II. Point processes

Consider processes defined by point like events on the real time axis:

Examples:
- particle counters, photon statistics
- electrons in a mesoscopic device
- traffic counters on a highway
- earthquakes
- neural spikes
- human activity

1° General formalism

a) State space and probability measure

The state space of the process is

$$\{ s, (\tau_1, \dots, \tau_s) : s = 0, 1, 2 \dots, \tau_i \in \mathbb{R}, \; -\infty < \tau_1 < \tau_2 < \dots < \tau_s < \infty \}$$

number of events

Probability measure is defined by function

$\tilde{Q}_s(\tau_1, .., \tau_s)$, $s = 0, 1, 2, ..$ with

$\tilde{Q}_s(\tau_1 .., \tau_s) = 0$ unless $\tau_1 < \tau_2 < .. < \tau_s$.

Normalization:

$$\tilde{Q}_0 + \int_{-\infty}^{\infty} d\tau_1 \tilde{Q}_1(\tau_1) + \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \, \tilde{Q}_2(\tau_1, \tau_2) + .. = 1$$

Simplification: We extend the domain of $\tilde{Q}_s$ to $\mathbb{R}^s$ and identify all sequences $\{\tau_1, .., \tau_s\}$ that are related by permutations

$$\Rightarrow \quad Q_0 + \sum_{s=1}^{\infty} \frac{1}{s!} \int_{-\infty}^{\infty} d\tau_1 . \int_{-\infty}^{\infty} d\tau_s \, Q_s(\tau_1, .., \tau_s) = 1$$

b) Observables

An observable A is defined by a set of functions $A = \{ A_0, A_1(\tau_1), A_2(\tau_1, \tau_2) .. \}$ such that the expectation of A is

$$\langle A \rangle = A_0 Q_0 + \int_{-\infty}^{\infty} d\tau_1 \, A_1(\tau_1) Q_1(\tau_1) +$$

$$+ \frac{1}{2} \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \, A_1(\tau_1, \tau_2) Q_2(\tau_1, \tau_2)$$

$$+ \ldots$$

Example: $N_{[a,b]}$ = # events in $[a,b]$

This can be defined by introducing the indicator function

$$\chi_{[a,b]}(t) = \begin{cases} 1 & t \in [a,b] \\ 0 & \text{else} \end{cases}$$

$$\Rightarrow N_{[a,b]} = \{ N_0, N_1(\tau_1), N_2(\tau_1, \tau_2) \dots \} =$$

$$= \{ 0, \chi_{[a,b]}(\tau_1), \chi_{[a,b]}(\tau_1) + \chi_{[a,b]}(\tau_2) \dots \}$$

$$\Rightarrow \langle N_{[a,b]} \rangle = \sum_{s=1}^{\infty} \frac{1}{s!} \int_{-\infty}^{\infty} d\tau_1 \dots \int_{-\infty}^{\infty} d\tau_s \, Q_s(\tau_1, \dots, \tau_s) \sum_{i=1}^{s} \chi_{[a,b]}$$

$$= \sum_{s=1}^{\infty} \frac{1}{(s-1)!} \int_a^b d\tau_1 \int_{-\infty}^{\infty} d\tau_2 \dots \int_{-\infty}^{\infty} d\tau_s \, Q_s(\tau_1, \dots, \tau_s)$$

due to the permutation symmetry of Q_s.

In particular, for $a \to -\infty$, $b \to \infty$ the probability to see exactly s events is

$$P(s) = \frac{1}{s!} \int_{-\infty}^{\infty} d\tau_1 \dots \int_{-\infty}^{\infty} d\tau_s \, Q_s(\tau_1, \dots, \tau_s).$$

c) n-point densities

The n-point densities $f_n(t_1, \ldots, t_n)$ are defined by

$$f_n(t_1, \ldots, t_n)\, dt_1 \ldots dt_n = \mathrm{Prob}[\text{one event each in } (t_1, t_1 + dt), (t_2, t_2 + dt_2), \ldots, (t_n, t_n + dt_n)]$$

