Point processes - tutorial

Rafał Kulik

Department of Mathematics and Statistics
University of Ottawa
and
Mathematical Institute
University of Wrocław

Workshop on Forest Fires and Point Processes

May 24-28, 2005
The Fields Institute

1 Examples and informal definitions

A **single jump point process** consists of a single point whose location can be defined by a positive random variable Y with distribution F. Let $N(B)$ be the number of points in a Borel set B. Then

$$N(B) = \mathbb{I}\{Y \in B\}\,.$$

More generally, a point process defined by a finite collection of random variables $\{Y_1, \ldots, Y_k\}$ is given by

$$N(B) = \sum_{i=1}^{k} \mathbb{I}\{Y_i \in B\}\,.$$

If Y is a single uniform random variable on $[0,1]$ then a **lattice point process** is defined as
$N(B)=\sum_{i=1}^{\infty}\mathbb{I}\{Y_i\in B\}$, where $Y_1:=Y$,
$Y_{n+1}=Y_n+1$, $n\geq 1$.

1.1 Poisson process on $\mathbb{R}_+$

Let N_t be the number of points in the time interval $(0,t]$. If for all $k\geq 1$ and any increasing sequence $0=:t_0<t_1<\cdots<t_k$, the random variables

$$N_{t_1}, N_{t_2}-N_{t_1}, \ldots, N_{t_k}-N_{t_{k-1}}$$

are independent and Poisson distributed with parameters $\lambda(t_i-t_{i-1})$, $i=1,\ldots,k$, respectively, for a known constant $\lambda>0$, then this point process is called a **(homogeneous) Poisson process**.

1.2 Cox process

Assume that $(\lambda_t, t \geq 0)$ is a nonnegative and locally integrable stochastic process (i.e. $\int_0^t \lambda_u du < \infty$ for every $t \geq 0$). Let $\Lambda_t = \int_0^t \lambda_u du$. Let $0 < t_1 < \cdots < t_k$, $k \in \mathbb{N}$. Now $(N_t, t \geq 0)$ is a **doubly stochastic Poisson process** (or **Cox process**) if conditionally on $\{\Lambda_{t_1} = a_1, \Lambda_{t_2} - \Lambda_{t_1} = a_2, \ldots, \Lambda_{t_k} - \Lambda_{t_{k-1}} = a_k\}$ the random variables $N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_k} - N_{t_{k-1}}$ are independent and Poisson distributed with parameters $a_1, \ldots, a_k$, respectively.

Special cases:

- If (λ_t) is deterministic and non-constant then we have a **inhomogeneous Poisson process**;
- If $\lambda_t \equiv \lambda$ then we have a homogeneous Poisson process.

2 Basic properties

2.1 Stochastic process

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For a given stochastic process $(X_t, t \geq 0)$ we define the **internal history** $\mathcal{F}_t^X := \sigma(X_s, 0 \leq s \leq t)$.

Example 2.1 • If N is the single jump point process then $\mathcal{F}_t^N = \sigma(\{Y \leq s\}, s \leq t)$;

Let $\{\tau_n\}$ be an increasing sequence of random variables such that $\tau_0 \leq 0$ and $\tau_i > 0$ for all $i \geq 1$. A **counting process** (or **point process**) is defined via

$$N_t = \sum_{n=1}^{\infty} \mathbb{I}\{\tau_n \leq t\}. \tag{1}$$

Note that such a process $(N_t, t \geq 0)$ is a random element of the space $D[0, \infty)$ (the space of all real-valued functions which are right-continuous and have left-hand limits). If the sequence $\{\tau_n\}$ is strictly increasing then the point process is **simple** and $\tau_1, \tau_2, \ldots$ are called the jump times of the process.

2.2 Point process as a random measure

Let

$$N := \sum_n \delta_{\tau_n} , \tag{2}$$

where $\delta.$ is a Dirac measure,

$$\delta_x(B) =$$

N is an integer-valued locally bounded (i.e. $N(B) < \infty$ for bounded B) **random measure**, i.e. for each Borel set B, $N(B)$ is a random variable and for each ω, $N(\omega, \cdot)$ is a measure. The space of all locally bounded integer valued measures will be denoted by $\mathcal{N}$. Therefore, a point process can be viewed as a random element with values in $\mathcal{N}$.

Example 2.2 If N is the Poisson point process then $N = \sum_{n=1}^{\infty} \delta_{\tau_n}$; $\tau_n = \sum_{j=1}^{n} X_j$, where $\{X_n\}$ is the sequence of i.i.d. exponential random variables.

