z_L(B_r) + Z_R(B_r) + 2W(B_r).$$

By the above-mentioned fact about L and R, the left-hand side of this equation converges almost surely to 2τ when normalized by $2r$. Hence

$$\frac{W(B_r)}{2r} \xrightarrow{\text{a.s.}} \tau.$$

Therefore

$$\frac{N(B_r)}{2r} \xrightarrow{\text{a.s.}} \tau. \tag{16}$$

By (10) and (16)

$$\frac{S(B_r)}{N(B_r)} \xrightarrow{\text{a.s.}} \nu. \tag{17}$$

Thus the ergodic definition of mean segment-length is consistent with ν.

To show that $G(l)$ $(l \geq 0)$ is the length distribution of a typical line segment in the ergodic sense, consider the random variables $N^l(B_r), L^l(B_r), R^l(B_r), \ldots$ (the superscript l is used to denote the variables corresponding to the 'segments in B_r with complete length $\leq l$'). The arguments shown above in this section can be reproduced without any difficulty to these variables.

Only notice

$$L^l(A) = \int_0^l \int_0^{\mathcal{L}} \tau dx\, dG(r) = \tau \mathcal{L} G(l),$$

for $A = (0, \mathcal{L})$, we can conclude that

$$\frac{L^l(B_r)}{2r} \xrightarrow{\text{a.s.}} \tau G(l) \qquad \text{as } r \to \infty.$$

Similarly $\frac{R^l(B_r)}{2r}$ converges to the same limit. Hence

$$\frac{N^l(B_r)}{2r} \xrightarrow{\text{a.s.}} \tau G(l).$$

The length distribution in ergodic sense of a typical line segment is the almost sure limit

$$\lim_{r\to\infty} \frac{N^l(B_r)}{N(B_r)} = G(l),$$

which proves the above claim that G is the desirable length distribution.

4 Applications and way forward

The one-dimensional line-segment process is a natural generalization of the one-dimensional point process. If interpreting $\mathbb{R}^1$ as the time axis, the line-segment process can be applied to mathematical modelling of some natural phenomena, for which an event occurs not just at a single time point, but exists for a certain period of time.

For example, the occurrence and extinction of species of living organisms (the line-segment signifies the existence of the species on earth), the lifespan of individuals of a certain kind of animals, the duration of investors holding a certain kind of stock etc. may be modelled by the theory. Of course, stationarity and ergodicity assumptions have to be satisfied before the relevant results can be applied.

Another direction of further study is to apply weights (to account for, say, the population of the species or the value of stock being held in the above examples), to the line-segments. The weights may be an independent random variable following a certain distribution. The method of study will be much similar to the discussions in this paper, except that the underlying space (the APP) should be raised to one higher dimension to capture the distribution of the weights. This deserves another study and will not be discussed in this paper.

References:

Coleman R. (1972) “Sampling procedures for the lengths of random straight lines”, *Biometrika*, 59, 415-426.

Coleman R. (1974) *The distance from a given point to the nearest end of one member of a random process of linear segments*. In Harding and Kendall (Eds.), Stochastic Geometry, pp.192-201. Wiley, London.

Cowan R. (1978) “The use of ergodic theorems in random geometry”, *Suppl. Adv. Appl. Prob.*, 10, 47-57.

Cowan R. (1979) “Homogeneous line-segment processes”, *Math. Proc. Camb. Phil. Soc.*, 86, 481-489.

Cowan R. (1980) “ Properties of ergodic random mosaic processes”, *Math. Proc. Camb. Phil. Soc.*, 86, 481-489.

Parker P. and Cowan R. (1976) “Some properties of line segment processes”, *J. Appl. Prob.*, 13, 96-107.

Wiener N. (1939) “ The ergodic theorem”, *Math. Nachr.*, 97, 89-102.