$$\Rightarrow f_1(t_1) = Q_1(t_1) + \int_{-\infty}^{\infty} d\tau_2\, Q_2(t_1, \tau_2) +$$

$$+ \frac{1}{2} \int_{-\infty}^{\infty} d\tau_2\, d\tau_3\, Q_3(t_1, \tau_2, \tau_3) + \ldots$$

account for permutations of τ_2, τ_3

$$\Rightarrow f_n(t_1, \ldots, t_n) = \sum_{s=n}^{\infty} \frac{1}{(s-n)!} \int_{-\infty}^{\infty} d\tau_{n+1} \ldots \int_{-\infty}^{\infty} d\tau_s \; \times$$

$$\times\, Q_s(\underbrace{t_1, \ldots, t_n}_{\text{fixed}}, \tau_{n+1}, \ldots, \tau_s)$$

d) Stationarity

A point process is called stationary, if its statistics is translationally invariant in time, i.e.

$$f_n(t_1+\tau, t_2+\tau, .., t_n+\tau) = f_n(t_1, .., t_n)$$

for all n, τ.

For stationary processes the event densities Q_s cannot be defined, because they are not normalizable.

2° Independent events

a) Poisson distribution

For independent events we have

$$\left.\begin{aligned} Q_s(\tau_1, .., \tau_s) &= C\, q(\tau_1)\, q(\tau_2) .. q(\tau_s) \\ Q_0 &= C \end{aligned}\right\}$$

$$\Rightarrow C \sum_{s=0}^{\infty} \frac{1}{s!} \underbrace{\int_{-\infty}^{\infty} d\tau_1 .. \int_{-\infty}^{\infty} d\tau_s \, q(\tau_1) .. q(\tau_s)}_{\left(\int_{-\infty}^{\infty} d\tau\, q(\tau)\right)^s =: \nu^s} = C e^{\nu} = 1$$

$$\Rightarrow \quad C = e^{-\nu}.$$

We consider the number of events $N = N_{[a,b]}$ and compute its generating function:

$$\langle e^{ikN} \rangle = \sum_{s=0}^{\infty} \frac{1}{s!} \int_{-\infty}^{\infty} d\tau_1 \ldots \int_{-\infty}^{\infty} d\tau_s \; e^{ik \sum_{i=1}^{s} \chi_{[a,b]}(\tau_i)} \times$$

$$\times e^{-\nu} \prod_{i=1}^{s} q(\tau_i) =$$

$$= e^{-\nu} \sum_{s=0}^{\infty} \frac{1}{s!} \left(\int_{-\infty}^{\infty} d\tau \; e^{ik\chi_{[a,b]}(\tau)} q(\tau) \right)^s =$$

$$= \exp\left(\int_{-\infty}^{\infty} d\tau \; \underbrace{\left(e^{ik\chi_{[a,b]}(\tau)} - 1 \right)}_{= \begin{cases} e^{ik} - 1 & \tau \in [a,b] \\ 0 & \text{else} \end{cases}} q(\tau) \right) =$$

$$= \exp\left(\left(e^{ik} - 1 \right) \underbrace{\int_a^b d\tau \; q(\tau)}_{\langle N \rangle} \right) = \exp\left(\langle N \rangle \left(e^{ik} - 1 \right) \right) =$$

as can be checked

$$= e^{-\langle N\rangle} \sum_{N=0}^{\infty} \frac{1}{N!}\left(\langle N\rangle e^{ik}\right)^N = \sum_{N=0}^{\infty} P_N\, e^{ikN}$$

with the Poisson distribution $P_N = e^{-\langle N\rangle} \frac{\langle N\rangle^N}{N!}$

This argument can be extended to sets of independent points in $\mathbb{R}^d$, characterized by probability densities

$$\left.\begin{aligned} Q_s(\vec{r}_1, \dots, \vec{r}_s) &= e^{-\nu} \prod_{i=1}^{s} q(\vec{r}_i), \quad \vec{r}_i \in \mathbb{R}^d \\ \nu &= \int_{\mathbb{R}^d} d^d\vec{r}\; q(\vec{r}) \end{aligned}\right\}$$