3 Convergence of point processes

3.1 Convergence in the vague topology

Let $\mu^{(n)}, \mu \in \mathcal{N}$. We say that a sequence $(\mu^{(n)})$ converges to μ in the **vague topology** ($\mu^{(n)} \xrightarrow{v} \mu$) if

$$\int f d\mu^{(n)} \to \int f d\mu \tag{3}$$

for all continuous functions $f : \mathbb{R}_+ \to \mathbb{R}$ with compact support.

We say that a sequence $(N^{(n)})$ of point processes converges **weakly** to a point process N ($N^{(n)} \xrightarrow{w} N$) if

$$\int f dP^{(n)} \to \int f dP$$

for all continuous bounded functions $f : \mathcal{B}(\mathcal{N}) \to \mathbb{R}$. Here, $P^{(n)}, P$ are the distributions of the corresponding point processes, i.e. $P(A) := P(N \in A)$, where $A \in \mathcal{B}(\mathcal{N})$.

Note that the continuous mapping theorem states that if g is continuous and $N^{(n)} \xrightarrow{w} N$ then $g(N^{(n)}) \xrightarrow{w} g(N)$.

Take $g : \mathcal{N} \to \mathbb{N}$ as $g(\mu) = \mu(B)$ for bounded $B \in \mathcal{B}(\mathbb{R}_+)$. This map is continuous, i.e. $\mu^{(n)} \xrightarrow{v} \mu$ implies $\mu^{(n)}(B) \to \mu(B)$, if $\mu(\partial B) = 0$. So, by the preceding mapping theorem $N^{(n)}(B) \xrightarrow{w} N(B)$.

Let's summarize the above types of convergence:

- $N^{(n)} \xrightarrow{w} N$ is the convergence of *random measures*;
- $N^{(n)}(B) \xrightarrow{w} N(B)$ is the convergence of *integer-valued random variables*.

Weak convergence in the vague topology is equivalent to convergence of the finite dimensional distributions on *continuity sets*:

Theorem 3.1 $N^{(n)} \xrightarrow{w} N$ *if and only if*

$$(N^{(n)}(B_1), \ldots, N^{(n)}(B_k)) \xrightarrow{w} (N(B_1), \ldots, N(B_k)) \quad (4)$$

for all $k \geq 1$ *and all bounded Borel sets such that* $P(N(\partial B_i) = 0) = 1$, $i = 1, \ldots, k$.

Counterexample 3.2 The assumption on the boundary ∂B_i in Theorem 3.1 is important. In order to see this consider $N^{(n)} := \delta_{1+\frac{1}{n}}$, $N := \delta_1$. For $B = [\frac{1}{2}, 1]$, $N(B) = 1$, $N^{(n)}(B) = 0$ for all $n \geq 1$. Hence the $0-1$-valued random variables $N^{(n)}(B)$ cannot converge to $N(B)$, but $N^{(n)} \xrightarrow{w} N$.

On the other hand, weak convergence is equivalent to convergence of jump points, i.e. $N^{(n)} \xrightarrow{w} N$ if and only if $(\tau_1^{(n)}, \ldots, \tau_k^{(n)}) \xrightarrow{w} (\tau_1, \ldots, \tau_k)$ for $k \geq 1$.

4 Palm measure

Example 4.1 Assume that in Toronto buses run exactly on schedule every 1 minute. If we arrive just after a bus departure then we have to wait exactly 1 minute for the next one. However, if we arrive at a random time then we have to wait a random time with a uniform distribution on $(0, 1)$.

This phenomenon can be explained using **Palm measures**.

We shall say that a random measure (point process) N is **time stationary** if the joint distribution of

$$(N(B_1 + t), \ldots, N(B_k + t)) \tag{5}$$

is independent of $t \in \mathbb{R}$, for all natural numbers k and all Borel sets B_j, $j = 1, \ldots, k$. Here,
$B + t = \{b + t : b \in B\}$.

We denote by $\lambda := \mathbb{E}[N((0,1])]$ and we call λ the **intensity** of the stationary point process N.

A point process is **event stationary** if its sequence of interpoint distances $\{\tau_{n+1} - \tau_n\}_{n\geq 1}$ is stationary.

Example 4.2 Assume that $N = \sum_{i=-\infty}^{\infty} \delta_i$. It is event stationary but not time stationary. Its *time stationary version* is the lattice point process.

Assume that a point process N is time stationary and has finite intensity $\lambda > 0$. Define a probability measure $\mathbb{P}_N^0$ as

$$\mathbb{P}_N^0(\cdot) = \lim_{h\to 0} \mathbb{P}(\cdot \mid N((-h,0]) > 0)\,.$$

Note that $\mathbb{P}_N^0(N(\{0\}) = 1) = 1$.