Then the number of points in an arbitrary volume is Poisson distributed.

b) Shot noise and Poisson process

If events are independent & the process is stationary, the 1-point density $f_1(t) = f_1(t+\tau) \equiv \rho$ which fully characterizes the process

The number of events in a interval of length l has distribution $P_N = e^{-\rho l} \frac{(\rho l)^N}{N!}$

This process is called shot noise (Schrotrauschen)

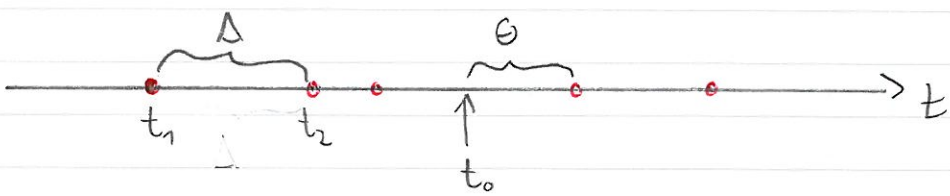
The Poisson process is the counting process of shot noise, defined by

$N(t)$ = number of events in $[0, t] = N_{[0,t]}$

$$\left.\begin{array}{ll} \text{with} \quad \langle N(t) - N(s) \rangle = \rho(t-s) & t > s \\ \mathrm{Var}(N(t) - N(s)) = \rho(t-s) & \end{array}\right\}$$

3° Waiting times and renewal theory

Consider a general point process:



Two questions are of interest:

Q1: What is the distribution of waiting times Θ for the next event after t_0?

Q2: What is the distribution of intervals Δ between subsequent events?

a) General expressions

For a general point process define

$$P_0(a, b) = \text{Prob}[\text{no event in } [a, b]]$$

which contains all information about waiting times.

Question (i): Density $f_\Theta(t|t_0)$ is

$$f_\Theta(t|t_0)\,dt = P_0(t_0, t_0+t) - P_0(t_0, t_0+t+dt)$$

$$= -\frac{d}{dt} P_0(t_0, t+t)\,dt$$

$$\Rightarrow f_\Theta(t|t_0) = -\frac{d}{dt} P_0(t_0, t_0+t)$$

Question (ii): Here both endpoints of the interval have to be specified, which leads to [see van Kampen]

$$f_\Delta(t|t_1) = -\frac{1}{f_1(t_1)} \left. \frac{\partial^2 P_0(t_1, t_2)}{\partial t_1 \, \partial t_2} \right|_{t_2 = t_1 + t}$$

For a stationary process $P_0(t_1, t_2) = P_0(t_2 - t_1)$ and

$$f_\Theta(t) = -\frac{dP_0}{dt}, \quad f_\Delta(t) = \frac{1}{\rho} \frac{d^2 P_0}{dt^2} \qquad (*)$$

---

b) The waiting time paradox (Zernike, 1929)

We want to know if there is a general relation between $\langle \Delta \rangle$ and $\langle \Theta \rangle$, and argue as follows:

- A randomly chosen time t_0 is on average halfway between two events
- Thus we expect $\langle \Theta \rangle = \frac{1}{2} \langle \Delta \rangle$

In view of the general relation $(*)$ this is not likely to be true.

Consider for concreteness the simple case of shot noise: $P_0(t) = e^{-\rho t}$

$$\Rightarrow \quad f_\Theta(t) = \rho\, e^{-\rho t}$$

$$f_\Delta(t) = -\frac{1}{\rho}\frac{d}{dt} f_\Theta(t) = \rho\, e^{-\rho t} = f_\Delta(t)$$

$$\Rightarrow \quad \langle \Theta \rangle = \langle \Delta \rangle = \frac{1}{\rho}$$

What is wrong with the Zernike argument? A randomly chosen time is more likely to lie in a longer interval.