If a point process is time stationary under $\mathbb{P}$ then it is event stationary under $\mathbb{P}_N^0$.

Palm-Khinchin equation:

$$\mathbb{P}(N((0,t]) > j) = \lambda \int_0^t \mathbb{P}_N^0(N((0,u]) = j)du\,.$$

A typical application of the Palm-Khinchin equation is to finding a formula for the distribution of a **forward recurrence time**, which is defined as

$$\eta_u \equiv \inf\{t > 0 : N((u, u+t]) > 0\}\,.$$

$\mathbb{P}$

$\mathbb{P}_N^0$

Note that for the stationary point processes we have $\eta_u \stackrel{\mathrm{d}}{=} \eta_0$ for all $u \geq 0$.

From the Palm-Khinchin equation we have

$$\begin{aligned}
\mathbb{P}(\eta_0 > x) &= \mathbb{P}(\tau_1 > x) \\
&= \mathbb{P}(N((0,x]) = 0) \\
&= 1 - \mathbb{P}(N((0,x]) > 0) \\
&= \lambda \int_x^\infty \mathbb{P}_N^0(N((0,u]) = 0)du \\
&= \lambda \int_x^\infty \mathbb{P}_N^0(\tau_1 > u)du
\end{aligned}$$

where the last equality follows from $\mathbb{P}_N^0(\tau_0 = 0) = 1$. Note that as a by-product we obtain that the distributions of η_0 $(= \tau_1)$ under $\mathbb{P}$ and $\mathbb{P}_N^0$ do not coincide.

Under $\mathbb{P}$ and $\mathbb{P}_N^0$ a Poisson process has the same distribution. This explains the following **waiting time paradox**. Assume that in Toronto buses run according to a Poisson process with rate 1. If we come at a random time then our expected waiting time is 1 although the expected interval between buses is one as well.

Little's Formula for queueing systems:

Consider a 1-server FIFO queue, i.e. customers arrive according to a point process $\{\tau_n\}$ and then they are served immediately if the server is empty or they join the queue and wait to be served in the order of arrival. Assume that a customer coming at time instant τ_n requires a service time S_n. Then we may define *sojourn time* sequence $\{D_n\}$, which is (under some assumptions) stationary. Let $(L_t, t \geq 0)$ be the queue length process, i.e.

$$L_0 = \sum_n \mathbb{I}\{\tau_n < 0 < \tau_n + D_n\}\,.$$

Then (**Little's Formula**):

$$\mathbb{E}[L_0] = \lambda \mathbb{E}_N^0[D_0]\,.$$

This connects the *mean queue length* at the time instant of an arrival of a *virtual customer* (i.e. a customer which would come at time t) and a *mean sojourn time* for a customer calculated *just after his arrival.* Note that in general $\mathbb{E}[L_0] \neq \mathbb{E}_N^0[L_0]$.

5 Compensator

Using an infinitesimal notation
$dN_t = \lim_{dt\to 0}[N_{t+dt} - N_t]/(dt)$, and

$$\mathbb{E}[dN_t \mid \mathcal{F}_t] = P(dN_t = 1 \mid \mathcal{F}_t) =: \lambda_t dt\,,$$

which for a simple point process gives a stochastic intensity as the conditional probability of having a point in the small time interval $(t, t+dt]$ given the history $\mathcal{F}_t$. A compensator: $\Lambda_t = \int_0^t \lambda_s ds$.

5.1 Single jump point process

Note that if up to time t the point has not occurred then the conditional probability of having a point in a small time interval $[t, t+dt)$ is just
$F(dt)/\mathbb{P}(\tau \geq t) = F(dt)/(1-F(t))$. On the other hand, if the point has occurred before t then this probability is equal to 0. Hence, the stochastic intensity should be

$$\lambda_t dt = \frac{F(dt)}{1-F(t)}\mathbb{I}\{Y \geq t\}\,. \tag{6}$$

5.2 Poisson process

By the definition of the Poisson process we have

$$\mathbb{E}[N_t - N_s \mid \mathcal{F}_s] = \mathbb{E}[N_t - N_s] = \lambda(t-s)\,.$$

Hence, the compensator has the form $\Lambda_t = \lambda t$.

Watanabe's characterization: A simple point process with deterministic compensator must be (inhomogeneous) Poisson.