Remarks: (i) The requirement that $f_\Theta = f_\Delta$ specifies the exponential distribution uniquely:

$$f_\Delta = -\frac{1}{\rho}\frac{df_\Theta}{dt} \overset{!}{=} f_\Theta(t) \quad \Rightarrow \quad f_\Theta = \rho\, e^{-\rho t}.$$

(ii) When the waiting time distribution decays more slowly than exponential, in general

$$\langle \Theta \rangle > \langle \Delta \rangle.$$

Example: $f_\Delta(t) = a(1+t)^{-(a+1)}$, $a > 1$

$$\Rightarrow \quad \rho = \frac{1}{\langle \Delta \rangle} = a-1 \quad \text{stationary process}$$

$$\Rightarrow \quad f_\theta(t) = (a-1)(1+t)^{-a}, \quad P_0(t) = (1+t)^{-(a-1)}$$

$$\Rightarrow \quad \langle \theta \rangle = \begin{cases} \frac{1}{a-2} > \frac{1}{a-1} = \langle \Delta \rangle & \text{if } a > 2 \\ \infty & \text{if } 1 < a < 2 \end{cases}$$

Example: statistics of classical light

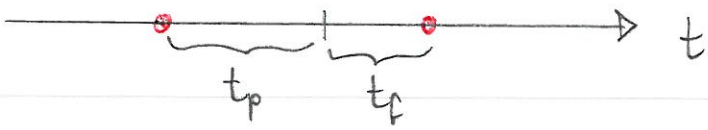
---

c) Renewal processes and residual waiting times

Definition: A renewal process is a point process in which the time intervals between events are non-negative i.i.d. RV's Δ_i drawn from a pdf $f_\Delta(t)$.

Example: Shot noise, $f_\Delta(t) = \rho e^{-\rho t}$.

Q3: We ask for the residual waiting time t_f, given that time t_p has passed since the previous event:



The naive expectation is that t_f decreases with increasing t_p, but this is not generally true

(Sornette, Knopoff, 1997)

Quantitative analysis:

$$\mathrm{Prob}[\, t_f > t \mid t_p \,] = \frac{1 - F_\Delta(t_p + t)}{1 - F_\Delta(t_p)} = P_{t_p}(t) ,$$

the excess probability of Sect. I. 4° e). In the case of independent events we have $1 - F_\Delta(t) = e^{-\rho t}$

$$\Rightarrow \quad P_{t_p}(t_f) = e^{-\rho t_f} = 1 - F_\Delta(t_f)$$

and the knowledge of t_p has no effect.

In the general case we use the peaks-over-threshold theorem for $t_p \to \infty$:

$$P_{t_p}(t) \longrightarrow H_\gamma(t/a(t_p)) , \quad H_\gamma(x) = (1 + \gamma x)^{-1/\gamma}$$

The corresponding density is

$$p_{t_p}(t) = -\frac{d}{dt} P_{t_p}(t) = -\frac{1}{a(t_p)} H_\gamma'(t/a(t_p))$$

$$\Rightarrow \langle t_f \rangle_{t_p} = \int dt\, t\, p_{t_p}(t) = -a(t_p) \int dx\, x\, H_\gamma'(x)$$

$$= \frac{1}{1-\gamma}\, a(t_p)$$

Examples: (i) Fréchet / Pareto $(\gamma > 0)$; $a(t) = \gamma t$

$$\Rightarrow \langle t_f \rangle = \frac{\gamma}{1-\gamma}\, t_p = \frac{1}{\alpha - 1}\, t_p \,, \quad \alpha = 1/\gamma$$

For $\alpha < 2$ we have $\langle t_f \rangle > t_p$.

(ii) Gauss: $a(t) = \sigma^2/t \Rightarrow \langle t_f \rangle = \sigma^2/t_p$

In general, t_f increases / decreases with t_p for distributions that are heavier / lighter than exponential.

Recurrence time distribution for earthquakes is heavy-tailed, but the process is not renewal.

(Corral 2005, Livina et al 2005)