Theorem 5.1 *Assume that $(N^{(n)})$ is a sequence of point processes with the corresponding histories $(\mathcal{F}_t^{(n)})$ and with the corresponding compensators $\Lambda^{(n)}$. If for each $t \geq 0$,*

$$\Lambda_t^{(n)} \xrightarrow{w} \Lambda_t\,,$$

where (Λ_t) is a continuous deterministic function, then $N^{(n)} \xrightarrow{w} N$, where N is a Poisson process with compensator (Λ_t).

6 Applications

1. **Inhomogeneous Poisson processes.**

   Assume that λ_t is a deterministic function and that for a $\lambda \in \mathbb{R}_+$, $\lambda_t < \lambda$. Simulate points of a homogeneous Poisson process with intensity λ. Given a jump point τ_n of the homogeneous process, accept it with the probability λ_{τ_n}/λ. Indeed, simulate independently at each jump point r.v's $U_n \sim U[0,1]$. If $U_n < \lambda_{\tau_n}/\lambda$ then accept the jump point and retain it. The point process defined by the retained points is inhomogeneous Poisson with intensity λ_t.

2. **Simulation of Cox processes.**

   We can use Watanabe's characterization to simulate a Cox process as follows. Simulate first a path of (λ_t). Then, given the path $(\lambda_t(\omega), t \in [0,T])$ we can simulate points of a Cox process as in the previous case.

3. **Radial simulation of Poisson process** on $B_0(r)$ - a ball located at the origin with radius r. Simulate a Poisson process on $[-r, r]$ and then independently simulate random variables U_n, $n = 1, \ldots, n_0$ from a uniform distribution on $\{x \in \mathbb{R}^d : ||x|| = 1\}$. Locations of the points are given by $\pi_1 U_1, \ldots, \pi_{n_0} U_{n_0}$.

4. **Poisson limit for empirical processes.** Consider a sequence $\{Y_n\}$ of i.i.d. non-negative random variables with distribution F and assume that $F'(0) \neq 0$ exists. Define $N^{(n)} = \sum_{i=1}^{n} \delta_{nY_i}$. Then the stochastic intensity of $N^{(n)}$ is given by (compare to the single jump point process)

$$\lambda_t^{(n)} = \sum_{i=1}^{n} \mathbb{I}\{nY_i \geq t\} \frac{F(dt/n)}{1 - F(t/n)} .$$

Note that this is the stochastic intensity w.r.t. to a larger filtration than $\mathcal{F}_t^N$, i.e. w.r.t. $\bigvee_{i=1}^{k} \sigma(\{Y_i \leq s\}, s \leq t)$. It is possible to show that $\Lambda_t^{(n)} \xrightarrow{w} F'(0)t$ for each $t \geq 0$ and hence a properly scaled sequence of empirical processes converges to a Poisson process.

5. **Statistics for spatial point processes.** Let N be a stationary point process indexed by $\mathbb{R}^d$.
   - **Empty space function** F - distance from the origin to the nearest point of N, i.e.

     $$F(r) := \mathbb{P}(N(B_0(r)) \neq 0)\,;$$

   - **Nearest neighbour function** G

     $$G(r) := \frac{1}{\lambda|A|}\mathbb{E}\left[\sum_{n:\tau_n\in A} \mathbb{I}\{N_{\tau_n}(B_{\tau_n}(r)) \neq 0\}\right]\,,$$

     where $A \subseteq \mathbb{R}^d$ and N_{τ_n} is the point process obtained from N by removing the point τ_n.

   Consider $d = 1$. Then $F(r) = \mathbb{P}(\tau_1 \leq r)$ is the distribution of the first point under the stationary measure $\mathbb{P}$ and G is the distribution of the first point under the Palm measure.

   The functions F and G have practical interpretations for $d \geq 2$. The inequality $F(r) < G(r)$ implies *clustering* of points, whereas $F(r) > G(r)$ denotes a *regularity* in the point pattern.

7 Long Range (count) Dependence

A stationary point process N is LRcd if

$$\limsup_{t\to\infty} \frac{\mathbf{Var}[N(t)]}{\mathbb{E}[N(t)]} = +\infty \,.$$

A renewal process N with interpoint distances $\{X_n, n \geq 1\}$ is LRcD if and only if $\mathbb{E}_N^0[X^2] = +\infty$ (Daley, 1997). Thus, intuitively, being LRcD means clustering of points. Daley's conjectured: any stationary, ergodic point process has such the property. This s not true.

Any stationary point process with $\mathbb{E}_N^0[X^2] = +\infty$ and some positive dependence between intervals is LRcD.

How to characterize LRcD for point processes on $\mathbb{R}^d$?